IDENTIFICATION OF VIRAL CONTAMINATION ON LETTUCE FROM THE FIELD TO POST-HARVEST PROCESSING

By

Samantha Lynn Wengert

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife - Master of Science

2015

ABSTRACT

IDENTIFICATION OF VIRAL CONTAMINATION ON LETTUCE FROM THE FIELD TO POST-HARVEST PROCESSING

By

Samantha Lynn Wengert

Viral foodborne outbreaks are a serious threat to public health and fresh produce is becoming increasingly recognized as a transmission vehicle. Potential pre- and post-harvest sources of contamination include irrigation and processing water, soil, manure, equipment, and human handling. Traditional detection methods limit studies on viruses in produce. New culture-independent metagenomic next generation sequencing (NGS) technologies present an opportunity for generating an improved understanding of the virus communities (virome) associated with foods. The goals of this study were to use NGS technology for the first time to identify the virome present in irrigation water and lettuce in the field and to investigate the efficacy of current post-harvest leafy green processing and disinfection practices during a contamination event. In this study, most viruses found in irrigation water and lettuce from the field environment could not be identified suggesting limited knowledge of the virome in these environments. Human enteric viruses such as rotavirus A and picobirnavirus were identified in field lettuce. On the processing side, the efficacy of a chlorine-based sanitizer against MS2 coliphage on fresh-cut romaine lettuce was assessed during simulated commercial production of fresh-cut lettuce. Flume washing lettuce in 25 ppm of free chlorine did not significantly reduce viral levels on romaine lettuce when compared to water without chlorine. Overall, this study suggests that metagenomic technology can be used as a potential tool for monitoring food safety. Viruses were present in field lettuce and resistant to current commercial chlorine disinfection techniques, posing a possible threat to public health.

ACKNOWLEDGEMENTS

This work was supported by a grant from the United States Department of Agriculture, National Institute for Food and Agriculture Grant # 11281370 Proposal 2012-04389. I would like to thank my main advisor, Dr. Joan B. Rose, for all her support and encouragement throughout my time as a graduate student at MSU. I would also like to thank Tiong G. Aw, Rebecca Ives, and other Rose Lab members for their help and support in carrying out this research project. Thank you to my committee members Dr. Mohamad Faisal and Dr. Elliot Ryser for their guidance, with special recognition to Dr. Elliot Ryser who provided the small-scale leafy green processor for our research project. I would also like to acknowledge the Research Technology Support Facility and the High Performance Computing Center at Michigan State University for sequencing and computational support as well as Channah Rock and Kurt Nolte (The University of Arizona) for helping with the Yuma field sampling. Finally, I would like to thank my family for all their love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

BLAST	Basic Local Alignment Search Tool
bp	Base Pair
CDC	Centers for Disease Control and Prevention
CFSAN	Center for Food Safety and Applied Nutrition
Contig	Contiguous Sequence
DNA	Deoxyribonucleic Acid
ds	Double-Stranded
eFORS	Electronic Foodborne Outbreak Reporting System
ELISA	Enzyme-Linked Immunosorbant Assay
E-Value	Expected Value
FDA	U.S. Food and Drug Administration
FDOSS	Foodborne Disease Outbreak Surveillance System
FOOD	Foodborne Outbreak Online Database
FoodNet	Foodborne Diseases Active Surveillance Network
HACCP	Hazard Analysis and Critical Control Points
HPAI	Highly Pathogenic Avian Influenza Virus
IDBA-UD	Iterative de Bruijn Graph Assembler
IFIC	International Food Information Foundation Council
kb	Kilobase
MEGAN	Metagenome Analyzer
MMWR	Morbidity and Mortality Weekly Report
MPN	Most Probable Number

NCBI	National Center for Biotechnology Information
NFPA	National Food Processors Association
NGS	Next-Generation Sequencing
NORS	National Outbreak Reporting System
PBS	Phosphate Buffered Saline
PBW	Phosphate Buffered Water
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PFU	Plaque Forming Unit
qPCR	Quantitative PCR
RefSeq	Viral Reference Sequence
RTE	Ready-to-Eat
RT-PCR	Reverse Transcriptase PCR
RNA	Ribonucleic Acid
ss	Single –Stranded
STEC	Shiga toxin- producing *Escherichia coli*
TSA	Trypticase Soy Agar
TSB	Tryptic Soy Broth
USDA	United States Department of Agriculture
WHO	World Health Organization

I. LITERATURE REVIEW

1. Burden of Viral Foodborne Illness Associated with Fresh Produce on Public Health

1.1 Foodborne Disease in the United States

Foodborne disease is a serious threat to public health and food safety worldwide. The World Health Organization (WHO) describes foodborne illness as "diseases, usually either infectious or toxic in nature, caused by agents that enter the body through the ingestion of food" (CDC, 2013a). Currently WHO estimates that diarrheal diseases many of which are attributed to contaminated food result in 2 million deaths annually (CDC, 2013a). In the United States, food safety monitoring efforts in the food supply chain have greatly increased since the early 1990's (Crutchfield & Roberts, 2000). With the implementation of Hazard Analysis and Critical Control Point (HACCP) programs and food safety initiatives including the National Food Safety Initiative of 1997 and the Food and Drug Administration (FDA) Food Safety Modernization Act of 2011, the United States is considered to have one of the world's safest food supplies (Crutchfield & Roberts, 2000, Oliver et al., 2009). However, despite recent advances in food monitoring, foodborne outbreaks remain a serious threat to public health and it has been suggested that consumer confidence in food safety has slowly begun to decline (IFIC, 2014). According to the Centers for Disease Control and Prevention (CDC), approximately 48 million people in the United States are expected to be affected by foodborne illnesses each year (CDC, 2014). Of the 48 million people affected, approximately 9.4 million (20%) cases of illness are caused by 31 known foodborne pathogens while a staggering 38.4 million (80%) are caused by unspecified agents (CDC, 2014). This remarkably high portion of unspecified agents suggests a current shortage of data collection from affected patients at local health agencies and ultimately a current knowledge gap concerning unknown or unidentified foodborne agents.

A foodborne outbreak occurs when a group of individuals consume contaminated food and two or more of them develop the same illness. In 2013 alone, an estimated 818 foodborne disease outbreaks resulted in 13,360 cases of illness, 1,062 hospitalizations, and 16 deaths in the United States (CDC, 2015b). Consequently, foodborne illness is estimated to be costly for the health care system. The approximate annual cost due to foodborne illness from pathogenic bacteria, parasites, and viruses combined in the U.S. ranges from $51.0 to $77.7 billion, varying with the U.S. Department of Agriculture (USDA) and U.S. Food and Drug Administration (FDA) cost-of-illness models utilized (Scharff, 2012).

Currently there are a large variety of bacterial, viral, and parasitic human pathogens associated with foodborne disease. The bacterial pathogens causing the highest number of reported foodborne outbreaks in 2013 were *Salmonella spp.,* Shiga toxin- producing *Escherichia coli* (STEC), *Clostridium perfringens*, *Campylobacter spp.*, and *Vibrio parahaemolyticus* accordingly (CDC, 2015b). Furthermore, viral foodborne pathogens were dominated by norovirus while *Cryptosporidium parvum*, *Giardia lambia* and *Trichinella spiralis* were the most commonly reported parasitic infections (CDC, 2015b). Despite recognition of these human pathogens in our food supply, foodborne infections continue to emerge and in some cases have risen in recent years (CDC, 2011). Viruses in particular are becoming increasingly recognized as foodborne pathogens, with an increasing number of outbreaks occurring between 1998 and 2008 (Gould et al., 2013). To date, numerous foodborne enteric viruses causing gastroenteritis in humans have been identified, however, there are still a variety of human viruses that are capable of replication within the intestinal tract and their role in our food systems is currently unknown. These data indicate the vulnerability of our food system to contamination and emphasize the

need to identify current knowledge gaps in food safety, particularly in relation to foodborne viruses.

1.2 Foodborne Enteric Viruses and Public Health

Human enteric viruses, which are commonly transmitted through the fecal-oral route, are a serious threat to public health and safety. Enteric viruses have high infectivity rates (10-100 particles result in high probability of infection) and many lack an envelope which allows for resistance to thermal inactivation, facilitating virus survival and maintaining infectivity in the environment (Fong & Lipp, 2005, Gibson & Schwab, 2011, Newell et al., 2010). These viruses are able to replicate inside the host's gastrointestinal tract and are shed at extremely high concentrations (10^5-10^{11} viral particles per gram of stool) into the environment through the feces or vomit of an infected individual (Fong & Lipp, 2005). Water and food can then become contaminated at both pre- and post-harvest levels of production.

Viruses are increasingly being recognized as water and foodborne etiological agents. There are hundreds of human pathogenic water and foodborne viruses of fecal origin including adenoviruses, astrovirus, norovirus (genotypes I and II), polioviruses, enteroviruses, hepatitis A and E virus, sapoviruses, reoviruses, and rotaviruses which can cause gastroenteritis, meningitis, liver disease, infantile diarrhea, respiratory illness, or neurological symptoms (Cook, 2013, Fong and Lipp, 2005, Koopmans et al. 2004). However, many symptoms associated with these human pathogens are generally mild or self-limiting and therefore many infections are underreported (Fong & Lipp, 2005, O'Brien, 2008, Seymour & Appleton, 2001). More severe cases that result in illness, hospitalization, and death are usually observed in children, the elderly, and immunocompromised individuals (Fong & Lipp, 2005, Rodríguez-Lázaro et al., 2012).

To date, the viruses most frequently associated with foodborne illness include human norovirus and hepatitis A virus (Koopmans & Duizer, 2004, Newell et al., 2010). Norovirus is a non-enveloped, non-culturable, single-stranded RNA virus in the *Calciviridae* family and is the leading cause of viral gastroenteritis worldwide (DiCaprio et al. 2004). This enteric virus has a short incubation period (12-48 hours) and is commonly associated with outbreaks on cruise ships, however, it is often difficult to confirm the specific food and water source of transmission (Isakbaeva et al., 2005, Kroneman et al., 2008). Hepatitis A is a non-enveloped, ssRNA virus of the *Picornaviridae* family that exhibits slow replication in culture and has a significantly longer incubation period (15-50 days) than norovirus. In the early 1990's the CDC listed hepatitis A as the sixth leading cause of foodborne disease in the United States, however introduction of the hepatitis A viral vaccine in 1995 led to a significant drop in hepatitis A infections (Seymour & Appleton, 2001).

In the United States, the CDC Morbidity and Mortality Weekly Report (MMWR) named viruses as the primary cause of foodborne outbreaks with known etiology in a survey of foodborne outbreaks from 1998 to 2008 (Gould et al., 2013). The CDC monitors and gathers data on foodborne outbreaks in the United States through the Foodborne Disease Outbreak Surveillance System (FDOSS), which relies on public health agency reporting through the electronic Foodborne Outbreak Reporting System (eFORS). In 2009, eFORS was replaced by the National Outbreak Reporting System (NORS) and data on foodborne outbreaks from 1998-2013 can now be viewed online through the Foodborne Outbreak Online Database (FOOD). Table 1 shows the updated viral foodborne outbreak data from 1998-2013. Data from 1998-2008 were obtained from Gould et al. (2013) and 2009-2013 data were added using viral outbreak data obtained in FOOD.

Table 1. Viral foodborne outbreaks and outbreak associated illnesses, hospitalizations, and deaths in the United States from 1998-2013*

	Outbreaks				Illnesses				Hospitalizations				Deaths			
			Total				Total				Total				Total	
Etiology**	CE	SE	#	%	CE	SE	#	%	CE	SE	#	%	CE	SE	#	%
Astrovirus	1	1	2	0	14	22	36	0	0	0	0	0	0	0	0	0
Hepatitis A	85	1	86	2	2370	4	2374	2	363	0	363	22	8	0	8	35
Norovirus	2,786	1,936	4,722	96	92,339	34,629	126,968	96	967	300	1,267	76	7	0	7	30
Rotavirus	4	8	12	0	204	110	314	0	0	5	5	0	7	1	8	35
Other viral	8	96	104	2	510	2,568	3,078	2	7	18	25	2	0	0	0	0
Total	2,884	2,042	4,926	100	95,437	37,333	132,770	100	1,337	323	1,660	100	22	1	23	100

*Data obtained from the CDC Foodborne Outbreak Online Database
**Definitions: CE = confirmed etiology; SE = suspected etiology. Etiologies are confirmed using laboratory and clinical guidelines (CDC, 2015a). Those that do not meet guidelines are labeled suspected etiology.
***Reproduced from Gould et al. 2013

From 1998-2013 foodborne viruses caused 4,926 outbreaks, 132,770 cases of illness, 1,660 hospitalizations and 23 deaths. Norovirus was the most frequently reported foodborne viral pathogen, accounting for 96% of viral outbreaks, 96% of illnesses and 76% of hospitalizations. Hepatitis A (86), rotavirus (12), and astrovirus (2) outbreaks were also reported.

Interestingly, recent foodborne outbreaks also suggest that viruses are commonly associated with fresh produce contamination and transmission. Studies on norovirus in particular have shown strong associations with fresh produce consumption (Dicaprio et al., 2012, Gould et al., 2013, Widdowson et al., 2005). It has been suggested that norovirus is responsible for over 40% of the annual fresh produce outbreaks in the United States (DiCaprio et al. 2012, Seymour & Appleton, 2001). Furthermore, CDC data from 1998-2008 labeled norovirus and leafy greens as the pathogen-commodity pair most likely to be associated with a foodborne outbreak, causing a total of 4,011 illnesses (Gould et al. 2013). Hepatitis A has also been linked to multiple outbreaks in green onion, blueberries, and strawberries (Calder et al., 2003, Hutin et al., 1999, Wheeler et al., 2005). Although norovirus and hepatitis A have been frequently associated with fresh produce outbreaks, there is currently limited knowledge on the role of other viruses in our food systems.

Despite increased awareness of viruses as foodborne disease agents, the Foodborne Diseases Active Surveillance Network (FoodNet) which actively monitors trends in foodborne illnesses and assesses food safety initiative impacts in the U.S. continues to only monitor bacteria (*Campylobacter, Listeria, Salmonella,* STEC O157 and non-O157, *Shigella*, *Vibrio*, and *Yersinia)* and parasites (*Cryptosporidium, Cyclospora*). However in addition to reporting through FDOSS, the CDC has begun to actively monitor viral outbreaks, namely norovirus, through the CaliciNet (2009) and NoroSTAT (2012) surveillance systems (CDC, 2013b).

Although these surveillance systems have led to increased awareness of the role of viruses in the nation's food supply, there is still limited knowledge on where contamination occurs in the supply chain and other viral pathogens of concern.

1.3 Significance of Fresh Produce as a Food Commodity and Vehicle of Pathogen Transmission

Fresh produce is a food commodity of increasing public health interest. Fruits and vegetables compose two of the five basic food groups recommended by the USDA as components of a healthy diet. Studies on food consumption trends in the United States have shown increased per capita consumption of fruits and vegetables. According to an agricultural economist with the USDA, average annual per capita fruit and vegetable consumption per pound increased by 25% between 1977-1979 and 1997-1999 (Pollack, 2001). In another study, per capita consumption of fruit and vegetables in the United States was estimated to increase 19% and 29% respectively between 1980 and 2001 (Clemens, 2004). Suggested drivers of increased fresh produce consumption include increased production, product convenience, improved technology that maintains produce quality, greater availability, and consumer desire to maintain a healthy lifestyle (Pollack, 2001). For example, from 1970 to 2012, the average amount of fresh vegetables and fruits available for consumption increased by 67 and 6 pounds, respectively (USDA, 2014). In response to the changing fresh produce supply and demand, traditional agricultural and post-harvest practices have been altered. Practices such as cutting and coring at harvest, increased importation and transportation, and large scale production facilities are now employed to support changing consumer habits (Heaton et al. 2008, Lynch et al. 2009).

Outbreaks associated with fresh produce are becoming increasingly recognized. In addition to intensive production and processing practices, fresh produce is commonly consumed raw, making it an ideal vehicle for pathogen transmission. An analysis of the FDOSS found an increasing number of foodborne outbreaks associated with the consumption of raw produce in the U.S., rising from 0.7% in 1970 to 6% in the 1990's (Sivapalasingam et al, 2004). Of particular interest, the food items most commonly associated with fresh produce outbreaks were leafy greens such as lettuce and salads in addition to melon, sprouts, and berries (Sivapalasingam et al., 2004). In a more recent study, Painter et al. (2013) used FDOSS to summarize data on foodborne outbreaks, illnesses, and hospitalizations attributed to 17 mutually exclusive food commodities. According to this study, from 1998 to 2008 an estimated 46% of annual illnesses, 38% of annual hospitalizations, and 23% of annual deaths acquired in the United States were attributed to fresh produce; whereas meat and poultry contributed 22% and 29% of the illnesses and deaths (dairy and eggs; fish and shellfish contributed 20% and 15% and 6.1% and 6.4%, respectively) (Painter et al., 2013). Furthermore, this study found that leafy green vegetables were the food commodity responsible for the highest number of foodborne illnesses (2.2 million) (Painter et al., 2013). More recently, in 2013 leafy vegetables were implicated in 9 foodborne outbreaks which resulted in 207 cases of illness (CDC, 2015b). Despite increasing recognition of fresh produce as a vehicle for pathogen transmission, foodborne outbreaks attributed to fresh produce remain a public health and food safety concern.

2. Methods for Studying Viruses in the Environment

A summary of the current viral detection methods in food and water is provided in Table 2. Historically, cell culture has served as the "gold standard" method for viral detection and

discovery. Bacteriophage were among the first viruses to be replicated *in vitro* due to easy laboratory based manipulation of bacterial hosts.

Table 2. Current methods used in virus detection

Detection Method	Method description	Current examples	Advantages	Disadvantages
Cell culture	Viruses infect and replicate in host specific cells	Continuous culture lines from animal cells	Direct isolation of a variety of cultivable viruses to high titers	Many viruses uncultivable, requires specific cell line, costly, time consuming
In vitro bacterial host culture	Viruses infect and lyse bacterial host cells	Bacteriophage plaque assay	Direct bacteriophage isolation, purification, and enumeration	Requires specific bacterial host, issues in reproducibility (diluting, plaque size, incomplete lysis, plaque aggregation)
Electron microscopy	Microscope that uses an electron beam to illuminate and magnify viruses in detail	Transmission and scanning electron microscopes	Does not require prior knowledge of organism DNA, provides high resolution image	Poor detection limit, need high concentrations, maintenance, cumbersome, training, cannot identify virus
ELISA	Antigen-antibody pathogen detection	Indirect and sandwich ELISA	Quick and rapid detection, by-pass cell culture	Require a specific probe for detection, not applicable for virus discovery
PCR; quantitative PCR (qPCR)	Viral DNA or RNA amplification and enumeration of a known sequence	Reverse transcriptase PCR (RT-PCR), integrated cell culture PCR	Fast, high throughput, quantitative data, sensitivity, repeatability	Must know the sequence. Cost of equipment and reagents, reaction inhibition, data analysis, training.
Next generation sequencing (metagenomics)	High throughput viral DNA or RNA sequencing all genomes in an environmental sample	Pyrosequencing, sequencing-by-synthesis, sequencing-by-ligation	Fast, high throughput, cost of sequencing, reliable identification of microbial communities	Short sequence read length, time and training required for bioinformatics analysis

To date, the most common method of bacteriophage isolation, purification, and enumeration is the plaque assay, which dates back to its original discovery in 1917 (d'Herelle, 1917, Kutter & Sulakvelidze, 2004). In this method, bacterial host cells are exposed to a virus that upon infection lyses the surrounding cells, resulting in a clear zone in the agar medium called a plaque which represents a single infectious viral particle. Although this detection method has proven

beneficial in the field of virology, cell culture and *in vitro* virus replication has many disadvantages. A major limitation of vial studies in cell culture or *in vitro* is that this method only targets viruses capable of replicating in cells that can be propagated, all of which require a specific cell line for virus proliferation. Such viruses include adenovirus, enteroviruses (poliovirus, coxsackie viruses, echoviruses), influenza A and B, Measles virus, Mumps virus, rhinovirus, Ebola, SARS-coV, VZV and hMPV (Leland & Ginocchio, 2007). Many other viruses, including those important in foodborne disease, cannot replicate in cell culture and therefore require a different method of detection.

Electron microscopy is another traditional method used in virus detection. Electron microscopes (scanning and transmission) use a beam of electrons to illuminate and magnify viruses in great detail. In 1939, the first virus (tobacco mosaic virus) was visualized using electron microscopy and this technology has since then aided in the discovery of viruses such as smallpox and poliovirus (Goldsmith & Miller, 2009). Although prior knowledge of organism DNA is not necessary for detection, disadvantages include the need for high viral concentrations, poor detection limits, and the inability to identify the virus beyond the family level.

To date, the serological method most commonly used in virus research is the Enzyme-Linked Immunosorbant Assay (ELISA). In this technique, a viral antigen immobilized to a solid surface binds to a specific antibody which is either linked to an enzyme or can be detected by a secondary antibody linked to an enzyme. After adding an enzymatic substrate, a visible signal such as color change is produced and the antigen can be quantified. This technology has been regularly applied in plant virus detection (as early as 1976) as well as food authenticity in the food industry (Voller et al. 1976, Asensio, González et al., 2008). Although serological methods

bypass the need for cell culture and are time and cost effective, they require a specific probe for virus detection and are not applicable to virus discovery.

To date advanced filtration and molecular detection methods have greatly improved virus detection and monitoring in environmental samples, especially for enteric viruses in water systems. Specifically tangential-flow, hollow fiber ultrafiltration allows for virus concentration based on size exclusion (molecular weight cutoff) from large volumes of water (Gibson & Schwab, 2011, Liu et al., 2012, Smith & Hill, 2009). This technology has been readily applied to concentrate viruses from water systems (reclaimed and surface) and when combined with molecular detection techniques, has provided a better understanding of the microbial quality of water (Gibson & Schwab, 2011, Liu et al., 2012). Viruses are traditionally further concentrated by passing the remaining filtrate through a 0.22 µm filter (bacteria removal), polyethylene glycol (PEG) precipitation, and ultracentrifugation prior to molecular detection (Croci et al., 2008, Liu et al., 2012, Rosario et al., 2009). The following sequence- and culture- dependent molecular methods have been used for direct and rapid detection of viruses in the environment 1) polymerase chain reaction (PCR), reverse transcription PCR (RT-PCR), and integrated cell culture PCR, 2) quantitative PCR (qPCR), 3) Sanger sequencing, 4) and whole genome sequencing.

Polymerase chain reaction was first developed in 1980's by Kary Mullis (Bartlett & Stirling, 2003). Since its discovery, traditional endpoint PCR has been modified to better detect viruses in clinical and environmental samples. Many foodborne enteric viruses (norovirus, hepatitis A virus, astrovirus, rotavirus, enterovirus) are composed of RNA and direct detection requires reverse transcription. To date, RT- PCR remains a gold standard for enteric virus detection, especially in foods (Bidawid et al., 2000, Hyeon et al., 2011, Leggitt & Jaykus, 2000,

Love et al., 2008). Integrated cell culture PCR combines culture with molecular methods to detect viruses in the environment (Reynolds, 2004). This technique has been frequently applied to study enteric viruses in multiple water types including drinking water, river water, and sewage, but has also been combined with qPCR technology to detect enteric viruses (Hepatitis A) in fresh produce (Greening et al., 2002, Hyeon et al., 2011, Lee & Jeong, 2004). Unlike the conventional end-point PCR, qPCR (developed in the early 1990's) uses fluorescent technology to monitor and quantify targeted nucleic acids as they are amplified (Aw & Rose, 2012, Fraga 2014). To date, numerous studies have used RT-qPCR to detect enteric viruses in water sources (Aw & Rose, 2012). More recently, studies have begun to investigate methods of enteric virus recovery and RT-qPCR detection in fresh produce, including vegetables (lettuce, chicory, spinach, mixed salads, green onion, basil) and fruits (strawberries, raspberries, blueberries, blackberries) (Butot et al., 2007, Dubois et al., 2007, Sánchez et al., 2012).

In the past decade, the development of new genomic technologies has been exceedingly important for the discovery of novel viruses. More specifically, culture- and sequence-independent sequencing technologies, known as next-generation sequencing (NGS), allow for the examination of entire microbial communities in an environmental sample (metagenomics) and do not require previous knowledge of viral nucleic acid sequences. These new technologies and metagenomic techniques are now being used to study viruses in the environment to gain insights into the virus world.

3. Viruses in the Environment and Pre-Harvest Sources of Fresh Produce Contamination

3.1 Viral Types, Characteristics, and Role in Food Industry

Viruses are intracellular, infectious agents that are ubiquitous in nature and replicate within the cells of living organisms to cause a wide range of diseases. The size, shape, chemical

structure, and genome composition are all characteristics used to classify viruses. In the Baltimore classification system, viruses are grouped into families based on mode of replication and genome composition, which includes double-stranded (ds) and single-stranded (ss) DNA and RNA. In addition, viruses are host specific and are often identified by the hosts they infect. Examples of viral hosts include bacteria (bacteriophage), algae, plants, animals (invertebrate and vertebrate), and humans. The significance of these viral hosts in the environment and popular examples of each virus type are included in Table 3.

Viruses that infect bacteria, called bacteriophage or phage for short, are diverse and widely distributed in the environment (Breitbart & Rohwer, 2005). Bacteriophage are of interest to the food industry because they provide insight into the bacterial host populations in the environment and the host specificity of viruses can be used as potential indicators of fecal contamination (Aw et al., 2014, Leclerc et al., 2000). In addition, male-specific, non-enveloped, RNA bacteriophage such as MS2 are shown to have similar resistance and survival characteristics as enteric viruses (including norovirus) in water and fresh produce and can therefore act as a surrogate for foodborne viruses that cannot be propagated in cell culture (Dawson, 2005, Havelaar et al., 1993) Furthermore, studies have shown that specific lytic bacteriophage are a promising tool for reducing bacterial pathogens on fresh produce (Sharma, 2013). Viruses that infect algae, which are aquatic chlorophyll containing unicellular and multicellular organisms, are diverse and prevalent in aquatic ecosystems. These viruses can act as mortality agents to control algal host populations and are possible constituents of irrigation water (varying with source). Pathogenic viruses that infect and cause disease in staple crops are of primary concern to the food industry. Plant viruses that cause physical and chemical alterations to fruits and vegetables are responsible for major losses in crop productivity, yield,

quality, and ultimately are costly to the fresh produce industry. Invertebrates (insects, crustaceans, arthropods, bivalves, mollusks, etc.) compose the largest proportion of Earth's

Table 3. Viral host examples and their significance in the environment and food industry

Host	Significance	Viral examples	Reference(s)
Bacteria (bacteriophage)	• Control pathogenic bacteria in the food chain • Biogeochemical cycling • Antibiotic resistance • Potential indicators of fecal contamination	• Somatic Phage: T even phages (*Myoviridae*), λ phage (*Siphoviridae*), P22 (*Podoviridae*), phi X174 (*Microviridae*) • F^+ specific phage: MS2 (*Leviviridae*), CTX (*Inoviridae*)	(Ackermann, 2009, Colomer-Lluch et al., 2011, Davis et al., 2000, Leclerc et al., 2000, Rodríguez-Lázaro et al., 2012)
Algae	• Regulate fresh water and marine food webs • Biogeochemical cycling • Assist in algal bloom reduction and formation	• dsDNA viruses of the Phycodnaviridae family (*Chlorovirus*, *Coccolithovirus*, *Prasinovirus*, *Prymnesiovirus*, *Phaeovirus* and *Raphidovirus)*	(Baudoux & Brussaard, 2005, Brussaard, 2004, Wilson et al., 2006)
Plants	• Plant disease: leaves and fruit spotting, ringspots, discoloration, reduced vegetative output, poor growth • High economic costs for fresh produce food industry due to reduced crop quality, productivity, and yield	• Mosaic viruses: Cucumber, tobacco, tomato (*Cucumovirus*, *Tobamovirus*) • Tomato spotted wilt virus (*Tospovirus*) • Potato virus X and Y (*Potexvirus, Potyvirus*) • Plum pox potyvirus (*Potyvirus*)	(Mehle & Ravnikar, 2012, Rybicki, 2015)
Invertebrate animals	• Vector borne virus transmission to animals and humans • Biological control agents for management of agroecosystems, stored products, and forestry • High economic costs for food industry and aquaculture (loss of productivity)	• Vector borne viruses: Yellow fever, Dengue fever, West Nile, Japanese encephalitis (*Flavivirus*) • Viruses as biological control agents: dsDNA *Baculoviruses* • Pathogenic marine invertebrate viruses: *Baculoviruses*, *Iridoviruses*, *Reoviruses*, *Rhabdoviruses*	(Johnson, 1984, Lacey et al., 2001)
Vertebrate animals	• Viral zoonotic illnesses • High economic costs due to livestock productivity loss	• Livestock pathogens: Foot-and-mouth disease, bovine viral diarrhea, Newcastle disease • Emerging viruses: Swine hepatitis E, Nipah virus, SARS Coronavirus ,highly pathogenic avian influenza virus (HPAI-H5N1)	(Chi et al., 2002, FAO/WHO, 2008, Pimentel et al., 2001)

Table 3 (cont'd)

Humans	• Threat to public health and food safety • High economic costs to public health system	• Food and waterborne viruses: Norovirus and Sapovirus (*Caliciviridae*), enterovirus and hepatitis A (*Picornaviridae*), adenovirus (*Adenoviridiae*), and astrovirus (*Astroviridae*)	(Fong & Lipp, 2005, Seymour & Appleton)

species and therefore are infected by a wide variety of viruses, many of which also infect mammal, bird, and plant species. Invertebrate viruses may directly or indirectly impact numerous food commodities by infecting and causing disease in agricultural pests as well as food fish and shrimp species. For vertebrate animals, viruses infecting livestock that ultimately impact food production and safety are of primary importance. Although many animal viruses that cause disease in humans (zoonotic) are transmitted by direct contact rather than through a food vehicle, emerging viruses such as swine hepatitis E, Nipah virus, SARS Coronavirus and highly pathogenic avian influenza virus (HPAI-H5N1) are currently suspected of foodborne transmission (FAO/WHO 2008). This shows the increasing importance of animal viruses in our food systems and public health.

3.2. Irrigation Water as an Environmental and Pre-Harvest Source of Fresh Produce Contamination

Fresh produce, which is subject to intense production practices, has many opportunities for human pathogen contamination from farm-to-fork. The HACCP principles and guidelines were established in 1997 to help guide the food industry in identifying, evaluating, and ultimately controlling chemical, biological, and physical hazards in foods throughout the food supply chain (FDA, 2014). Recognizing that sources of contamination vary with hazard type, food commodity, and stage of production (pre- vs. post-harvest), the HACCP system provides the food industry with recommendations on how to control possible hazards and critical control points based on individual practices. For the fresh produce industry at the pre-harvest level,

irrigation water and runoff, soil, fertilizer, animals, insects, and food handlers have all been implicated as sources of human pathogen contamination in fresh produce (Heaton & Jones, 2008, Lynch et al., 2009).

The microbial quality of irrigation water is now recognized as one of the primary pre-harvest factors influencing the microbial quality of fresh produce. In 2006, two separate outbreaks of *E.coli* O157:H7 were attributed to leafy greens (spinach and iceberg lettuce) grown in California. An investigation of the watersheds surrounding the farms suggested that water used to irrigate the crops was the likely source of contamination (Gelting & Baloch, 2012). Regarding viruses, sewage contaminated irrigation water has previously been implicated as the source of Hepatitis A outbreaks attributed to fresh produce, however in many cases this could not be proven (Seymour & Appleton, 2001). Surveillance and on-sight field investigations of the environmental sources of fresh produce contamination, such as irrigation water, during foodborne viral outbreaks is limited by current methods for virus detection.

Microbial quality standards for irrigation water, when they exist, vary greatly between countries (and states) as well as by water source. Compared to groundwater, reclaimed and surface irrigation water sources are more susceptible to human pathogen contamination and therefore may have recommended guidelines for agricultural use (Steele & Odumeru, 2004). Common microbial indicators used in irrigation water quality guidelines include coliform bacteria (total and fecal), *E. coli*, enterococci, and nematode eggs (Steele & Odumeru, 2004, EPA, 2012). For reclaimed water used in the irrigation of food crops intended for human consumption, the EPA currently recommends daily monitoring of fecal coliforms, with no detectable fecal coliforms per 100 mL of water (EPA, 2012). For surface waters, the EPA recommends fewer than 1000 fecal coliforms per 100 mL of irrigation water for use on crops

(Steele & Odumeru, 2004). However studies have shown that viral pathogens do not correlate with traditional indicators in water and the application of enteric viruses or coliphage as indicators of fecal contamination has not yet been utilized for monitoring irrigation water quality (Harwood et al., 2005).

A few studies have investigated the relationship between foodborne bacterial pathogens present in irrigation water and recipient fresh produce (Heaton & Jones, 2008). For viruses, studies that focus on detecting enteric viruses in irrigation water or in foods exist, however few studies have investigated the relationship between viruses in irrigation water and recipient produce. In one hydroponic study, DiCaprio et al. (2012) found that human norovirus and animal caliciviruses (Tulane virus and murine norovirus) were capable of efficient internalization and dissemination in romaine lettuce when introduced into the feed water (Dicaprio et al., 2012). Here internalization refers to virus entry into the plant interior tissues (not virus infection of plant cells). This suggests that viruses present in irrigation water can also occupy plant tissues of recipient crops and ultimately impact their viral composition.

A study by Stine et al. (2005) determined the concentration of hepatitis A in irrigation water needed for a 1:10,000 yearly risk of infection from consuming irrigated fresh produce to ultimately help guide microbial standards in irrigation water. In this study field-grown cantaloupe, iceberg lettuce, and bell peppers were drip (target plant roots) or furrow (flooded channels) irrigated with coliphage PRD1 (surrogate for Hepatitis A) and a quantitative microbial risk assessment was conducted. Results suggest that risk of infection varies with crop type, irrigation method, and time between irrigation and harvest for consumption. Specifically, direct targeting of the plant roots using subsurface drip irrigation was found to reduce the risk of crop contamination (Stine et al., 2005). The use of subsurface drip irrigation as a way to mitigate

fresh produce contamination of viral pathogens has been supported in other studies (Alum et al., 2011, Song et al., 2006). However in a study by Choi et al. (2004) where MS2 and PRD1 phage inoculated irrigation water was used to cultivate field lettuce, lettuce virus levels were higher using the subsurface drip method when compared to the furrow method. This study suggested direct contact of irrigation water with lettuce stems in addition to shallow drip irrigation depth as likely causes of increased contamination (Choi et al., 2004).

More recently, a study by Cheong et al. (2009) used reverse transcription and cell culture PCR to detect norovirus, enteroviruses, adenoviruses, and rotaviruses in surface applied irrigation groundwater and recipient fresh produce (cherry tomato, chicory, cabbage, beet, lettuce, spinach). Although a clear relationship between enteric viruses in irrigation water and recipient fresh produce was not detected, this study found that 1) virus occurrence did not relate to coliform (total and fecal) or enterococci levels traditionally used to assess microbial quality and 2) irrigation water and fresh produce samples positive for enteric viruses were collected during the same time period (irrespective of sampling location). These results suggest that bacterial indictors do not accurately represent all microbial hazards in water and enteric virus presence in irrigation water and fresh produce may vary seasonally. The relationship between irrigation water and fresh produce viral contamination could be better understood if greater assessments could be undertaken to provide more resolution on the types of viruses present. This is now possible using novel metagenomics techniques which study entire microbial communities.

4. Viral Metagenomics

4.1 Viral Metagenomic Technology

Metagenomics is defined as the study of all genetic material from a mixed community of organisms (Handelsman, 2004). This approach is a sequence- and culture-independent method

for studying entire microbial communities in an environmental sample. In 1998, the first commercially available high-throughput instruments the GE Healthcare MegaBACE 1000 and ABI Prism 3700 DNA Analyzer, used a combination of traditional Sanger and capillary sequencing to create these large DNA datasets (Kircher & Kelso, 2010). However, in the past decade, innovative sequencing systems, including Roche 454 Genome Sequencers (Junior, Junior+, and FLX Titanium), Illumina (Genome analyzer (I, IIx, IIe), Miseq, and Hiseq), and the Applied Biosystems SOLiD sequencing platforms, have resulted in greater daily throughput and significant cost-reductions (Kircher & Kelso, 2010, Liu et al., 2012). These NGS technologies are the current metagenomic tools most commonly used to identify microbial communities from a wide variety of environmental samples and are compared in Table 4.

Table 4. Advantages and disadvantages of current metagenomic next-generation sequencing technologies* and examples applied to the food industry

Sequencing technology	Sequencing method**	Advantages	Disadvantages	Current metagenomic research (food industry application)	References
Roche 454 Genome Sequencer (junior and FLX systems)	Pyrosequencing	Larger read length	Lower throughput, higher error rate, high reagent costs	Food microbiota (cheese, fresh produce, fermented foods); changes in microbiota during food processing, fermentation, and storage; microbiota in irrigation water	(Ercolini, 2013, Leff & Fierer, 2013, Lopez-Velasco et al., 2011, Ottesen et al., 2013, Park et al., 2011)
Illumina (Genome Analyzer, Hiseq, Miseq)	Sequencing-by-synthesis	High throughput, lower sequencing cost	Short read length	Alcohol fermentation, rumen microbiome and virome in cattle	(Ercolini, 2013, Ross et al., 2012, Ross et al., 2013)
Applied Biosystems SOLiD	Sequencing-by-ligation	Increased accuracy	Short read length	None***	

*Reproduced from Liu et al. 2012

**Definitions: Pyrosequencing is a method based on the detection of pyrophosphates released during DNA polymerase synthesis (nucleotide incorporation). Illumina sequencing-by-synthesis relies on the detection of single bases (base-by-base sequencing) as they are incorporated into DNA strands by DNA polymerase. Sequencing-by-ligation uses a DNA ligase enzyme to identify fluorescently labeled oligonucleotide probes and perform the sequencing reaction.

***Could not find examples of this technology applied to foods or food systems

The Roche 454 and Illumina systems are the NGS technologies most frequently applied to viral metagenomics (Mokili et al., 2012). The Roche 454 GS FLX system has been called the "gold standard in NGS". In contrast to the GS junior system which fits on a benchtop, the FLX is a large system for accurate, high-throughput sequencing that results in long read lengths of DNA (up to 1 kb) (Roche, 2015). This system is well suited for large genomic projects and can be used for pathogen detection in complex environmental samples. To date, the Genome Analyzer (IIx), Miseq, NextSeq, Hiseq, and HiSeqx all use the latest Illumina sequencing technologies, with Illumina Hiseq providing the necessary sequencing power for studying large scale production genomics. Using base-by-base sequencing, these NGS systems provide more information than traditional Sanger sequencing and can be used to sequence whole genomes, target regions, RNA, and entire microbial communities in humans and the environment (Illumina, 2015). All of the NGS technologies result in millions of reads consisting of short fragments of nucleic acids (the building blocks of DNA) which need to undergo bioinformatics analysis (computer programs that identify the DNA) in order to determine the associated microbial communities in the sample. Bioinformatics uses a combination of known databases that have to be built, algorithms, computational techniques, and statistical tools to analyze and match the complex genetic and genome sequence data generated by NGS technologies to known sequences so that organisms can be identified. The primary aims of bioinformatics include the organization of data for research access and entry, development of tools that help analyze all of the complex data, and the use of these tools to analyze biological data in a meaningful manner (Luscombe et al., 2001). Viral metagenomics integrates bioinformatics tools into pipelines in order to analyze and characterize entire viral communities (Aw et al., 2014). The basic steps

during bioinformatics analysis of viral metagenomic data include sequence read preprocessing (quality control and trimming), assembly, and annotation (Kunin et al., 2008). Common viral bioinformatics tools used for Illumina sequencing reads and their uses are described in Table 5.

Table 5. Bioinformatic computational techniques and tools used for viral metagenomics

Bioinformatics step	Description	Tool examples
Quality control check	Check raw data to ensure sequence quality	FastQC
Quality trimming	Adapter removal	AdapterRemoval, Cutadapt, and Trimmomatic
Assembly	Fragmented nucleotide sequences assembled into overlapping segments of nucleic acids (contigs)	Velvet and IDBA-UD
Annotation	Compare contigs with sequence databases to identify genes and assign biological information	tBLASTx or BLASTn
Taxonomic classification	Identification and organization of the virus species present	MEGAN

Bioinformatic analysis of the viral sequences generated using NGS technology includes an initial quality control check followed by sequence trimming, assembly, annotation, and taxonomic classification. FastQC is a common control tool used to check the quality of raw sequence data. Quality trimming of the sequences is then performed to trim sequences to desired lengths and remove contaminant adapters (Lindgreen, 2012). Next, fragmented nucleotide sequences are assembled into contigs or overlapping segments of nucleic acids in a process called sequence assembly. To date there are numerous metagenomic sequence assembly tools, however these tools can vary in suitability depending on the sequencing technology used. A few

examples of the assembly tools currently used for viral metagenomics include Celera software, IDBA-UD, MetaVelvet, and Velvet (Vázquez-Castellanos et al. 2014, Wylie et al. 2013), however Velvet and IDBA-UD are generally used for Illumina sequence analysis (Aw et al., 2014b, Vázquez-Castellanos et al., 2014, Wylie et al., 2013). Once assembled, annotation and taxonomic classification of the assembled sequence reads is performed. Genome annotation is where elements of the genome are identified (gene prediction) and biological information is linked to specific sequences. For viral metagenomics, the most commonly used method of annotation includes comparing sequences to Genbank using tBLASTx or BLASTn (Bibby et al., 2011; Leclerc et al., 2001; Mokili et al., 2012; Wylie et al., 2013). Following annotation, viruses can be grouped into taxonomic classifications and phylogenetic trees to analyze the virus communities present. To date, the most common software used for analysis of virus communities is the Metagenome Analyzer (MEGAN) (Aw et al., 2014, Kim et al., 2011, Moore et al., 2015, Park et al., 2011).

