IMPACT AND RECOVERY OF GALVESTON BAY MICROBIAL COMMUNITIES FOLLOWING HURRICANE HARVEY

A Thesis

by

ALAINA CHRISTINE WOODS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Jessica M. Labonté
Committee Members,	Karl Kaiser
	Jeffrey Turner
	Antonietta Quigg
Head of Department,	Daniel Roelke

December 2020

Major Subject: Marine Biology

ABSTRACT

Microorganisms account for ~70% of the total biomass of the ocean and fill many necessary roles within the ecosystem. However, little is known about how the marine microbial communities respond and recover from environmental stressors such as hurricanes. Hurricanes bring sudden changes to the marine ecosystem such as drops in salinity and temperature and influx of terrestrial sediment. All of these changes may impact the marine microbial community structure and genomic composition. Hurricane Harvey hit the Texas Gulf Coast on August 25th, 2017 as a category 4 hurricane and brought a record-breaking amount of rainfall to the Houston area. Storm water run-off from Hurricane Harvey brought excess freshwater and sediment to Galveston Bay. Here, we characterized how Hurricane Harvey altered the microbial communities within Galveston Bay. This was achieved through the following three objectives: 1) describe the change in microbial community structure, 2) identify how the genomic composition and metabolic potential of the microbial community changed and 3) demonstrate the role of viruses in ecosystem recovery by showing how the viral community adapts and changes with host abundances following Hurricane Harvey. These objectives were accomplished through sampling four stations along a transect in Galveston Bay once a week for five weeks. The microbial communities were identified using 16S ribosomal RNA gene sequencing for the hosts and metagenomics for the hosts and viruses. Viral production experiments were used to characterize viral activity. The results show that the microbial community transitioned from an ecosystem that was dominated by marine

microorganisms (e.g., Cyanobacteria and Acidimicrobiia) prior to Hurricane Harvey to a system dominated by microorganisms that were terrestrially derived (e.g., Actinobacteria and Verrucomicrobiota) after landfall. The genomic composition and metabolic potential of the community changed after Hurricane Harvey with an increase in genes involved in nitrogen and sulfur metabolisms and a decrease in genes involved in photosynthetic metabolisms. Common marine viruses such as *Podoviridae* and *Myoviridae* were removed from the ecosystem and there was an increase in auxiliary metabolic genes associated with nitrogen, sulfur and methane metabolisms. While the prokaryotic community almost recovered, the viral community did not recover to pre-Hurricane Harvey conditions within five weeks, when comparing to pre-Harvey conditions. With climates changing, it is predicted that hurricane and rainfall intensities are going to increase over the years; therefore, it is important to understand how pulse disturbances like large rain event impact the marine microbial community, how the ecosystem recovers, and the impact of the changes on the global ecosystem.