4.2 Metagenomic Insights into the Virome and Food Safety

Current metagenomic research has provided insights into microbial communities associated with a number of environmental samples. With reference to viral metagenomics, it is suggested that less than 1% of viral diversity has been explored with unknown (novel) sequences ranging between 60 and 99% in human or environmental samples (Mokili et al., 2012). The first environmental viral metagenomics study investigated viruses in marine waters through shotgun library sequencing (Mokili et al., 2012, Rosario & Breitbart, 2011). The results suggested that most viral community diversity is currently undescribed and supported the conclusion that the majority of identifiable viruses in marine environments are phages (Breitbart et al., 2002). Using NGS technology, it has been suggested that not only do marine waters have high viral diversity,

but also diversity varies with geographic region and consists of a large proportion of single-stranded DNA viruses (Angly et al., 2006). Next generation viral sequencing of human feces and wastewater revealed a large proportion of phage, many of which belong to ds DNA bacteriophage of the *Caudovirales* order (Aw et al., 2014, Kim et al., 2011). Furthermore, wastewater contained a wide array of human viruses which were dominated by three adenovirus species (B, C and F), Enterovirus B, polyomaviruses and papillomavirus (Aw et al., 2014). In reclaimed water, eukaryotic viral sequences belonging to plant pathogens of agriculturally important crops were dominated by viruses in the ss DNA *Geminiviridae* and *Nanoviridae* families, however viruses infecting numerous animal species (vertebrate and invertebrate) were also identified (Rosario et al., 2009).

Next-generation sequencing technology is now being used as a tool to detect and track pathogen outbreaks and transmission routes (Bergholz et al., 2014). For example, metagenomic sequencing and analysis of fecal samples collected from individuals involved in previous gastroenteritis outbreaks of unknown etiology in New Zealand were able to identify eight viruses including human enteric adenovirus, rotavirus, and sapovirus (Moore et al., 2015). In addition NGS is now applied as a diagnostic tool in plant viral disease and in the discovery of insect viruses (Adams et al., 2009, Liu et al., 2011). These are just a few examples of how NGS technology has already provided knowledge of the virus world.

Recognizing that NGS technology could be used as a tool for monitoring food safety, scientists are now beginning to investigate the bacterial and viral communities associated with foods. To date, a number of studies have focused on the microbial communities associated with fermented foods. In a study by Park et al. (2011), the viral ds DNA in fermented kimchi, sauerkraut, and shrimp was amplified and sequenced using Roche 454 pyrosequencing. This

study found that 99.3% to 99.9% of viral reads showed the greatest sequence similarity to phages, with 99.9% of phages belonging to bacteriophages in the *Caudovirales* order (Park et al., 2011). Although fermented food viral communities were less diverse than other environmental habitats, these samples contained a large proportion of unidentified viral sequences suggesting a lack of data and understanding of viral genomes associated with these samples (Park et al., 2011). These data advance our current knowledge on the diversity of viruses and ultimately the ecological roles that these viruses play in food systems.

To date, metagenomic approaches to study the genetic material in fresh produce and sources of fresh produce contamination have focused primarily on bacterial communities. In a study by Ottesen et al. (2013), NGS was used to characterize the tomato microbiome by sampling different parts of the tomato plant (fruit, flowers, leaves, stems, and roots) to identify ecological contributors to *Salmonella* persistence. This study observed 10 phyla from bacterial, eukaryotic, and viral domains and identified *Pseudomonas* and *Xanthomonas* as the most common bacterial taxa across all plant regions. Although *Salmonella* was not detected, this is one of the first studies to investigate the microbiome in fresh produce, concluding that microbial diversity decreases as distance from the soil increases and bacterial diversity varies between different parts of the tomato plant (Ottesen et al., 2013). In addition, metagenomics has been used to study the impact of suggested sources of pathogen contamination, including irrigation water, on the microbial surface communities of fresh produce. Telias et al. (2011) discovered major differences in the bacterial composition between ground and surface irrigation water, however the surface microbial communities of tomatoes irrigated with these waters were not significantly impacted. In a more extensive study, Leff & Fierer (2013) used NGS technology to study bacterial communities on numerous fruits (grapes, strawberries, apples, peaches, tomatoes)

and vegetables (lettuce, spinach, mushrooms, sprouts, peppers) at the point of sale, investigating how community structure differs between produce types and if farming practices contribute to composition. This study revealed that bacterial communities: (1) are highly diverse and vary with produce type, (2) on average, are similar between produce types grown in similar environments (tree vs. ground), and (3) differ significantly in composition based on farming practices (conventional vs. organic) (Leff & Fierer, 2013).

Metagenomic technology has also been applied to study the change of bacterial communities in spinach during packaging and storage. Results suggest that bacterial diversity, richness, and evenness significantly decreased when spinach was packaged and stored at 4°C and 10°C with the entire microbiome reduced from 11 to 5 phyla after 1 day of storage at 4°C (Lopez-Velasco et al., 2011). These studies serve as examples of how metagenomics can enhance our knowledge of microbial communities associated with foods, especially fresh produce, from farm-to-fork and lead to the identification of possible control points for enhanced safety.

5. Post-Harvest Leafy Green Processing and Viral Contamination

5.1 Leafy Green Commercial Processing Practices

As stated previously, fresh produce production practices are becoming more intensive to accommodate growing consumer demand. Following harvest, fresh produce is subject to a number of processing techniques which may vary between location and food type. In addition, many fruits and vegetables are often combined (before and after sale) to make up complex foods or beverages such as salads mixes, smoothies, and sandwiches. The complexity of the food supply chain provides multiple opportunities for human pathogen contamination, making it difficult to determine the single vehicle of transmission or source of contamination. Leafy green

vegetables, including lettuce (iceberg, romaine, red leaf, butter, etc.), escarole, endive, spring mix and spinach, are an example of a food commodity that is commonly consumed raw.

Leafy greens are provided to the consumer either as bulk products to be washed (e.g., head of lettuce) or as ready-to-eat (RTE) salads (a "value-added" product). To harvest lettuce intended for bulk sale, the entire head is usually cut manually with a harvesting knife (with wrapper leaflets removed), placed onto a processing platform for packaging, and transported for vacuum (iceberg) or spray-vacuum (romaine) cooling and cold storage. Lettuce may be packaged "naked" or in perforated plastic and cellophane bags depending on lettuce type. Rapid vacuum cooling followed by cold storage ($\leq 5°C$) are essential steps for preserving the quality and shelf-life of porous leafy greens. Following cold storage, lettuce is packed into shipping containers and transported to distribution centers that further transport the food product to retailers or foodservice establishments (CFSAN, 2009).

The supply chain for packaged fresh-cut lettuce is even more complex. Practices such as coring, rinsing, and outer leaflet removal at the time of harvest are now implemented in the field in an attempt to provide cleaner, safer lettuce to processing facilities (NFPA, 2001). Following harvest and cooling, lettuce intended for RTE salads undergoes a number of additional processing steps to reduce foodborne pathogen transmission and improve overall food quality. Currently, standard steps in post-harvest leafy greens processing for RTE salads include lettuce shredding, washing, shaker table dewatering, centrifugal drying, and packaging prior to cold storage and transportation to distribution centers and end-users.

Although due diligence is necessary at all production levels to ensure food product safety, wash water disinfection and monitoring of sanitizer levels during processing are seen as essential to minimize foodborne disease (CFSAN, 2014). To date, numerous physical (ultrasound high

pressure, ultraviolet, ionizing radiation) and chemical (chlorine dioxide, sodium chlorite, quaternary ammonium compounds, peroxyacetic acid) disinfection methods have been investigated for use in foods, however, washing with a chlorine-based sanitizer is the disinfection method most commonly used in the fresh-cut produce industry due to the low cost (CFSAN, 2014). Fresh produce is generally placed into a large tank and washed with recirculated water containing a sanitizer, a process known as fluming. Guidelines for washing fresh produce in chlorinated water include a maximum free chlorine (hypochlorite) concentration of 200 ppm and a 1 to 2 minute contact time (CFSAN, 2014). Currently, the National Advisory Committee on Microbiological Criteria for Foods uses a 5-log pathogen reduction performance standard for fruit and vegetable juice production (CFSAN, 2001). However, this standard is primarily targeted towards bacterial pathogens since there is no defined criterion for antiviral disinfectants, a 3-log reduction has been generally accepted for virus efficacy testing (Allwood et al, 2004; Gulati et al., 2001).

5.2 Post-Harvest Fresh Produce Contamination and Viral Survival on Foods

Human pathogen contamination of leafy greens can occur at any stage of post-harvest production including processing, packing, storage, and transportation. Critical control points include the quality of water used (cooling, washing), worker hygiene, and the condition or overall cleanliness of processing equipment, cooling facilities, storage and packaging containers, and transportation vehicles (CFSAN, 2006). To date, numerous studies have investigated bacterial pathogen survival (primarily *E. coli* O157:H7) and contamination during farm-to-fork production (Beuchat, 2002). Interestingly, research has shown that bacterial pathogens are easily transferred between lettuce and processing equipment (such as coring knifes and processing lines) and preferentially attach to fresh cut lettuce surfaces allowing for survival and persistence

during processing (Buchholz et al., 2012a, Buchholz et al. 2012b, Seo & Frank, 1999, Taormina et al., 2009).

To date numerous studies have investigated virus attachment and survival in fresh produce, all of which affect virus recovery and disinfection. Viruses are thought to use physicochemical forces, specifically electrostatic forces, for nonspecific attachment to solid surfaces such as fresh produce which can be disrupted by a high pH (Deboosere et al., 2012, Vega et al., 2008). For example Vega et al. (2005) observed maximum adsorption of MS2 to butterhead lettuce at a pH of 3.0, while a basic pH of 8.0 led to almost complete dissociation of the virus from the lettuce surface. This information is critical for understanding how to better recover, remove, and inactivate viruses from lettuce surfaces during processing.

Temperature is a well-known factor influencing both the survival and internalization (virus entry into plant interior) of pathogens in lettuce. Regardless of the specific fruit or vegetable, survival studies consistently show that non-enveloped viruses (rotavirus, MS2 phage, poliovirus, adenovirus) are able to survive for long periods of time (25 -76 days) at traditional storage temperatures (4°C) (Badawy et al., 1985, Dawson et al., 2005, Ward & Irving, 1987). These data stress the importance of limiting viral pathogen contamination of fresh produce at the field level prior to harvest and storage. Changes in storage conditions such as increased temperatures and CO_2 levels have shown to significantly decrease virus survival in fresh produce and the time of survival has also been shown to vary with the food commodity and virus type (Dawson et al., 2005, Rzezutka & Cook, 2004). For example, Allwood et al. (2004) found that the decimal reduction time (the time needed at a given temperature to kill 90% of the organisms) for MS2 on iceberg lettuce was reduced from 5 days at 4°C to 3 days at 37°C. Interestingly, studies addressing multiple food commodities have observed greater virus survival in lettuce

compared to other fresh produce commodities (Badawy et al., 1985, Croci et al., 2002). Currently there is limited knowledge as to why virus survival is greater on lettuce in comparison to other food commodities, however suggestions include protection from the rough surface of lettuce, resistance to naturally occurring antimicrobials, or protection due to internalization (virus entry into plant interior) through roots or cut surfaces (Badawy et al., 1985, Seymour & Appleton, 2001, Wei et al., 2010). Despite survival study insights into virus persistence in the food supply chain, there is currently limited research investigating where in the supply chain contamination occurs.

There have been numerous studies investigating bacterial pathogen reduction on fresh produce using chlorine sanitizers. Pilot-scale studies have found that chlorine-based sanitizers generally reduce bacterial pathogen populations on lettuce between 1 and 3 logs (Davidson et al., 2013, Gil et al., 2009). Laboratory studies investigating the effects of chlorine on viruses when inoculated onto fresh produce have shown that viral (MS2, feline calicivirus, murine norovirus) reduction generally does not exceed 3 logs when exposed to a variety of free chlorine (15-800 ppm) levels (Allwood et al., 2004, Dawson et al., 2005, Fraisse et al., 2011, Gulati et al., 2001). In a study by Fraisse et al. (2011), feline calicivirus, murine norovirus, and hepatitis A virus populations on inoculated lettuce decreased 1.9, 1.4, and 1.4 logs, respectively, when washed for 2 min in 15 ppm of free chlorine. This can be compared to an overall 1 log reduction using water alone and a 3.2 (feline calicivirus), 2.4 (murine norovirus), and 0.7 (hepatitis A) log reduction using 100 ppm of a peroxide-based disinfectant (Fraisse et al., 2011). In another study, MS2 and hepatitis A virus experienced slightly higher inactivation rates on lettuce ($\geq$1.7 log) than strawberries (1-1.2 log) when exposed to 20 ppm free chlorine for 3 to 5 min (Casteel et al., 2008). Currently, there are a limited studies investigating bacterial and viral reduction on fresh

produce during simulated commercial processing. Davidson et al. (2013) showed that *E. coli* O157:H7 populations on lettuce were not significantly reduced when washed with either water or 30 ppm of free chlorine during simulated commercial processing. In addition, the use of 30 ppm of available chlorine in the wash water at both a pH of 7.85 and 6.5 were found to result in significant population reductions on the lettuce (Davidson et al., 2013). In a viral study by Casteel et al. 2009, an industrial-scale processing unit consisting of a washing compartment, grates, and conveyor with tap water spray was used to study chlorine inactivation of MS2 on strawberries. This study found that processing strawberries with wash water containing 20 and 200 ppm free chlorine inactivated 92% and 96% (~1 log PFU) of the MS2, respectively, compared to 68% using water alone. More research is needed to provide a better understanding of virus inactivation on fresh produce during simulated commercial processing.

II. RESEARCH GOALS AND OBJECTIVES

This thesis was split into two primary studies and had a related set of goals and objectives. The first study goal (Part III) was to use NGS technology and metagenomic techniques for the first time to identify the virus communities (virome) present in irrigation water and lettuce and use this information to better understand contamination in the field. The specific objectives were to:

1) Evaluate a method of virus recovery from lettuce
2) Identify and evaluate the diversity of virus communities present in irrigation water and lettuce (iceberg and romaine) at the field level

The goal of the second study (Part IV) was to investigate the efficacy of current post-harvest leafy green processing and disinfection practices to better understand viral risks from farm-to-fork during a contamination event. Specifically, the goal was to assess the efficacy of a chlorine-based sanitizer against coliphage MS2 (an enteric virus surrogate) on romaine lettuce during simulated commercial processing. The specific objectives were to:

1) Evaluate MS2 reduction on romaine lettuce during and following small-scale commercial leafy green processing (shredding, flume washing, shaker table dewatering, centrifuge drying) with and without a sanitizer wash treatment
2) Determine MS2 levels in the flume wash water and centrifugation water following processing

III. METAGENOMIC IDENTIFICATION OF VIRUS COMMUNITIES ASSOCIATED WITH LETTUCE AND IRRIGATION WATER

1. Introduction

New scientific methods and genomics tools can help us take a broad view of food safety like never before, particularly for hazards such as viruses where traditional methods have limited our ability to monitor. Food safety monitoring efforts in the United States food supply chain have greatly increased since the early 1990's (Crutchfield & Roberts, 2000). However despite recent advances in monitoring, foodborne illness remains a serious threat to public health with approximately 48 million people in the United States affected each year (CDC, 2014). In addition fresh produce, which is commonly consumed raw, is becoming increasingly recognized as a vehicle of human pathogen transmission. Specifically, leafy green vegetables were the food commodity responsible for the highest number of foodborne illnesses (2.2 million) between 1998 and 2008 in the United States (Painter et al., 2013).

Recent foodborne outbreaks also suggest that viruses play a larger role than previously thought and studies on norovirus in particular have shown strong associations with leafy green consumption (Gould et al. 2013). Viruses are host-specific, obligate intracellular infectious agents that are ubiquitous in nature and can cause a wide range of diseases. Examples of viral hosts include bacteria (bacteriophage), algae, plants, animals (invertebrate and vertebrate), and humans. Human enteric viruses, which are commonly transmitted through the fecal-oral route, are of particular public health concern. These viruses are able to replicate inside the host's gastrointestinal tract and are shed at extremely high concentrations into the environment through the feces or vomit of an infected individual. Food and water can then become contaminated at both pre- and post-harvest levels of production. Common examples of food and waterborne

viruses include adenoviruses, astroviruses, enteroviruses, hepatitis A and E viruses, norovirus, reoviruses and rotaviruses which can cause gastroenteritis, meningitis, liver disease, infantile diarrhea, respiratory illness, or neurological symptoms (Cook, 2013, Fong and Lipp, 2005, Koopmans et al. 2004).

The microbial quality of irrigation water is now recognized as one of the primary pre-harvest factors influencing the microbial quality of fresh produce. Currently, there is no universal standard for irrigation water microbial quality and traditional bacterial indicators of fecal pollution (coliforms, *E. coli*, enterococci) fail to identify enteric virus hazards in water (Harwood et al., 2005). In addition, surveillance and on-sight field investigations of the environmental sources of fresh produce contamination, such as irrigation water, during foodborne viral outbreaks is limited by current methods for virus detection. To date the relationship between irrigation water and fresh produce viral contamination is poorly understood and greater assessments are needed to provide resolution on the types of viruses present.

Current methods for virus detection include cell culture, electron microscopy, PCR, and metagenomics. Although cell culture is still one of the most often used methods today, virus detection is difficult due to the host specificity of viruses which requires the correct cell line for proliferation and isolation. Another major disadvantage is that many viruses such as human noroviruses are currently unable to grow in any of the known cell lines. Molecular methods such as PCR, RT- PCR (for the detection of RNA viruses), and qPCR are sequence-dependent detection methods which have allowed for virus detection and quantification but first require knowledge of the viral nucleic acid (DNA or RNA) sequence and therefore discovery of novel viruses is not possible (Mokili et al. 2012). An exciting emerging field of science and technology includes metagenomics and next generation sequencing.

Metagenomics is a sequence- and culture-independent approach for studying entire microbial communities in an environmental sample using NGS technology. Current NGS tools include Roche 454 Genome Sequencers, Illumina (Genome Analyzer, Hiseq, Miseq), and the Applied Biosystems (AB) SOLiD sequencing platforms which use pyrosequencing, sequencing-by-synthesis, and sequencing-by-ligation technology, respectively. These parallel sequencing technologies result in millions of reads per run and have led to significant cost reductions. To date, viral metagenomic research using Roche 454 or Illumina sequencing technology has primarily focused on the virus communities in water (marine, reclaimed, and wastewater) and human feces (Angly et al., 2006, Aw et al., 2014, Kim et al. 2011, Mokili et al. 2012, Rosario et al., 2009). Data suggests that less than 1% of viral diversity has been explored and a large proportion of viruses in both water and clinical samples belong to double-stranded DNA bacteriophage of the *Caudovirales* order. In addition, NGS is now being used as a tool to detect and track viral pathogen outbreaks and transmission routes (Moore et al. 2015). These are just a few examples of how NGS technology has already provided knowledge of the virus world.

Recognizing that NGS technology could be used as a tool for monitoring food safety, scientists are now beginning to investigate the viral communities associated with foods such as fermented kimchi, sauerkraut, and shrimp (Park et al. 2011). Such studies can help advance our current knowledge on the diversity of viruses and ultimately the ecological roles these viruses play in our food systems. To date, current application of metagenomic approaches to study the genetic material in fresh produce and sources of fresh produce contamination have focused primarily on the bacterial communities (Leff & Fierer 2013, Lopez-Velasco et al. 2011, Ottesen et al. 2013, Telias et al. 2011). There is currently limited research on the virus communities associated with fresh produce and how the food production chain affects the viral microbiome.

Applying NGS technology to studying the virus communities associated with irrigation water and fresh produce is a promising tool for enhancing knowledge of viruses in our food systems and can lead to the identification of possible control points for enhanced safety.

2. Research Goals and Objectives

The goal of this portion of the thesis was to use NGS technology and metagenomic techniques for the first time to identify the virus communities (virome) present in irrigation water and lettuce and use this information to better understand contamination in the field.

The specific objectives were to:

1) Evaluate a method of virus recovery from lettuce
2) Identify and evaluate the diversity of virus communities present in irrigation water and lettuce (iceberg and romaine) at the field level

3. Materials and Methods

3.1 Virus Recovery from Lettuce

3.1.1 Bacterial Host and Bacteriophage Preparation

Bacteriophage and respective bacteria hosts were prepared as seed cultures following a standard procedure for plaque assays adapted from EPA Method 1602 (EPA, 2001). The bacteriophage used in this study included P22 (provided by Dr. Charles Gerba, University of Arizona, AZ, USA) and F+ specific coliphage MS2 (ATCC#15597-B1, ATCC, Manassas, VA, USA) which infect *Salmonella* (LT2 pLM2 1217 HER #1023, Félix d'Hérelle Reference Center for bacterial viruses of the Université Laval, Quebec, Canada) and *Escherichia coli* (*E. coli* Famp ATCC#700891, ATCC) hosts, respectively. *Salmonella* LT2 and *E. coli* Famp host stock cultures were prepared by rehydrating lyophilized cultures in tryptic soy broth (TSB, Becton,

Dickinson and Company, Sparks, MD, USA) and incubating overnight at 37°C before adding 10 to 20% glycerol by volume. A 1-mL aliquot was then added to a cryovial and stored at -80°C. *Escherichia coli* Famp was prepared in TSB containing a 1% volume/volume (v/v) solution of ampicillin-streptomycin prepared by dissolving 0.15 g ampicillin sodium salt (Sigma-Aldrich, St. Louis, MO, USA) and 0.15 g streptomycin sulfate (Sigma-Aldrich) in 100 mL reagent grade water, which was then filtered through a 0.22 µm filter and stored in 5-mL vials at -20°C.

A working stock solution of bacteriophage P22 or MS2 was prepared by first rehydrating lyophilized phage in TSB and then diluting in phosphate buffered saline (PBS). A double-agar overlay method was used for phage replication and enumeration. To perform the double agar overlay, 1 mL of bacterial host grown to log phase (see below) and 1 mL of each phage suspension (dilutions) were added to 2.5 mL, 1.5% trypticase soy agar (TSA, Becton, Dickinson and Company) overlays (tempered in a 48^{o} C water bath). After addition of host and sample the tube was gently mixed by rolling the tube between the hands and immediately poured onto solidified TSA plates for each dilution series. Plates were allowed to solidify, inverted to avoid condensation and then incubated at 37°C for 24 h to allow the growth of the monolayer of bacterial host cells and plaques formation (areas where the phage replicated and lysed the bacterial cells) . Plates with high plaque counts (exhibiting a lacey pattern with approximately 1000 Plaque Forming Units (PFU)/plate) were flooded with ~10 mL of TSB and incubated with shaking at 4°C for 1 hr. Finally, the TSB was recovered using a pipette and bacteria were removed by filtering through a 0.22 µm filter. The resulting P22 (10^{10} PFU/mL) and MS2 (10^{9}-10^{10}) bacteriophage working stock cultures were stored at 4°C.

Bacteriophage working stocks were then used to maintain future working stocks. To maintain P22 and MS2 working stocks, overnight log phase cultures of *Salmonella* LT2 and

E.coli Famp were prepared by adding 1 mL of host stock culture from the freezer to 9 mL of TSB followed by overnight incubation at 37°C. Log-phase host cells were prepared by inoculating 1-mL of the overnight host into a desired working stock volume (500 mL) followed by 4 h of incubation at 37°C with shaking. Once the bacterial host cells reached log phase, 1 mL of bacteriophage working stock was added to the TSB bacterial suspension and the mixture was incubated overnight at 37°C. *E. coli* Famp was grown in TSB containing a 1% v/v solution of amplicillin-streptomycin, as described for host stock preparation. The majority of the bacteria were infected and lysed by the phage, releasing virus into the broth. Afterwards, the samples were filtered through 0.45 µm and then through 0.22 µm filters to remove bacteria and the new working bacteriophage stocks were stored at 4°C (10^9-10^{10} PFU/mL).

To prepare bacterial hosts for the plaque assay, 1 mL of host stock culture from the freezer was added to 9 mL of TSB and incubated overnight at 37°C. Log-phase host cells were prepared by inoculating 1-mL of the overnight host into 30-40 mL of TSB and incubating for 4 hr at 37°C. Again, *E. coli* Famp was grown in TSB containing a 1% v/v solution of amplicillin-streptomycin, as described for host stock preparation.

3.1.2 Lettuce Inoculation, Elution, and Plating of Bacteriophage

Bacteriophage recovery experiments were performed to determine the effectiveness of virus elution from lettuce. Both bacteriophage MS2 and P22 were used as inoculants for this pilot recovery assay. Unpackaged romaine lettuce heads were purchased from local grocery stores and stored at 4°C for a maximum of 48 h. On a sterile surface, 50 g of the outer leaflets cut with a scalpel 2.5 to 5 cm from the core was inoculated with 1 mL of P22 or MS2 diluted to 10^4 PFU/mL by hand-pipetting droplets (60-70) from a 1000 µL pipette evenly onto the leaflet surfaces in a biosafety cabinet. For the inoculation suspension, bacteriophage stock containing

approximately 10^9-10^{10} PFU/mL was diluted 10-fold to 10^4 PFU/mL in 1x Phosphate Buffered Water (PBW). Inoculated leaflets were then allowed to air-dry at room temperature for 20 min (following procedure by Dubois et al., 2007). A ~50-g, uninoculated lettuce sample was processed as a negative control.

Inoculated lettuce samples were then eluted. Each sample was placed into a Whirl-pak bag (Nasco filter Whirl-pak 19 x 30 cm, Fort Atkinson, WI, USA) and soaked in 250 mL of 100 mM Tris (UltraPure Tris, Invitrogen, Carlsbad, CA, USA) – 50 mM glycine (Tris-glycine) elution buffer at pH 9.5, a method adapted from Dubois et al., 2007. A high pH was used to disrupt the electrostatic forces that viruses use for surface attachment (Vega, Garland, & Pillai, 2008). The samples were then placed on a rocking platform shaker (Model 100, VWR, Radnor, PA, USA) at full speed for 20 min. Samples were inverted half-way through the shaking process (10 min) to soak both sides of the lettuce leaflets equally. The solution was then recovered and transferred to a sterile 500 mL plastic container (Nalgene wide mouth environmental sample bottle with lid) while recording the recovery volume. Then 6.0 M HCl was used to adjust the pH of the eluent to 7.2 ± 0.2 using a pH probe disinfected between samples by immersion in 10% bleach for 10 to 15 min and neutralized with sterile 5% sodium thiosulfate. The eluent solution was then vortexed, diluted 10-fold in PBW, and plated both undiluted and at a 10^{-1} dilution (2 mL sample and 0.5 mL host) on TSA plates using a double agar overlay method adapted from EPA Method 1602 (EPA, 2001) as described above. Positive (spot plate), negative (2.5 mL of host), TSA media, overlay, and PBW controls were included in each experiment. In addition, the 10^{-8}, 10^{-9}, and 10^{-10} dilutions from the inoculation suspension were assayed by the same double agar overlay method to confirm the initial concentration upon inoculation and determine percent

recovery. After the overlay, plates were allowed to solidify for 10 min, inverted and incubated at 36 ± 1.0°C for 16-24 h prior to enumeration of plaques in the monolayer of the bacterial lawn.

3.1.3 Plaque Concentration Assay and Percent Recovery Calculation

To calculate the phage concentration, the plaques were counted on TSA overlay plates within each dilution series containing a countable range (10-300 PFU). After choosing the appropriate dilution, the number of plaques were summed for each of the three plates and divided by the volume of sample on each plate (2 mL). This value was divided by the number of plates with plaques in the countable range and then multiplied by the dilution factor to obtain average PFU/mL in the eluent as shown in Formula 1. If two dilutions were in the countable range for an individual sample, the PFU/mL was calculated within each dilution series then averaged to obtain a single plaque concentration (value used in percent recovery calculation). To calculate the concentration of phage per g of lettuce (PFU/g), the calculated phage concentration from Formula 1 was multiplied by the eluent volume recovered and divided by the lettuce weight of each sample, as shown in Formula 2. The estimated initial phage concentration was calculated by multiplying the inoculum concentration (PFU/mL) by the volume of inoculum concentrate added (1 mL) and dividing by the total weight of lettuce added to the suspension, as shown in Formula 3. The formula used to calculate the percent recovery of phage from the virus elution recovery protocol is shown in Formula 4.

Formulas:

(1) $$Phage\ Concentration\ in\ the\ Eluent\ \left(\frac{PFU}{mL}\right) = \frac{\left(\frac{Sum\ of\ Plaques}{Volume\ of\ Sample(mL)}\right)}{\#\ of\ Plates\ with\ Plaques\ in\ Countable\ Range} * \ Dilution\ Factor$$

(2) $$Lettuce\ Phage\ Concentration\ \left(\frac{PFU}{g}\right) = (Plaque\ Concentration\ in\ the\ eluent\ (PFU/mL) * Eluent\ volume\ Recovered\ (mL)\ )/Sample\ Weight\ (g)$$

$$(3)\ Inoculated\ Lettuce\ Phage\ Concentration\ \left(\frac{PFU}{g}\right) = (Inoculated\ phage\ concentrate\ (PFU/mL) * Volume\ of\ Inoculated\ Concentrate\ (mL)\ )/Sample\ Weight\ (g)$$

$$(4) Percent\ Recovery = \frac{Lettuce\ Phage\ Concentration\ \left(\frac{PFU}{g}\right)}{Inoculated\ Lettuce\ Phage\ Concentration\ \left(\frac{PFU}{g}\right)} * 100$$

3.2 Virus Communities from Lettuce and Irrigation Water

3.2.1 Irrigation Water and Lettuce Sample Collection

Lettuce and irrigation water samples were collected in Yuma, AZ, during December of 2013. Irrigation water was sampled from six sites along the Yuma main canal. At each location temperature, pH, conductivity and turbidity were measured using two portable meters (HACH HQ40d Portable pH, Conductivity, DO Multi Parameter Meter, and HACH 2100Q Portable Turbidity Meter respectively, Loveland, CO, USA). Images of the Yuma irrigation water sampling locations are provided in Figure A1, Appendix. Sample collection consisted of lowering a 5-L disinfected bucket into the irrigation canal from an overpass or other accessible site, avoiding contact with the shoreline or bottom sediment to minimize turbidity. Prior to use, the bucket was sanitized by exposing to 10% bleach for 10 to 15 min and neutralizing with 5% sodium thiosulfate for a maximum of 5 min. For each irrigation water sample, five 20-L collapsible containers (Cole-Parmer, Vernon Hills, IL, USA) were filled for a total of 100 L. In addition, 500 mL of water were collected for Colilert and Enterolert testing to detect the presence of the indicator bacteria *E. coli* and enterococci, respectively. Colilert and Enterolert testing was performed using a Quanti- Tray/2000 test kit (IDEXX laboratories Inc., Westbrook, ME, USA).

Briefly, the water sample was vortexed and 100 mL was transferred to a sterile graduated bottle. The reagent provided in the kit was then added to the sample and mixed before the entire volume was poured into a Quanti- Tray/2000. Once sealed, the samples were incubated at 41 ± 0.5°C and 35 ± 2.0°C for the Enterolert and Colilert tests, respectively. Yellow wells (total coliform) and ultraviolet fluorescent wells (enterococci and *E. coli*) were counted and the IDEXX most probable number (MPN) generator (version 3.2) was used to obtain the MPN/100mL. Enterolert testing included Nanopure water, *Streptococcus bovis* and *E.coli* as negative controls and *Enterococci faecium* as the positive control (IDEXX-QC Enterococci). Colilert testing included Nanopure water, *Pseudomonas aeruginosa*, and *Klebsiella pneumonia* as negative controls and *E. coli* as the positive control (IDEXX-QC Coliform and *E. coli*).

A total of 42 (21 iceberg and 21 romaine) lettuce heads were collected at different stages of farm-level production. A total of 8 iceberg and 8 romaine samples were hand cut at ground level by the research crew using gloves and a sterile harvesting knife as a control. The outer leaflets were removed before placing the heads in Whirl-pak bags. A total of 8 iceberg and 5 romaine lettuce samples were hand cut by workers and placed on a packaging machine for bagging prior to sampling. In addition, 5 iceberg lettuce samples that were harvested and bagged by workers were collected following a 30 min worker break. Finally, a total of 8 samples were collected post worker chop and wash (n=3) and from buckets containing mixed romaine head lettuce samples to be packed as salad (n=5). Chop and wash is a harvesting process used in bagged salad production in which the core of the romaine head is removed and the lettuce leaflets are conveyed and rinsed on a processing machine in the field prior to collection in large storage containers. Lettuce leaflets will undergo further post-harvest processing (shredding and disinfection) prior to packaging and sale.

3.2.2 Virus Concentration and Purification

Viruses in irrigation water samples were concentrated using a tangential flow, low cost, disposable hollow fiber ultrafiltration system (AsahiKasei, Tokyo, Japan) (Figure A3, Appendix). To prime the system, 1 L of 0.01% sodium polyphosphate (NaPP) solution was recirculated through the ultrafilter at a rate of 1700 mL/min for 15 min. Samples were then added to the reservoir and pumped at 2900 mL/min. Then the flow regulator was adjusted to a flow rate of 1200-1300 mL/min with the system pressure maintained below 10 psi until the final concentrate was between 250 and 300 mL. In addition, 300 mL of surfactant solution (0.01% NaPP, 0.5% Tween 80, and 0.001% Antifoam) was circulated at a rate of 600 mL/min for 5 min and combined with the sample for a final volume of about 500 mL. The final concentrate was then shipped on ice to the laboratory at MSU for further processing.

The same elution procedure was used as described in Part III, Section 3.1.2, pages 37-38. Each ~50-g lettuce sample was eluted with 250 mL of buffer at pH 9.5. Following pH adjustment using 1.0 M HCl to 7.2 ± 0.2, 3.0 M NaCl was added to the recovered solution (volume varied with recovery) for a final concentration of 0.3M NaCl. The same procedure was used to adjust the final concentration of irrigation water samples to 0.3M NaCl. After thorough mixing, the sample was cooled for 30 min at 4°C. Polyethylene glycol precipitation was then used to further concentrate viruses in irrigation water and lettuce samples. Molecular biology grade PEG 8000 powder (Promega Corporation, Madison, WI, USA) was gradually added to each sample and mixed thoroughly for a final concentration of 10% (w/v). After 18 h of incubation at 4°C the sample was centrifuged at 10,800 x *g* (8000 rpm) for 30 min at 4°C (Beckman Coulter JS-HS centrifuge, Brea, CA, USA). The resulting pellet was dissolved in 20 mL of PBS at room temperature for 1 hr. The walls of the container were rinsed before adding

and mixing an equal volume of chloroform (10-20 mL) to the PBS to remove the PEG and purify the sample. The solution was then vortexed gently for 30 s and centrifuged at 3000 x *g* (4300 rpm) for 15 min at 4° C to collect the supernatant containing virus particles. The remaining supernatant was then passed through 0.45 and 0.22 μm filters and the final concentrates were stored at -80°C until further concentrated by Amicon centrifugal ultrafiltration.

3.2.3 Final Concentration, Purification, and Nucleic Acid Extraction

Following filtration, the concentrate was added to an Amicon Ultra 30kDA centrifugation column (Millipore, Billerica, MA, USA) and centrifuged at 1000 x *g* at 4°C in a swinging-bucket rotor until about 1 mL of sample was left in the filter (~5-7 min). The sample was removed from the reservoir and the filter was rinsed with 1.5 mL of the filtrate, vortexed, and added to the final concentrate. Samples were treated with DNase-I (Roche, Indianapolis, IN, USA) before nucleic acid extraction to remove free nucleic acids from the concentrated virus samples. Viral nucleic acids were extracted using a QIAGEN QIAamp MinElute Virus Spin Kit, following the manufacturer's instructions (QIAGEN, Maryland, USA). Following extraction, the samples were checked for 16s contamination by PCR using a 50 μL cocktail consisting of 25 μL GoGreen master mix (Promega), 1 μL forward (5'-AGAGTTTGATCCTGGCTCAG-3') and reverse (5'-GGTTACCTTGTTACGACTT-3') primer, 18 μL of water, and 5μL of template. PCR conditions included a 95°C denaturation for 5 min, followed by 30 cycles at 94°C for 30 sec, 55°C for 30 s, 72°C for 1 min, and a final extension at 72°C for 10 min.

3.2.4 Random Amplification

Viral nucleic acid was amplified using a protocol adapted from Wang et al., (2003). In this method, two rounds of enzymatic reactions were used to randomly amplify the viral nucleic acid. In the first round, two cycles of first strand cDNA synthesis were performed with reverse

transcriptase SuperScript III (Invitrogen) and 40 pmol/µL primer A (GTT TCC CAG TCA CGA TCN NNN NNN NN, Eurofins Genomics, Huntsville, AL, USA). For the first cycle, 5 µL of template was added to 1 µL of primer A and 4 µL of RNase-free H_20 (QIAGEN) for a 10 µL reaction and incubated at 65°C for 5 min followed by 5 min at room temperature. The second cycle of first strand synthesis consisted of a 20 µL reaction containing 4 µL RT buffer (5x, Invitrogen), 0.5 µL of RNA-free H_20, 1.5 µL of DTT (0.1 M, Invitrogen), 1 µL dNTP (10 mM, Promega), 1 µL RNAse OUT (Invitrogen), 2 µL of SSIII Reverse Transcriptase (Invitrogen), and 10 µL of the first cycle template. PCR conditions for the second cycle included 50°C incubation for 60 min, followed by 94°C for 2 min and a 10°C hold for 5 min. In addition, second strand synthesis was performed using Sequenase (bacteriophage T7 DNA polymerase) (Affymetrix, Santa Clara, CA, USA). The Sequenase enzyme can be used for dideoxy-sequencing and is useful in that it is not impeded by secondary structures and allows strand displacement and lacks exonuclease activity. Ten microliters of Sequenase mix consisting of 2 µL of 5x Sequenase buffer, 7.7 µL of RNA-free H_20, and 0.3 µL of Sequenase were added for a total reaction volume of 30 µL. PCR conditions for second strand synthesis included an 8 min ramp (temperature cycling over a 8 min time period) from 10°C to 37°C, a 37°C hold for 8 min, rapid ramp to 94°C for 2 min, and a 10°C hold for 5 min during which 1.2 µL of diluted Sequenase (1:4) was added. The samples were then ramped to 37°C for 8 min, held at the same temperature for an additional 8 min, incubated at 94°C for 8 min, and cooled to10°C to complete the first round PCR template preparation. The second round used the previously generated template and Primer B (GTT TCC CAG TCA CGA TC, Eurofins Genomics) to amplify the viral nucleic acid. Six microliters of first round template was added to 10 µL of 10X PCR buffer, 2 µL of dNTP (10 mM), 1 µL of 100 pmol/µL Primer B, 6 µL of $MgCl_2$, 80 µL H_20, and 1 µL of Amplitaq Gold for a 100 µL

reaction volume (Applied Biosystems, Austin, TX, USA). Random amplification PCR conditions included a 95°C incubation for 15 min, followed by 40 cycles of 94°C for 30 s, 40°C for 30 s, 50°C for 30 s, and 72°C for 1 min. For each sample, three second round PCR reactions were amplified and combined for purification. Each PCR run was confirmed by running 5 µL of PCR product on a 1-2% agarose gel in which a visible smear of DNA appeared between 500 base pairs (bp) to 1 kilobase (kb). A negative control was included for the entire amplification process. The amplified products were purified using a Promega Wizard SV Gel and PCR Clean-Up System according to the manufacturer's instructions (Promega).

3.2.5 Sequencing and Bioinformatics

Sequencing was performed at the Research Technology Support Facility at Michigan State University. For irrigation water samples, libraries were prepared using an Illumina TruSeq kit (Illumina, San Diego, CA, USA) while lettuce sample libraries were prepared using a Rubicon Tenomics ThruPLEX DNA-seq kit (Rubicon Tenomics, Ann Arbor, MI, USA). Both sample types were sequenced using Illumina HiSeq 2500 Rapid Run flow cell in a 2x100 bp paired end format. Following sequencing, FastQC was used to check the quality of the Illumina sequencing reads (Babraham Bioinformatics, 2015). Cutadapt software was used for Primer B removal (Cutadapt, 2015). Cutadapt parameters included a maximum error rate of 0.2 and minimum overlap of 10 bases. Trimmomatic was used for sequencing adapter removal and quality trimming with the following parameters included: a maximum mismatch count value of 2 allowed for a full match (seed mismatch), a palindrome clip threshold of 30, a simple clip threshold of 10, a minimum adapter length of 8 with both the forward and reverse read kept, removal of low quality leading and trailing bases below a quality of 3, a 4-base sliding window scan that cuts when the average quality is below 15, and removal of reads less than 30 bases long

(Bolger, Lohse, & Usadel, 2014). Khmer script, interleave-reads.py was used to interleave the paired-end reads prior to assembly (Crusoe et al. 2014). Assembly was performed using Iterative De Bruijn Graph Assembler (IDBA-UD, Hong Kong, China) software which is an algorithm based on De Bruijn Graph *de novo* assembly and is used for sequencing data with short reads and uneven sequencing depth (Peng, Leung, Yiu, & Chin, 2012). Khmer script, extract-long-sequences.py was used to extract contigs larger than 200 base pairs (Crusoe et al. 2014). The Biopieces analyze assembly (Danish Agency for Science, Technology and Innovation) tool was used to analyze the N50, maximum, minimum, mean, total, and number of contigs for each sequence assembled (Biopieces, 2015). The N50 is a statistical measure and is defined as the contig length where using equal or longer contigs produces half the bases of the genome (Biopieces, 2015). Assembly results were analyzed using the protein database Basic Local Alignment Search Tool (BLASTx) against the National Center for Biotechnology Information (NCBI) Viral Reference Sequence (RefSeq) database. BLASTx paramaters included an expected value (E-value) of $< 10^{-5}$. The E-value is a parameter describing the number of hits that are expected by chance when searching a database. The BLASTx hits were assigned to NCBI taxonomy using the MEGAN program (version 5.10.0) with the following parameters for the Lowest Common Ancestor algorithm: minimum score 50, top percent 10, and minimum support 1 (Huson et al., 2007).

4. Results

4.1 Virus Recovery Efficiency from Lettuce

Raw data, the calculated eluent, lettuce, and inoculated phage concentrations used for calculating percent recovery for P22 and MS2 are presented in Tables 6 and 7, respectively. The calculated phage concentration in the eluent (PFU/mL), eluent volume recovered, and weight of

lettuce were used to calculate the phage concentration on the lettuce. For trials where lettuce was inoculated with 1 mL of P22 at 10^4 PFU/mL, final lettuce phage concentrations ranged from 4.4 x 10^2 PFU/g to 6.4 x 10^2 PFU/g with an l average concentration of 5.4 x 10^2 PFU/g for 3 trials. MS2 phage concentrations on romaine lettuce ranged 1.5 x 10^2 PFU/g to 6.2 x 10^2 PFU/g with an average of 4.2 x 10^2 PFU/g for the 3 trials. The inoculated phage concentration (PFU/g) was calculated by multiplying the inoculum concentration (PFU/ml) by the volume of inoculum applied and dividing by the lettuce weight. The eluent calculated PFU/g was then divided by the inoculum calculated PFU/g to determine percent recovery. Phage recovery from romaine lettuce ranged 43.3%-77.4% and 2.8%-91.1% for P22 and MS2 respectively. The average percent recovery for all P22 and MS2 inoculated lettuce trials was 57.9% and 39.0% respectively.

Table 6. Raw data and calculated eluent concentrations, lettuce concentrations, and percent recovery for P22 from inoculated romaine lettuce*

Trial	Replicate	Lettuce weight (g)	Eluent volume (mL)	Dilution	Plate 1 plaque count	Plate 2 plaque count	Plate 3 plaque count	Eluent plaque concentration (PFU/mL)**	Eluent plaque concentration sample average (PFU/mL)	Phage concentration on lettuce (PFU/g)**	Inoculated phage concentration on lettuce (PFU/g)**	% Recovery **
1	1	48.15	240.00	10^{0}	152	194	186	$8.9x10^{1}$	$1.0x10^{2}$	$5.2x10^{2}$	$1.0x10^{3}$	51.5
				10^{-1}	30	18	24	$1.2x10^{2}$				
	2	51.05	240.00	10^{0}	245	235	260	$1.2x10^{2}$	$1.3x10^{2}$	$6.2x10^{2}$	$9.5x10^{2}$	64.5
				10^{-1}	20	29	34	$1.4x10^{2}$				
	3	52.27	240.00	10^{0}	199	178	201	$9.6x10^{1}$	$1.0x10^{2}$	$4.7x10^{2}$	$9.3x10^{2}$	50.1
				10^{-1}	24	23	17	$1.1x10^{2}$				
2	1	53.15	240.00	10^{0}	174	241	214	$1.1x10^{2}$	$1.3x10^{2}$	$5.9x10^{2}$	$1.0x10^{3}$	57.9
				10^{-1}	31	27	36	$1.6x10^{2}$				
	2	53.11	240.00	10^{0}	219	237	207	$1.1x10^{2}$	$9.8x10^{1}$	$4.4x10^{2}$	$1.0x10^{3}$	43.3
				10^{-1}	22	15	14	$8.5x10^{1}$				
	3	48.09	240.00	10^{0}	217	229	217	$1.1x10^{2}$	$1.1x10^{2}$	$5.3x10^{2}$	$1.1x10^{3}$	46.6
				10^{-1}	23	19	18	$1.0x10^{2}$				
3	1	50.41	242.00	10^{0}	150	163	156	$7.8x10^{1}$	$9.8x10^{1}$	$4.7x10^{2}$	$7.9x10^{2}$	59.7
				10^{-1}	20	26	25	$1.2x10^{2}$				
	2	48.43	240.00	10^{0}	210	220	232	$1.1x10^{2}$	$1.3x10^{2}$	$6.4x10^{2}$	$8.2x10^{2}$	77.4
				10^{-1}	29	21	38	$1.5x10^{2}$				
	3	54.52	240.00	10^{0}	193	202	188	$9.7x10^{1}$	$1.2x10^{2}$	$5.9x10^{2}$	$8.4x10^{2}$	69.9
				10^{-1}	28	33	20	$1.4x10^{3}$				
Avg	NA***	NA***	NA***	NA***	NA***	NA***	NA***	NA***	$1.1x\ 10^{2}$	$5.4x\ 10^{2}$	$9.5x\ 10^{2}$	57.9

*Inoculum concentrate was $4.9x\ 10^{4}$ PFU/mL, $5.4x\ 10^{4}$ PFU/mL, and $4.0x\ 10^{4}$ PFU/mL for samples Trials 1, 2, and 3 respectively

** Calculated using equations in Part III, Section 3.1.3, pages 39-40

*** NA: not applicable

Table 7. Raw data and calculated eluent concentration, lettuce concentration, and percent recovery of MS2 from inoculated romaine lettuce*

Trial	Replicate	Lettuce weight (g)	Eluent volume (mL)	Dilution	Plate 1 plaque count	Plate 2 plaque count	Plate 3 plaque count	Eluent plaque concentration (PFU/mL)**	Eluent plaque concentration sample average (PFU/mL)	Phage concentration on lettuce** (PFU/g)	Inoculated phage concentration on lettuce** (PFU/g)	% Recovery **
1	1	51.02	237.00	10^{0}	57	70	66	$3.2x10^{1}$	$3.2x10^{1}$	$1.5x10^{2}$	$5.1x10^{2}$	29.1
	2	45.00	235.00	10^{0}	78	67	61	$3.4x10^{1}$	$3.4x10^{1}$	$1.8x10^{2}$	$5.8x10^{2}$	30.8
	3	48.36	235.00	10^{0}	79	61	80	$3.7x10^{1}$	$3.7x10^{1}$	$1.8x10^{2}$	$5.4x10^{2}$	32.9
2	1	50.00	231.00	10^{0}	171	184	179	$8.9x10^{1}$	$1.3x10^{2}$	$5.9x10^{2}$	$1.5x10^{4}$	4.0
				10^{-1}	34	33	32	$1.7x10^{2}$				
	2	50.79	231.00	10^{0}	177	180	191	$9.1x10^{1}$	$1.3x10^{2}$	$5.9x10^{2}$	$1.5x10^{4}$	4.0
				10^{-1}	35	34	32	$1.7x10^{2}$				
	3	50.26	234.00	10^{0}	166	159	145	$7.8x10^{1}$	$8.9x10^{1}$	$4.2x10^{2}$	$1.5x10^{4}$	2.8
				10^{-1}	21	23	16	$1.0x10^{2}$				
3	1	50.05	230.00	10^{0}	168	156	185	$8.5x10^{1}$	$1.0x10^{2}$	$4.8x10^{2}$	$6.9x10^{2}$	69.7
				10^{-1}	23	27	24	$1.2x10^{2}$				
	2	50.06	231.00	10^{0}	35	28	38	$9.0x10^{1}$	$1.3x10^{2}$	$6.0x10^{2}$	$6.9x10^{2}$	87.0
				10^{-1}	165	193	183	$1.7x10^{2}$				
	3	50.59	232.00	10^{0}	226	231	231	$1.6x10^{2}$	$1.4x10^{2}$	$6.2x10^{2}$	$6.8x10^{2}$	91.1
				10^{-1}	22	39	32	$1.2x10^{2}$				
Avg	NA***	NA***	NA***	NA***	NA***	NA***	NA***	NA***	$9.0x10^{1}$	$4.2x10^{2}$	$5.3x10^{3}$	39.0

* Inoculum concentrate was $2.6x\ 10^{4}$ PFU/mL, $7.4x\ 10^{5}$ PFU/mL, $3.4x\ 10^{4}$ PFU/mL for samples Trials 1, 2, and 3 respectively

** Calculated using equations in Part III, Section 3.1.3, pages 39-40

*** NA: not applicable

4.2 Yuma Irrigation Water Quality

Yuma irrigation water sampling site descriptions as well as respective temperature, pH, turbidity, and conductivity measurements are provided in Table 8. Overall, water conditions remained relatively stable between all sampling sites. Tables 9 and 10 present the Yuma irrigation water results for enterococci and *E. coli* indicator bacteria, respectively. The average enterococci concentration for all six samples was 5.2 MPN/100mL with a standard deviation of 1.9 MPN/100 mL. Average *E. coli* and total coliform levels for all six samples were 151.6 MPN/100mL and 2.6 MPN/100mL, with a standard deviation of 109.6 and 1.4 respectively.

Table 8. Yuma irrigation water sample location descriptions and conditions

Sample ID*	Location description	Temperature (^{O}C)	pH	Turbidity (NTU)	Conductivity (μS/cm)
YW1	Opening of main canal from Imperial Dam. Carries Colorado River water into the Yuma Valley.	11.3	8.41	3.4	1134
YW2	Yuma main canal at Picacho Road. Water has been carried through multiple agricultural areas.	11.5	8.44	3.6	1136
YW3	West main canal at West 2nd street and Ave B. Canal water previously suspected of septic tank contamination.	12.5	8.50	2.0	1137
YW4	Yuma main canal at West 1st street and Ave A, residential area.	13.2	8.45	1.6	1113
YW5	Yuma East Canal at Co. 18th street and Ave D, residential area.	11.8	8.41	2.3	1129
YW6	Yuma East Canal at 14th street and Ave B, near palm tree and agricultural fields.	12.3	8.44	1.8	1147

* Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6

Table 9. Yuma irrigation water enterococci levels							
Sample ID*	Dilution	Volume (mL)	# Large wells that fluoresce blue	# Small wells that fluoresce blue	Enterococci MPN/100mL	95% confidence limit**	
						Lower	Upper
YW1	10^0	100	5	3	8.4	3.7	15.3
YW2	10^0	100	3	2	5.1	1.7	10.6
YW3	10^0	100	4	0	4.1	1.7	9.5
YW4	10^0	100	6	0	6.3	2.9	13.7
YW5	10^0	100	4	0	4.1	1.7	9.5
YW6	10^0	100	3	0	3.1	0.7	8.9
Average**	NA***	NA***	4.2	2.5	5.2	3.3	7.1

* Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6
**Standard deviation = 1.91
***NA: not applicable

Table 10. Yuma irrigation water total coliform and *E. coli* levels												
Sample ID*	Dilution	Sample volume (mL)	# Large wells that are yellow	# Small wells that are yellow	# Large wells that fluoresce blue	# Small wells that fluoresce blue	Coliform MPN/100 mL	95% confidence limit**		*E.coli* MPN/100mL	95% confidence limit**	
								Lower	Upper		Lower	Upper
YW1	10^0	100	47	12	4	0	172.3	119.5	242.2	4.1	1.7	9.5
YW2	10^0	100	49	21	3	0	365.4	231.9	555.5	3.1	0.7	8.9
YW3	10^0	100	38	11	4	0	91.0	66.6	121.0	4.1	1.7	9.5
YW4	10^0	100	43	3	2	0	96.0	68.5	132.1	2.0	0.3	7.1
YW5	10^0	100	40	4	1	0	83.3	59.4	114.6	1.0	0.1	5.5
YW6	10^0	100	41	9	1	0	101.4	74.3	136.1	1.0	0.1	5.5
Average**	NA***	NA***	43	10	2.5	0	151.6	49.1	261.2	2.6	1.1	4.0

* Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6
** Coliform standard deviation = 109.62; *E.coli* standard deviation = 1.43
*** NA: not applicable

4.3 Irrigation Water Metagenomic Statistics

Table 11 shows the number of sequence reads following quality trimming, N50 statistic, and assembled contiguous sequence (contig) information for each Yuma irrigation water sample. The irrigation water virome resulted in 21.2 to 66.0 million sequence reads following quality trimming and 18.5 to 127.8 thousand assembled contigs larger than 200 bp. Average total assembly length for all samples was 36.2 million base pairs with a maximum contig size range of 12,172 to 61,978 bp.

Table 11. Yuma irrigation water sequence statistics following trimming and assembly

Sample*	# Sequence reads following trimming	# Contigs (> 200bp)	Total assembly length (M bp)	Max contig size (bp)	N50**
YW1	65,987,608	127,880	83.4	21,123	717
YW2	53,935,130	22,383	14.3	12,651	680
YW3	32,092,020	18,536	10.6	12,172	573
YW4	41,080,774	27,580	16.6	14,587	627
YW5	21,258,660	77,348	52.9	61,978	761
YW6	24,880,404	59,759	39.5	19,794	730
Average	39,872,433	55,581	36	23,718	681

*Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6
**N50: genome assembly contig statistic

4.4 Irrigation Water Virome

Table 12 and Figures 1-6 show the distribution of viral contigs (≥200 bp) for each irrigation water virome. For all six samples, the majority of contigs could not be annotated against the NCBI RefSeq viral database. The percentage of contigs with no hits ranged 64.5-84.5% while the percentage of those not assigned ranged 4.3-7.4% (Figures 1-6). Of the assigned viral sequence, the average majority (69.6%) of contigs were assigned as dsDNA viruses while no more than 0.2% shared sequence similarities with dsRNA viruses (Figures 1-6). Single-

stranded DNA and ssRNA viral sequences composed 5.3-31.3 and 1.7-12.3% of the irrigation water virome respectively.

Table 12. Distribution of contigs larger than 200 bp for Yuma irrigation water virome

Genome*	Samples**												
	YW1 # contigs	YW 1 (%)	YW2 # contigs	YW 2 (%)	YW3 # contigs	YW 3 (%)	YW4 # contigs	YW 4 (%)	YW5 # contigs	YW5 (%)	YW6 # contigs	YW 6 (%)	Avg (%)
No Hits	107,829	84.5	14,533	65.4	13,022	70.5	19,023	69.1	49,695	64.5	40,505	67.9	70.3
Not Assigned	5,491	4.3	1,477	6.7	1,144	6.2	1,791	6.5	5,695	7.4	4,145	6.9	6.3
Assigned	14,337	11.2	6,196	27.9	4,308	23.3	6,711	24.4	21,680	28.1	14,991	25.4	23.4
dsDNA	9,260	64.6	2,593	41.8	3,294	76.5	5,498	81.9	14,938	68.9	12,597	84.0	69.6
dsRNA	21	0.1	13	0.2	2	0.0	3	0.0	1	0.0	9	0.1	0.1
Retro-transcribing	6	0.0	0	0.0	3	0.1	0	0.0	2	0.0	3	0.0	0.0
Satellites	34	0.2	27	0.4	11	0.3	11	0.2	29	0.1	17	0.1	0.2
ssDNA	2,381	16.6	1,936	31.2	444	10.3	359	5.3	3,618	16.7	792	5.3	14.3
ssRNA	880	6.1	764	12.3	175	4.1	240	3.6	854	3.9	253	1.7	5.3
Unclassified	1,755	12.2	863	13.9	379	8.8	600	8.9	2,238	10.3	1,320	8.8	10.5

*Contigs assigned to viral taxa but did not meet the selected MEGAN parameters are “Not assigned,” and contigs without any hits to known sequences in the databases were placed under “No hits.”

**Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6

Figure 1. Distribution of contigs larger than 200bp for Yuma irrigation water 1 virome

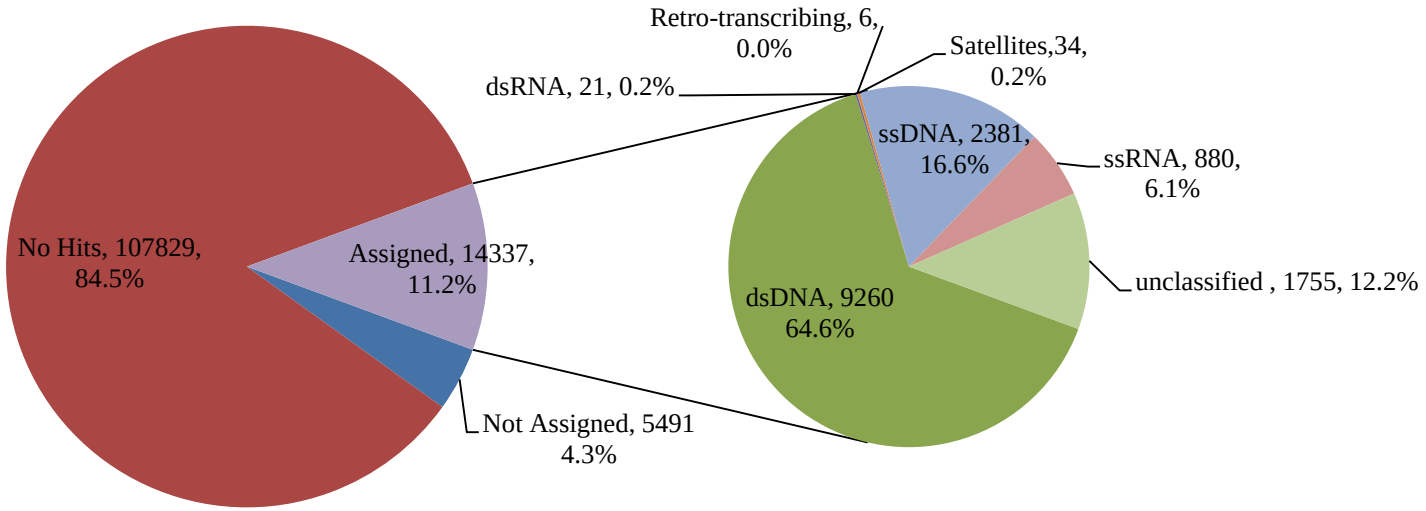


Figure 2. Distribution of contigs larger than 200bp for Yuma irrigation water 2 virome

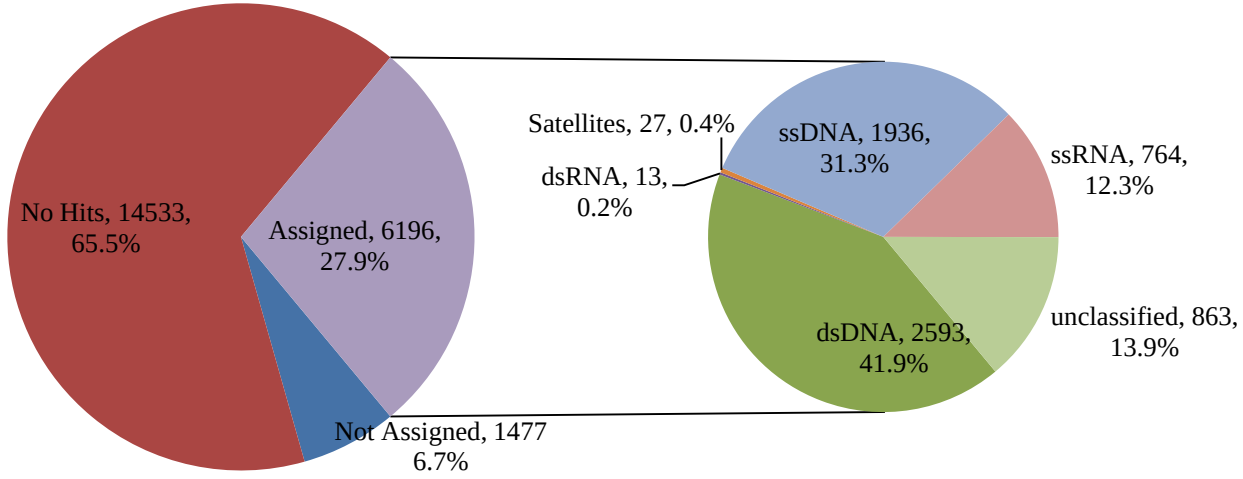


Figure 3. Distribution of contigs larger than 200bp for Yuma irrigation water 3 virome

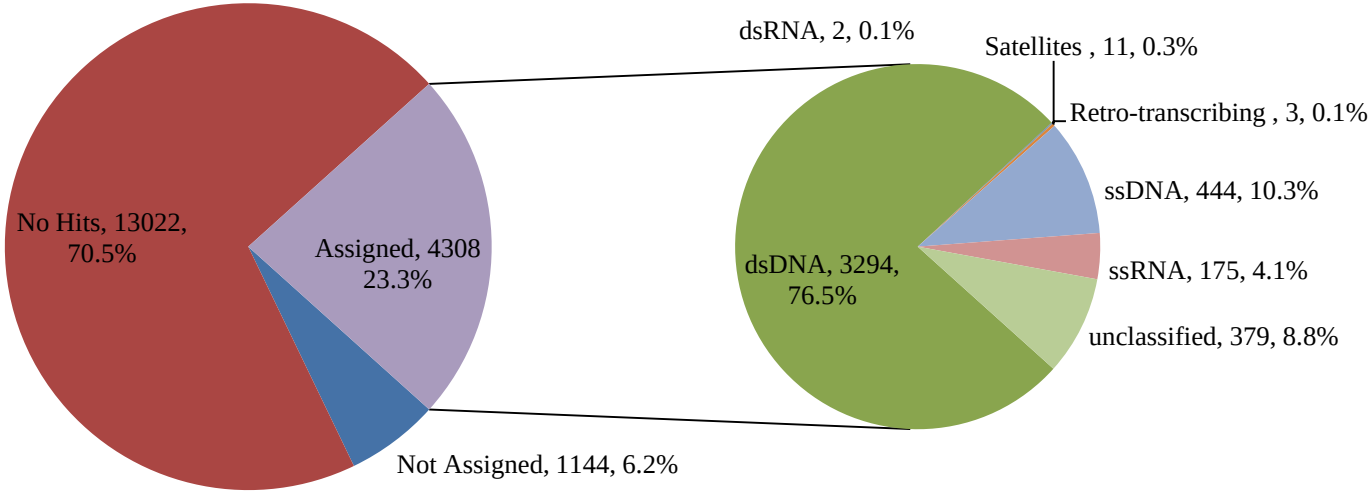


Figure 4. Distribution of contigs larger than 200bp for Yuma irrigation water 4 virome

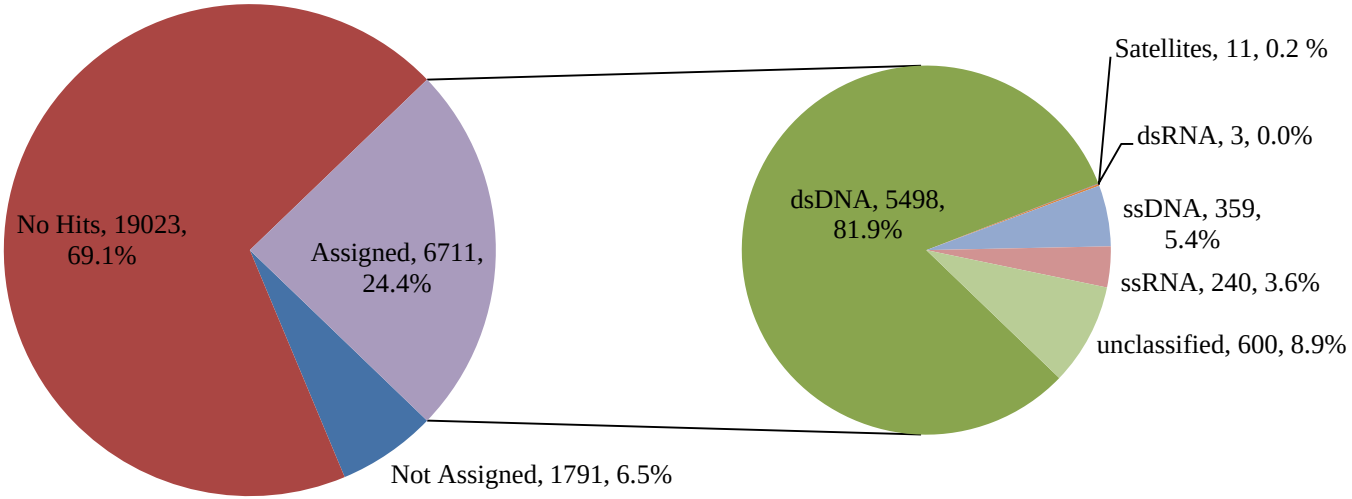


Figure 5. Distribution of contigs larger than 200bp for Yuma irrigation water 5 virome

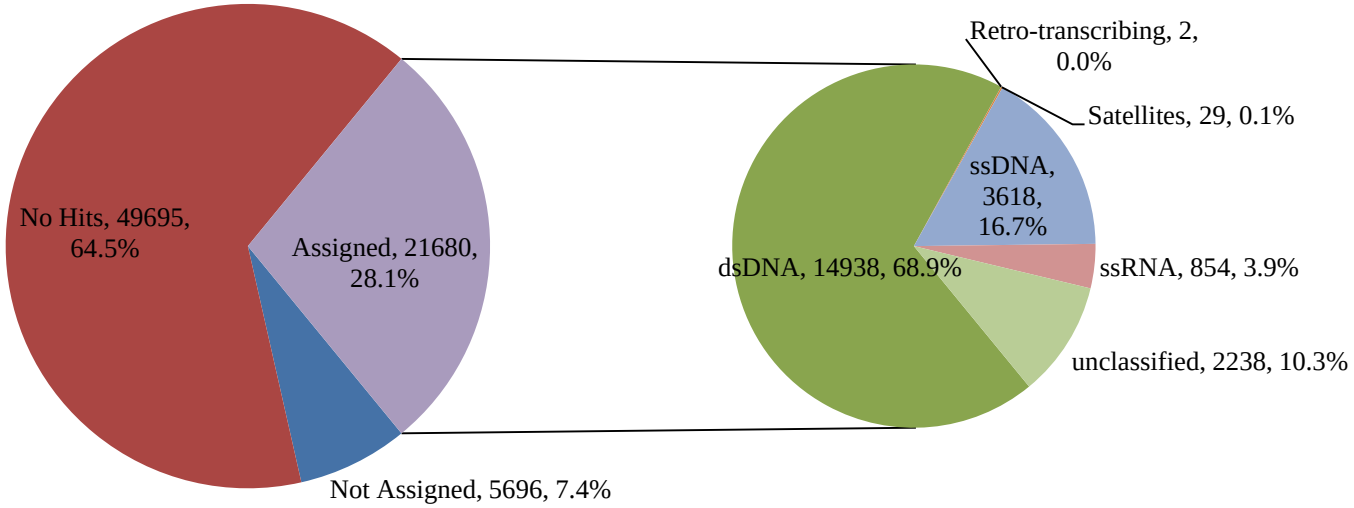
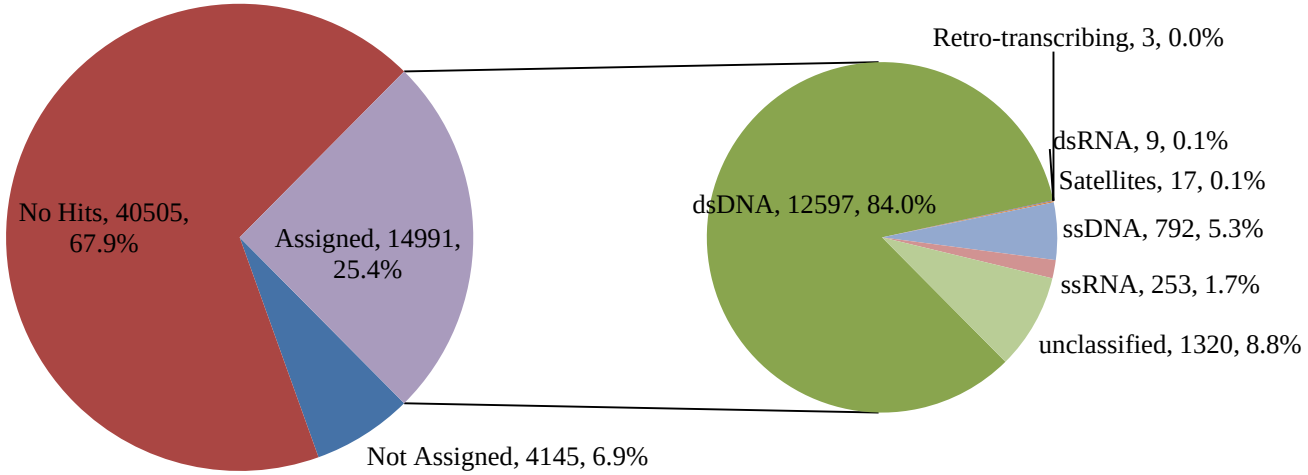


Figure 6. Distribution of contigs larger than 200bp for Yuma irrigation water 6 virome

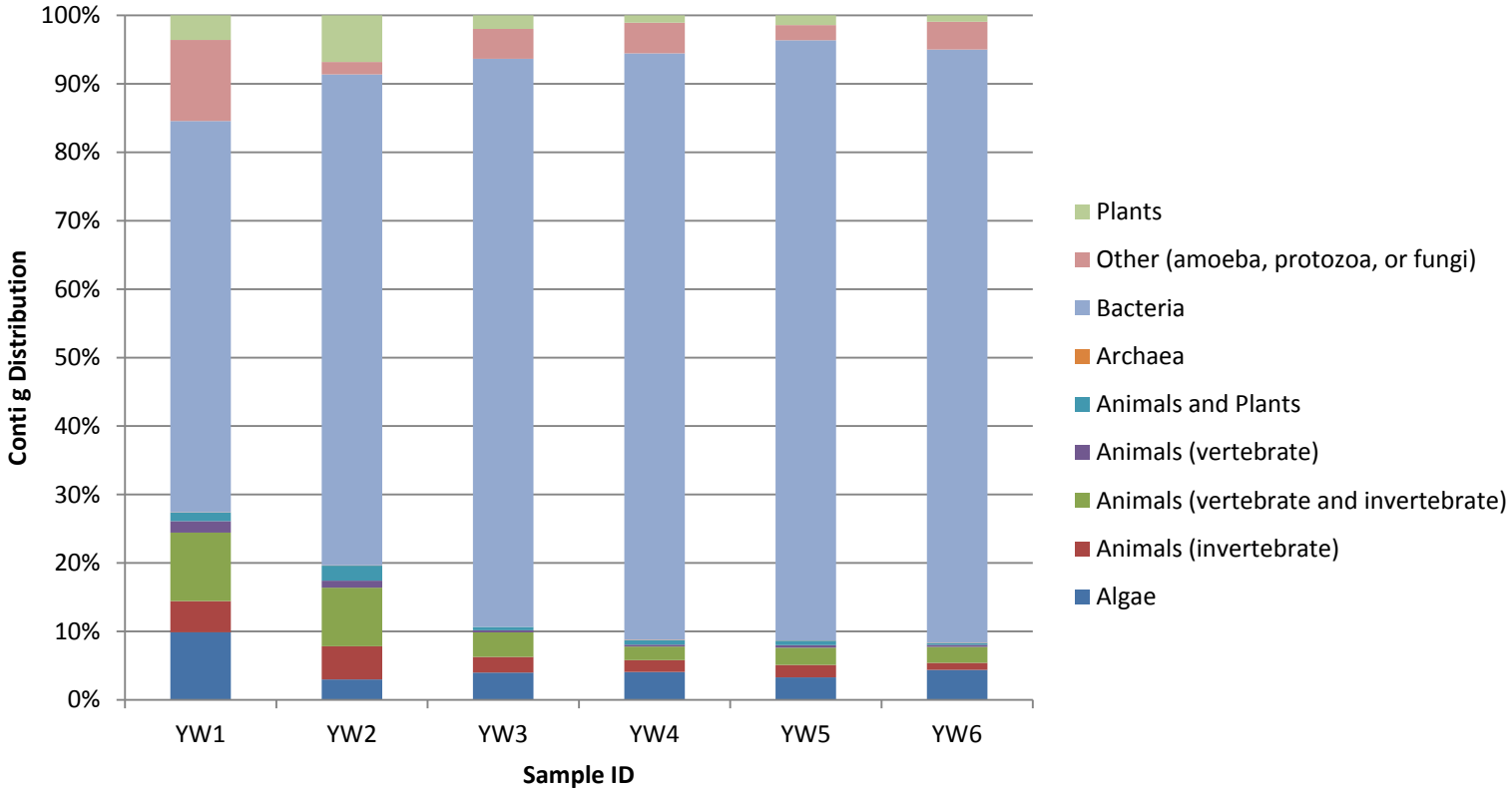


The distribution of viral host species for each irrigation water virome is shown in Table 13 and Figure 7. Contigs assigned to viral orders and families were classified as algae, animals (vertebrates only, invertebrates only, vertebrate and invertebrates), animals and plants, archaea, bacteria, or as other (amoeba, protozoa, and fungi). Bacteriophage dominated in all 6 irrigation water samples, comprising on average 78.7% (range 57.2-87.7%) of the irrigation water virome. Of the remaining single host categories, algae (4.8%), plants (2.6%), invertebrate animals (2.7%), and vertebrate animals (0.7%) comprised a far smaller average percentage of the irrigation water virome. The average percentage of viral sequences assigned to eukaryotic, multiple host categories including vertebrate and invertebrate animals, animals and plants, as well as eukaryotic amoeba, protozoa, and fungi (classified as other) was 4.8%, 0.9%, 4.8%, respectively.

Table 13. Yuma irrigation water viral host distribution													
	Sample*												
Host	YW1 # contigs	YW1 (%)	YW2 # contigs	YW2 (%)	YW3 # contigs	YW3 (%)	YW4 # contigs	YW4 (%)	YW5 # contigs	YW5 (%)	YW6 # contigs	YW6 (%)	Avg (%)
Algae	1,004	9.9	134	3.0	135	4.0	219	4.1	566	3.3	523	4.4	4.8
Animals (invertebrate)	460	4.5	217	4.8	77	2.3	92	1.7	310	1.8	122	1.0	2.7
Animals (vertebrate and invertebrate)	1,017	10.0	384	8.6	121	3.6	107	2.0	433	2.5	279	2.3	4.8
Animals (vertebrate)	170	1.7	46	1.0	11	0.3	15	0.3	62	0.4	38	0.3	0.7
Animals and Plants	125	1.2	99	2.2	16	0.5	34	0.6	113	0.7	30	0.3	0.9
Archaea	4	0.0	2	0.0	0	0.0	3	0.1	6	0.0	4	0.0	0.0
Bacteria	5,804	57.2	3,216	71.7	2,806	83.0	4,598	85.7	15,046	87.7	10,349	86.7	78.7
Other (amoeba, protozoa, or fungi)	1,203	11.9	82	1.8	147	4.3	240	4.5	380	2.2	485	4.1	4.8
Plants	364	3.6	305	6.8	67	2.0	58	1.1	242	1.4	112	0.9	2.6

*Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6

Figure 7. Viral host distribution in six Yuma irrgation water samples



A total of 64 viral families were observed for all 6 irrigation water samples. The viral family contig distribution can be viewed in a phylogenetic tree constructed for all six irrigation water samples using MEGAN software (Figure 8). The number of assigned contigs in this figure is represented by circle size and colors show the proportion belonging to each sample. A large proportion of viruses for all 6 samples were assigned to *Myoviridae, Podoviridae, Siphoviridae,* and *Microviridae* bacteriophage families. The distribution of viral contigs associated with families composing ≥ 3.0% across the six irrigation water viromes was further analyzed (Table 14 and Figure 9). Of the viral contigs composing ≥3.0% of the irrigation water virome, a majority were assigned to *Myoviridae* (28.6%), *Podoviridae* (16.5%) and *Siphoviridae* (20.5%) dsDNA bacteriophage families of the *Caudovirales* order, however the *Microviridae* (15.7%) bacteriophage family was prevalent in the irrigation water virome as well. Other viral families that were well represented in irrigation water included those infecting animals (*Circoviridae*, *Dicistroviridae*, and *Poxviridae*), algae (*Phycodnaviridae*), and amoeba (*Mimiviridae*). The contig distribution of phage host species is provided in Table 15. Analysis of the bacteriophage host species revealed a total of 88 bacterial hosts. Host species representing ≥3.0% across the 6 irrigation water viromes included *Pelagibacter* (11.9-16.5%), *Cellulophaga* (10.3-13.6%), *Synechococcus* (12.7-13.2%), *Bdellovibrio* (0.9-8.3%), *Bacillus* (3.1-4.4%), *Puniceispirillium* (2.6-3.8%), *Mycobacterium* (2.0-3.7%), and *Prochlorococcus* (2.8-3.2%). Phage infecting potential bacterial foodborne pathogens including *Campylobacter* (0.7%), *Enterococcus* (0.3%), *Escherichia* (1.9%), *Listeria* (0.1%), *Salmonella* (2.1%), and *Shigella* (0.0%) on average composed a much lower percentage of the bacterial host species.

Figure 8. Yuma irrigation water viral family distribution

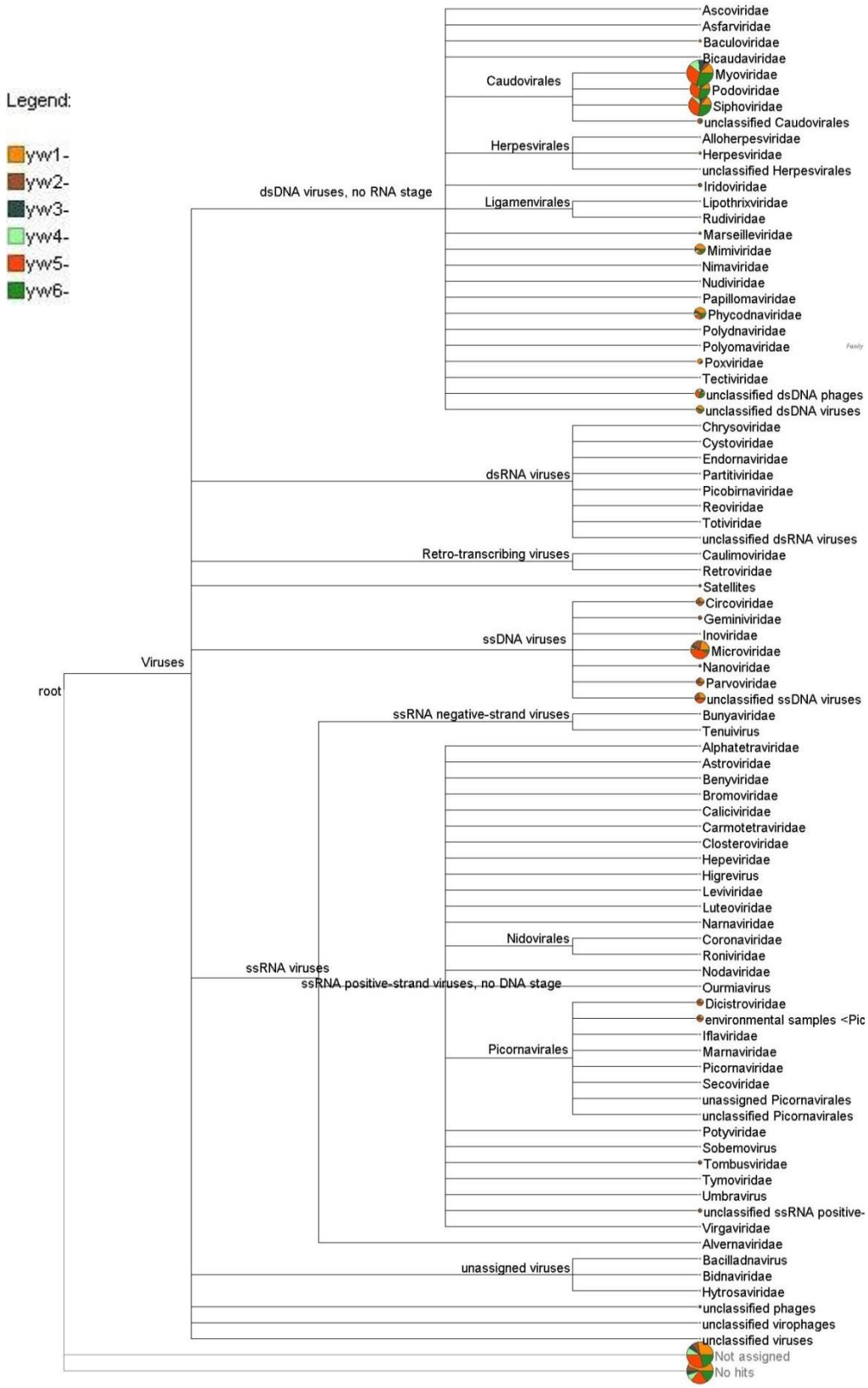


Table 14. Irrigation water viral family distribution representing greater than 3.0% of the virome													
	Sample*												
Family	YW1 # contigs	YW1 (%)	YW2 # contigs	YW2 (%)	YW3 # contigs	YW3 (%)	YW4 # contigs	YW4 (%)	YW5 # contigs	YW5 (%)	YW6 # contigs	YW6 (%)	Avg (%)
Circoviridae	350	4.4	304	8.8	75	2.8	58	1.4	257	1.9	147	1.6	3.5
Dicistroviridae	198	2.5	188	5.4	60	2.3	56	1.3	234	1.7	65	0.7	2.3
Microviridae	1,471	18.5	1,174	33.9	255	9.6	233	5.6	2,900	21.6	440	4.8	15.7
Mimiviridae	1,080	13.6	67	1.9	136	5.1	225	5.4	354	2.6	458	5.0	5.6
Myoviridae	1,460	18.4	683	19.7	894	33.8	1,498	36.1	3,595	26.7	3,400	36.8	28.6
Phycodnaviridae	997	12.6	131	3.8	133	5.0	216	5.2	560	4.2	517	5.6	6.1
Podoviridae	770	9.7	416	12.0	473	17.9	842	20.3	2,488	18.5	1,820	19.7	16.3
Poxviridae	503	6.3	7	0.2	12	0.5	16	0.4	45	0.3	32	0.3	1.3
Siphoviridae	1,111	14.0	489	14.1	607	22.9	1,008	24.3	3,019	22.4	2,352	25.5	20.5

*Yuma irrigation water (YW) sample locations 1, 2, 3, 4, 5, and 6

Figure 9. Irrigation water viral family distribution representing greater than 3.0% of the virome

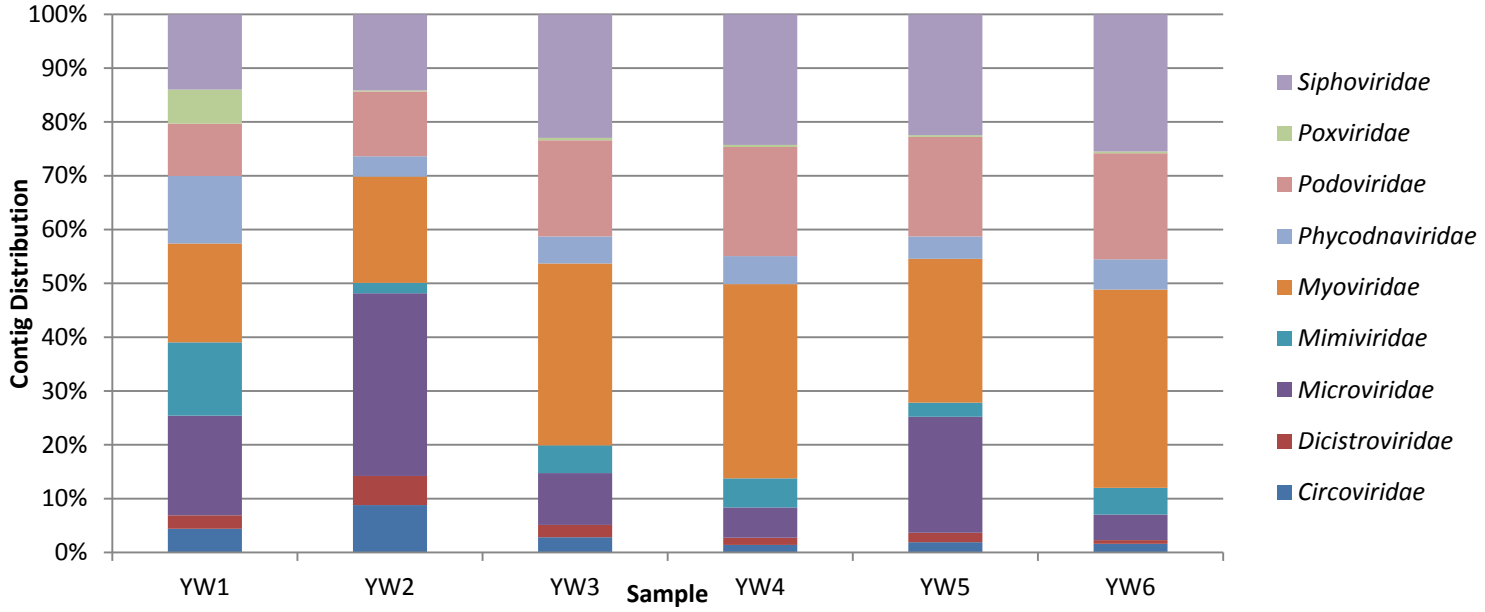


Table 15. Bacteriophage host distribution of Yuma irrigation water samples							
Host	YW1 (%)	YW2 (%)	YW3 (%)	YW4 (%)	YW5 (%)	YW6 (%)	Average (%)
Acinetobacter	0.3	0.3	0.5	0.5	0.6	0.6	0.5
Actinomyces	0.1	0.1	0.1	0.3	0.0	0.0	0.1
Actinoplanes	0.2	0.5	0.3	0.0	0.3	0.2	0.2
Aeromonas	0.2	0.2	0.4	0.3	0.4	0.3	0.3
Aggregatibacter	0.1	0.2	0.1	0.2	0.1	0.1	0.1
Agrobacterium	0.2	0.0	0.0	0.1	0.1	0.1	0.1
Altermonas	0.6	0.4	0.2	0.5	0.5	0.5	0.4
Arthrobacter	0.1	0.2	0.0	0.2	0.2	0.2	0.1
Azospirillum	0.2	0.4	0.4	0.2	0.4	0.2	0.3
Bacillus	**3.7**	**3.7**	**4.4**	**4.2**	**3.1**	**3.5**	**3.8**
Bacteroids	0.1	0.0	0.1	0.0	0.1	0.1	0.1
Bdellovibrio	**5.9**	**8.3**	**1.3**	**1.1**	**6.0**	**0.9**	**3.9**
Brochothrix	0.1	0.1	0.1	0.1	0.0	0.0	0.1
Burkholderia	1.6	1.5	1.4	1.1	1.6	1.7	1.5
Candidatus	0.1	0.0	0.0	0.1	0.0	0.0	0.0
Campylobacter	0.9	0.9	0.7	0.6	0.6	0.4	0.7
Caulobacter	1.2	0.4	0.7	0.6	0.9	0.7	0.8
Cellulophaga	**11.2**	**13.6**	**10.3**	**13.2**	**11.4**	**10.5**	**11.7**
Chlamydia	0.7	0.8	0.1	0.1	0.5	0.1	0.4
Clavibacter	0.1	0.0	0.1	0.0	0.3	0.2	0.1
Clostridium	0.7	0.7	1.2	0.7	0.7	0.8	0.8
Colwellia	0.1	0.2	0.3	0.1	0.1	0.1	0.1
Corynebacterium	0.2	0.4	0.3	0.2	0.2	0.2	0.2
Croceibacter	0.3	0.3	0.3	0.3	0.3	0.4	0.3
Cronobacter	2.4	2.2	2.1	2.3	1.7	2.1	2.1
Deftia	0.3	0.3	0.2	0.3	0.2	0.3	0.2
Dickeya	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Edwardsiella	0.1	0.0	0.1	0.1	0.1	0.2	0.1
Enterobacter	0.1	0.3	0.2	0.1	0.1	0.1	0.1
Enterococcus	0.3	0.4	0.1	0.2	0.5	0.2	0.3
Erwinia	0.6	0.1	0.2	0.2	0.3	0.2	0.2
Escherichia	2.5	1.4	1.6	2.0	2.0	1.8	1.9
Flavobacterium	1.2	0.9	1.3	1.6	1.4	1.4	1.3
Gordonia	0.1	0.0	0.0	0.0	0.1	0.0	0.0
Haemophilus	0.0	0.1	0.0	0.0	0.0	0.1	0.0
Iodobacter	0.1	0.1	0.2	0.1	0.1	0.1	0.1
Klebsiella	0.9	0.4	0.3	0.6	0.5	0.6	0.5
Lactobacillus	0.7	0.6	0.6	0.5	0.9	0.8	0.7
Lactococcus	1.7	0.8	1.5	1.0	0.8	1.1	1.2
Listeria	0.1	0.1	0.1	0.0	0.3	0.2	0.1
Listonella	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Methanobacterium	0.1	0.0	0.1	0.0	0.0	0.0	0.0
Methanothermobacter	0.0	0.0	0.0	0.1	0.0	0.0	0.0
Microbacterium	0.5	0.5	0.4	0.4	0.2	0.4	0.4
Mycobacterium	**3.7**	**2.1**	**3.6**	**2.5**	**2.0**	**2.2**	**2.7**
Myxococcus	1.0	1.3	0.7	0.6	1.6	0.8	1.0

Table 15 (cont'd)							
Natrialba	0.1	0.0	0.3	0.2	0.2	0.3	0.2
Nocardia	0.1	0.1	0.1	0.1	0.0	0.1	0.1
Paenibacillus	0.0	0.1	0.2	0.1	0.1	0.1	0.1
Pantoea	0.1	0.1	0.3	0.2	0.1	0.1	0.1
Pasteurella	0.0	0.0	0.1	0.0	0.0	0.0	0.0
Pectobacterium	0.1	0.1	0.3	0.2	0.2	0.1	0.2
Pelagibacter	**11.9**	**13.2**	**16.5**	**14.9**	**12.5**	**14.7**	**13.9**
Phormidium	0.0	0.0	0.1	0.0	0.1	0.0	0.0
Planktothrix	1.9	1.8	2.2	3.1	1.6	3.0	2.3
Prochlorococcus	**3.5**	**2.8**	**3.1**	**3.6**	**3.3**	**3.9**	**3.4**
Propionibacterium	0.0	0.1	0.1	0.1	0.0	0.0	0.1
Pseudoalteromonas	0.4	0.3	0.3	0.3	0.3	0.4	0.3
Pseudomonas	**2.5**	**2.6**	**3.0**	**3.2**	**2.8**	**2.9**	**2.8**
Psychrobacter	0.4	0.6	0.9	0.5	0.7	0.6	0.6
Puniceispirillum	**3.8**	**2.6**	**3.7**	**3.5**	**4.4**	**4.7**	**3.8**
Ralstonia	0.7	0.3	1.2	1.0	1.4	1.3	1.0
Rhizobium	0.5	1.0	1.1	1.1	0.8	1.0	0.9
Rhodobacter	0.0	0.1	0.1	0.1	0.0	0.0	0.1
Rhodococcus	1.0	0.7	0.5	0.7	0.8	0.8	0.8
Rhodothermus	2.2	3.4	2.5	2.2	2.0	2.0	2.4
Riemerella	0.3	0.5	0.4	0.4	0.6	0.4	0.4
Roseobacter	0.2	0.5	0.5	0.5	0.5	0.4	0.4
Salinivibrio	0.1	0.1	0.3	0.2	0.2	0.1	0.2
Salmonella	1.7	2.1	2.5	2.7	1.8	1.8	2.1
Serratia	0.1	0.0	0.0	0.2	0.0	0.0	0.1
Shigella	0.1	0.0	0.0	0.1	0.1	0.0	0.0
Sodalis	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Sphingomonas	1.1	0.3	1.2	1.4	0.5	0.7	0.9
Spiroplasma	0.7	0.5	0.1	0.0	0.2	0.0	0.3
Staphylococcus	0.4	0.5	0.1	0.3	0.1	0.1	0.2
Stenotrophomonas	0.1	0.0	0.2	0.2	0.1	0.1	0.1
Streptococcus	0.7	0.5	0.5	0.6	0.7	0.9	0.7
Streptomyces	0.8	0.5	0.5	0.6	0.5	0.9	0.6
Synechococcus	**13.2**	**12.7**	**14.4**	**15.6**	**15.3**	**19.0**	**15.0**
Thalassomonas	0.2	0.5	0.5	0.7	0.7	0.3	0.5
Thermoanaerobacterium	0.7	0.3	0.4	0.5	1.1	0.5	0.6
Thermus	0.1	0.0	0.0	0.1	0.0	0.0	0.0
Tsukamurella	0.0	0.0	0.1	0.2	0.1	0.1	0.1
Vibrio	1.6	2.2	2.5	2.4	2.6	2.4	2.3
Xanthomonas	0.4	0.5	0.5	0.2	0.2	0.4	0.4
Xylella	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Yersinia	0.3	0.5	0.9	0.7	0.4	0.7	0.6
Unknown host genus*	3.5	4.0	1.6	1.4	2.0	1.1	2.3

* Viral contigs assigned as phage but are not assigned or classified into single host category

** Bolded numbers represent phage species that compose ≥3.0% of any individual Yuma irrigation water virome

For all six irrigation water samples, species of interest to the food industry belonging to 30 virus families were grouped into 5 categories based on the hosts infected including agricultural insect pests, crops, commercial fish or shrimp (fresh and salt water), livestock, and human pathogens. A summary of virus families, hosts, possible species of interest to the food industry, and % contig identities are presented in Table 16. Virus families that infect agricultural insect pests included *Alphatetraviridae*, *Ascoviridae*, *Baculoviridae*, *Dicistroviridae*, *Iflaviridae*, *Nudiviridae*, *Parvoviridae*, and *Poxviridae*. Recovered virus hosts including the beet armyworm, corn earworm, green peach aphid, cabbage looper, and alfalfa looper are all known agricultural insect pests that damage lettuce. Many popular plant and crop viruses belonged to single-stranded RNA families including *Benyviridae*, *Bromoviridae*, *Closteroviridae*, *Secoviridae*, *Tombusviridae*, *Tymoviridae*, and *Virgaviridae*. In addition, irrigation water contained numerous crop single-stranded DNA viral pathogens belonging to the *Geminiviridae* and *Nanoviridae* families. Viruses infecting commercial fish and shrimp were also recovered, represented by the *Circoviridae*, *Hepeviridae*, *Iridoviridae*, *Nimaviridae*, *Roniviridae*, and *Totiviridae* virus families. Relevant livestock pathogens infecting pigs (*African swine fever virus*, pseudorabies, *swinepox* virus, and *porcine circovirus* (type 1/2a), *parvovirus*, *astrovirus*, and *teschovirus*), cow (*cowpox*, *bovine papular stomatitis virus*, *bovine hungarovirus 1*, *Enterovirus F*), and poultry (*fowlpox* and *turkey hepatitis virus 2993D*) belonged to the *Asfarviridae*, *Astroviridae*, *Circoviridae*, *Herpesviridae*, *Parvoviridae*, *Picornaviridae*, *and Poxviridae* virus families. It is important to note that the *Iridoviridae*, *Poxviridae*, *Circoviridae*, and *Parvoviridae* viral families infect a wide variety of both vertebrate and invertebrate animal hosts; therefore host infection will vary with viral species. Human viral pathogens including *picobirnavirus*, *coronavirus*, *parechovirus*, and *rhinovirus* as well as hepatitis A and E virus

which are commonly transmitted through fresh produce were identified in the irrigation samples as well. A majority of virus species had a low percent contig identity with the exceptions of *Spodoptera exigua* iflavirus 1 (93-100%), Tobacco necrosis virus D (93-97%), and Cucumber green mottle mosaic virus (96%).

Table 16. Viral families and species of interest identified in irrigation water

Species relevance	Species of interest	Virus family	Species contig identity range (%)
Agricultural insect pest	*Helicoverpa armigera* stunt virus (Old World cotton bollworm)	*Alphatetraviridae*	21
	Spodoptera frugiperda ascovirus 1a (fall armyworm)	*Ascoviridae*	29
	Trichoplusia ni ascovirus 2c (cabbage looper)		22-44
	Heliothis virescens ascovirus 3a (tobacco budworm)		18
	Agrotis segetum granulovirus (turnip moth)	*Baculoviridae*	22
	Anticarsia gemmatalis nucleopolyhedrovirus (velvetbean caterpillar)		29
	Autographa californica nucleopolyhedrovirus (alfalfa looper)		29
	Clanis bilineata nucleopolyhedrovirus (soybean pest)		33-36
	Cryptophlebia leucotreta granulovirus (false codling moth)		22-27
	Cydia pomonella granulovirus (codling moth)		40
	Phthorimaea operculella granulovirus (potato tuber moth)		21-31
	Plutella xylostella granulovirus (diamondback moth)		45
	Pseudaletia unipuncta granulovirus (white-speck moth)		29
	Pieris rapae granulovirus (cabbage butterfly)		46
	Spodoptera litura nucleopolyhedrovirus (oriental leafworm moth)		29
	Trichoplusia ni single nucleopolyhedrovirus (cabbage looper)		23
	Himetobi P virus (Small brown planthopper)	*Dicistroviridae*	21-46
	Homalodisca coagulata virus-1 (glassy-winged sharpshooter)		21-68
	Rhopalosiphum padi virus (bird cherry oat aphid)		19-98
	Solenopsis invicta virus-1 (imported fire ant)		23-45
	Brevicoryne brassicae picorna-like virus (cabbage aphid)	*Iflaviridae*	27-34

Table 16 (cont'd)

	Nilaparvata lugens honeydew virus-2 (brown planthopper)		26-29
	Spodoptera exigua iflavirus 1 (beet armyworm)		93-100
	Helicoverpa zea nudivirus 2 (corn earworm)	*Nudiviridae*	21-60
	Helicoverpa armigera densovirus (Old World cotton bollworm)	*Parvoviridae**	26-40
	Myzus persicae densovirus (green peach aphid)		28
	Planococcus citri densovirus (citrius mealybug)		38-86
	Adoxophyes honmai entomopoxvirus 'L' (smaller tea tortix)	*Poxviridae**	26
	Amsacta moorei entomopoxvirus 'L' (tiger moth)		23-42
	Choristoneura rosaceana entomopoxvirus 'L' (oblique banded leafroller)		26,29
	Melanoplus sanguinipes entomopoxvirus (migratory grasshopper)		21-47
	Mythimna separata entomopoxvirus 'L' (northern armyworm)		24-30
Crop virus	Beet soil-borne mosaic virus	*Benyvirus* (unassigned family)	23-32
	Prune dwarf virus	*Bromoviridae*	36
	Grapevine rootstock stem lesion associated virus	*Closteroviridae*	30
	Persea americana endornavirus (avacado)	*Endornaviridae*	25
	Chickpea chlorosis	*Geminiviridae*	24
	Chickpea redleaf virus		24-32
	Citrus chlorotic dwarf associated virus		35-42
	Maize streak Reunion virus		24
	Melon chlorotic mosaic virus		34
	Papaya leaf curl China virus		38
	Pepper golden mosaic virus		34
	Tomato mild mosaic virus		49
	Tomato yellow leaf curl virus		32
	Soybean chlorotic spot virus		28
	Watermelon chlorotic stunt virus		31
	Wheat dwarf virus		27-37
	Abaca bunchy top virus	*Nanoviridae*	41
	Banana bunchy top virus		39-43
	Pea necrotic yellow dwarf virus		33
	Cassava virus C	*Ourmiavirus* (unassigned family)	23-36
	Epirus cherry virus		30-46
	Ourmia melon virus		21-43
	Blackcurrant reversion virus	*Secoviridae*	32

Table 16 (cont'd)

	Maize chlorotic dwarf virus		35
	Satsuma dwarf virus		41
	Strawberry latent ringspot virus		33
	Rice tungro spherical virus		25-34
	Tomato ringspot virus		25
	Southern bean mosaic virus	*Sobemovirus* (unassigned family)	79
	Rice grassy stunt virus	*Tenuvirus* (unassigned family)	32
	Beet black scorch virus	*Tombusviridae*	24
	Cucumber bulgarian virus		50-72
	Cucumber leaf spot virus		27-54
	Cucumber necrosis virus		80
	Maize chlorotic mottle virus		34-46
	Maize white line mosaic virus		29
	Melon necrotic spot virus		57,59
	Oat chlorotic stunt virus		32-71
	Olive mild mosaic virus		87
	Soybean yellow mottle mosaic virus		34
	Tobacco necrosis virus D		93,97
	Grapevine fleck virus	*Tymoviridae*	28
	Carrot mottle mimic virus	*Umbravirus* (unassigned family)	33
	Cucumber green mottle mosaic virus	*Virgaviridae*	96
Fish and shrimp pathogens	*Penaeus monodon* circovirus VN11	*Circoviridae**	29-54
	Cutthroat trout virus	*Hepeviridae*	19-34
	Lymphocystis disease virus 1 (European flounder and plaice)	*Iridoviridae**	18
	Lymphocystis disease virus - isolate China (flounder)		25-50
	Infectious spleen and kidney necrosis virus (fish)		20-48
	Singapore grouper iridovirus		20-42
	Shrimp white spot syndrome virus	*Nimaviridae*	26-57
	Gill-associated virus (black tiger prawn)	*Roniviridae*	27-30
	Penaeid shrimp infectious myonecrosis virus	*Totiviridae*	40
Human pathogen	Human coronavirus HKU1	*Coronaviridae*	39
	Hepatitis E virus	*Hepeviridae*	23-34
	Human Picobirnavirus	*Picobirnaviridae*	26-34
	Hepatitis A virus	*Picornaviridae*	26-34
	Human parechovirus		26

Table 16 (cont'd)

	Rhinovirus C		22
Livestock pathogen	African swine fever virus	*Asfarviridae*	27-30
	Porcine astrovirus 3	*Astroviridae*	24
	Porcine circovirus type 1/2a	*Circoviridae**	31-47
	Suid herpesvirus 1 (pseudorabies)	*Herpesviridae*	27
	Porcine parvovirus	*Parvoviridae**	32,43
	Bovine hungarovirus 1	*Picornaviridae*	50
	Enterovirus F (bovine)		34
	Porcine teschovirus		36
	Turkey hepatitis virus 2993D		36
	Bovine papular stomatitis virus	*Poxviridae**	39
	Cowpox virus		23-35
	Fowlpox virus		25-37
	Swinepox virus		24-30

*Viral families with wide host range, host depends on specific viral species

4.5 Yuma Lettuce Virome Statistics

Table 17 shows the number of sequence reads following quality trimming, N50 statistic, and assembled contiguous sequence (contig) information for Yuma lettuce samples. The romaine lettuce virome resulted in ~14.3 to 40.3 million sequence reads following quality trimming and 887 to 3,491 assembled contigs larger than 200 bp. Total assembly length for all romaine lettuce samples was 1.2 million base pairs with a maximum contig size range of 5,875 to 59,171 bp. The iceberg lettuce virome resulted in ~15.5 to 45.2 million sequence reads following quality trimming and 703 to 5,577 assembled contigs larger than 200 bp. Average total assembly length for all iceberg lettuce samples was 1.3 million base pairs with a maximum contig size range of 2,623 to 58,547 bp.

Table 17. Yuma lettuce sequence statistics following trimming and assembly							
Lettuce Type	Sample ID	Sample Description	# Sequence reads following trimming	# contigs (>200bp)	Total assembly length (bp)	Max contig size (bp)	N50
Romaine	YL1	Worker harvest	22,192,992	1,620	962,920	8,632	580
	YL2	Worker harvest	17,200,830	2,275	1,334,015	10,091	582
	YL3	Worker harvest	16,977,540	1,086	722,002	20,312	652
	YL4	Control	29,424,106	2,251	1,247,443	6,476	543
	YL5	Control	35,041,916	1,663	964,267	12,624	553
	YL6	Control	29,026,710	2,002	1,035,832	7,519	492
Iceberg	YL7	Control	29,742,342	2,720	1,330,578	12,556	474
	YL8	Control	25,153,498	2,549	1,176,162	5,526	440
	YL9	Control	26,246,806	1,981	1,075,789	5,135	526
	YL10	Worker harvest	21,406,636	1,262	841,841	13,516	694
	YL11	Worker harvest	16,549,704	3,475	1,854,371	8,344	537
	YL12	Worker harvest	20,684,240	784	511,741	8,381	662
	YL13	Worker harvest	18,585,596	3,315	2,236,049	19,370	703
	YL14	Worker harvest	20,012,550	1,292	1,182,495	58,547	1,092
	YL15	Worker harvest	23,122,222	1,194	911,149	11,376	825
	YL16	Worker harvest	24,312,410	703	596,082	11,488	1,064
	YL17	Worker harvest	25,204,382	2,435	1,694,477	11,427	717
	YL18	Control	29,990,912	2,201	1,255,819	37,446	570
	YL19	Control	28,146,472	1,699	897,536	5,206	528
	YL20	Control	45,206,036	2,980	1,689,182	38,197	566
	YL21	Control	38,761,924	3,312	2,022,747	9,904	610
	YL22	Control	36,458,692	2,072	1,093,235	8,085	527
	YL23	After worker 30 min break	15,534,726	5,577	3,602,163	21,540	692
	YL24	After worker 30 min break	17,791,284	1,641	987,166	4,163	639
	YL25	After worker 30 min break	17,898,636	713	605,017	12,873	974

Table 17 (cont'd)

	YL26	After worker 30 min break	16,128,066	845	477,225	2,623	583
	YL27	After worker 30 min break	18,140,228	2,644	1,901,198	13,928	791
Romaine	YL28	Control	37,243,138	1,446	941,058	27,199	680
	YL29	Control	34,020,132	1,223	873,414	59,171	781
	YL30	Control	40,302,750	1,229	841,510	16,353	728
	YL31	Control	37,703,770	1,637	1,062,459	41,708	675
	YL32	Control	39,795,730	1,767	1,670,543	58,805	1,307
	YL33	Worker harvest	14,261,170	1,562	1,321,166	16,909	938
	YL34	After chop and wash	15,191,256	887	776,029	12,523	996
	YL35	After chop and wash	18,912,828	3,491	2,411,520	20,547	707
	YL36	After chop and wash	16,904,810	3,342	2,384,372	21,488	795
	YL37	Mixed salad	20,698,078	992	772,902	12,779	820
	YL38	Mixed salad	19,025,938	1,168	836,064	5,875	788
	YL39	Mixed salad	18,618,602	925	671,693	6,841	765
	YL40	Mixed salad	17,769,542	1,766	1,280,403	12,422	759
	YL41	Mixed salad	18,109,612	3,145	2,085,159	15,874	692
	YL42	Worker harvest	20,578,110	2,078	1,418,464	18,786	686

4.6 Yuma Lettuce Virome

To analyze the iceberg (N=21) and romaine (N=21) lettuce virome, the viral contigs belonging to multiple samples collected from the same stage of field production were combined into a single sample. The resulting categories include iceberg (N=8) and romaine (N=8) lettuce controls (harvested by research crew), iceberg (N=8) and romaine (N=5) lettuce worker harvest, iceberg harvested post worker break (N=5), romaine chop and wash (N=3), and romaine mixed salad (N=5). The viral genome distribution for each of these categories is shown in Table 18 and Figures 10-16. The majority of contigs belonging to both iceberg and romaine lettuce could not be annotated against the NCBI RefSeq viral database.

Table 18. Distribution of contigs larger than 200 bp for iceberg and romaine lettuce virome																
Genome*	Sample Type**															
	IC # contigs	IC (%)	IWH # contigs	IWH (%)	IPWB # contigs	IPWB (%)	Ice Avg (%)	RC # contigs	RC (%)	RWH # contigs	RWH (%)	RCW # contigs	RCW (%)	RMS # contigs	RMS (%)	Rom Avg (%)
No Hits	17,004	87.1	12,563	86.9	10,288	90.1	88.0	10,611	80.3	7,298	84.7	6,962	90.2	6,884	86.1	85.3
Not Assigned	388	2.0	252	1.7	179	1.6	1.8	239	1.8	148	1.7	96	1.2	116	1.5	1.6
Assigned	2,122	10.9	1,644	11.4	953	8.4	10.2	2,366	17.9	1,175	13.6	662	8.6	994	12.4	13.1
dsDNA	1,215	57.3	613	37.3	397	41.7	45.4	1,370	57.9	385	32.8	124	18.7	386	38.8	37.1
dsRNA	59	2.8	141	8.6	51	5.4	5.6	250	10.6	237	20.2	98	14.8	160	16.1	15.4
Retro-transcribing	371	17.5	600	36.5	369	38.7	30.9	275	11.6	298	25.4	329	49.7	282	28.4	28.8
Satellites	0	0.0	0	0.0	0	0.0	0.0	2	0.1	0	0.0	0	0.0	0	0.0	0.0
ssDNA	132	6.2	53	3.2	29	3.0	4.2	59	2.5	38	3.2	18	2.7	22	2.2	2.6
ssRNA	178	8.4	140	8.5	55	5.8	7.6	193	8.2	134	11.4	55	8.3	80	8.1	9.0
Unclassified	167	7.9	97	5.9	52	5.5	6.4	217	9.2	83	7.1	38	5.7	64	6.4	7.1

* Contigs assigned to viral taxa but did not meet the selected MEGAN parameters are "Not assigned," and contigs without any hits to known sequences in the databases were placed under "No hits."

** Iceberg control (IC), iceberg worker harvested (IWH), iceberg post worker break (IPWB), iceberg average (ice avg), romaine control (RC), romaine worker harvested (RWH), romaine chop and wash (RCW), romaine mixed salad (RMS), romaine average (rom avg).

Figure 10. Iceberg lettuce - control genome distribution

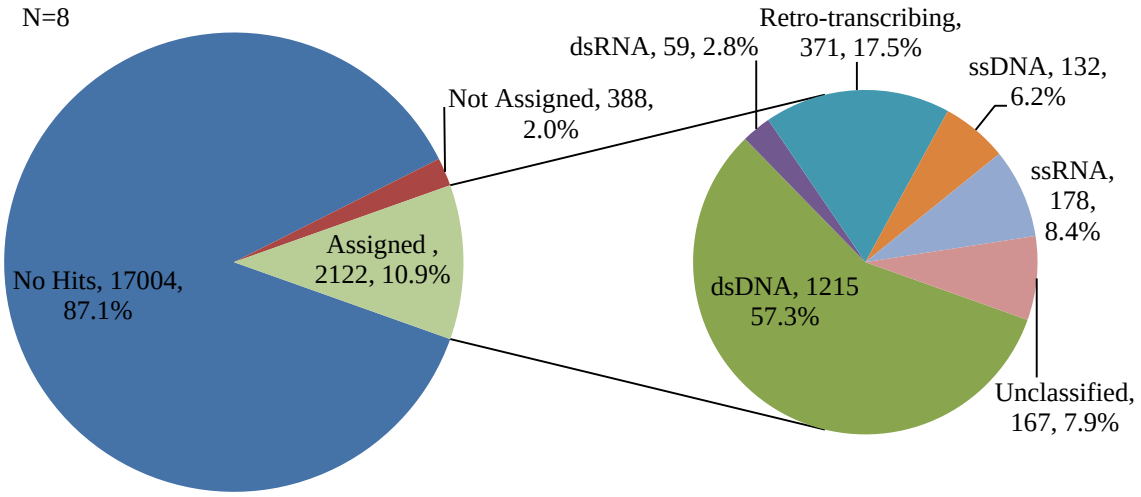


Figure 11. Iceberg lettuce - worker harvest genome distribution

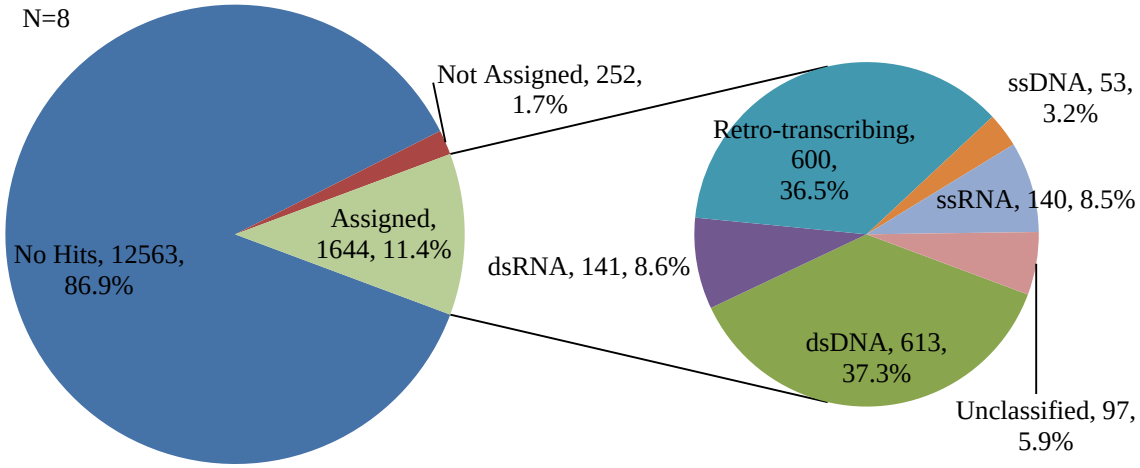


Figure 12. Iceberg lettuce - post worker break genome distribtuion

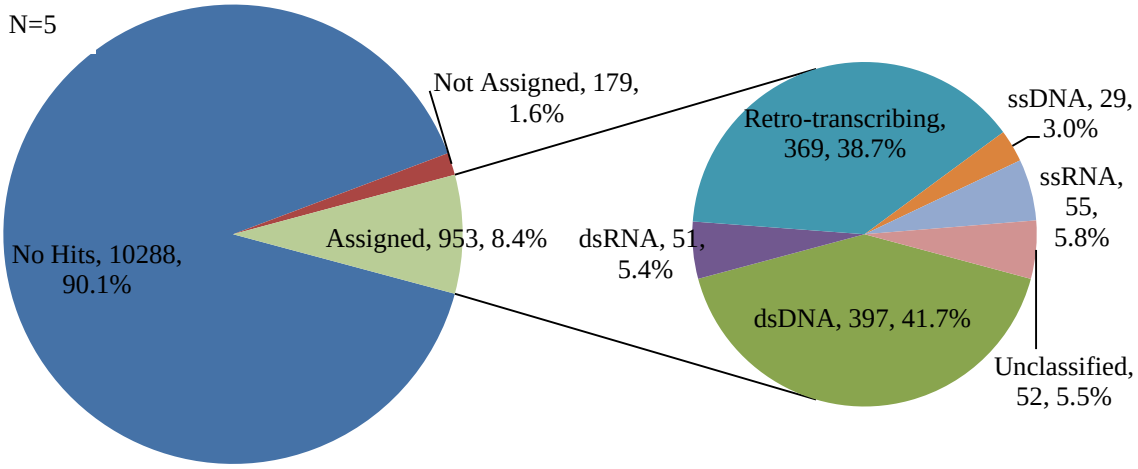


Figure 13. Romaine lettuce - control genome distribution

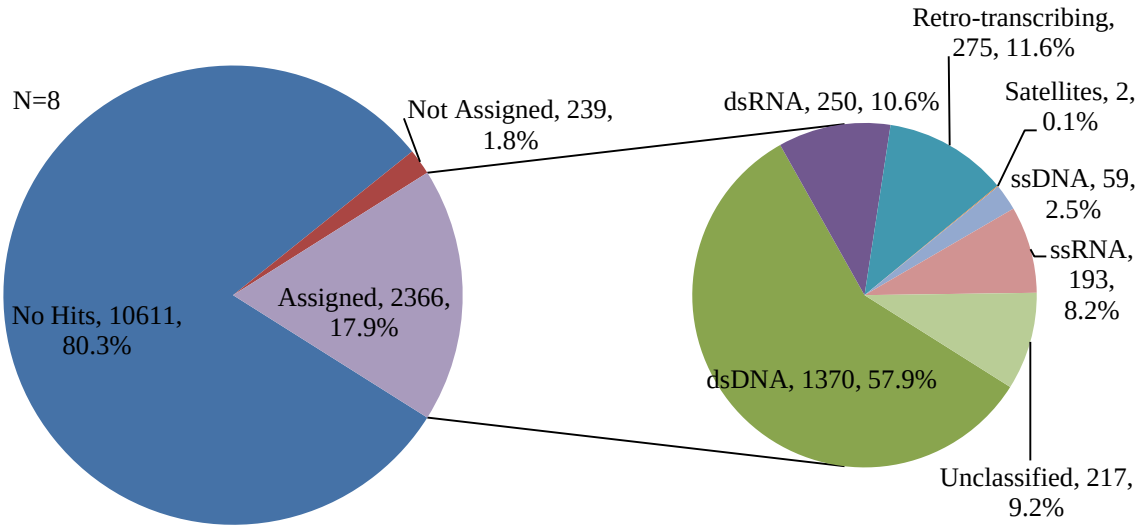


Figure 14. Romaine lettuce - worker harvest genome distribution

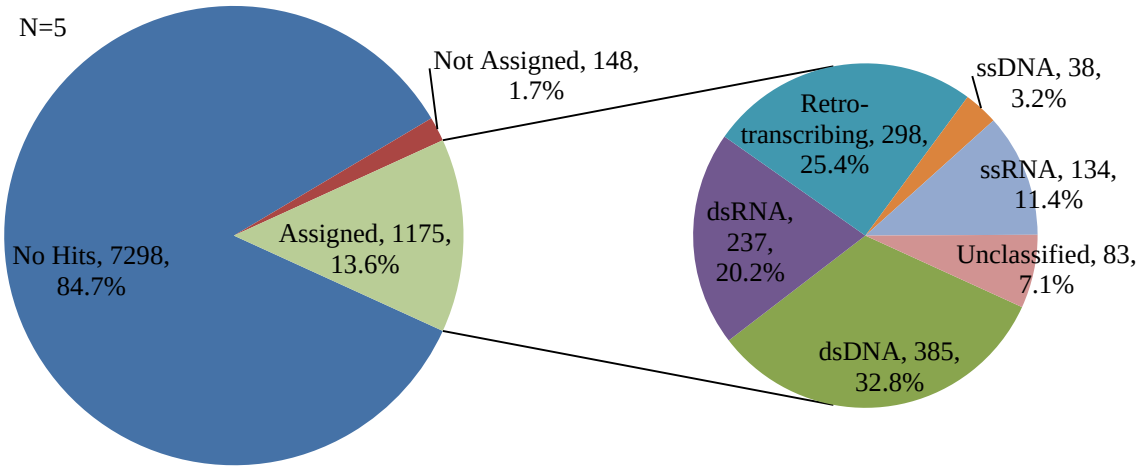


Figure 15. Romaine lettuce - chop and wash genome distribution

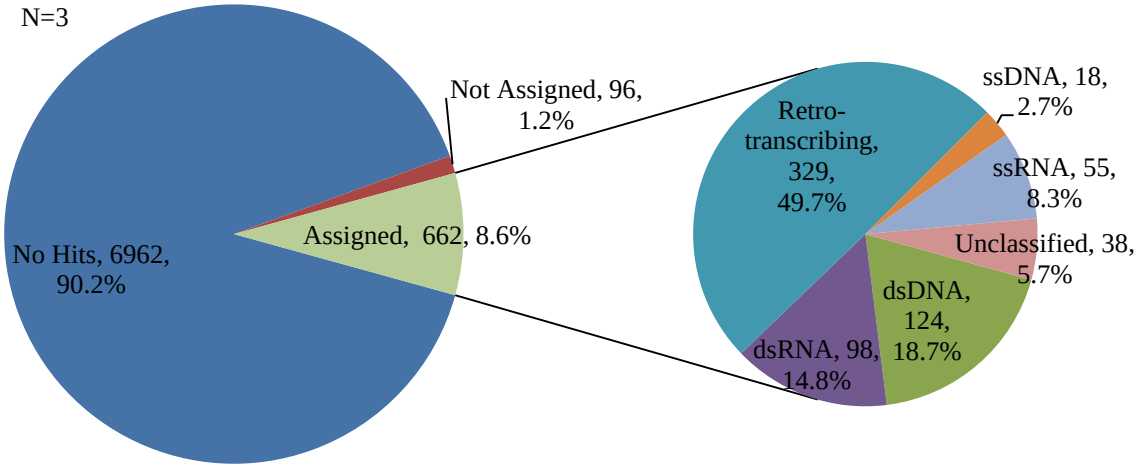
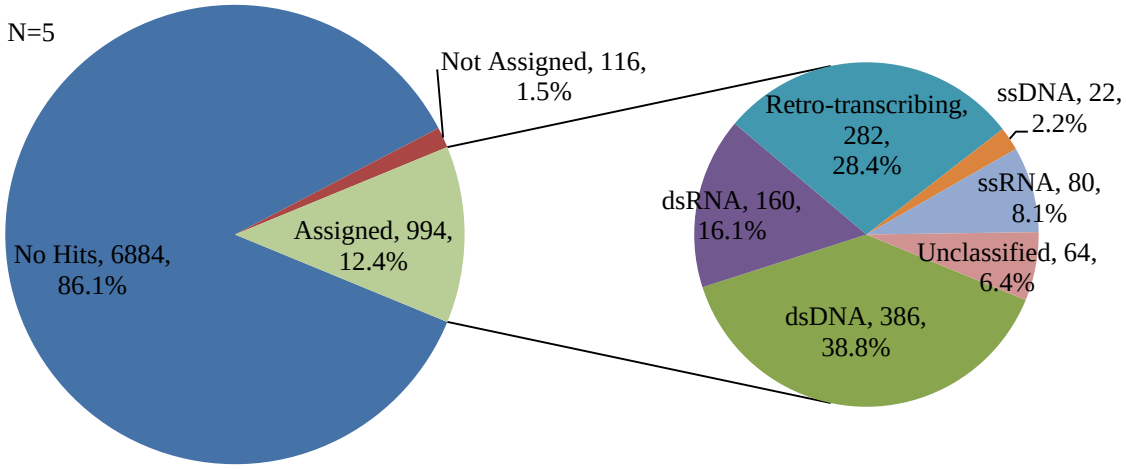


Figure 16. Romaine lettuce - mixed salad genome distribution



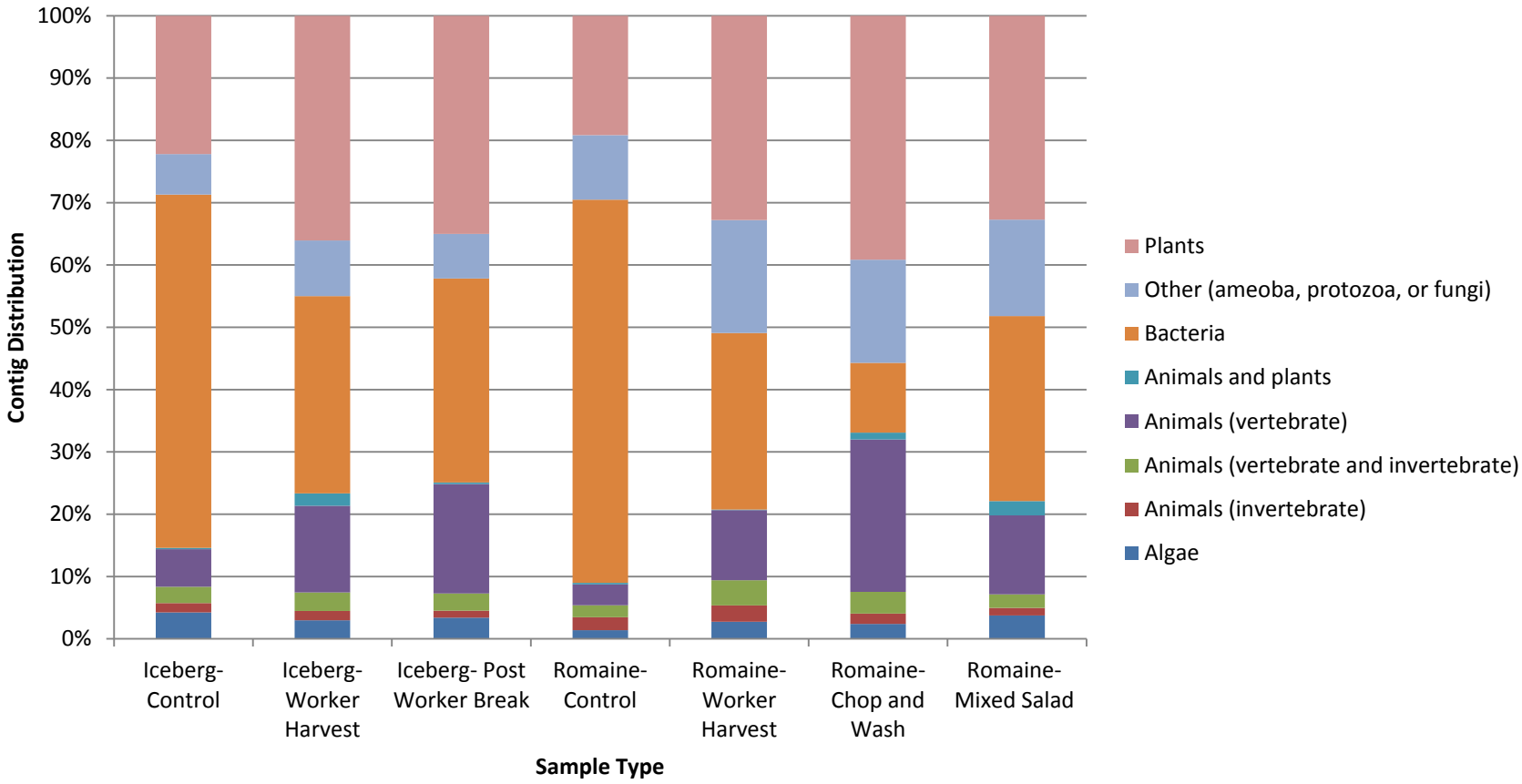
The average number of assigned contigs >200 bp was 10.2% and 13.1% for iceberg and romaine lettuce samples respectively. Of the assigned viral sequences belonging to iceberg lettuce, the majority of contigs shared similarities with dsDNA (average of 45.4%) and retro-transcribing (average of 30.9%) viruses. Single-stranded DNA, dsRNA, and ssRNA viral sequences composed on average 4.2%, 5.6%, and 7.6% of the iceberg lettuce virome, respectively. Similar results were observed for romaine lettuce viral contigs, with dsDNA and retro-transcribing viruses representing an average of 37.1% and 28.8% of the virome, respectively. The remaining romaine lettuce viral sequences were assigned to ssDNA (2.6%), dsRNA (15.4%), and ssRNA (9.0%) viruses.

The distribution of viral contigs based on virus hosts is shown in Table 19 and Figure 17. Contigs assigned to viral orders and families were classified into the same host categories as Yuma irrigation water samples. For iceberg lettuce samples (control, worker harvest, and post worker break), the majority of viral contigs were associated with bacteriophage and plant hosts, representing an average of 40.4% and 30.7% of the iceberg virome respectively. Of the remaining host categories, algae, invertebrates, vertebrates, and multiple hosts (invertebrates and vertebrates, animals and plants, and other) composed on average 3.5%, 1.4%, 12.5%, and 11.6% of the iceberg lettuce virome, respectively. Similar to iceberg lettuce samples, the majority of viral contigs belonging to romaine lettuce harvested by researchers (control), workers, as well as sampled from mixed salad were associated with bacteria (average of 39.9%) and plant hosts (average of 28.2%). However for romaine lettuce sampled after chop and wash, the majority of contigs were associated with plants (39.2%), vertebrate animals (24.5%), and hosts classified as "other" (16.5%), with bacteriophage comprising only (11.2%) of the virome (Table 19). Algae, invertebrates, vertebrates, and multiple hosts (invertebrates and vertebrates, animals and plants, and other) composed on average 2.6%, 1.9%, 12.8%, and 19.1% of the romaine lettuce virome, respectively.

Table 19. Iceberg and romaine lettuce viral host distribution																
Host		Sample Type*														
	IC # contigs	IC (%)	IWH # contigs	IWH (%)	IPWB # contigs	IPWB (%)	Ice Avg (%)	RC # contigs	RC (%)	RWH # contigs	RWH (%)	RCW # contigs	RCW (%)	RMS # contigs	RMS (%)	Rom Avg (%)
Algae	76	4.2	42	3.0	27	3.4	3.5	29	1.4	27	2.7	13	2.4	31	3.7	2.6
Animals (invertebrate)	27	1.5	21	1.5	9	1.1	1.4	43	2.1	26	2.6	9	1.7	10	1.2	1.9
Animals (vertebrate and invertebrate)	47	2.6	42	3.0	22	2.8	2.8	40	1.9	21	2.1	19	3.5	19	2.3	2.5
Animals (vertebrate)	108	6.0	196	13.9	140	17.6	12.5	70	3.4	107	10.8	133	24.5	105	12.7	12.8
Animals and plants	4	0.2	28	2.0	12	1.5	1.2	5	0.2	24	2.4	6	1.1	18	2.2	1.5
Bacteria	1,017	56.7	446	31.7	261	32.7	40.4	1,281	61.5	280	28.3	61	11.2	246	29.7	32.7
Other (ameoba, protozoa, or fungi)	117	6.5	126	8.9	57	7.2	7.5	216	10.4	179	18.1	90	16.5	128	15.5	15.1
Plants	398	22.2	508	36.1	269	33.8	30.7	399	19.2	324	32.8	213	39.2	271	32.7	31.0

* Iceberg control (IC), iceberg worker harvested (IWH), iceberg post worker break (IPWB), iceberg average (ice avg), romaine control (RC), romaine worker harvested (RWH), romaine chop and wash (RCW), romaine mixed salad (RMS), romaine average (rom avg).

Figure 17. Yuma lettuce viral host distribution



A total of 53 viral families were observed for the Yuma lettuce virome. The viral family contig distribution for iceberg and romaine lettuce can be viewed in a phylogenetic tree constructed using MEGAN software (Figure 18). The number of assigned contigs in this figure is represented by circle size and colors show the proportion belonging to each sample type. A large proportion of viruses were assigned to dsDNA *Myoviridae, Podoviridae,* and *Siphoviridae* bacteriophage families of the *Caudovirales* order as well as DNA *Caulimoviridae* and RNA *Retroviridae* retro-transcribing virus families which infect plant and vertebrate animal hosts respectively. The distribution of viral contigs associated with families composing ≥ 3.0% across the 7 Yuma lettuce viromes was further analyzed (Table 20 and Figure 19). Of the viral families representing ≥3.0% of the Yuma lettuce virome, *Caulimoviridae* and *Retroviridae* families composed on average 25.6% and 14.7% the iceberg virome respectively. Similar results were observed for the romaine lettuce virome, with a large proportion of sequences assigned to *Caulimoviridae* (21.8%) and *Retroviridae* (14.3%) viral families. *Myoviridae, Podoviridae, and Siphoviridae* bacteriophage families of the *Caudovirales* order represented 11.6-14.6%, 5.4-12.5%, and 11.3-24.3% of the iceberg lettuce (control, worker harvest, and post worker break) virome, respectively. Although the *Caudovirales* order represented 59.5% of the romaine control lettuce virome, only 2.5-8.7%, 2.3-7.5%, and 3.2-11.1% of viral contigs were assigned to *Myoviridae*, *Podoviridae,* and *Siphoviridae* viral families respectively for the remaining romaine lettuce samples (worker harvest, chop and wash, mixed salad). Other well represented viral families include other bacteriophage families (*Microviridae*) and viruses that infect plants (*Closteroviridae* and *Endornaviridae*), fungi (*Partitiviridae* and *Totiviridae*), algae (*Phycodnaviridae*), and amoeba (*Mimiviridae*).

Figure 18. Yuma lettuce viral family distribution

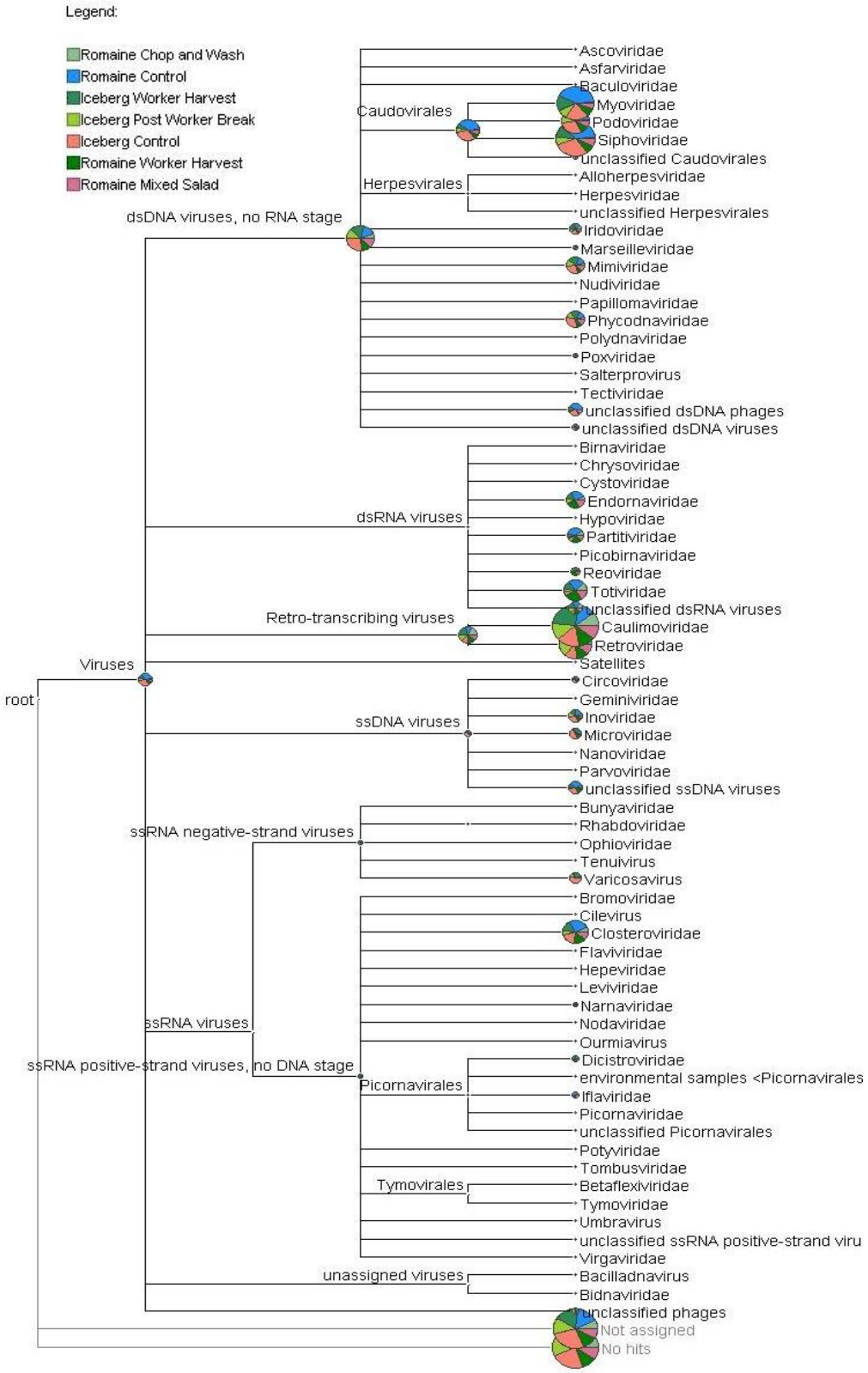
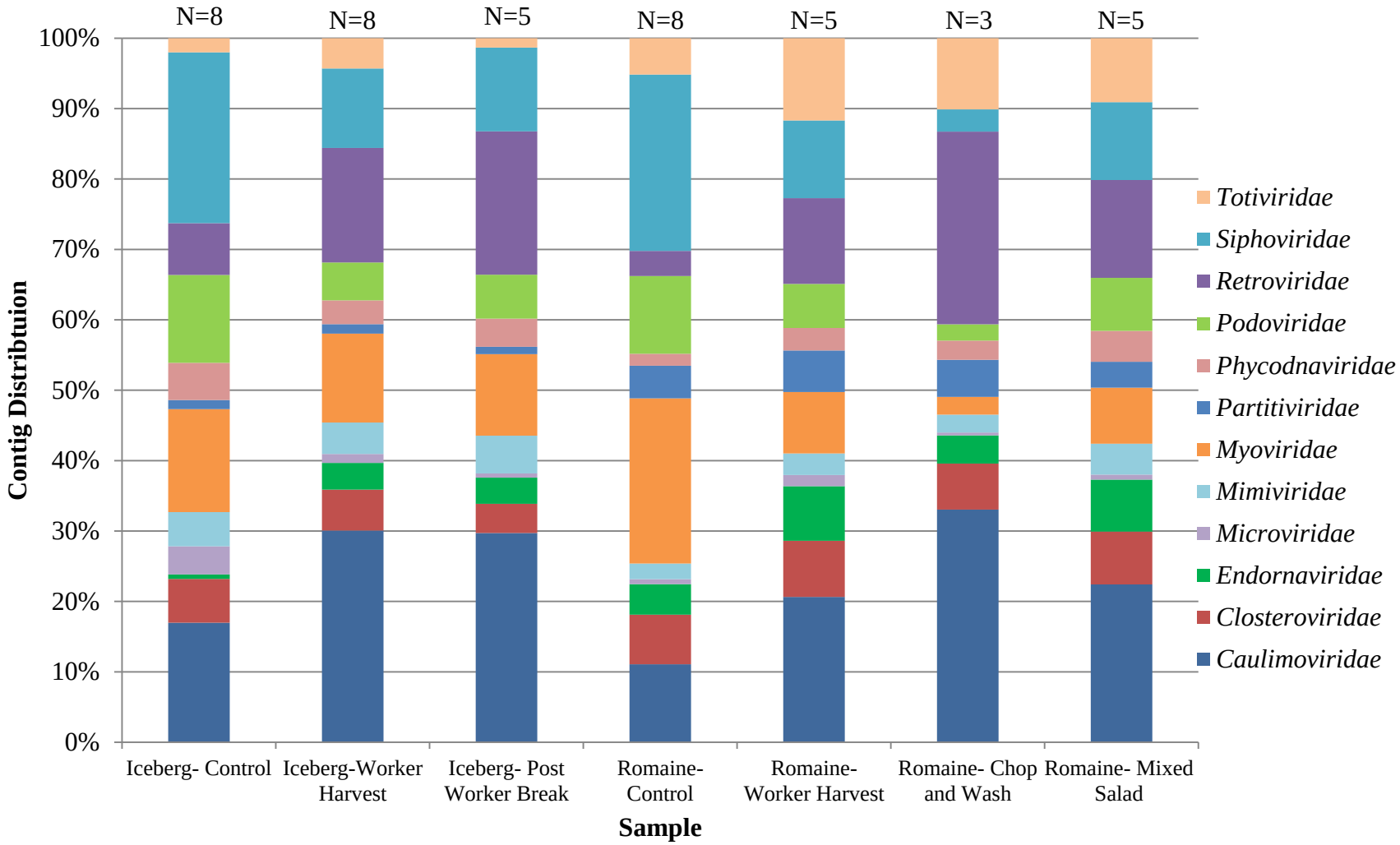


Table 20. Lettuce viral family distribution representing greater than 3.0% of the virome																
Host	Sample Type*															
	IC # contigs	IC (%)	IWH # contigs	IWH (%)	IPWB # contigs	IPWB (%)	Ice Avg (%)	RC # contigs	RC (%)	RWH # contigs	RWH (%)	RCW # contigs	RCW (%)	RMS # contigs	RMS (%)	Rom Avg (%)
Caulimoviridae	238	17.0	357	30.1	200	29.7	25.6	189	11.1	168	20.6	157	33.1	158	22.4	21.8
Closteroviridae	87	6.2	69	5.8	28	4.2	5.4	120	7.0	65	8.0	31	6.5	53	7.5	7.3
Endornaviridae	9	0.6	45	3.8	25	3.7	2.7	74	4.3	63	7.7	19	4.0	52	7.4	5.9
Microviridae	56	4.0	15	1.3	4	0.6	2.0	12	0.7	13	1.6	2	0.4	5	0.7	0.9
Mimiviridae	68	4.9	53	4.5	36	5.3	4.9	38	2.2	25	3.1	12	2.5	31	4.4	3.1
Myoviridae	205	14.6	150	12.6	78	11.6	13.0	400	23.5	71	8.7	12	2.5	56	7.9	10.7
Partitiviridae	18	1.3	16	1.3	7	1.0	1.2	79	4.6	48	5.9	25	5.3	26	3.7	4.9
Phycodnaviridae	74	5.3	40	3.4	27	4.0	4.2	29	1.7	26	3.2	13	2.7	31	4.4	3.0
Podoviridae	175	12.5	64	5.4	42	6.2	8.0	188	11.0	51	6.3	11	2.3	53	7.5	6.8
Retroviridae	103	7.4	193	16.3	137	20.4	14.7	61	3.6	99	12.2	130	27.4	98	13.9	14.3
Siphoviridae	340	24.3	134	11.3	80	11.9	15.8	427	25.0	90	11.1	15	3.2	78	11.1	12.6
Totiviridae	28	2.0	51	4.3	9	1.3	2.5	88	5.2	95	11.7	48	10.1	64	9.1	9.0

* Iceberg control (IC), iceberg worker harvested (IWH), iceberg post worker break (IPWB), iceberg average (ice avg), romaine control (RC), romaine worker harvested (RWH), romaine chop and wash (RCW), romaine mixed salad (RMS), romaine average (rom avg).

Figure 19. Lettuce viral family distribution representing greater than 3.0% of the virome



Yuma lettuce samples were further investigated at the viral species level to identify human enteric viruses. Rotavirus A and rotavirus C were identified in YL10 (iceberg worker harvest) and YL4 (romaine lettuce control) samples, respectively. In addition, picobirnavirus was identified in the iceberg control (YL18), romaine worker harvest (YL33 and YL42), romaine chop and wash (YL35), and romaine mixed salad (YL37) lettuce samples. The contigs belonging to these enteric viruses were further blasted against the NCBI nucleotide database (BLASTn) to determine the gene, contig length, % identity, Query coverage (% of sequence that overlaps the subject sequence), E-value (number of hits expected by chance when searching a particular database) and viral host (Table 21). The rotavirus A contig shared a 99% sequence identity similarity to the VP1 gene of a human isolate. Rotavirus C identified in the romaine lettuce control sample shared a 99% sequence identity similarity to the NSP2 gene of a bovine isolate. A majority of the picobirnavirus contigs had no significant similarity to any sequences in the BLASTn database. Two contigs belonging to the YL33 sample shared a 80% and 95% sequence identity similarity to the RDRP gene of human and porcine isolates respectively, however, the query coverage for both picobirnavirus contigs was low (28-33%).

Table 21. Yuma lettuce enteric virus nucleotide BLAST results

Virus	Host	Sample	Contig Number	Gene	Contig length (bp)	% Identity	% Query coverage	E-value	Score
Rotavirus C	Bovine	YL4	824	NSP2	483	99	79	0	686
Rotavirus A	Human	YL10	539	VP1	476	99	99	0	837
Picobirnavirus	Human	YL33	1487	RDRP	315	80	28	6.00E-21	111
Picobirnavirus	Porcine	YL33	1133	RDRP	436	95	33	5.00E-58	235

5. Discussion

Before the development of NGS technology, virus detection in food matrices was primarily performed using one of two methods: cell culture or reverse transcriptase PCR (RT-PCR). Improvements in NGS technology have led to the ability to view entire microbial communities in a wide range of environments, providing the opportunity to investigate the role of viruses in various ecosystems including foods. However despite our current knowledge of fresh produce as a vehicle for human pathogenic virus transmission, there are no studies that have examined the virus communities in leafy greens or irrigation water which is a suggested pre-harvest source of fresh produce contamination. In this study we developed a metagenomic pipeline to investigate the virome from iceberg and romaine lettuce from the field as well as from irrigation water used to cultivate these crops.

Virus elution and concentration prior to extraction and sequencing are critical steps in the metagenomic pipeline that ultimately impact virus recovery. In our study we evaluated a method of virus recovery from lettuce using an elution protocol adapted from Dubois et al., (2006) and enumeration by double agar overlay plating. Viruses such as MS2 are thought to use physicochemical forces, specifically electrostatic forces, for nonspecific attachment to solid surfaces such as fresh produce and a higher pH has been shown to effectively dissociate viruses from lettuce surfaces (Deboosere et al., 2012, Vega et al., 2008, Vega et al. 2005). P22 recovery from lettuce ranged 43.3%-77.4%, while MS2 had a broader range of recovery between experiments (2.8%-91.1%). The low recovery was due to trial 2 and could have been due to issues during enumeration using double agar overlay plaque assay, which is subjective to diluting, plaque size, incomplete lysis, or plaque aggregation leading to inaccurate measurements ultimately variation in virus recovery (as described in Kropinski et al., 2009). Other possibilities

include incorrect dilution spiking suspension or a potential poorly prepared virus stock. However, MS2 recovery in trials 1 and 3 in this study (29.1-91.1%) was similar to a study by Dubois et al., (2006) who achieved a 22.9-96.2% recovery of MS2 from inoculated butter lettuce using the same elution protocol. For comparison, studies investigating enteric (norovirus, hepatitis A, poliovirus) and enteric surrogate (murine norovirus) virus recovery from fresh produce using a real-time RT-PCR method have discovered virus recovery can range anywhere between 8.2% and 100%, and can vary within and between viral species (Hennechart et al., 2002; Sánchez et al., 2012).

A total of six 100-L irrigation water samples as well as 21 iceberg and 21 romaine lettuce samples were collected for metagenomic analysis. Yuma irrigation water had low turbidity and low levels of total coliforms as well as enterococci and *E. coli* which are the typical fecal indicator bacteria used in water quality analysis. Viral communities were sequenced using Illumina Hiseq technology and were analyzed using a variety of bioinformatics tools targeted to short sequencing reads. For both the Yuma irrigation water and lettuce viromes, a large proportion of viruses had no hits against the viral NCBI viral database. This observation is consistent with most other viral metagenomic studies to date (Angly et al., 2006, Aw et al., 2014, Breitbart et al., 2002, Mokili et al., 2012, Rosario et al., 2009). No hits could represent either novel viruses or generally poor global data bases for viral genome comparison. These data suggest, however that we have limited knowledge of the viral genome diversity in these environments and that there is a strong need to develop and improve current metagenomic techniques and viral databases.

Of the irrigation water contigs that were assigned, a total of 64 viral families were identified with dsDNA viruses dominating the virome. This is largely due to the high recovery

of *Myoviridae, Podoviridae,* and *Siphoviridae* bacteriophage virus families of the *Caudovirales* order. Viral metagenomic studies of sewage, reclaimed water, human feces, and fermented foods have shown that these viromes also contain a large proportion of bacteriophage viruses belonging to the *Caudovirales* order using either 454 pyrosequencing or Illumina technology (Aw et al., 2014, Kim et al., 2011, Park et al., 2011, Rosario et al., 2009). Bacteriophage that infect *E. coli* are promising indicators of fecal pollution and have even been suggested as possible bacterial control agents in foods, however little is known about their role within virus communities in irrigation water and fresh produce systems (Dawson et al., 2005; Havelaar et al., 1993; Sharma, 2013). Investigation of the host distribution of these phage species showed that most of the bacterial host species were common to freshwater or marine environments (*Pelagibacter, Cellulophaga, Synechococcus, Bdellovibrio, Bacillus, Puniceispirillium, Mycobacterium,* and *Prochlorococcus*) with low relative abundance of phage infecting the genus *Escherichia* or *Enterococcus* which are typical bacterial fecal indicators. This is similar to our enterolert and colilert results and may indicate low levels of fecal pollution in irrigation water; however, these bacterial indicators fail to recognize human enteric viral hazards which have been shown not to correlate with typical bacterial indicators in water (Harwood et al., 2005).

This is one of the first studies to use metagenomics to investigate viral species in irrigation water and lettuce and there is interest in how these data may be of importance to the food industry. Analysis of viral species in irrigation water revealed numerous viruses infecting agricultural insect pests, crops, fish and shrimp, livestock animals, and even humans. Although crop viruses identified were not specific to lettuce, many plant viral species belonging to *Geminiviridae, Nanoviridae,* and *Tombusviridae* families can cause disease in multiple food crops and are of interest to the fresh produce industry due to high economic costs from reduced

quality, productivity, or yield (Boulila 2011, Varma and Malathi 2003). Animal virus families were well represented by the *Circoviridae* and *Poxviridae* families. These plant and animal virus families have also been identified in wastewater and reclaimed water using next-generation sequencing (Aw et al., 2014, Rosario et al., 2009).

The identification of human enteric viruses is of particular importance to public health and monitoring food safety. In this study *picobirnavirus*, *coronavirus*, *parechovirus*, and *rhinovirus* as well as hepatitis A and E were human pathogens identified in irrigation water. Hepatitis A is an enteric viral pathogen that has been linked to fresh produce outbreaks while hepatitis E is just recently gaining recognition as a food safety hazard. Hepatitis E is transmitted through the fecal-oral route and genotypes 3 and 4, which are zoonotic, are now suspected of foodborne transmission (e.g. pig meats) (Meng, 2011). The picobirnavirus identified is also of interest because this dsRNA virus is likely zoonotic, is a suggested cause of diarrhea in pigs, and has been also isolated from the feces of immunocompromised and healthy individuals (Ganesh et al., 2012). However, the low percentage of sequence identity similarities for picobirnavirus as well as a majority of the plant, animal, and insect viral species detected suggests these are novel viruses based on our limited knowledge of the viral diversity in irrigation water.

For the iceberg and romaine lettuce viomes, dsDNA viruses also composed a large percentage of assigned contigs with a majority belonging to bacteriophage families of the *Cuadovirales* order as well as *Phycodnaviridae* and *Mimiviridae* families which infect algae and amoeba, respectively. Interestingly, iceberg and romaine lettuce contained of a greater proportion of retro-transcribing families, including *Caulimoviridae* (DNA) and *Retroviridae* (RNA). Retro-transcribing viruses use a reverse transcription step in order to replicate forming a DNA from an RNA strand. For viruses in the *Retroviridae* family, infection occurs following

incorporation of viral DNA into the host's cell genome however viruses of the *Caulimoviridae* family do not require integration and replication is instead in the cytoplasm. The *Caulimoviridae* family is a group of viruses that is widely distributed but known to cause serious crop and plant disease in tropic regions (Geering, 2007). *Retroviridae* species are known to infect humans and animals; however viruses within this family are not recognized as foodborne pathogens. Interestingly, control iceberg and romaine lettuce samples collected by researchers had a larger proportion of phage viruses when compared to subsequent stages of field harvest. It is unclear as to why control samples had more viruses infecting bacteria. Assuming phage correspond with bacterial presence, bacteria could have been removed from romaine lettuce during chop and wash resulting in lower populations in these samples. However, this does not explain the lower levels observed following worker harvest or in iceberg lettuce (control vs. worker harvest) which is not washed prior to packaging.

Metagenomics is a random sequencing approach and combined with a small sample size, variations in virus communities between samples are likely. Further studies with a larger sample size as well as use of a quantification assay (qPCR) could be helpful to investigate differences between these sample types. Interestingly, dsRNA viruses including Rotavirus A and human picobirnavirus were discovered in the lettuce virome. Rotavirus A, which can be transmitted through food including fresh produce, is a common cause of gastroenteritis in children but has also been shown to cause disease in adults (Fletcher et al., 2000, Newell et al., 2010). Rotavirus A was found in an iceberg lettuce sample collected by workers and found to have a high percentage identity match to a human isolate, which suggests this enteric virus was present on the lettuce either prior or following field harvest. In addition, rotavirus A is a well-known waterborne enteric pathogen and possible sources of contamination could be from fecally

contaminated irrigation water or poor worker hygiene. Bovine rotavirus C was identified in a romaine lettuce control sample, indicating fecal contamination from cattle with possible sources including contaminated irrigation water or runoff or improperly treated manure used as fertilizer. However the location of cattle farms near the farm as well as whether manure was used to fertilize the fields is unknown. Human picobirnavirus was also identified in both iceberg and romaine lettuce. These data suggest that dsRNA enteric viruses pose a threat to food safety beginning at the pre-harvest stage of production. However, a majority of picobirnavirus contigs were not significantly similar to sequences in the NCBI nucleotide database suggesting limited knowledge of this virus genome and the role it plays in water and food systems. In addition, metagenomics detects DNA or RNA and it is unknown as to whether these organisms are viable and infectious.

Metagenomic analysis revealed that most of the viruses in both irrigation water and lettuce are novel or unknown. Of those assigned, a total of 47 viral families were shared between the irrigation water and lettuce viromes. Analysis of viral families representing ≥3.0% of the irrigation water and lettuce viromes revealed six viral families *(Microviridae, Mimiviridae, Myoviridae, Phycodnaviridae, Podoviridae, Siphoviridae*) in common. As stated previously, viral families of the *Caudovirales* order are numerous in a variety of environmental samples. Interestingly, the *Phycodnaviridae* viral family infects algae or aquatic chlorophyll containing organisms. This indicates the presence of algal viruses in the lettuce virome, which may be a result of the irrigation water used during cultivation. However, the role of these viruses in lettuce and the lettuce virome is currently unknown. Viruses infecting animals (vertebrate and/or invertebrate) and plants in irrigation water and lettuce varied in proportion and viral family type. Yuma lettuce contained a larger proportion of plant viruses (19.16-39.15%) compared to

irrigation water (0.95-7.17%), largely due to the *Endornaviridae*, *Closteroviridae*, and *Caulimoviridae* viral families. In addition, of the animal viral families representing ≥3.0% of the irrigation water virome, a large proportion belonged to those infecting both vertebrates and invertebrates (*Poxviridae* and *Circoviridae* families) while in the lettuce virome a majority of viruses belonged to the *Retroviridae* family which infects vertebrates alone. Of the human viruses identified, dsRNA picobirnavirus was the only enteric virus identified in both lettuce and irrigation water. This virus, which had low percentage identities, needs to be further investigated for further potential involvement in water and food systems. These data show that the same families of algal viruses, phage, and human enteric viruses can be present in both the lettuce and irrigation water viromes, while animal viruses tend to differ between these two environments.

IV. VIRAL CROSS-CONTAMINATION OF LETTUCE DURING SMALL-SCALE LEAFY GREEN PROCESSING

1. Introduction

Fresh produce consumption in the United States has increased in recent years as consumers desire to maintain a healthy lifestyle. Consequently, foodborne outbreaks associated with fresh produce are becoming increasingly recognized. Leafy greens in particular are often consumed raw and considered a food commodity of high risk, accounting for the highest number of foodborne illnesses (2.2 million) between 1998 and 2008 (Painter et al., 2013). Furthermore foodborne outbreak data from this same time period suggests viruses are commonly associated with leafy green contamination and transmission, with norovirus and leafy greens the pathogen-commodity pair most likely to be associated with a foodborne outbreak (Gould et al. 2013). Although human norovirus has been frequently associated with leafy green outbreaks, there is currently limited knowledge on viral contamination in the supply chain and other viral pathogens of concern.

In response to the changing fresh produce supply and demand, traditional agricultural and post-harvest practices have been altered which ultimately increase food supply chain complexity. Practices such as cutting and coring at harvest, increased importation and transportation, and large scale production facilities are now employed to support changing consumer habits (Heaton et al. 2008, Lynch et al. 2009). These practices contribute to the increasing number of leafy green outbreaks by providing multiple opportunities for human viral pathogen contamination and ultimately make it difficult to determine the single source of contamination. At the post-harvest level of leafy green production, contamination can occur during any stage of processing, packing, storage, or transportation. Critical control points include the quality of water used

(cooling, washing), worker hygiene, and the condition or overall cleanliness of processing equipment, cooling facilities, storage and packaging containers, and transportation vehicles (CFSAN, 2006).

Leafy greens are now provided to the consumer either as bulk products to be washed (e.g., head of lettuce) or as ready-to-eat (RTE) salads. Many consumers today desire RTE salads which are typically shredded, washed, dewatered, dried, and packaged prior to cold storage and transportation to distribution centers. Leafy green washing is an important step for improving the quality of the food product by removing contaminants (soil, debris, microorganisms) and prolonging shelf life. In addition, proper disinfection and monitoring of sanitizer levels during processing of RTE salads are essential steps to minimize foodborne disease (CFSAN, 2014). Although numerous physical and chemical disinfection methods have been investigated for use in foods, washing with a chlorine-based sanitizer is the disinfection method most commonly used in the fresh-cut produce industry due to the low cost and limited negative impact on product quality (CFSAN, 2014). Guidelines for washing fresh produce in chlorinated water include a maximum free chlorine (hypochlorite) concentration of 200 ppm and generally a 1 to 2 minute contact time (CFSAN, 2014). Currently a 5-log pathogen reduction standard is suggested for fresh produce production, however, this standard primarily targets bacterial pathogens and to date there is no defined criterion for antiviral disinfectants (Allwood, Malik, Hedberg, & Goyal, 2004; Gulati, Allwood, Hedberg, & Goyal, 2001).

There have been numerous studies investigating bacterial pathogen reduction on fresh produce using chlorine sanitizers. Pilot-scale studies have found that chlorine-based sanitizers generally reduce bacterial pathogen populations on lettuce between 1 and 3 logs (Davidson et al., 2013; Gil et al., 2009). Several laboratory studies investigating the effects of chlorine on viruses

when inoculated onto fresh produce have also shown that viral (MS2, feline calicivirus, murine norovirus) reduction generally does not exceed 3 logs when exposed to a variety of free chlorine (15-800 ppm) levels (Allwood et al., 2004, Dawson et al., 2005, Fraisse et al., 2011, Gulati et al., 2001). However, it is difficult to deduce how these data relate to large-scale leafy green processing and currently there is limited research investigating the bacterial and viral reduction on fresh produce during simulated commercial processing.

2. Research Goals and Objectives

The ultimate goal of the second half of the thesis research was to investigate the efficacy of current post-harvest leafy green processing and disinfection practices to better understand viral risks from farm-to-fork during a contamination event. Specifically, the goal was to assess the efficacy of a chlorine-based sanitizer against coliphage MS2 (an enteric virus surrogate) on romaine lettuce during simulated commercial processing. Having found a wide variety of viruses on lettuce from the field, there was interest in how removal might be achieved during processing after harvest. The study objectives were as followed:

1. Evaluate MS2 reduction on romaine lettuce during and following small-scale commercial leafy green processing (shredding, flume washing, shaker table dewatering, centrifuge drying) with and without a sanitizer wash treatment
2. Determine MS2 levels in the flume wash water and centrifugation water following processing

3. Materials and Methods

3.1 Bacteriophage Inactivation and Free Chlorine Demand

3.1.1 MS2 Inactivation during Sanitizer Exposure

Prior to small-scale commercial leafy green processing trials, experiments were performed on the bench to determine the extent of MS2 inactivation when exposed to a chlorine-based sanitizer (XY12 Ecolab, St. Paul, MN, USA) containing 25 ppm of available chlorine as recommended on the manufacturer's label for fruit and vegetable washing. Previous viral inactivation studies have used similar concentrations of free chlorine (Casteel et al., 2008; Casteel, Schmidt, & Sobsey, 2009; Fraisse et al., 2011). To achieve a level of 25 ppm available chlorine in the solution, 31.3 µL of XY12 were added to 100 mL of distilled water in a 250 mL glass bottle and the pH was adjusted to 7 using 6 N HCl. The sanitizer solution was mixed and the available chlorine concentration was confirmed using a chlorine color disc test kit (Hach test kit, 0.0-3.5 mg/l Model CN-66, Loveland, CO, USA) according to the manufacturer's instructions. Since the available chlorine concentration was out of the test kit measurement range, the solution was diluted 10-fold in PBW before analysis. Following free chlorine confirmation, 1 mL of high titer MS2 (10^{10} PFU/mL) was added to the sanitizer solution and vortexed. After 30 s, 60 s, and 120 s of exposure, 1 mL of sample was transferred to a 500 mL plastic bottle (one for each time point) containing 250 mL of tris-glycine buffer (pH 9.5) and 2% sodium thiosulfate (mimicking the virus elution recovery assay). The pH was adjusted to 7.2 ± 0.2 and the samples were diluted in PBW before plating using the double agar overlay method described in Part III, Section 3.1.2, pages 37-38. The free chlorine level in the neutralized samples was then measured using the same color disc test kit described above. As a control, 1 mL of MS2 (10^{10} PFU/mL) was added to 100 mL of distilled water without sanitizer. The

control was processed following the same protocol as the sanitizer challenge experiment. In addition, the MS2 suspension used for inoculation was diluted in PBW and plated to determine the inoculum level. The plaque concentration for the inoculated suspension, control and challenge samples was calculated using Formula 1 from Part III, Section 3.1.3, page 39.

3.1.2 Coliphage MS2 Chlorine Demand

To test the chlorine demand of the bacteriophage during chlorine inactivation, MS2 was exposed to a stock solution containing 25-ppm available chlorine. One mL of stock, TSB-based MS2 (10^{10} or 10^{7} PFU/mL) or PBW diluted MS2 (10^{8} or 10^{6} PFU/mL) was added to a 250-mL glass bottle containing 25 ppm of available chlorine (achieved by adding 31.3 µL of XY12 to 100 mL of distilled water). The available chlorine was measured before the addition of phage and after 0.5 min, 1 min, 3 min, 5 min, 10 min, and 15 min of exposure using the color disc test kit described in the bacteriophage inactivation experiment above. If necessary, the solution was diluted 10x in PBW before free chlorine measurement and analysis.

3.2 Bacteriophage Reduction during Small-Scale Leafy Green Processing With and Without a Chlorine-Based Sanitizer

3.2.1 Lettuce and Processing Line Preparation

Heads of romaine lettuce were purchased from a local produce supplier the day before processing. Upon arrival, the core of each romaine lettuce head was removed by cutting 2.5 to 5 cm from the core using a sterile scalpel. Remaining romaine lettuce leaflets were weighed until a total of 6-kg was achieved. The 6-kg batch was placed into Whirl-pak bags and stored at 4°C until processing the following day. A small-scale commercial leafy green processing line located in the Department of Food Science and Human Nutrition at Michigan State University was used (under the supervision of Dr. Elliot Ryser). The processing line included a lettuce shredder

(TranSlicer 2500, Urschel, Valparaiso, IN), conveyor (model 736018 mc series, Dorner Manufacturing, Hartland, WI), flume tank (3.6 m; Heinzen Manufacturing, Inc., Gilroy, CA), mechanical shaker table (Baldor Electric Co., Ft. Smith, AR), and centrifugal dryer (model SD50-LT, Heinzen Manufacturing, Inc.) as described by Buchholz et al. (2012). For lettuce washing, water was recirculated through a stainless steel water recirculation tank (1,000 L volume) connected by a hard plastic discharge hose to the stainless steel flume tank containing two overhead spray jets using a centrifugal pump (model XB754FHA, Sterling Electric, Inc., Irvine, CA) at 15 liters/s (Buchholz et al. 2012). The processing line sampling locations are shown in Appendix, Figure A4. The day before use, the entire processing line as well as the 121-L collection and inoculation bins were sanitized by spraying with 200 ppm of Quorum Clear (sodium hypochlorite active ingredient) (Ecolab, St. Paul, MN, USA). After at least a 1 min exposure all equipment was rinsed with tap water and air-dried overnight.

3.2.2 Lettuce Inoculation and Sampling

Prior to running the processing experiment, un-inoculated lettuce was sent through the shredder in order to prime the machine. Three ~50-g samples of the un-inoculated shredded lettuce were taken as negative controls. Prior to lettuce inoculation the water recirculation tank was filled with 800 L of municipal tap water (Michigan State University <0.05 ppm of free chlorine). For sanitizer experiments, approximately 946.4 mL of XY12 was added to the 800 L of water to achieve a free chlorine concentration of 25 ppm. The free chlorine level in the flume wash water was confirmed using the same chlorine color disc test kit mentioned previously and the pH was measured and adjusted to ≤7.5 by adding small volumes of hydrochloric acid to the flume wash water as it was recirculated through the flume and holding tank. For inoculation, a 6-kg batch of lettuce was submerged for 15 min in a sanitized 121-L plastic bin containing 80 L

of municipal tap water (Michigan State University <0.05 ppm of free chlorine) and 100 mL of MS2 previously grown to 10^9-10^{10} PFU/mL. The lettuce was then placed into the dewatering centrifuge to remove excess water prior to processing and three ~50-g samples were collected. The remaining virus inoculated lettuce was then sent through the entire processing line with three ~50-g samples taken at four different stages of processing. Samples were collected after shredding, flume washing (2 min), shaker table dewatering, and centrifugal drying following processing. At each station, designated workers wearing latex gloves placed a handful of lettuce into a Whirl-pak bag then weighed out 50-g samples which were transferred to a filter Whirl-pak bag. In addition, samples of centrifugation water and flume wash water were collected post-processing. Approximately 500 mL of water was collected from the centrifuge drain during the last stage of lettuce drying for both sanitizer-free and sanitizer experiments. For sanitizer experiments alone a total of two 20 L wash water samples were collected from the flume tank and further concentrated using hollow-fiber ultrafiltration. The flume tank wash water samples were collected to confirm efficacy of the sanitizer.

3.2.3 Lettuce Processing and Plating

Lettuce was processed to recover the phage using the elution procedure described in Part III, Section 3.1.2, pages 37-38. Each 50-g lettuce sample was eluted with 250 mL of tris-glycine buffer at pH 9.5 and shaken for 20 min with the eluent pH adjusted to 7.2 ± 0.2. For the sanitizer experiment, lettuce was immediately transferred to Whirl-pak bags containing 250 mL of buffer and 2% sodium thiosulfate following flume washing, shaking, and final centrifugal drying sampling. In addition, 2% sodium thiosulfate was added to flume wash water and centrifugal water sampled during sanitizer processing. The lettuce eluent and water samples were plated using the double agar overlay method described in Part III, Sections 3.1.1 and 3.1.2, pages 35-38.

3.2.4 Calculations and Statistical Analysis

The plaque concentrations for the inoculation suspension and recovered eluent from each ~50-g lettuce sample collected immediately after inoculation, after shredding, flume washing, shaking, and centrifuging were calculated using formula 1 (Part III, Section 3.1.3, page 39). In addition, the concentration of phage per g of romaine lettuce (PFU/g) was calculated using formula 2 (Part III, Section 3.1.3, page 39). To estimate the plaque concentration of the 80 L inoculation suspension, the inoculated concentrate (PFU/mL) was multiplied by the volume of inoculant concentrate added (100mL) which was then divided by the total volume of water (80000 mL) as shown in formula 5. The estimated concentration per g of lettuce following inoculation (but prior to centrifugal drying) was calculated by multiplying the estimated plaque concentration in the inoculation suspension by the total volume of the suspension (total phage in inoculation suspension) and dividing by the total weight of lettuce added to the suspension (6000g) (formula 6). An unpaired two sample t-test was used to compare virus removal during the various stages of processing ($P \leq 0.05$). For spent centrifuge water and flume water samples, the MS2 concentration (PFU/mL) was calculated using formula 1 (Part III, Section 3.1.3, page 39). The total viral reduction (PFU) in the flume water before and after washing was estimated by subtracting the total MS2 entering the flume wash water by the total MS2 exiting the flume wash water. Total MS2 entering the flume wash water was calculated by first calculating the total PFU on lettuce entering the flume (multiply the average PFU/g after shredding by the total lettuce (5700 g)) and subtracting the total PFU on lettuce exiting the flume (multiply the average PFU/g after flume washing by total lettuce (5700 g)). Total MS2 remaining in the flume wash water was calculated by first estimating the total PFU/mL in the 20 L sample (PFU/mL multiplied by eluent concentrate volume). This value was then divided by our sample volume

(20 L) and multiplied by the total wash water volume (800 L) to get the remaining PFU in the flume wash water (formula 7).

(5) $Phage\ Concentration\ in\ Inoculation\ Suspension\ \left(\frac{PFU}{mL}\right) =$

$$\frac{Seeded\ Concentration\left(\frac{PFU}{mL}\right) * Volume\ of\ Seed\ Concentrate\ (mL)}{Volume\ of\ Inoculation\ Suspension\ (mL)}$$

(6) Inoculated Lettuce MS2 Concentration $\left(\frac{\text{PFU}}{\text{g}}\right) = \frac{Total\ phage\ in\ inoculation\ suspension}{Total\ lettuce\ weight\ (g)}$

(7) Total MS2 reduction in flume water (PFU)=

$$\left(\text{Total MS2 entering flume on lettuce (PFU)-Total MS2 exiting flume on lettuce(PFU)}\right)-$$

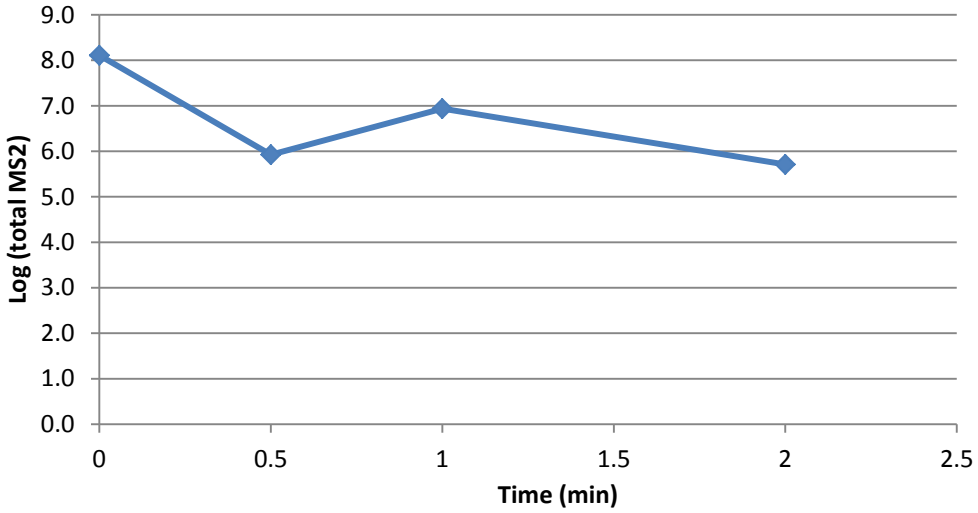
$$\left(\left(\frac{Total\ MS2\ in\ 20\ L}{20\ L}\right) * flume\ wash\ volume\right)$$

4. Results

4.1 MS2 Inactivation during Sanitizer Exposure

A benchtop inactivation experiment was conducted prior to small-scale leafy green processing trials to determine the relative inactivation of MS2 upon exposure to 25 ppm of free chlorine. One mL of high titer (10^{10} PFU/mL) TSB-based MS2 was added to the sanitizer solution (100 mL) and samples were processed after 30 s, 1 min, and 2 minutes of exposure. The time series was chosen based on standard commercial lettuce washing and exposure practices. For the liquid suspension challenge experiment, all samples started with 25 ppm of free chlorine and were immediately neutralized (0 ppm) with 2% sodium thiosulfate after the specific contact times. Figure 20 shows the log transformed total phage values at time 0 (represented by the control) and after 30 s, 1 min, and 2 min of exposure to 25 ppm free chlorine. Compared to the initial population, the number of phage decreased 2.2, 1.2, and 2.4 logs after 0.5, 1, and 2 min time points, respectively, with an overall average reduction of 1.9 logs.

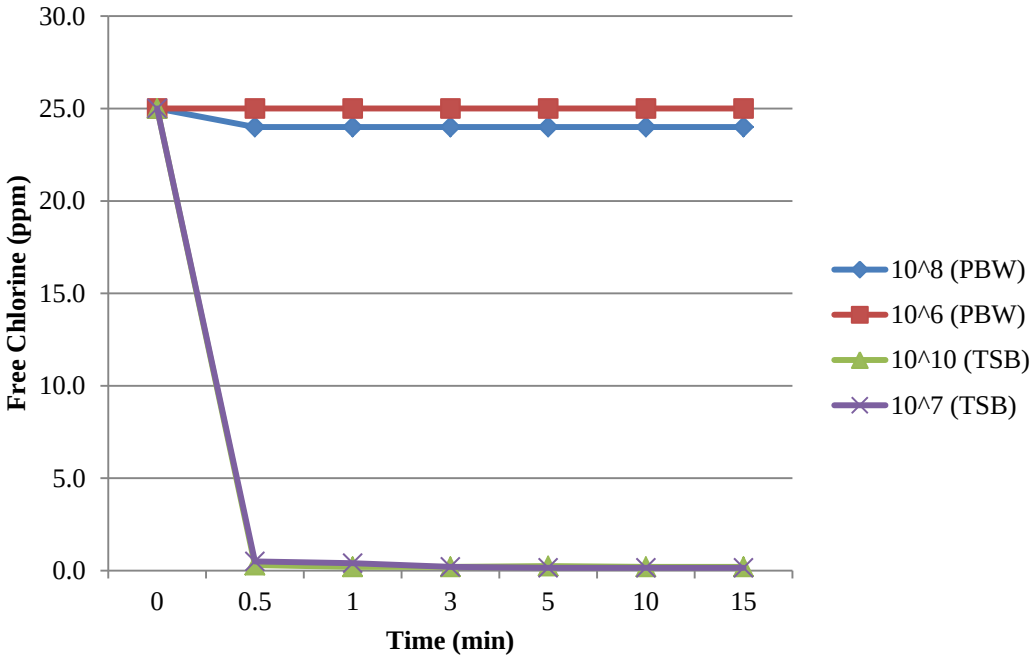
Figure 20. Coliphage MS2 inactivation after exposure to 25 ppm free chlorine in liquid suspension



4.2 Coliphage MS2 Chlorine Demand

The phage preparation was not purified and thus may carry a chlorine demand. This was evaluated and Figure 21 shows the free chlorine demand after adding 1 mL of MS2 phage stock (suspended in either TSB or PBW) to 100 mL of the wash solution (~2 fold dilution). After 15 min, the free chlorine level remained unchanged or only dropped 1 ppm when exposed to 1 ml of high titer (10^6 and 10^8) PBW-based MS2 in 100 mL of sanitizer. After 30 s of exposure to 1 mL TSB-based MS2 at 10^{10} and 10^7 (PFU/mL), free chlorine levels were reduced to 0.3 and 0.5 ppm respectively. After 15 min, free chlorine levels were reduced to 0.2 and 0.1 ppm, using 10^{10} and 10^7 (PFU/mL) TSB-based MS2 respectively. These data show that TSB exhibited a demand on for free chlorine. This could be similar to an actual virus contamination event in which the virus is associated with fecal matter.

Figure 21. MS2 free chlorine demand



* Green line representing 10^{10} TSB trial is located behind purple line representing 10^{7} TSB trial

4.3 Virus Reduction during Leafy Green Processing

Small-scale experiments were conducted with and without sanitizer to determine virus reduction on lettuce during simulated commercial processing. Lettuce was inoculated by immersion in 80 L of municipal tap water containing 100 mL of TSB-based MS2 (10^{9}-10^{10} PFU/mL) and examined prior to processing (after inoculation and drying) as well as following shredding, 2 min of flume washing, shaker table dewatering, and centrifugal drying. TSB-based MS2 (100 mL) was diluted ~3 fold when added to tap water (80 L) and was therefore not diluted in PBW prior to inoculation. For lettuce washing, a chlorine based sanitizer was used according to the manufacturer's instructions. In addition, lettuce was sampled prior to MS2 inoculation as a negative control. The first sanitizer trial was the only trial (including trials with and without sanitizer) to have a positive result in the negative control lettuce. A positive result was observed

for all three lettuce samples (triplicates) when plated undiluted on TSA overlay plates, with no more than 5 plaques on a plate and an average concentration of 3.5 PFU/g. This indicates either cross contamination during sampling or plating or the presence of *E. coli* Famp on the lettuce prior to processing.

The MS2 concentrations in the original culture, the inoculation suspension, and estimated MS2 concentration on the lettuce prior to drying are provided for both sanitizer and non-sanitizer experiments in Table 22. The MS2 concentration in the inoculation suspension was approximately 10^7 and 10^6 PFU/mL with average MS2 levels on romaine lettuce of 8.5 log and 7.8 log (PFU/g) for sanitizer free and sanitizer trials, respectively. The average MS2 concentrations on romaine lettuce (PFU/g) for triplicate lettuce samples collected before processing and after shredding, flume washing, shaker table dewatering, and centrifugal drying for both sanitizer and non-sanitizer experiments are presented in Table 23. For all three trials, the average MS2 concentration on romaine lettuce before processing (following inoculation and drying) was 6.0 and 4.9 log (PFU/g) for the sanitizer-free and sanitizer experiments, respectively. Following shredding, flume washing, shaker table dewatering, and centrifugal drying, the average MS2 concentration decreased to 5.0 and 4.2 log (PFU/g) for sanitizer-free and sanitizer experiments, respectively. Free chlorine levels tested following processing showed a reduction of 10 ppm, 5 ppm, and 5 ppm for trials 1, 2, and 3 respectively.

Table 22. Estimated MS2 concentration in inoculation suspension and on romaine lettuce following inoculation (prior to centrifugal drying)					
Wash water	Trial	Original culture concentrate (Log PFU/mL)	Estimated MS2 concentration in inoculation suspension (Log PFU/mL)	Estimated MS2 concentration on lettuce (Log PFU/g)	Average estimated MS2 concentration on lettuce (Log PFU/g)
Sanitizer-free	1	10.4	7.5	8.7	8.5
	2	10.4	7.4	8.6	
	3	10.1	7.2	8.3	
Sanitizer	1	9.5	6.6	7.7	7.8
	2	9.5	6.6	7.7	
	3	9.7	6.8	7.9	

Table 23. Average MS2 concentration on romaine lettuce for triplicate lettuce samples collected following various stages of small-scale leafy greens processing						
Flume wash water	Trial	Average Concentration of MS2 on romaine lettuce (log PFU/g)				
		Before processing	Shredder	Flume wash	Shaker	Centrifuge
Sanitizer-free	1	6.1	5.8	5.2	4.9	5.1
	2	5.9	5.8	4.9	4.7	5.0
	3	5.9	5.7	4.8	4.9	4.9
	Average	6.0	5.8	5.0	4.8	5.0
Sanitizer	1	5.0	4.9	4.4	4.4	4.5
	2	4.9	5.0	3.9	3.9	4.0
	3	4.8	4.9	4.2	4.1	4.0
	Average	4.9	4.9	4.2	4.2	4.2

The MS2 concentrations on romaine lettuce at various stages of leafy green processing without and with the sanitizer treatment are shown in Figures 22 and 23, respectively. For sanitizer-free experiments, the average lettuce phage concentrations before processing were 6.1, 5.9, and 5.9 log PFU/g for trials 1, 2 and 3 respectively. Following shredding, 2 min of flume washing (tap water alone), shaker table dewatering, and centrifugal drying, MS2 phage populations decreased to 5.0, 5.0, and 4.9 log PFU/g for trials 1, 2 and 3 respectively, with an overall average reduction of 1.0 log PFU/g. For experiments with 25 ppm of free chlorine included in the flume water, the average lettuce phage concentrations before processing were 5.1, 4.9, and 4.8 log PFU/g for trials 1, 2 and 3 respectively. Following shredding, 2 min of flume washing (with sanitizer), shaker table dewatering, and centrifugal drying, MS2 phage populations decreased to 4.5, 4.0, and 4.0 log PFU/g for trials 1, 2 and 3 respectively, with an overall average reduction of 0.8 log PFU/g. Although centrifugal drying was the final lettuce processing step, the largest average reduction occurred following shaking, reducing phage populations by a total of 0.8 and 1.1 log PFU/g for experiments with and without sanitizer, respectively. In addition, coliphage MS2 numbers remained relatively stable following flume washing with or without 25 ppm free chlorine. No statistical difference was observed for the total phage reduction (log PFU/g) between water and sanitizer treatments ($P>0.05$).

Figure 22. MS2 reduction on inoculated romaine lettuce during small-scale leafy green processing without sanitizer

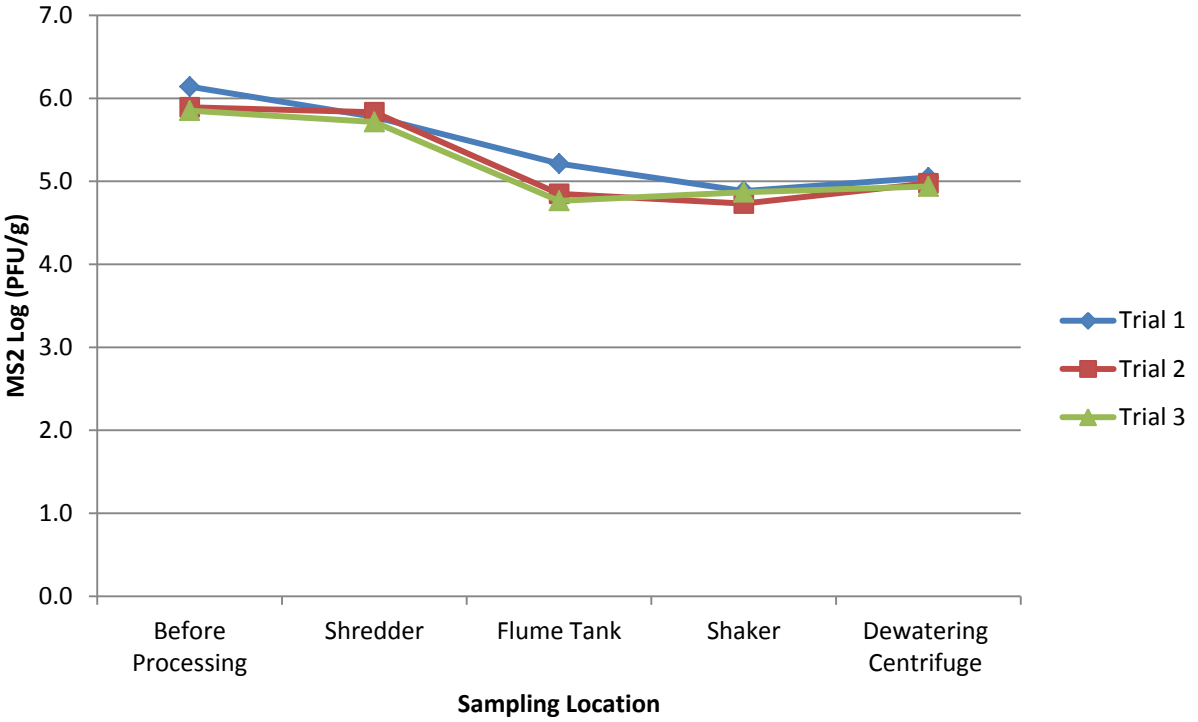
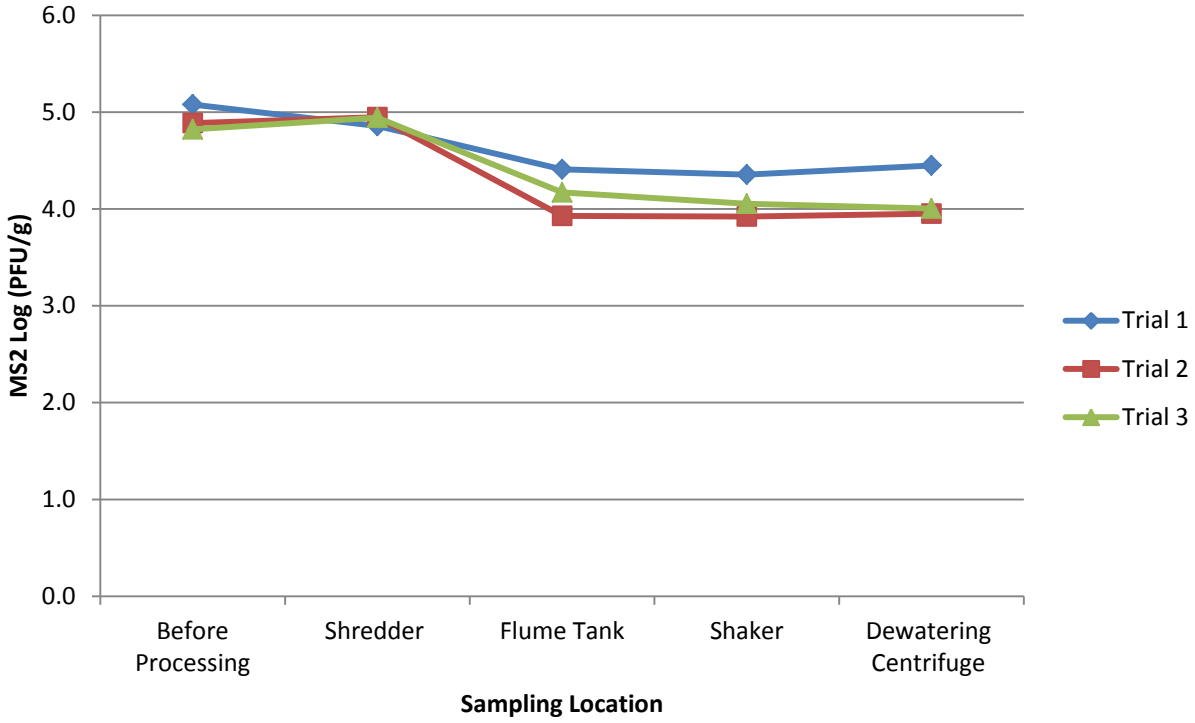


Figure 23. MS2 reduction on inoculated romaine lettuce during small-scale leafy green processing with sanitizer



For both sanitizer and sanitizer-free treatments, the reduction of phage populations on romaine lettuce between consecutive stages of food processing was compared (Table 24). The largest reduction occurred between shredding and flume washing, ranging from 0.5 to 1.0 log PFU/g and 0.6 to 1.0 log PFU/g for sanitizer and sanitizer-free treatments, respectively. For sanitizer-free experiments, phage populations decreased an average of 0.2 log PFU/g between before processing and shredding and 0.1 PFU/g between fluming and shaking. This is comparable to experiments with sanitizer where little reduction of MS2 on lettuce occurred before or following flume washing. Interestingly, lettuce phage populations for both sanitizer and sanitizer-free experiments increased slightly between shaking and centrifugation in all trials but one (trial 3 sanitizer).

Table 24. Log reduction of MS2 on romaine lettuce between consecutive processing stages

Flume Wash Water	Trial	Log removal (PFU/g)			
		Before processing-Shredding	Shredding-Flume	Flume-Shaker	Shaker-Centrifuge
Sanitizer-free	1	0.4	0.6	0.3	-0.2
	2	0.1	1.0	0.1	-0.3
	3	0.1	1.0	-0.1	-0.1
	Average	0.2	0.8	0.1	-0.2
Sanitizer	1	0.2	0.5	0.1	-0.1
	2	-0.1	1.0	0.0	-0.0
	3	-0.1	0.8	0.1	0.1
	Average	0.0	0.8	0.1	0.0

The MS2 concentration (PFU/mL) in centrifugation water decreased during lettuce drying following processing, ranging from 5.0 log PFU/mL to 5.5 log PFU/mL (average 5.2 log PFU/mL) and 3.8 log PFU/mL to 4.2 log PFU/mL (average 4.0 log PFU/mL) for the sanitizer-free and sanitizer experiments, respectively. Two, 20 L flume water samples were also collected after sanitizer processing of lettuce. The average MS2 concentration in 20 L of flume wash water was 3.2 PFU/mL and did not exceed 4.6 PFU/mL for all three sanitizer trials. MS2 reduction due

to flume washing with sanitizer was estimated and included in Table 25. For all three trials, the average total MS2 log (PFU) reduction in the flume was water was about 4 logs.

Table 25. Flume water MS2 levels and reduction following processing

Trial	PFU				log (PFU)		
	Total MS2 on lettuce entering flume	Total MS2 on lettuce exiting flume	Total MS2 lost on lettuce during flume wash	Total PFU in 20 L of flume water	Total MS2 in flume water before washing	Total MS2 in flume water remaining	MS2 Removal
1	$4.1x10^8$	$1.5 x10^8$	$2.6 x10^8$	$8.0 x10^2$	8.4	4.5	3.9
2	$5.1 x10^8$	$4.8 x10^7$	$4.6 x10^8$	$1.2 x10^3$	8.7	4.7	4.0
3	$5.0 x10^8$	$8.4 x10^7$	$2.9 x10^8$	$8.1 x10^2$	8.6	4.5	4.1
Avg	$4.7 x10^8$	$9.3 x10^7$	$3.4 x10^8$	$9.3 x10^2$	8.6	4.6	4.0

5. Discussion

Viral foodborne diseases are a growing concern for raw produce. With consumer demand increasing for fresh-cut lettuce, this study was focused on determining if the use of chlorine-based sanitizers might provide some ability to decrease virus on incoming lettuce. MS2 coliphage was used as an enteric virus surrogate on romaine lettuce during simulated commercial processing. This phage is non-enveloped, composed of ssRNA, and shares similar resistance characteristics to enteric viruses including norovirus and can thus be used to represent foodborne viruses that cannot be propagated in cell culture (Dawson et al., 2005).

A small-scale laboratory experiment was conducted to determine the efficacy of a chlorine-based sanitizer against MS2 in liquid suspension. This experiment demonstrated a range in virus inactivation from 1.2-2.4 logs (PFU) and was comparable to a study by Doultree et al. (1999). They examined the effect of 100- 500 ppm of free chlorine (Det-Sol 5000 sanitizer) using feline calicivirus and reported 1.75 to 2.75 log inactivations in liquid suspension following a 1 min exposure. This suggests that the inactivation efficacy of 25 ppm free chlorine against MS2 virus in liquid suspension as observed in this study is comparable to inactivation rates for

other enteric virus surrogates at higher free chlorine levels (some exceeding maximum levels allowed in the food industry).

One of the issues is the demand for the halogen within the food processing environment which leads to the depletion of free chlorine. A small-scale laboratory experiment conducted to examine the chlorine demand of high titer, TSB-cultured, MS2 demonstrated this phenomenon, finding that TSB-based MS2 had a high demand for free chlorine immediately following exposure to 25 ppm while MS2 diluted in PBW did not. Previous studies have shown that organic compounds, including culture media and microbial suspensions, have a high demand for chlorine and reduce free chlorine levels following exposure (Kotula et al., 1997, Shang & Blatchley, 2001). These data suggest that the presence of organic substrates in liquid suspensions strongly impacts sanitizer efficacy against viruses. Although TSB-based MS2 was shown to reduce free chlorine demand, we used MS2 TSB stock solutions in our inoculums for simulated commercial processing. Following inoculation, TSB-based cultures were diluted 3-fold and further removal occurred following lettuce drying and shredding prior to flume washing. Therefore, the effect of TSB-cultured virus on free chlorine in the flume wash water suspension was expected to be minimal. However, this could represent the type of organic material that might be associated with naturally occurring viruses which are found on lettuce from fecal contamination.

Processing and washing lettuce with water alone reduced MS2 populations on lettuce an average of 1 log (PFU/g). This is comparable to simulated commercial processing and laboratory studies where virus reductions of ~0.5 to 1 log were reported for strawberries and lettuce, respectively, following washing with tap water alone (Casteel et al., 2009, Fraisse et al., 2011). In the present processing experiments, a 2 min wash in 25 ppm of free chlorine led to an

average reduction following processing of 0.8 log (PFU/g), which was not significantly different from processing with tap water (<0.05 ppm of free chlorine) alone. Hence, these manufacturer recommended free chlorine levels do not provide enhanced inactivation against non-enveloped, RNA viruses when present on fresh produce. However following processing, free chlorine levels were reduced by 5 to 10 ppm which, as already stated, reduces the efficacy of the disinfection process. It is likely that this is a result of the organic load introduced by the lettuce during fluming. Although chlorine levels were reduced, previous laboratory studies have shown that RNA virus reduction on fresh produce (strawberries, lettuce, or tomato) ranged from 1 to 3 logs when exposed to 15-20 ppm of free chlorine for up to 10 minutes (Casteel et al., 2009; Fraisse et al., 2011). In the present study, results were similar to a simulated commercial processing study where 20 ppm of free chlorine in wash water reduced MS2 populations on strawberries by ~1 log (Casteel et al., 2009). Furthermore, laboratory studies using higher levels of free chlorine ranging between 50 ppm and 800 ppm could not achieve more than a 3 log reduction of non-enveloped RNA viruses (feline calicivirus, hepatitis A, MS2, murine norovirus) when inoculated on fresh produce (Allwood et al., 2004, Dawson et al., 2005, Fraisse et al., 2011, Gulati et al., 2001). These data suggest that enteric viruses and enteric virus surrogates are very stable and resistant to current commercial chlorine disinfection techniques used in food processing.

Many studies investigating virus survival on fresh produce have also observed enhanced virus survival when attached to leafy greens and exposed to varying environmental conditions or disinfection techniques (Badawy et al., 1985, Croci et al., 2002). To date there is limited knowledge as to why virus survival is greater on lettuce compared to other food commodities. It is possible that enteric viruses are protected due to the rough or rigid surfaces of lettuce or resistant to naturally occurring antimicrobials (Badawy et al., 1985, Seymour & Appleton, 2001).

In addition, a study where cut romaine lettuce was dip inoculated into biosolids containing murine norovirus has shown that viruses have the ability to internalize (access inner tissues) cut edges of leaflet surfaces which may offer protection from chemical disinfection (Wei et al., 2010).

The main reason for including a sanitizer in the wash water is to inactivate viruses after they have been removed from fresh produce surfaces to ultimately prevent cross contamination from viruses recirculating in the flume wash. Therefore, the flume water was sampled following processing with the sanitizer. MS2 was present in all flume water samples processed and was concentrated using ultrafiltration. Estimation of the reduction of MS2 in flume wash water suggested that 25 ppm free chlorine had greater efficacy in the wash water than on lettuce (~4 logs). However, the initial MS2 concentration in the wash water was calculated based on the difference in total MS2 entering and exiting the flume tank on lettuce and therefore our log reduction is an estimation, assuming all viruses were removed into the 800 L of wash water. These data together suggest that viruses are very resistant to chlorine when present on the surface of romaine lettuce. In addition, since MS2 acts as a surrogate for foodborne viruses (e.g. norovirus and hepatitis A), enteric viruses are also likely to persist on lettuce given current fresh produce processing practices.

V. CONCLUSIONS, LIMITATIONS, AND FUTURE NEEDS

Although innovative next-generation sequencing technologies have resulted in greater throughput and significant cost-reductions for sequencing entire microbial communities, there are still limitations when applied to viral metagenomics (Kircher & Kelso, 2010; L. Liu et al., 2012). For example, there is currently no standard method for studying the virome in environmental samples. In addition, metagenomics must overcome bias associated with each stage of virus recovery and detection including elution, concentration, purification, extraction, and PCR amplification (Aw et al., 2014). Current computational tools used for bioinformatics analysis are intensive and the analysis of entire communities is difficult and time consuming. Novel bioinformatics tools are desirable for more rapid analysis of the large amount of data obtained from next-generation sequencing.

Despite these limitations next-generation sequencing and metagenomics provides a broad view of food safety like never before, particularly for hazards such as viruses where traditional methods have limited the ability to monitor. Metagenomic techniques can be applied to a wide range of environmental samples including foods and the use and development of new genomics tools will aide in identifying and assessing better monitoring targets from harvesting to food processing. This is the first study to use next-generation sequencing and metagenomics to characterize the virome in irrigation water and lettuce. The large number of no hits in each sample demonstrates the importance of building upon current viral databases to gain knowledge into the virus world. In addition, these results suggest that metagenomic techniques can be used to successfully identify specific viruses of importance to the food industry and public health, including foodborne human enteric viruses. Specifically, dsRNA enteric viruses including rotavirus A and human picobirnavirus were identified in irrigation water and lettuce, suggesting

these viruses are a food safety concern beginning at the farm-level stage of production. For future metagenomic studies on fresh produce, the author recommends a larger sample size at each stage of field production. In addition, in order to determine the viral risks from farm-to-fork, lettuce needs to be sampled throughout the supply chain including post-harvest production. Currently, viral metagenomic analysis of lettuce collected from distribution centers is being conducted and will hopefully provide a better view of the lettuce virome and viral hazards following postharvest.

Having found a wide variety of viruses, including human enteric viruses, on lettuce from the field using metagenomics there was interest in how removal might be achieved during processing after harvest. In this study, simulated commercial leafy green processing experiments demonstrated that MS2 is very resistant to chlorine when present on the surface of romaine lettuce. Since MS2 acts as a surrogate for foodborne viruses, enteric viruses on lettuce are also likely to be very stable and resistant to current commercial chlorine disinfection techniques used with food processing systems. This suggests current recommended commercial production practices are unable to effectively decrease virus on leafy green during processing

These data advocate that disinfection methods other than washing with a chlorine-based sanitizer should be investigated further for leafy green processing. To date, numerous physical (ultrasound high pressure, ultraviolet, ionizing radiation) and chemical (chlorine dioxide, sodium chlorite, quaternary ammonium compounds, hydrogen peroxide, peroxyacetic acid) disinfection methods have been investigated for use in foods (CFSAN, 2014). Many viral disinfection studies on leafy greens in particular have investigated peroxyacetic acid and hydrogen peroxide as sanitizers. These chemicals have achieved 1-3 log reductions for feline calicivirus on leafy greens in laboratory studies. However, these same treatments can negatively impact end product

quality. The human and environmental safety of these compounds also warrants further investigation (Allwood et al., 2004, Baert et al., 2009, CFSAN, 2014, Fraisse et al., 2011, Gulati, et al., 2001).

This study demonstrates the importance of minimizing viral contamination on leafy greens prior to processing. There are many critical control points from farm-to-fork that aid in the reduction of fresh produce contamination. To date, the FDA has developed laws, rules, and guidelines to combat human pathogen contamination and transmission in fresh produce, all of which recommend clean, safe water at all levels of production. At the pre-harvest level, potential sources of microbial contamination include irrigation water, pesticide and fertilizer application, cooling, and frost control (CFSAN, 2009). In respect to irrigation water, recognition of the water source, historical use of the land, and existing human or agricultural practices are all important considerations that help identify potential sources of microbial contamination (CFSAN, 2009). Recommended practices for limiting microbial hazards in irrigation water include restricting livestock and wildlife access to water to prevent fecal contamination, well and septic tank maintenance, treating and testing, as well as conservation practices such as sod waterways, vegetative buffers, and runoff control structures that help limit runoff pollution (CFSAN, 2009). Worker hygiene is a major control point at both pre- and postharvest levels of production. Suggested practices to provide safe produce include proper hand washing, use of disposable gloves, and not coming to work when ill as well as employee training to address these expectations. These practices combined with proper disinfection techniques are critical for supplying safe, quality leafy greens to the consumer and ultimately protecting public health and food safety.

APPENDIX

Figure A1. Yuma irrigation water sampling locations

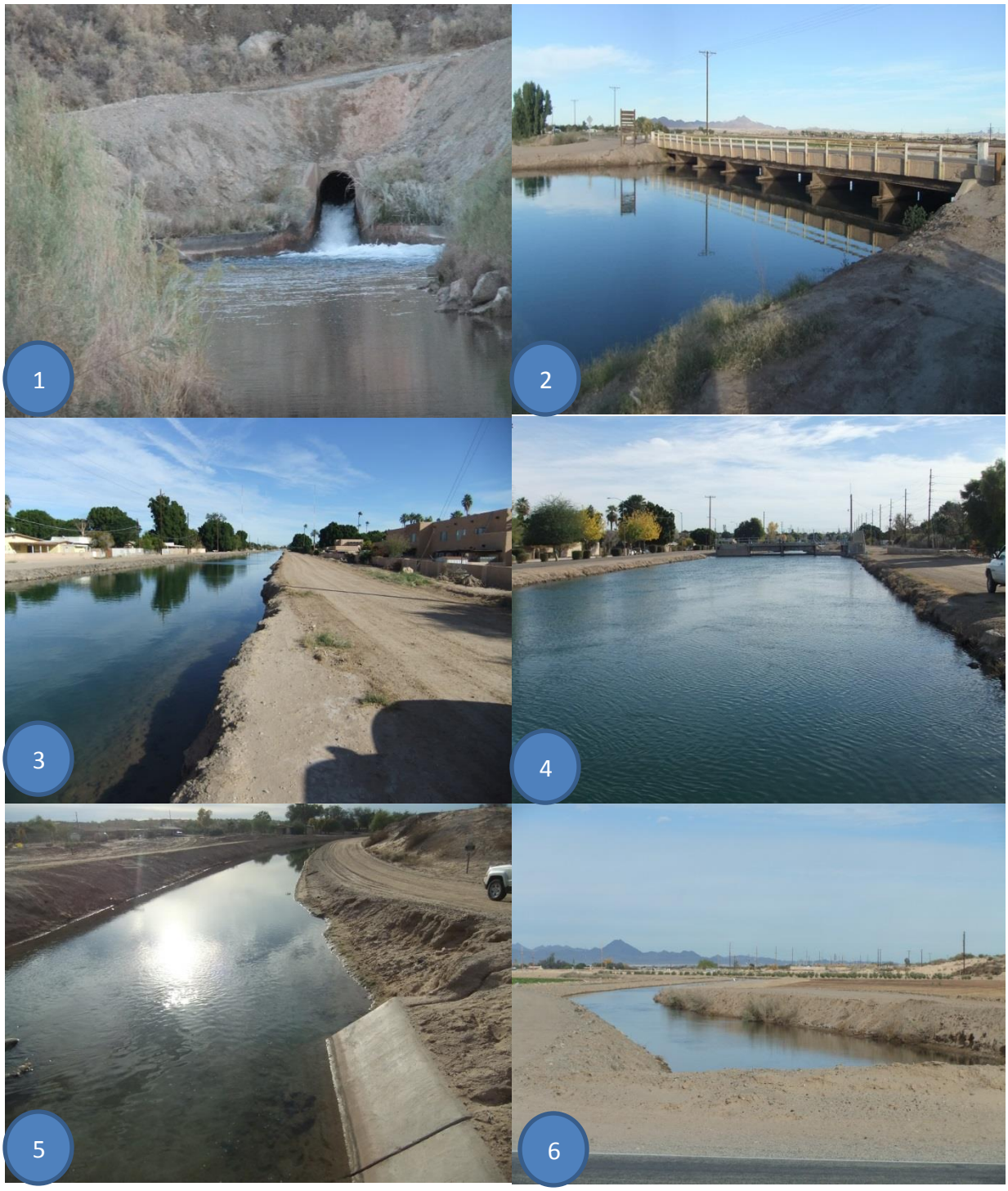


Figure A2. Examples of Yuma furrow irrigation water (1), iceberg lettuce following harvest (2), and romaine lettuce following harvest (3)

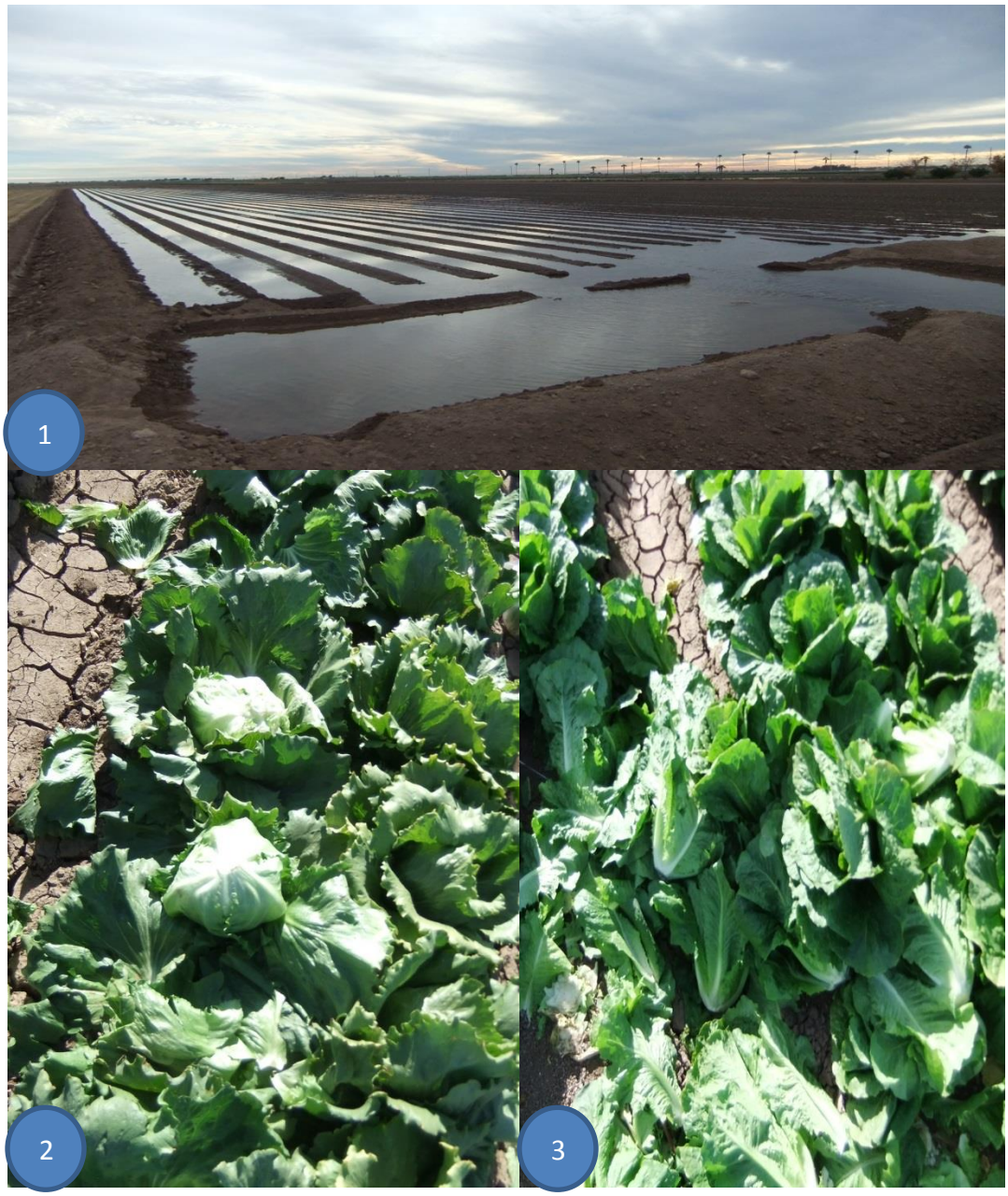


Figure A3. Ultrafiltration system for the primary concentration of irrigation water samples

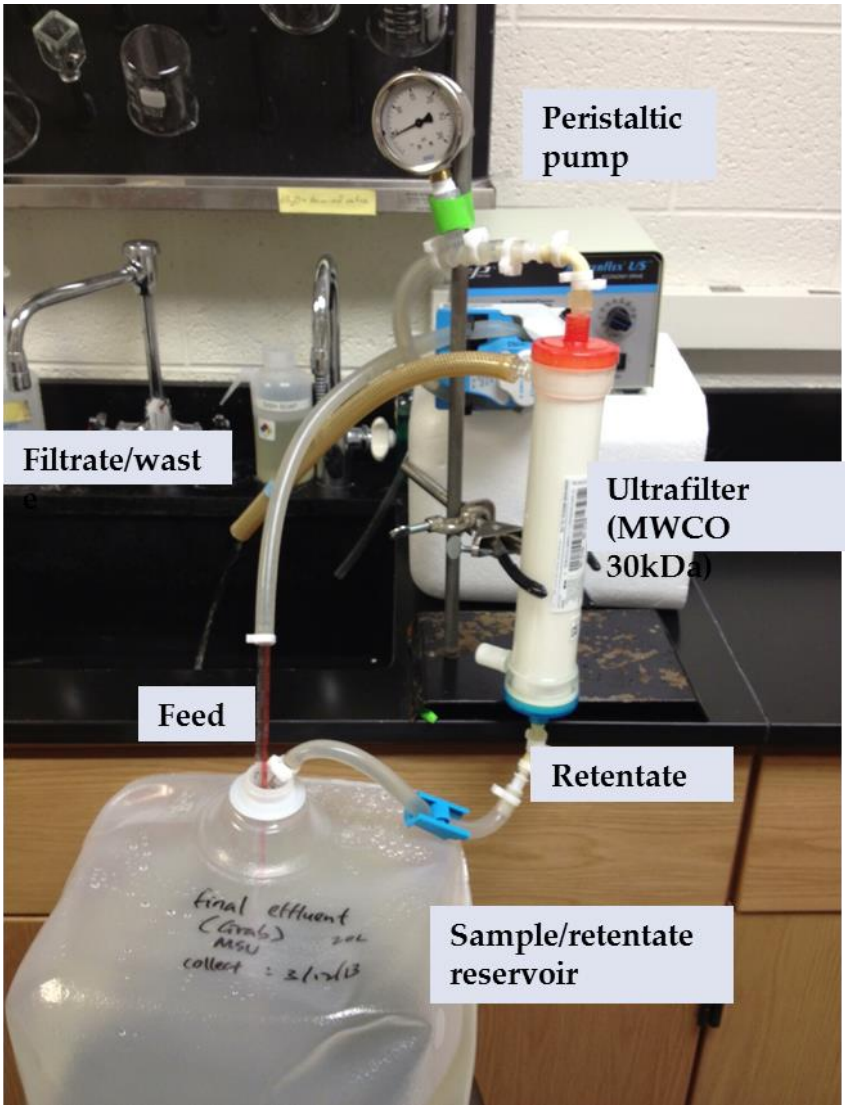


Figure A4. Lettuce centrifugal dryer (1), shredder (2), flume tank (3), and mechanical shaker table (4) sampling locations

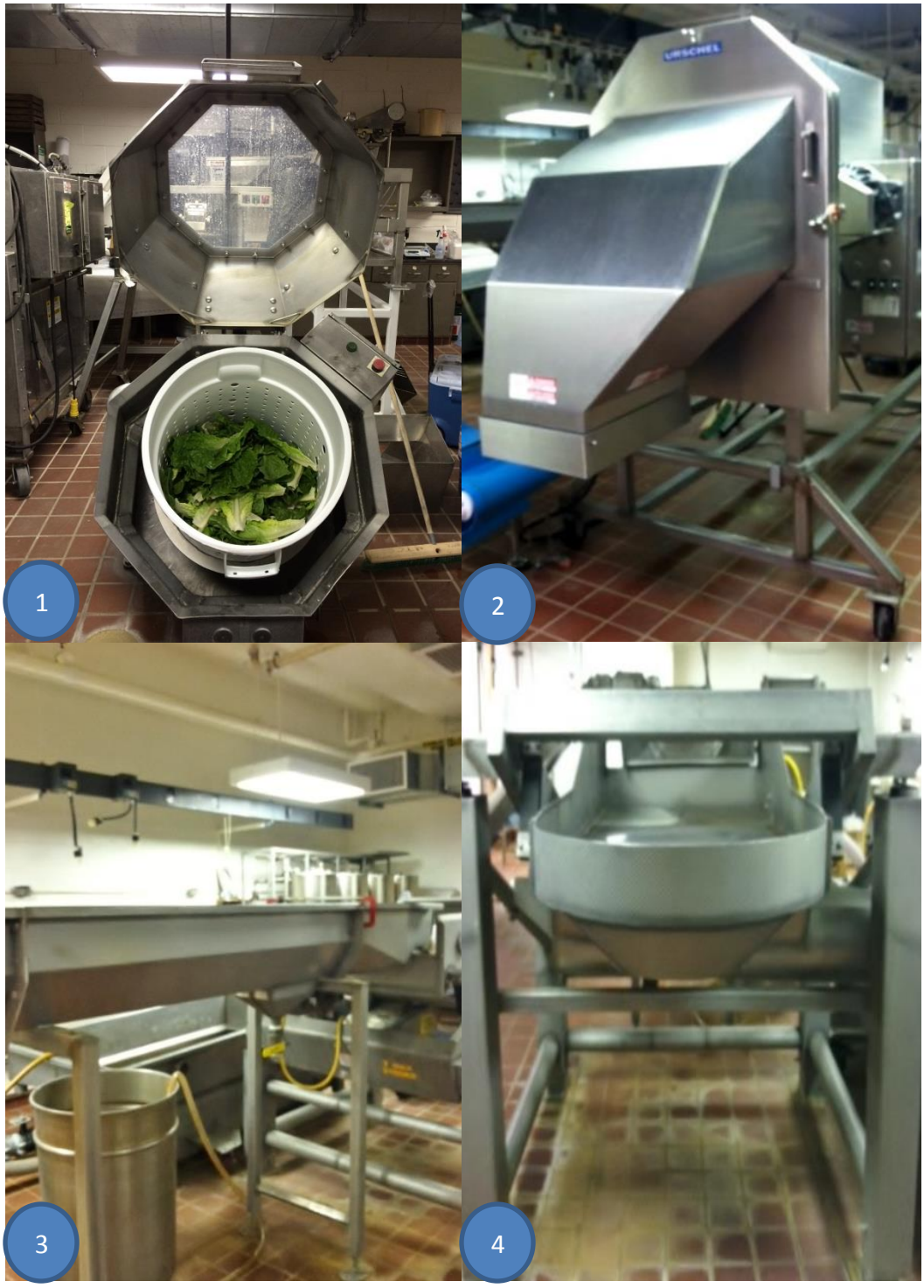


Table A1. Raw contig sequence data used for Yuma irrigation water virome analysis								
Families	Viral types	Host	YW1	YW2	YW3	YW4	YW5	YW6
Viruses	Unclassified	NA*	353	179	88	131	482	293
dsDNA viruses, no RNA stage	dsDNA	NA *	1993	397	534	852	2303	2038
Ascoviridae	dsDNA	Animals (invertebrate)	24	2	5	16	12	18
Asfarviridae	dsDNA	Animals (vertebrate)	2	0	0	0	0	0
Baculoviridae	dsDNA	Animals (invertebrate)	179	8	2	2	14	7
Bicaudaviridae	dsDNA	Archaea	2	0	0	0	0	0
Caudovirales	dsDNA	Bacteria	713	319	410	682	2106	1637
Myoviridae	dsDNA	Bacteria	1460	683	894	1498	3595	3400
Podoviridae	dsDNA	Bacteria	770	416	473	842	2488	1820
Siphoviridae	dsDNA	Bacteria	1111	489	607	1008	3019	2352
unclassified *Caudovirales*	dsDNA	Bacteria	53	31	38	69	251	166
Herpesvirales	dsDNA	Animals (vertebrate)	3	0	1	1	0	1
Alloherpesviridae	dsDNA	Animals (vertebrate)	17	2	3	3	10	10
Herpesviridae	dsDNA	Animals (vertebrate)	101	2	4	8	20	17
unclassified *Herpesvirales*	dsDNA	Animals (invertebrate)	2	0	1	5	2	4
Iridoviridae	dsDNA	Animals (vertebrate and invertebrate)	88	16	23	26	86	70
Ligamenvirales	dsDNA	Archaea	0	0	0	0	0	0
Lipothrixviridae	dsDNA	Archaea	2	2	0	3	5	3
Rudiviridae	dsDNA	Archaea	0	0	0	0	1	1
Marseilleviridae	dsDNA	Other (ameoba, protozoa, or fungi)	116	8	9	13	26	21
Mimiviridae	dsDNA	Other (ameoba, protozoa, or fungi)	1080	67	136	225	354	458
Nimaviridae	dsDNA	Animals (invertebrate)	7	0	0	0	2	0
Nudiviridae	dsDNA	Animals (invertebrate)	19	2	5	2	12	9
Papillomaviridae	dsDNA	Animals (vertebrate)	1	0	0	0	0	0
Phycodnaviridae	dsDNA	Algae	997	131	133	216	560	517
Polydnaviridae	dsDNA	Animals (invertebrate)	10	6	3	9	25	14

Table A1 (cont'd)

Polyomaviridae	dsDNA	Animals (vertebrate)	1	0	0	0	0	0
Poxviridae	dsDNA	Animals (vertebrate and invertebrate)	503	7	12	16	45	32
Tectiviridae	dsDNA	Bacteria	6	5	1	2	2	2
unclassified dsDNA phages	Unclassified	Bacteria	166	70	102	230	633	485
unclassified dsDNA viruses	Unclassified	NA*	485	35	63	108	235	240
dsRNA viruses	dsRNA	NA *	0	0	0	0	0	0
Chrysoviridae	dsRNA	Other (ameoba, protozoa, or fungi)	0	1	0	0	0	1
Cystoviridae	dsRNA	Bacteria	0	1	0	0	0	0
Endornaviridae	dsRNA	Plants	2	0	0	0	0	0
Partitiviridae	dsRNA	Other (ameoba, protozoa, or fungi)	4	4	2	2	0	5
Picobirnaviridae	dsRNA	Animals (vertebrate)	2	1	0	0	0	0
Reoviridae	dsRNA	Animals and plants	10	5	0	1	1	3
Totiviridae	dsRNA	Other (ameoba, protozoa, or fungi)	3	1	0	0	0	0
unclassified dsRNA viruses	Unclassified	NA*	1	0	0	0	0	1
Retro-transcribing viruses	Retro-transcribing	NA*	1	0	1	0	1	0
Caulimoviridae	Retro-transcribing	Plants	1	0	0	0	1	2
Retroviridae	Retro-transcribing	Animals (vertebrate)	4	0	2	0	0	1
Satellites	Satellites	NA*	34	27	11	11	29	17
ssDNA viruses	ssDNA	NA*	306	250	64	38	301	109
Circoviridae	ssDNA	Animals (vertebrate and invertebrate)	350	304	75	58	257	147
Geminiviridae	ssDNA	Plants	136	113	21	11	90	41
Inoviridae	ssDNA	Bacteria	5	2	7	7	2	6
Microviridae	ssDNA	Bacteria	1471	1174	255	233	2900	440
Nanoviridae	ssDNA	Plants	44	37	11	5	23	19
Parvoviridae	ssDNA	Animals (vertebrate and invertebrate)	69	56	11	7	45	30
Parvovirus NIH-CQV**	NA*	NA*	233	177	62	55	278	118
unclassified ssDNA viruses	Unclassified	NA *	602	475	92	71	672	184

Table A1 (cont'd)

ssRNA viruses	ssRNA	NA*	0	0	0	0	0	0
ssRNA negative-strand viruses	ssRNA	NA*	1	0	0	0	0	0
Bunyaviridae	ssRNA	Animals (vertebrate and invertebrate)	6	1	0	0	0	0
Tenuivirus (genus)	ssRNA	Plants	0	1	0	0	0	0
ssRNA positive-strand viruses, no DNA stage	ssRNA	NA*	44	38	7	5	50	9
Alphatetraviridae	ssRNA	Animals (invertebrate)	1	2	0	0	1	1
Astroviridae	ssRNA	Animals (vertebrate)	1	1	0	0	1	1
Benyviridae	ssRNA	Plants	6	5	6	3	4	4
Bromoviridae	ssRNA	Plants	1	1	0	2	1	1
Caliciviridae	ssRNA	Animals (vertebrate)	2	1	0	0	1	0
Carmotetraviridae	ssRNA	Animals (invertebrate)	0	0	0	0	1	1
Closteroviridae	ssRNA	Plants	1	1	0	0	1	0
Hepeviridae	ssRNA	Animals (vertebrate)	6	10	1	0	7	5
Higrevirus (genus)	ssRNA	Plants	1	0	0	0	0	0
Leviviridae	ssRNA	Bacteria	30	17	9	13	0	1
Luteoviridae	ssRNA	Plants	2	0	0	0	1	0
Narnaviridae	ssRNA	Other (ameoba, protozoa, or fungi)	0	1	0	0	0	0
Nidovirales	ssRNA	Animals (vertebrate and invertebrate)	1	0	0	0	0	0
Coronaviridae	ssRNA	Animals (vertebrate)	4	0	0	0	0	0
Roniviridae	ssRNA	Animals (invertebrate)	2	0	0	0	0	0
Nodaviridae	ssRNA	Animals (vertebrate)	14	17	0	0	15	0
Ourmiavirus (genus)	ssRNA	Plants	38	32	4	7	27	7
Picornavirales	ssRNA	Animals and Plants	115	94	16	33	112	27
Dicistroviridae	ssRNA	Animals (invertebrate)	198	188	60	56	234	65
environmental samples <*Picornavirales*>	ssRNA	NA*	228	201	39	81	269	82
Iflaviridae	ssRNA	Animals (invertebrate)	7	5	0	0	3	1

Table A1 (cont'd)

Marnaviridae	ssRNA	Plants	4	3	2	2	6	2
Picornaviridae	ssRNA	Animals (vertebrate)	10	12	0	3	8	3
Secoviridae	ssRNA	Plants	4	7	2	1	3	1
unassigned *Picornavirales*	ssRNA	NA*	23	20	8	5	22	6
unclassified *Picornavirales*	ssRNA	Animals (vertebrate)	2	0	0	0	0	0
Potyviridae	ssRNA	Plants	9	5	1	3	7	5
Sobemovirus(genus)	ssRNA	Plants	10	7	1	1	3	3
Tombusviridae	ssRNA	Plants	99	90	19	23	74	27
Tymovirales	ssRNA	Plants	0	0	0	0	0	0
Tymoviridae	ssRNA	Plants	1	0	0	0	0	0
Umbravirus(genus)	ssRNA	Plants	1	1	0	0	1	0
unclassified ssRNA positive-strand viruses	ssRNA	NA*	102	86	18	32	117	47
Virgaviridae	ssRNA	Plants	4	2	0	0	0	0
unassigned ssRNA viruses	Unclassified	NA*	0	0	0	0	0	0
Alvernaviridae	ssRNA	Algae	4	1	0	2	2	1
unassigned viruses	Unclassified	NA*	0	0	0	0	0	0
Bacilladnavirus	Unclassified	Algae	3	2	2	1	4	5
Bidnaviridae	Unclassified	Animals (invertebrate)	2	3	1	1	1	2
Hytrosaviridae	Unclassified	Animals (invertebrate)	9	1	0	1	3	0
unclassified phages	Unclassified	Bacteria	19	9	10	14	50	40
unclassified virophages	Unclassified	NA*	10	1	2	11	37	20
unclassified viruses	Unclassified	NA*	3	2	1	0	4	3
Not assigned	NA*	NA*	5491	1477	1144	1791	5695	4145
No hits	NA*	NA*	107829	14533	13022	19023	49695	40505

* NA. Not applicable, data was not used in category analysis

** Contaminant removed from analysis

Table A2. Raw contig sequence data used for Yuma lettuce virome analysis									
Viral Assignment	Viral Type	Host	Iceberg Control	Iceberg Worker Harvest	Iceberg Post Worker Break	Romaine Control	Romaine Worker Harvest	Romaine Chop and Wash	Romaine Mixed Salad
Viruses	Unclassified	NA*	47	17	11	49	14	3	8
dsDNA viruses, no RNA stage	dsDNA	NA*	152	94	74	93	66	33	76
Ascoviridae	dsDNA	Animals (invertebrate)	4	1	2	0	0	1	1
Asfarviridae	dsDNA	Animals (vertebrate)	0	0	0	0	1	0	0
Baculoviridae	dsDNA	Animals (invertebrate)	2	3	1	1	1	0	1
Caudovirales	dsDNA	Bacteria	137	34	27	159	31	9	33
Myoviridae	dsDNA	Bacteria	205	150	78	400	71	12	56
Podoviridae	dsDNA	Bacteria	175	64	42	188	51	11	53
Siphoviridae	dsDNA	Bacteria	340	134	80	427	90	15	78
unclassified *Caudovirales*	dsDNA	Bacteria	9	2	0	2	2	0	4
Herpesvirales	dsDNA	Animals (vertebrate)	0	0	0	0	1	0	1
Alloherpesviridae	dsDNA	Animals (vertebrate)	0	0	1	0	0	0	2
Herpesviridae	dsDNA	Animals (vertebrate)	3	2	2	2	0	0	1
unclassified *Herpesvirales*	dsDNA	Animals (vertebrate)	0	1	1	0	0	0	0
Iridoviridae	dsDNA	Animals (vertebrate and invertebrate)	27	27	14	22	12	15	12
Marseilleviridae	dsDNA	Other (ameoba, protozoa, or fungi)	2	3	4	1	1	2	3
Mimiviridae	dsDNA	Other (ameoba, protozoa, or fungi)	68	53	36	38	25	12	31

Table A2 (cont'd)

Nudiviridae	dsDNA	Animals (invertebrate)	0	0	0	0	1	0	0
Papillomaviridae	dsDNA	Animals (vertebrate)	1	0	0	0	0	0	0
Phycodnaviridae	dsDNA	Algae	74	40	27	29	26	13	31
Polydnaviridae	dsDNA	Animals (invertebrate)	4	0	2	3	2	0	0
Poxviridae	dsDNA	Animals (vertebrate and invertebrate)	6	5	5	4	4	1	1
Salterprovirus (genus)	dsDNA	Archaea	1	0	0	0	0	0	1
Tectiviridae	dsDNA	Bacteria	5	0	1	1	0	0	1
unclassified dsDNA phages	Unclassified	NA*	38	15	4	56	6	2	17
unclassified dsDNA viruses	Unclassified	NA*	15	10	4	10	8	0	5
dsRNA viruses	dsRNA	NA*	0	1	0	1	3	1	0
Birnaviridae	dsRNA	Animals (vertebrate and invertebrate)	1	0	0	0	0	0	0
Chrysoviridae	dsRNA	Other (ameoba, protozoa, or fungi)	0	2	0	4	1	0	1
Cystoviridae	dsRNA	Bacteria	2	0	0	0	0	0	0
Endornaviridae	dsRNA	Plants	9	45	25	74	63	19	52
Hypoviridae	dsRNA	Other (ameoba, protozoa, or fungi)	0	0	0	1	0	0	0
Partitiviridae	dsRNA	Other (ameoba, protozoa, or fungi)	18	16	7	79	48	25	26
Picobirnaviridae	dsRNA	Animals (vertebrate)	1	0	0	0	4	1	1
Reoviridae	dsRNA	Animals and plants	0	26	10	3	23	4	16

Table A2 (cont'd)

Totiviridae	dsRNA	Other (ameoba, protozoa, or fungi)	28	51	9	88	95	48	64
unclassified dsRNA viruses	Unclassified	NA*	15	31	21	31	23	21	22
Retro-transcribing viruses	Retro-transcribing	NA*	30	50	32	25	31	42	26
Caulimoviridae	Retro-transcribing	Plants	238	357	200	189	168	157	158
Retroviridae	Retro-transcribing	Animals (vertebrate)	103	193	137	61	99	130	98
Satellites	Satellites	NA*	0	0	0	2	0	0	0
ssDNA viruses	ssDNA	NA*	17	5	2	6	3	2	2
Circoviridae	ssDNA	Animals (vertebrate and invertebrate)	12	5	2	10	3	2	3
Geminiviridae	ssDNA	Plants	2	1	0	0	1	1	0
Inoviridae	ssDNA	Bacteria	43	25	21	27	18	10	11
Microviridae	ssDNA	Bacteria	56	15	4	12	13	2	5
Nanoviridae	ssDNA	Plants	0	0	0	0	0	0	1
Parvoviridae	ssDNA	Animals (vertebrate and invertebrate)	0	2	0	4	0	1	0
Parvovirus NIH-CQV**	NA*	NA*	2	1	0	2	0	0	2
unclassified ssDNA viruses	Unclassified	NA*	43	17	4	54	25	11	7
ssRNA viruses	ssRNA	NA*	0	0	0	0	0	0	0
ssRNA negative-strand viruses	ssRNA	NA*	0	5	4	1	3	1	1
Bunyaviridae	ssRNA	Animals (vertebrate and invertebrate)	0	2	1	0	1	1	3
Mononegavirales	ssRNA	Animals and plants	1	0	0	0	0	1	0

Table A2 (cont'd)

Rhabdoviridae	ssRNA	Animals and plants	0	0	0	0	0	0	1
Ophioviridae	ssRNA	Plants	0	0	0	2	1	0	0
Tenuivirus (genus)	ssRNA	Plants	3	2	2	1	4	1	1
Varicosavirus	ssRNA	Plants	50	26	12	3	1	0	0
ssRNA positive-strand viruses, no DNA stage	ssRNA	NA*	3	5	0	6	4	3	1
Bromoviridae	ssRNA	Plants	0	3	0	2	4	0	3
Cilevirus (genus)	ssRNA	Plants	0	0	0	1	0	0	0
Closteroviridae	ssRNA	Plants	87	69	28	120	65	31	53
Flaviviridae	ssRNA	Animals (vertebrate and invertebrate)	1	1	0	0	1	1	0
Hepeviridae	ssRNA	Animals (vertebrate)	0	0	0	2	0	0	0
Leviviridae	ssRNA	Bacteria	2	0	0	1	0	0	0
Narnaviridae	ssRNA	Other (ameoba, protozoa, or fungi)	1	3	1	5	9	3	3
Nodaviridae	ssRNA	Animals (vertebrate)	0	0	0	4	2	0	3
Ourmiavirus (genus)	ssRNA	Plants	5	2	1	2	4	0	0
Picornavirales	ssRNA	Animals and Plants	3	2	2	1	0	0	0
Dicistroviridae	ssRNA	Animals (invertebrate)	7	10	0	11	20	1	7
environmental samples <*Picornavirales*>	ssRNA	NA*	1	0	0	1	0	0	0
Iflaviridae	ssRNA	Animals (invertebrate)	10	7	3	23	2	7	0
Picornaviridae	ssRNA	Animals (vertebrate)	0	0	0	1	0	0	0

Table A2 (cont'd)

unclassified *Picornavirales*	ssRNA	Animals (vertebrate)	0	0	0	1	1	1	1
Potyviridae	ssRNA	Plants	0	2	1	2	3	1	1
Tombusviridae	ssRNA	Plants	3	0	0	3	7	1	0
Tymovirales	ssRNA	Plants	0	0	0	0	0	0	0
Betaflexiviridae	ssRNA	Plants	0	1	0	0	0	0	0
Tymoviridae	ssRNA	Plants	1	0	0	0	0	0	0
Umbravirus (genus)	ssRNA	Plants	0	0	0	0	2	2	0
unclassified ssRNA positive-strand viruses	Unclassified	NA*	2	0	0	4	1	0	0
Virgaviridae	ssRNA	Plants	0	0	0	0	1	0	2
unassigned viruses	Unclassified	NA*	0	0	0	0	0	0	0
Bacilladnavirus (genus)	Unclassified	Other (ameoba, protozoa, or fungi)	2	2	0	0	1	0	0
Bidnaviridae	Unclassified	Animals (invertebrate)	0	0	0	5	0	0	0
unclassified phages	Unclassified	Bacteria	5	5	8	8	4	1	5
Not assigned	NA*	NA*	388	252	179	239	148	96	116
No hits	NA*	NA*	17004	12563	10288	10611	7298	6962	6884

* NA. Not applicable, data was not used in category analysis

** Contaminant removed from analysis

Table A3. Raw data and the calculated MS2 eluent plaque concentration, lettuce concentration, and average lettuce concentration for samples collected during trial 1 leafy green processing without sanitizer									
Sample ID*	Lettuce weight (g)	Eluent volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Eluent plaque concentration (PFU/mL)**	Phage concentration on lettuce (PFU/g)**	Average phage concentration on lettuce for each sample location (PFU/g)
BP1	50.00	234.50	10^{-4}	108	101	92	$5.02x10^{5}$	$2.35x10^{6}$	$1.38x10^{6}$
BP2	50.00	240.00	10^{-4}	49	37	43	$2.15x10^{5}$	$1.03x10^{6}$	
BP3	50.00	233.00	10^{-4}	33	25	40	$1.63x10^{5}$	$7.61x10^{5}$	
SR1	52.10	235.00	10^{-4}	20	26	29	$1.25x10^{5}$	$5.64x10^{5}$	$5.97x10^{5}$
SR2	54.20	235.00	10^{-4}	38	37	20	$1.58x10^{5}$	$6.87x10^{5}$	
SR3	59.30	235.00	10^{-4}	38	34	10	$1.37x10^{5}$	$5.42x10^{5}$	
F1	58.00	235.00	10^{-3}	91	64	82	$3.95x10^{4}$	$1.60x10^{5}$	$1.64x10^{5}$
F2	51.20	240.00	10^{-3}	37	75	66	$2.97x10^{4}$	$1.39x10^{5}$	
F3	53.30	235.00	10^{-3}	85	114	63	$4.37x10^{4}$	$1.93x10^{5}$	
SA1	50.30	240.00	10^{-3}	13	14	55	$1.37x10^{4}$	$6.52x10^{4}$	$7.61x10^{4}$
SA2	57.20	235.00	10^{-3}	40	22	20	$1.37x10^{4}$	$5.61x10^{4}$	
SA3	53.20	237.00	10^{-3}	45	51	48	$2.40x10^{4}$	$1.07x10^{5}$	
C1	51.60	240.00	10^{-3}	NR***	NR***	NR***	NR***	NR***	$1.11x10^{5}$
C2	51.20	240.00	10^{-3}	X	X	53	$2.65\ x10^{4}$	$1.24x10^{5}$	
C3	55.40	237.00	10^{-3}	40	42	55	$2.28x10^{4}$	$9.77x10^{4}$	
CW	NR***	500.00	10^{-4}	64	52	55	$2.85x10^{5}$	NA****	NA****
IC	NR***	NR***	10^{-9}	65	48	53	$2.77\ x10^{10}$	NA****	NA****

*Sample ID descriptions: lettuce sampled before processing (BP) and after shredding (SR), flume washing (F), shaking (SA), and centrifugal drying (C). Centrifuge water (CW) and the Inoculated concentrate (IC) are also included. 1, 2 and 3 are triplicate samples

**Calculated using equations in Part III, Section 3.1.3, pages 39-40

*** NR no result, data were unavailable due no plaques in countable range (dilution error)

****NA not applicable

Table A4. Raw Data and the calculated MS2 eluent plaque concentration, lettuce concentration, and average lettuce concentration for samples collected during trial 2 leafy green processing without sanitizer *

Sample ID*	Lettuce Weight (g)	Eluent Volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Eluent plaque concentration (PFU/mL)**	Phage concentration on lettuce (PFU/g)**	Average Phage concentration on Lettuce (PFU/g)
BP1	49.9	241.0	10^{-4}	30	38	29	$1.62x10^{5}$	$7.81x10^{5}$	$7.81x10^{5}$
SR1	50.8	238.0	10^{-4}	20	40	23	$1.38x10^{5}$	$6.48x10^{5}$	$6.81x10^{5}$
SR2	50.7	237.0	10^{-4}	21	26	33	$1.33x10^{5}$	$6.23x10^{5}$	
SR3	50.5	236.0	10^{-4}	34	23	42	$1.65x10^{5}$	$7.71x10^{5}$	
F1	50.2	250.0	10^{-3}	30	34	34	$1.63x10^{4}$	$8.13x10^{4}$	$7.07x10^{4}$
F2	50.2	249.0	10^{-3}	27	30	40	$1.62x10^{4}$	$8.02x10^{4}$	
F3	50.1	205.0	10^{-3}	22	23	29	$1.23x10^{4}$	$5.05x10^{4}$	
SA1	52.0	248.0	10^{-3}	13	34	24	$1.18x10^{4}$	$5.64x10^{4}$	$5.39x10^{4}$
			10^{-2}	167	205	191	$9.38x10^{3}$	$4.48x10^{4}$	
SA2	52.1	249.0	10^{-3}	25	27	20	$1.20x10^{4}$	$5.74x10^{4}$	
			10^{-2}	242	257	160	$1.10x10^{4}$	$5.25x10^{4}$	
SA3	54.6	250.0	10^{-3}	34	33	19	$1.43x10^{4}$	$6.56x10^{4}$	
			10^{-2}	203	182	229	$1.02x10^{4}$	$4.69x10^{4}$	
C1	53.7	240.0	10^{-3}	34	45	28	$1.40x10^{4}$	$6.26x10^{4}$	$9.48x10^{4}$
C2	53.4	241.0	10^{-3}	56	36	50	$2.50x10^{4}$	$1.13x10^{5}$	
C3	55.5	240.0	10^{-3}	47	36	53	$2.27x10^{4}$	$9.80x10^{4}$	
CW	NA***	500.00	10^{-4}	37	33	21	$1.52x10^{5}$	NA***	NA***
IC	NA***	NA***	10^{-9}	54	43	38	$2.25x10^{10}$	NA***	NA***

*Sample ID descriptions: lettuce sampled before processing (BP) and after shredding (SR), flume washing (F), shaking (SA), and centrifugal drying (C). Centrifuge water (CW) and the Inoculated concentrate (IC) are also included. 1, 2 and 3 are triplicate samples.
** Calculated using equations in Part III, Section 3.1.3, pages 39-40
***NA not applicable

Table A5. Raw data and the calculated MS2 eluent plaque concentration, lettuce concentration, and average lettuce concentration for samples collected during trial 3 leafy green processing without sanitizer									
Sample ID*	Lettuce Weight (g)	Eluent Volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Eluent plaque concentration (PFU/mL)**	Phage concentration on lettuce (PFU/g)**	Average Phage concentration on Lettuce (PFU/g)
BP1	49.9	238.0	10^{-4}	39	25	48	$1.87x10^{5}$	$8.90x10^{5}$	$7.12x10^{5}$
BP2	54.6	236.0	10^{-4}	21	30	20	$1.18x10^{5}$	$5.11x10^{5}$	
BP3	50.3	236.0	10^{-4}	30	32	32	$1.57x10^{5}$	$7.35x10^{5}$	
SR1	51.1	237.0	10^{-4}	22	14	28	$1.07x10^{5}$	$4.94x10^{5}$	$5.20x10^{5}$
			10^{-3}	219	275	224	$1.20x10^{5}$	$5.55x10^{5}$	
SR2	51.6	235.0	10^{-4}	21	28	25	$1.23x10^{5}$	$5.62x10^{5}$	
			10^{-3}	224	211	230	$1.11x10^{5}$	$5.05x10^{5}$	
SR3	52.0	235.0	10^{-4}	24	24	30	$1.30x10^{5}$	$5.88x10^{5}$	
			10^{-3}	205	192	154	$9.18x10^{4}$	$4.15x10^{5}$	
F1	53.4	241.0	10^{-3}	29	26	26	$1.35x10^{4}$	$6.10x10^{4}$	$5.80x10^{4}$
F2	53.0	242.0	10^{-3}	28	21	24	$1.22x10^{4}$	$5.55x10^{4}$	
F3	53.2	245.0	10^{-3}	31	23	21	$1.25x10^{4}$	$5.76x10^{4}$	
SA1	50.3	240.0	10^{-2}	305	311	314	$1.55x10^{4}$	$7.39x10^{4}$	$7.38x10^{4}$
SA2	51.4	220.0	10^{-2}	296	344	265	$1.51x10^{4}$	$6.46x10^{4}$	
SA3	51.1	240.0	10^{-2}	386	318	354	$1.76x10^{4}$	$8.29x10^{4}$	
C1	52.2	235.0	10^{-2}	381	400	329	$1.85x10^{4}$	$8.33x10^{4}$	$8.66x10^{4}$
C2	50.2	235.0	10^{-2}	461	406	371	$2.06x10^{4}$	$9.67x10^{4}$	
C3	51.0	235.0	10^{-2}	371	345	324	$1.73x10^{4}$	$7.99x10^{4}$	
CW	NA***	500.00	10^{-4}	22	21	18	$1.02x10^{5}$	NA***	NA***
			10^{-3}	186	208	200	$9.90x10^{4}$	NA***	NA***
IC	NA***	NA***	10^{-9}	30	23	26	$1.32\ x10^{10}$	NA***	NA***

*Sample ID descriptions: lettuce sampled before processing (BP) and after shredding (SR), flume washing (F), shaking (SA), and centrifugal drying (C). Centrifuge water (CW) and the Inoculated concentrate (IC) are also included. 1, 2 and 3 are triplicate samples.

** Calculated using equations in Part III, Section 3.1.3, pages 39-40

***NA not applicable

Table A6. Raw data and the calculated MS2 eluent plaque concentration, lettuce concentration, and average lettuce concentration for samples collected during trial 1 leafy green processing with 25 ppm free chlorine*									
Sample ID*	Lettuce Weight (g)	Eluent Volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Eluent plaque concentration (PFU/mL)**	Phage concentration on lettuce (PFU/g)**	Average Phage concentration on Lettuce (PFU/g)
Control	51.4	200.0	10^{0}	2	2	5	$1.50x10^{0}$	$5.84x10^{0}$	$3.51x10^{0}$
Control	49.8	235.0	10^{0}	1	0	0	$5.00x10^{-1}$	$2.36x10^{0}$	
Control	50.6	235.0	10^{0}	1	0	1	$5.00x10^{-1}$	$2.32x10^{0}$	
BP1	50.2	233.0	10^{-3}	37	44	33	$1.90x10^{4}$	$8.82x10^{4}$	$1.20x10^{5}$
BP2	50.7	226.0	10^{-3}	46	58	53	$2.62x10^{4}$	$1.17x10^{5}$	
BP3	49.5	236.0	10^{-3}	69	59	66	$3.23x10^{4}$	$1.54x10^{5}$	
SR1	49.9	238.0	10^{-3}	41	39	20	$1.67x10^{4}$	$7.95x10^{4}$	$7.19x10^{4}$
SR2	51.2	237.0	10^{-3}	30	26	33	$1.48x10^{4}$	$6.87x10^{4}$	
SR3	49.3	238.0	10^{-3}	24	27	33	$1.40x10^{4}$	$6.76x10^{4}$	
F1	49.1	249.0	10^{-2}	101	91	109	$5.02x10^{3}$	$2.54x10^{4}$	$2.56x10^{4}$
F2	49.9	250.0	10^{-2}	85	99	97	$4.68x10^{3}$	$2.35x10^{4}$	
F3	50.0	250.0	10^{-2}	134	102	100	$5.60x10^{3}$	$2.80x10^{4}$	
SA1	50.9	250.0	10^{-2}	80	81	106	$4.45x10^{3}$	$2.19x10^{4}$	$2.26x10^{4}$
SA2	49.6	246.0	10^{-2}	75	70	87	$3.87x10^{3}$	$1.92x10^{4}$	
SA3	49.8	247.0	10^{-2}	109	114	102	$5.42x10^{3}$	$2.69x10^{4}$	
C1	50.0	245.0	10^{-2}	101	119	129	$5.82x10^{3}$	$2.85x10^{4}$	$2.81x10^{4}$
C2	50.0	245.0	10^{-2}	119	113	120	$5.87x10^{3}$	$2.87x10^{4}$	
C3	50.4	243.0	10^{-2}	106	107	124	$5.62x10^{3}$	$2.71x10^{4}$	
CW	NA***	500	10^{-3}	33	32	38	$1.72x10^{4}$	NA***	NA***
IC	NA***	NA***	10^{-8}	63	64	59	$3.10x10^{9}$	NA***	NA***

*Sample ID descriptions: lettuce sampled before processing (BP) and after shredding (SR), flume washing (F), shaking (SA), and centrifugal drying (C). Centrifuge water (CW) and the Inoculated concentrate (IC) are also included. 1, 2 and 3 are triplicate samples.
** Calculated using equations in Part III, Section 3.1.3, pages 39-40
***NA not applicable

Table A7. Raw data and the calculated MS2 eluent plaque concentration, lettuce concentration, and average lettuce concentration for samples collected during trial 2 leafy green processing with 25 ppm free chlorine									
Sample ID*	Lettuce Weight (g)	Eluent Volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Eluent plaque concentration (PFU/mL)**	Phage concentration on lettuce (PFU/g)**	Average Phage concentration on Lettuce (PFU/g)
BP1	49.8	235.0	10^{-3}	27	34	42	$1.72x10^{4}$	$8.10x10^{4}$	$7.70x10^{4}$
BP2	50.0	233.5	10^{-3}	33	38	47	$1.97x10^{4}$	$9.18x10^{4}$	
BP3	50.7	226.5	10^{-3}	23	29	26	$1.30x10^{4}$	$5.81x10^{4}$	
SR1	50.1	236.0	10^{-3}	45	40	34	$1.98x10^{4}$	$9.34x10^{4}$	$8.87x10^{4}$
SR2	50.0	235.0	10^{-3}	29	28	58	$1.92x10^{4}$	$9.01x10^{4}$	
SR3	49.8	235.0	10^{-3}	34	28	43	$1.75x10^{4}$	$8.26x10^{4}$	
F1	50.0	244.0	10^{-2}	35	45	59	$2.32x10^{3}$	$1.13x10^{4}$	$8.49x10^{3}$
F2	50.9	238.0	10^{-2}	30	24	35	$1.48\ x10^{3}$	$6.94x10^{3}$	
F3	49.8	245.0	10^{-2}	23	28	37	$1.47x10^{3}$	$7.22x10^{3}$	
SA1	50.1	246.0	10^{-2}	31	45	42	$1.97x10^{3}$	$9.66x10^{3}$	$8.36x10^{3}$
SA2	49.9	246.0	10^{-2}	32	30	29	$1.52x10^{3}$	$7.48x10^{3}$	
SA3	50.1	244.0	10^{-2}	24	40	34	$1.63x10^{3}$	$7.95x10^{3}$	
C1	50.0	226.0	10^{-2}	36	30	34	$1.67x10^{3}$	$7.53x10^{3}$	$8.96x10^{3}$
C2	50.1	237.0	10^{-2}	36	51	47	$2.23x10^{3}$	$1.06x10^{4}$	
C3	50.5	211.0	10^{-2}	53	39	34	$2.10x10^{3}$	$8.77x10^{3}$	
CW	NA***	500.00	10^{-3}	9	17	19	$9.00x10^{3}$	NA***	NA***
			10^{-2}	171	130	187	$8.13x10^{3}$	NA***	NA***
IC	NA***	NA***	10^{-8}	54	82	49	$3.08x10^{9}$	NA***	NA***

*Sample ID descriptions: lettuce sampled before processing (BP) and after shredding (SR), flume washing (F), shaking (SA), and centrifugal drying (C). Centrifuge water (CW) and the Inoculated concentrate (IC) are also included. 1, 2 and 3 are triplicate samples.

** Calculated using equations in Part III, Section 3.1.3, pages 39-40

***NA not applicable

Table A8. Raw data and the calculated MS2 eluent plaque concentration, lettuce concentration, and average lettuce concentration for samples collected during trial 3 leafy green processing with 25 ppm free chlorine									
Sample ID*	Lettuce Weight (g)	Eluent Volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Eluent plaque concentration (PFU/mL)**	Phage concentration on lettuce (PFU/g)**	Average Phage concentration on Lettuce (PFU/g)
BP1	50.3	235.0	10^{-3}	42	44	42	$2.13x10^{4}$	$9.96x10^{4}$	$6.64x10^{4}$
BP2	50.6	235.0	10^{-3}	8	32	28	$1.50x10^{4}$	$6.97x10^{4}$	
			10^{-2}	228	224	275	$1.21x10^{4}$	$5.63x10^{4}$	
BP3	49.9	235.0	10^{-3}	20	20	28	$1.13x10^{4}$	$5.33x10^{4}$	
			10^{-2}	205	210	261	$1.13x10^{4}$	$5.30x10^{4}$	
SR1	50.0	236.0	10^{-3}	30	39	41	$1.83x10^{4}$	$8.65x10^{4}$	$8.77x10^{4}$
SR2	50.1	239.0	10^{-3}	47	34	25	$1.77x10^{4}$	$8.43x10^{4}$	
SR3	50.1	235.0	10^{-3}	39	34	45	$1.97x10^{4}$	$9.23x10^{4}$	
F1	50.1	250.0	10^{-2}	53	44	43	$2.33x10^{3}$	$1.16x10^{4}$	$1.48x10^{4}$
F2	50.4	250.0	10^{-2}	55	58	56	$2.82x10^{3}$	$1.40x10^{4}$	
F3	50.6	250.0	10^{-2}	60	58	111	$3.82x10^{3}$	$1.89x10^{4}$	
SA1	50.0	250.0	10^{-2}	50	66	67	$3.05x10^{3}$	$1.52x10^{4}$	$1.13x10^{4}$
SA2	50.2	250.0	10^{-2}	31	39	36	$1.77x10^{3}$	$8.79x10^{4}$	
SA3	50.3	250.0	10^{-2}	34	43	43	$2.00x10^{3}$	$9.94x10^{3}$	
C1	50.0	243.0	10^{-2}	36	53	38	$2.12x10^{3}$	$1.03x10^{4}$	$1.01x10^{4}$
C2	50.0	243.0	10^{-2}	28	28	51	$1.78x10^{3}$	$8.66x10^{3}$	
C3	50.0	243.0	10^{-2}	35	59	46	$2.33x10^{3}$	$1.13x10^{4}$	
CW	NA***	720.00	10^{-2}	94	109	92	$4.92x10^{3}$	NA***	NA***
			10^{-3}	14	20	12	$7.67x10^{3}$	NA***	NA***
IC	NA***	NA***	10^{-8}	94	109	92	$4.92x10^{9}$	NA***	NA***

*Sample ID descriptions: lettuce sampled before processing (BP) and after shredding (SR), flume washing (F), shaking (SA), and centrifugal drying (C). Centrifuge water (CW) and the Inoculated concentrate (IC) are also included. 1, 2 and 3 are triplicate samples.

** Calculated using equations in Part III, Section 3.1.3, pages 39-40

***NA not applicable

Table A9. Flume water raw data and MS2 concentration for sanitizer experiments													
Trial	Sample ID	Concentrate volume (mL)	Dilution	Plate 1	Plate 2	Plate 3	Plate 4	Plate 5	Plate 6	Plate 7	Plate 8	PFU/mL	Average PFU/mL
1	FW1	350	0	2	2	6	6	8	8	6	5	2.7	2.3
	FW2	350	0	5	3	4	4	1	7	4	2	1.9	
2	FW1	330	0	11	12	8	9	6	NR	NR	NR	4.6	3.8
	FW2	250	0	4	6	7	4	8	NR	NR	NR	2.9	
3	FW1	250	0	5	2	6	5	0	NR	NR	NR	2.3	3.4
	FW2	230	0	9	5	10	10	12	NR	NR	NR	4.6	

*NR no result

REFERENCES

REFERENCES

Ackermann, H.-W. (2009). Phage classification and characterization. *Methods in Molecular Biology (Clifton, N.J.), 501*, 127–40.

Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., & Boonham, N. (2009). Next-generation sequencing and metagenomic analysis: A universal diagnostic tool in plant virology. *Molecular Plant Pathology*, *10*(4), 537–545.

Allwood, P. B., Malik, Y. S., Hedberg, C. W., & Goyal, S. M. (2004). Effect of temperature and sanitizers on the survival of feline calicivirus, *Escherichia coli*, and F-specific coliphage MS2 on leafy salad vegetables. *Journal of Food Protection*, *67*(7), 1451–6.

Alum, A., Enriquez, C., & Gerba, C. P. (2011). Impact of drip irrigation method, soil, and virus type on tomato and cucumber contamination. *Food and Environmental Virology*, *3*(2), 78–85.

Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., Chan, A. M., Haynes, M., Kelley, S., Liu, H., Mahaffy, J. M., Mueller, J. E., Nulton, J., Olson, R., Parsons, R., Rayhawk, S., Suttle, C. A., & Rohwer, F. (2006). The marine viromes of four oceanic regions. *PLoS Biology*, *4*(11), e368.

Asensio, L., González, I., García, T., & Martín, R. (2008). Determination of food authenticity by enzyme-linked immunosorbent assay (ELISA). *Food Control*, *19*(1), 1–8.

Aw, T. G., & Rose, J. B. (2012). Detection of pathogens in water: from phylochips to qPCR to pyrosequencing. *Current Opinion in Biotechnology*, *23*(3), 422–30.

Aw, T. G., Howe, A., & Rose, J. B. (2014). Metagenomic approaches for direct and cell culture evaluation of the virological quality of wastewater. *Journal of Virological Methods*, *210*, 15–21.

Babraham Bioinformatics. (2015). FastQC a quality control tool for high throughput sequence data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed on 3 August, 2015.

Badawy, A. S., Gerba, C. P., & Kelley, L. M. (1985). Survival of rotavirus SA-11 on vegetables. *Food Microbiology*, *2*(3), 199–205.

Baert, L., Debevere, J., & Uyttendaele, M. (2009). The efficacy of preservation methods to inactivate foodborne viruses. *International Journal of Food Microbiology*, *131*(2), 83–94.

Bartlett, J. M. S., & Stirling, D. (2003). A short history of the polymerase chain reaction. *Methods in molecular biology (Clifton, N.J.)* (3-6). Humana Press.

Baudoux, A.-C., & Brussaard, C. P. D. (2005). Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virology*, *341*(1), 80–90.

Bergholz, T. M., Moreno Switt, A. I., & Wiedmann, M. (2014). Omics approaches in food safety: fulfilling the promise? *Trends in Microbiology*, *22*(5), 275–281.

Beuchat, L. R. (2002). Ecological factors influencing survival and growth of human pathogens on raw fruits and vegetables. *Microbes and Infection*, *4*(4), 413–423.

Bibby, K., Viau, E., & Peccia, J. (2011). Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Letters in Applied Microbiology*, *52*(4), 386–92.

Bidawid, S., Farber, J. M., & Sattar, S. A. (2000). Rapid concentration and detection of hepatitis A virus from lettuce and strawberries. *Journal of Virological Methods*, *88*(2), 175–85.

Biopieces. (2015). Biopieces. Retrieved from https://code.google.com/p/biopieces/. Accessed on 3 August 2015.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford University Press)*, btu 170.

Boulila, M. (2011). Positive selection, molecular recombination structure and phylogenetic reconstruction of members of the family *Tombusviridae*: implication in virus taxonomy. *Genetics and Molecular Biology*, *34*(4), 647–60.

Breitbart, M., & Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology*, *13*(6), 278–84.

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J. M., Segall, A. M., Mead, D., Azam, F., & Rohwer, F. (2002). Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences of the United States of America*, *99* (22), 14250–14255.

Brussaard, C. (2004). Viral control of phytoplankton populations-a review. *The Journal of Eukaryotic Microbiology*, *51*(2), 125–138.

Buchholz, A. L., Davidson, G. R., Marks, B. P., Todd, E. C. D., & Ryser, E. T. (2012a). Quantitative transfer of *Escherichia coli* O157:H7 to equipment during small-scale production of fresh-cut leafy greens. *Journal of Food Protection*, *75*(7), 1184–1197.

Buchholz, A. L., Davidson, G. R., Marks, B. P., Todd, E. C. D., & Ryser, E. T. (2012b). Transfer of *Escherichia coli* O157:H7 from equipment surfaces to fresh-cut leafy greens during processing in a model pilot-plant production line with sanitizer-free water. *Journal of Food Protection*, *75*(11), 1920–9.

Butot, S., Putallaz, T., & Sánchez, G. (2007). Procedure for rapid concentration and detection of enteric viruses from berries and vegetables. *Applied and Environmental Microbiology*, *73*(1), 186–92.

Calder, L., Simmons, G., Thornley, C., Taylor, P., Pritchard, K., Greening, G., & Bishop, J. (2003). An outbreak of hepatitis A associated with consumption of raw blueberries. *Epidemiology and Infection*, *131*(1), 745-51.

Casteel, M. J., Schmidt, C. E., & Sobsey, M. D. (2008). Chlorine disinfection of produce to inactivate hepatitis A virus and coliphage MS2. *International Journal of Food Microbiology, 125* (3), 267–273.

Casteel, M. J., Schmidt, C. E., & Sobsey, M. D. (2009). Chlorine inactivation of coliphage MS2 on strawberries by industrial-scale water washing units. *Journal of Water and Health*, *7*(2), 244–250.

Center for Agricultural and Rural Development (CARD). (2004). The expanding U.S. market for fresh produce. Retrieved from http://www.agmrc.org/media/cms/ expandingusmarketproduce_07C66B1CA5F70.pdf. Accessed on June 26, 2014.

Centers for Disease Control and Prevention (CDC). (2011). Incidence and trends of infection with pathogens transmitted commonly through food—foodborne diseases active surveillance network, 10 U.S. sites, 1996-2012. *Morbidity and Mortality Weekly Report, 62*(15), 283-287.

Centers for Disease Control and Prevention (CDC). (2013a). Global foodborne infections network, formerly known as WHO global salm-surv (GFN). Retrieved from http://www.cdc.gov/ncezid/dfwed/international/gfin.html. Accessed on June 22 2015.

Centers for Disease Control and Prevention (CDC). (2013b). Reporting and surveillance for norovirus. Retrieved from http://www.cdc.gov/norovirus/reporting/index.html. Accessed on 26 March 2015.

Centers for Disease Control and Prevention (CDC). (2014). 2011- Estimates of foodborne illness in the United States- CDC 2011 estimates. Retrieved from http://www.cdc.gov/foodborneburden/2011-foodborne-estimates.html . Accessed on 23 June 2014.

Centers for Disease Control and Prevention (CDC). (2015a). Guide to confirming a diagnosis in foodborne disease. Retrieved from http://www.cdc.gov/foodsafety/outbreaks/investigatingoutbreaks/confirming_diagnosis.html. Accessed on 26 June 2015.

Centers for Disease Control and Prevention (CDC). (2015b). Surveillance for foodborne disease outbreaks, United States, 2013, annual report. Atlanta, Georgia: US Department of Health and Human Services.

Center for Food Safety and Applied Nutrition (CFSAN). (2001). Juice - guidance for industry: the juice HACCP regulation - questions & answers. Retrieved from http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/Juice/ucm072602.htm. Accessed on 26 June 2015.

Center for Food Safety and Applied Nutrition (CFSAN). (2006). Produce & plant products - commodity specific food safety guidelines for the lettuce and leafy greens supply chain - 1st edition. Retrieved from http://www.fda.gov/downloads/Food/GuidanceRegulation/UCM169008.pdf. Accessed on 26 June 2015.

Center for Food Safety and Applied Nutrition (CFSAN). (2009). Produce & plant products - draft guidance for industry: guide to minimize microbial food safety hazards of leafy greens. Retrieved from http://www.fda.gov/Food/GuidanceRegulation/GuidanceDocumentsRegulatoryInformation/ProducePlantProducts/ucm174200.htm. Accessed on 26 June 2015.

Center for Food Safety and Applied Nutrition (CFSAN). (2014). Safe practices for food processes - chapter V. Methods to reduce/eliminate pathogens from produce and fresh-cut produce. Retrieved from http://www.fda.gov/Food/FoodScienceResearch/SafePracticesforFoodProcesses/ucm091363.htm. Accessed on 26 June 2015.

Center for Food Safety and Applied Nutrition (CFSAN). (2015). Safe practices for food processes - chapter IV. Outbreaks associated with fresh and fresh-cut produce. Incidence, growth, and survival of pathogens in fresh and fresh-cut produce. Retrieved from http://www.fda.gov/Food/FoodScienceResearch/SafePracticesforFoodProcesses/ucm091265.htm. Accessed on 26 June 2015.

Cheong, S., Lee, C., Song, S. W., Choi, W. C., Lee, C. H., & Kim, S.-J. (2009). Enteric viruses in raw vegetables and groundwater used for irrigation in South Korea. *Applied and Environmental Microbiology*, *75*(24), 7745–51.

Chi, J., VanLeeuwen, J. A., Weersink, A., & Keefe, G. P. (2002). Direct production losses and treatment costs from bovine viral diarrhoea virus, bovine leukosis virus, Mycobacterium avium subspecies paratuberculosis, and Neospora caninum. *Preventive Veterinary Medicine*, *55*(2), 137–153.

Choi, C., Song, I., Stine, S., Pimentel, J., & Gerba, C. (2004). Role of irrigation and wastewater reuse: comparison of subsurface irrigation and furrow irrigation. *Water Science and Technology : A Journal of the International Association on Water Pollution Research*, *50*(2), 61–8.

Clemens, R. (2004). The expanding U . S . market for fresh produce. *Iowa Ag Review*, *10*(1), 1-4.

Colomer-Lluch, M., Jofre, J., & Muniesa, M. (2011). Antibiotic resistance genes in the bacteriophage DNA fraction of environmental samples. *PloS One*, *6*(3), e17549.

Cook, N. (2013). Viruses in food and water: risks, surveillance and control, 1st Ed., Woodhead Publishing Limited, Cambridge, UK.

Croci, L., De Medici, D., Scalfaro, C., Fiore, A., & Toti, L. (2002). The survival of hepatitis A virus in fresh produce. *International Journal of Food Microbiology*, *73*(1), 29–34.

Croci, L., Dubois, E., Cook, N., Medici, D., Schultz, A. C., China, B., Rutjes, S. A., Hoorfar, J., & Van der Poel, W. H. M. (2008). Current methods for extraction and concentration of enteric viruses from fresh fruit and vegetables: towards international standards. *Food Analytical Methods*, *1*(2), 73–84.

Crutchfield, S. R., & Roberts, T. (2000). Food safety efforts accelerate in the 1990's. *Food Review*, *23*(202), 44–49.

Crusoe, M., Edvenson, G., Fish, J., Howe, A., McDonald, E., Nahum, J., Nanlohy, K., Ortiz-Zuazaga, H., Pell, J., Simpson, J., Scott, C., Srinivasan, R. R., Zhang, Q., & Brown, C. T. (2014). The khmer software package: enabling efficient sequence analysis. Retrieved from http://figshare.com/articles/The_khmer_software_package_enabling _efficient_sequence_analysis/979190. Accessed on 3 August 2015.

Cutadapt. (2015). Cutadapt- A tool that removes adapter sequences from DNA sequencing reads. Retrieved from https://code.google.com/p/cutadapt/. Accessed 3August, 2015.

d'Herelle, F. (1917). An invisible microbe that is antagonistic to the dysentery bacillus. *Comptes Rendus l'Académie des Sciences Paris*.

Davidson, G. R., Buchholz, A. L., & Ryser, E. T. (2013). Efficacy of commercial produce sanitizers against nontoxigenic *Escherichia coli* O157:H7 during processing of iceberg lettuce in a pilot-scale leafy green processing line. *Journal of Food Protection*, *76*(11), 1838–1845.

Davis, B. M., Moyer, K. E., Boyd, E. F., & Waldor, M. K. (2000). CTX prophages in classical biotype Vibrio cholerae: functional phage genes but dysfunctional phage genomes. *Journal of Bacteriology*, *182*(24), 6992–8.

Dawson, D. J., Paish, A., Staffell, L. M., Seymour, I. J., & Appleton, H. (2005). Survival of viruses on fresh produce, using MS2 as a surrogate for norovirus. *Journal of Applied Microbiology*, *98*(1), 203–9.

Deboosere, N., Pinon, A., Caudrelier, Y., Loutreul, J., Merle, G., Delobel, A., Merle, G., Perelle, S., Temmam, S., Loutreul, J., Mornin, T., Estienney, M., Belliot, G., Pothier, P., Gantzer, C., Vialette, M. (2012). Adhesion of human pathogenic enteric viruses and surrogate viruses to inert and vegetal food surfaces. *Food Microbiology*, *32* (1), 48-56.

Dicaprio, E., Ma, Y., Purgianto, A., Hughes, J., & Li, J. (2012). Internalization and dissemination of human norovirus and animal caliciviruses in hydroponically grown romaine lettuce. *Applied and Environmental Microbiology*, *78*(17), 6143–52.

Doultree, J. C., Druce, J. D., Birch, C. J., Bowden, D. S., & Marshall, J. A. (1999). Inactivation of feline calicivirus, a Norwalk virus surrogate. *Journal of Hospital Infection*, *41*(1), 51–57.

Dubois, E., Hennechart, C., Deboosère, N., Merle, G., Legeay, O., Burger, C., Le Calvé, M., Lombard, B., Ferré, V., & Traoré, O. (2006). Intra-laboratory validation of a concentration method adapted for the enumeration of infectious F-specific RNA coliphage, enterovirus, and hepatitis A virus from inoculated leaves of salad vegetables. *International Journal of Food Microbiology*, *108*(2), 164–71.

Dubois, E., Hennechart, C., Merle, G., Burger, C., Hmila, N., Ruelle, S., Perelle, S., & Ferré, V. (2007). Detection and quantification by real-time RT-PCR of hepatitis A virus from inoculated tap waters, salad vegetables, and soft fruits: characterization of the method performances. *International Journal of Food Microbiology*, *117*(2), 141–9.

Ercolini, D. (2013). High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology. *Applied and Environmental Microbiology*, *79*(10), 3148–55.

Food and Agriculture Organization of the United Nations/World Health Organization (FAO/WHO). (2008). Viruses in food: scientific advice to support risk management activities, meeting report. Microbiological Risk Assessment Series No. 14. Rome. 151pp.

Fletcher, M., Levy, M. E., & Griffin, D. D. (2000). Foodborne outbreak of group A rotavirus gastroenteritis among college students - District of Columbia, March-April 2000. *Morbidity and Mortality Weekly Report*, *49*(50), 1131–1133.

Fong, T.-T., & Lipp, E. K. (2005). Enteric viruses of humans and animals in aquatic environments: health risks, detection, and potential water quality assessment tools. *Microbiology and Molecular Biology Reviews*, *69*(2), 357–71.

Fraisse, A., Temmam, S., Deboosere, N., Guillier, L., Delobel, A., Maris, P., Vialette, M., Morin, T., & Perelle, S. (2011). Comparison of chlorine and peroxyacetic-based disinfectant to inactivate Feline calicivirus, Murine norovirus and Hepatitis A virus on lettuce. *International Journal of Food Microbiology*, *151*(1), 98–104.

Gallagher, S. R., & Wiley, E. A. (2008). Current protocols essential laboratory techniques. *Hoboken, NJ, USA: John Wiley & Sons, Inc.*

Ganesh, B., Bányai, K., Martella, V., Jakab, F., Masachessi, G., & Kobayashi, N. (2012). Picobirnavirus infections: viral persistence and zoonotic potential. *Reviews in Medical Virology*, *22*(4), 245–56.

Geering, A. (2007). *Caulimoviridae* (Plant Pararetroviruses). *eLS*.

Gelting, R. J., & Baloch, M. A. (2012). The food-water nexus: irrigation water quality, risks to food safety, and the need for a systems-based preventive approach. *Journal of Environmental Health*, *75*(3), 40–1.

Gibson, K. E., & Schwab, K. J. (2011). Tangential-flow ultrafiltration with integrated inhibition detection for recovery of surrogates and human pathogens from large-volume source water and finished drinking water. *Applied and Environmental Microbiology*, *77*(1), 385–91.

Gil, M. I., Selma, M. V, López-Gálvez, F., & Allende, A. (2009). Fresh-cut product sanitation and wash water disinfection: problems and solutions. *International Journal of Food Microbiology*, *134*(1-2), 37–45.

Goldsmith, C. S., & Miller, S. E. (2009). Modern uses of electron microscopy for detection of viruses. *Clinical Microbiology Reviews*, *22*(4), 552–63.

Gould, L. H., Walsh, K. A, Vieira, A. R., Herman, K., Williams, I. T., Hall, A. J., & Cole, D. (2013). Surveillance for foodborne disease outbreaks - United States, 1998-2008. *Morbidity and Mortality Weekly Report. Surveillance Summaries (Washington, D.C. : 2002)*, *62*(2), 1–34.

Greening, G. E., Hewitt, J., & Lewis, G. D. (2002). Evaluation of integrated cell culture-PCR (C-PCR) for virological analysis of environmental samples. *Journal of Applied Microbiology*, *93*(5), 745–750.

Gulati, B. R., Allwood, P. B., Hedberg, C. W., & Goyal, S. M. (2001). Efficacy of commonly used disinfectants for the inactivation of calicivirus on strawberry, lettuce, and a food-contact surface. *Journal of Food Protection*, *64*(9), 1430–4.

Handelsman, J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Molecular Biology Reviews*, *68*(4), 669–85.

Harwood, V. J., Levine, A. D., Scott, T. M., Chivukula, V., Lukasik, J., Farrah, S. R., & Rose, J. B. (2005). Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Applied and Environmental Microbiology*, *71*(6), 3163–70.

Havelaar, A. H., van Olphen, M., & Drost, Y. C. (1993). F-specific RNA bacteriophages are adequate model organisms for enteric viruses in fresh water. *Appl. Envir. Microbiol.*, *59*(9), 2956–2962.

Heaton, J. C., & Jones, K. (2008). Microbial contamination of fruit and vegetables and the behaviour of enteropathogens in the phyllosphere: a review. *Journal of Applied Microbiology*, *104*(3), 613–26.

Hennechart, C., Merle, G., & Dubois, E. (2002). Modified concentration method for the detection of enteric viruses on fruits and vegetables by reverse transcriptase- polymerase chain reaction or cell culture, *65*(12), 1962–1969.

Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Research*, *17*(3), 377–86.

Hutin, Y., Pool, V., Cramer, E., Nainan, O. V., Weth, J., Williams, I. T., Goldstein, S. T., Gensheimer, K. F., Bell, B. P., Shapiro, C. N., Alter, M. J., & Margolis, H. (1999). A multistate, foodborne outbreak of hepatitis A. *The New England Journal of Medicine*, *340*(8), 595–602.

Hyeon, J.-Y., Chon, J.-W., Park, C., Lee, J.-B., Choi, I.-S., Kim, M.-S., & Seo, K.-H. (2011). Rapid detection method for hepatitis A virus from lettuce by a combination of filtration and integrated cell culture-real-time reverse transcription PCR. *Journal of Food Protection*, *74*(10), 1756–61.

Illumina, Inc. (2015). Next-generation sequencing (NGS). Retrieved from http://www.illumina.com/technology/next-generation-sequencing.html. Accessed on 26 March 2012

International Food Information Foundation Council (IFIC Foundation). (2014). More Americans choosing foods, beverages based on healthfulness. Retrieved from http://www.foodinsight.org/articles/more-americans-choosing-foods-beverages-based-healthfulness. Accessed on 26 March 2015.

Isakbaeva, E. T., Widdowson, M.-A., Beard, R. S., Bulens, S. N., Mullins, J., Monroe, S. S., Bresee, J., Sassano, P., Cramer, E. H., & Glass, R. I. (2005). Norovirus transmission on cruise ship. *Emerging Infectious Diseases*, *11*(1), 154–8.

Johnson, P. T. (1984). Viral diseases of marine invertebrates. *Helgoländer Meeresuntersuchungen*, *37*(1-4), 65–98.

Kim, M.-S., Park, E.-J., Roh, S. W., & Bae, J.-W. (2011). Diversity and abundance of single-stranded DNA viruses in human feces. *Applied and Environmental Microbiology*, *77*(22), 8062–8070.

Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing - concepts and limitations. *BioEssays*, *32*(6), 524–536.

Koopmans, M., & Duizer, E. (2004). Foodborne viruses: an emerging problem. *International Journal of Food Microbiology*, *90*(1), 23–41.

Kotula, K. L., Kotula, A. W., Rose, B. E., Pierson, C. J., & Camp, M. (1997). Reduction of aqueous chlorine by organic material. *Journal of Food Protection*, *60*(3), 276–282.

Kroneman, A., Verhoef, L., Harris, J., Vennema, H., Duizer, E., van Duynhoven, Y., Gray, J., Iturriza, M., Böttiger, B., Falkenhorst, G., Johnsen, C., von Bonsdorff, C. –H., Maunula, L., Kuusi, M., Pothier, P., Gallay, A., Schreier, E., Höhne, M., Koch, J., Szücs, G., Reuter, G., Krisztalovics, K., Lynch, M., McKeown, P., Foley, B., Coughlan, S., Ruggeri, F. M., Di Bartolo, I., Vainio, K., Isakbaeva, E., Poljsak-Prijatelj, M., Hocevar Grom, A., Zimsek Mijovski, J., Bosch, A., Buesa, J., Sanchez Fauquier, A., Hernandéz-Pezzi, G., Hedlund, K.-O., & Koopmans, M. (2008). Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the foodborne viruses in Europe network from 1 July 2001 to 30 June 2006. *Journal of Clinical Microbiology*, *46*(9), 2959–65.

Kropinski, A. M., Mazzocco, A., Waddell, T. E., Lingohr, E., & Johnson, R. P. (2009). Enumeration of bacteriophages by double agar overlay plaque assay. *Bacteriophages, Methods in Molecular Biology (Clifton, N.J.)*, (pp. 69–76). Humana Press.

Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., & Hugenholtz, P. (2008). A bioinformatician's guide to metagenomics. *Microbiology and Molecular Biology Reviews*, *72*(4), 557–78.

Kutter, E., & Sulakvelidze, A. (2004). Bacteriophages: biology and applications. *CRC Press*.

Lacey, L., Frutos, R., Kaya, H., & Vail, P. (2001). Insect pathogens as biological control agents: do they have a future? *Biological Control*, *21*(3), 230–248.

Leclerc, H., Edberg, S., Pierzo, V., & Delattre, J. M. (2000). Bacteriophages as indicators of enteric viruses and public health risk in groundwaters. *Journal of Applied Microbiology*, *88*(1), 5–21.

Leclerc, H., Edberg, S., Pierzo, V., & Delattre, J. M. (2001). Bacteriophages as indicators of enteric viruses and public health risk in groundwaters. *Journal of Applied Microbiology*, *88*(1), 5–21.

Lee, H. K., & Jeong, Y. S. (2004). Comparison of total culturable virus assay and multiplex integrated cell culture-PCR for reliability of waterborne virus detection. *Applied and Environmental Microbiology*, *70*(6), 3632–6.

Leff, J. W., & Fierer, N. (2013). Bacterial communities associated with the surfaces of fresh fruits and vegetables. *PLoS ONE*, *8*(3), 1–9.

Leggitt, P. R., & Jaykus, L.-A. (2000). Detection methods for human enteric viruses in representative foods. *Journal of Food Protection*, *63*(12), 1738–44.

Leland, D. S., & Ginocchio, C. C. (2007). Role of cell culture for virus detection in the age of technology. *Clinical Microbiology Reviews*, *20*(1), 49–78.

Lindgreen, S. (2012). AdapterRemoval: easy cleaning of next-generation sequencing reads. *BMC Research Notes*, *5*(1), 337.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., & Law, M. (2012). Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology*, 2012.

Liu, P., Hill, V. R., Hahn, D., Johnson, T. B., Pan, Y., Jothikumar, N., & Moe, C. L. (2012). Hollow-fiber ultrafiltration for simultaneous recovery of viruses, bacteria and parasites from reclaimed water. *Journal of Microbiological Methods*, *88*(1), 155–61.

Liu, S., Vijayendran, D., & Bonning, B. C. (2011). Next generation sequencing technologies for insect virus discovery. *Viruses*, *3*(10), 1849–1869.

Lopez-Velasco, G., Welbaum, G. E., Boyer, R. R., Mane, S. P., & Ponder, M. A. (2011). Changes in spinach phylloepiphytic bacteria communities following minimal processing and refrigerated storage described using pyrosequencing of 16S rRNA amplicons. *Journal of Applied Microbiology*, *110* (5), 1203–1214.

Love, D. C., Casteel, M. J., Meschke, J. S., & Sobsey, M. D. (2008). Methods for recovery of hepatitis A virus (HAV) and other viruses from processed foods and detection of HAV by nested RT-PCR and TaqMan RT-PCR. *International Journal of Food Microbiology*, *126*(1-2), 221–6.

Luscombe, N. M., Greenbaum, D., & Gerstein, M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods of Information in Medicine*, *40*(4), 346–58.

Lynch, M. F., Tauxe, R. V, & Hedberg, C. W. (2009). The growing burden of foodborne outbreaks due to contaminated fresh produce: risks and opportunities. *Epidemiology and Infection*, *137*(3), 307–15.

Mehle, N., & Ravnikar, M. (2012). Plant viruses in aqueous environment - Survival, water mediated transmission and detection. *Water Research*, *46* (16), 4902–4917.

Meng, X.-J. (2011). From barnyard to food table: the omnipresence of hepatitis E virus and risk for zoonotic infection and food safety. *Virus Research*, *161*(1), 23–30.

Mokili, J. L., Rohwer, F., & Dutilh, B. E. (2012). Metagenomics and future perspectives in virus discovery. *Current Opinion in Virology*, *2*(1), 63–77.

Moore, N. E., Wang, J., Hewitt, J., Croucher, D., Williamson, D. A., Paine, S., Yen, S., Greening, G. E., & Hall, J. (2015). Metagenomic analysis of viruses in feces from unsolved outbreaks of gastroenteritis in humans. *Journal of Clinical Microbiology*, *53*(1), 15–21.

National Food Processors Association (NFPA). (2001). International Fresh-cut Produce Association, United Fresh Fruit and Vegetable Association. Field Cored Lettuce Best Practices, NFPA, Washington, DC.

Newell, D. G., Koopmans, M., Verhoef, L., Duizer, E., Aidara-Kane, A., Sprong, H., Opsteeh, M., Langelaar, M., Threfall, J., Scheutz, F., van der Giessen, J., & Kruse, H. (2010). Food-borne diseases - the challenges of 20 years ago still persist while new ones continue to emerge. *International Journal of Food Microbiology*, *139*, 3–15.

O'Brien, S. J. (2008). The challenge of estimating the burden of an underreported disease. *Food-Borne Viruses: Progress and Challenges* (pp. 87–115). American Soceity of Microbiology, 1752 N Street NW, Washington, DC USA.

Oliver, S. P., Patel, D. A, Callaway, T. R., & Torrence, M. E. (2009). ASAS centennial paper: developments and future outlook for preharvest food safety. *Journal of Animal Science*, *87*(1), 419–37.

Ottesen, A. R., González Peña, A., White, J. R., Pettengill, J. B., Li, C., Allard, S., Rideout, S., Allard, M., Hill, T., Evans, P., Strain, E., Musser, S., Knight, R., & Brown, E. (2013). Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC Microbiology*, *13*(1), 114.

Painter, J. A., Hoekstra, R. M., Ayers, T., Tauxe, R. V., Braden, C. R., Angulo, F. J., & Griffin, P. M. (2013). Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998-2008. *Emerging Infectious Diseases*, *19*(3), 407–415.

Park, E.-J., Kim, K.-H., Abell, G. C. J., Kim, M.-S., Roh, S. W., & Bae, J.-W. (2011). Metagenomic analysis of the viral communities in fermented foods. *Applied and Environmental Microbiology*, *77*(4), 1284–91.

Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, *28* (11), 1420–1428.

Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J., Aquino, T., & Tsomondo, T. (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, *84*(1), 1–20.

Pollack, S. L. (2001). Consumer demand for fruit and vegetables: the U.S. example. *Changing Structure of Global Food and Trade WRS* (49–54).

Reynolds, K. A. (2004). Integrated cell culture/PCR for detection of enteric viruses in environmental samples. *Methods in Molecular Biology (Clifton, N.J.)*. (pp. 69–78). Humana Press.

Roche Diagnostics Corporation. (2015). GS FLX+ System. Retrieved from http://454.com/products/gs-FLX-system/index.asp. Accessed on 26 March 2015.

Rodríguez-Lázaro, D., Cook, N., Ruggeri, F. M., Sellwood, J., Nasser, A., Nascimento, M. S. J., D'Agostino, M., Santos, R., Saiz, J. C., Rzezutka, A., Bosch, A., Gironés, R., Carducci, A., Muscillo, M., Kovač, K., Diez-Valcarce, M., Vantarakis, A., von Bonsdorff, C.-H., de Roda Husman, A. M., Hernández, M., & van der Poel, W. H. M. (2012). Virus hazards from food, water and other contaminated environments. *FEMS Microbiology Reviews*, *36*(4), 786–814.

Rosario, K., & Breitbart, M. (2011). Exploring the viral world through metagenomics. *Current Opinion in Virology*, *1*(4), 289–297.

Rosario, K., Nilsson, C., Lim, Y. W., Ruan, Y., & Breitbart, M. (2009). Metagenomic analysis of viruses in reclaimed water. *Environmental Microbiology*, *11*(11), 2806–20.

Ross, E. M., Moate, P. J., Bath, C. R., Davidson, S. E., Sawbridge, T. I., Guthridge, K. M., Cocks, B. G., & Hayes, B. J. (2012). High throughput whole rumen metagenome profiling using untargeted massively parallel sequencing. *BMC Genetics*, *13*(1), 53.

Ross, E. M., Petrovski, S., Moate, P. J., & Hayes, B. J. (2013). Metagenomics of rumen bacteriophage from thirteen lactating dairy cattle. *BMC Microbiology*, *13*(1), 242.

Rybicki, E. P. (2015). A top ten list for economically important plant viruses. *Archives of Virology*, *160*(1), 17–20.

Rzezutka, A., & Cook, N. (2004). Survival of human enteric viruses in the environment and food. *FEMS Microbiology Reviews*, *28*(4), 441–53.

Sánchez, G., Elizaquível, P., & Aznar, R. (2012). A single method for recovery and concentration of enteric viruses and bacteria from fresh-cut vegetables. *International Journal of Food Microbiology*, *152*(1-2), 9–13.

Scharff, R. L. (2012). Economic burden from health losses due to foodborne illness in the United States. *Journal of Food Protection*, *75*(1), 123–31.

Seo, K., & Frank, J. (1999). Attachment of *Escherichia coli* O157 : H7 to lettuce leaf surface and bacterial viability in response to chlorine treatment as demonstrated by using confocal scanning laser microscopy. *Journal of Food Protection*, *62*(1), 3–9.

Seymour, I. J., & Appleton, H. (2001). Foodborne viruses and fresh produce. *Journal of Applied Microbiology*, *91*(5), 759–73.

Shang, C., & Blatchley, E. R. (2001). Chlorination of pure bacterial cultures in aqueous solution. *Water Research*, *35*(1), 244–254.

Sharma, M. (2013). Lytic bacteriophages: Potential interventions against enteric bacterial pathogens on produce. *Bacteriophage*, *3*(2), e25518.

Sivapalasingam, S., Friedman, C. R., Cohen, L., & Tauxe, R. V. (2004). Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. *Journal of Food Protection*, *67*(10), 2342–53.

Smith, C. M., & Hill, V. R. (2009). Dead-end hollow-fiber ultrafiltration for recovery of diverse microbes from water. *Applied and Environmental Microbiology*, *75*(16), 5284–9.

Song, I., Stine, S. W., Choi, C. Y., & Gerba, C. P. (2006). Comparison of crop contamination by microorganisms during subsurface drip and furrow irrigation. *Journal of Environmental Engineering*, *132* (10), 1243–1248.

Steele, M., & Odumeru, J. (2004). Irrigation water as source of foodborne pathogens on fruit and vegetables. *Journal of Food Protection*, *67*(12), 2839–2849.

Stine, S. W., Song, I., Choi, C. Y., & Gerba, C. P. (2005). Application of microbial risk assessment to the development of standards for enteric pathogens in water used to irrigate fresh produce. *Journal of Food Protection*, *68*(5), 913–918.

Taormina, P. J., Beuchat, L. R., Erickson, M. C., Ma, L., Zhang, G., & Doyle, M. P. (2009). Transfer of *Escherichia coli* O157:H7 to iceberg lettuce via simulated field coring. *Journal of Food Protection*, *72*(3), 456–684.

Telias, A., White, J. R., Pahl, D. M., Ottesen, A. R., & Walsh, C. S. (2011). Bacterial community diversity and variation in spray water sources and the tomato fruit surface. *BMC Microbiology*, *11*(1), 81.

United States Department of Agriculture (USDA). (2014). Economic research service: food availability (per capita) data system: summary findings. *United States Department of Agriculture website.* Retrieved from http://www.ers.usda.gov/data-products/food-availability-(per-capita)-data-system/summary-findings.aspx. Accessed on 26 March 2015.

United States Environmental Protection Agency (EPA) (2001). Method 1602: male-specific (F+) and somatic coliphage in water by single agar layer procedure. *EPA 821-R-98-029.* Office of Water, Washington, DC.

United States Environmental Protection Agency (EPA) (2012). Guidelines for Water Reuse. *EPA-600-R-12-618.* Office of Water, Washingtion DC.

Varma, A., & Malathi, V. G. (2003). Emerging geminivirus problems: a serious threat to crop production. *Annals of Applied Biology*, *142*(2), 145–164.

Vázquez-Castellanos, J. F., García-López, R., Pérez-Brocal, V., Pignatelli, M., & Moya, A. (2014). Comparison of different assembly and annotation tools on analysis of simulated viral metagenomic communities in the gut. *BMC Genomics*, *15* (1), 37.

Vega, E., Garland, J., & Pillai, S. D. (2008). Electrostatic forces control nonspecific virus attachment to lettuce. *Journal of Food Protection*, *71*(3), 522–9.

Vega, E., Smith, J., Garland, J., Matos, A., & Pillaii, S. D. (2005). Variability of virus attachment patterns to butterhead lettuce. *Journal of Food Protection*, *68*(10), 2112–7.

Voller, A., Bartlett, A., Bidwell, D. E., Clark, M. F., & Adams, A. N. (1976). The detection of viruses by enzyme-linked immunosorbent assay (ELISA). *Journal of General Virology*, *33*(1), 165–167.

Wang, D., Urisman, A., Liu, Y. T., Springer, M., Ksiazek, T. G., Erdman, D. D., Mardis, E. R., Hickenbotham, M., Magrini, V., Eldred, J., Latreille, J. P., Wilson, R. K., Ganem, D., & DeRisi, J. L. (2003). Viral discovery and sequence recovery using DNA microarrays. *PLoS Biology*, *1*(2).

Ward, B. K., & Irving, L. G. (1987). Virus survival on vegetables spray-irrigated with wastewater. *Water Research*, *21*(1), 57–63.

Wei, J., Jin, Y., Sims, T., & Kniel, K. E. (2010). Manure- and biosolids-resident murine norovirus 1 attachment to and internalization by Romaine lettuce. *Applied and Environmental Microbiology*, *76*(2), 578–83.

Wheeler, C., Vogt, T. M., Armstrong, G. L., Vaughan, G., Weltman, A., Nainan, O. V., Dato, V., Xia, G., Waller, K., Amon, J., Lee, T. M., Highbaugh-Battle, A., Hembree, C., Evenson, S., Ruta, M. A., Williams, I. T., Fiore, A. E., & Bell, B. P. (2005). An outbreak of hepatitis A associated with green onions. *The New England Journal of Medicine*, *353*(9), 890–897.

Widdowson, M.-A., Sulka, A., Bulens, S. N., Beard, R. S., Chaves, S. S., Hammond, R., Salehi, E. D. P., Swanson, E., Totaro, J., Woron, R., Mead, P. S., Bresee, J. S., Monroe, S. S., & Glass, R. I. (2005). Norovirus and foodborne disease, United States, 1991-2000. *Emerging Infectious Diseases*, *11*(1), 95–102.

Wilson, A. E., Sarnelle, O., & Tillmanns, A. R. (2006). Effects of cyanobacterial toxicity and morphology on the population growth of freshwater zooplankton: Meta-analyses of laboratory experiments. *Limnology and Oceanography*, *51*(4), 1915–1924.

Wylie, K. M., Weinstock, G. M., & Storch, G. A. (2013). Virome genomics: A tool for defining the human virome. *Current Opinion in Microbiology*, *16*(4), 479–484.

Yépiz-Gómez, M. S., Gerba, C. P., & Bright, K. R. (2013). Survival of respiratory viruses on fresh produce. *Food and Environmental Virology*, *5*(3), 150–156.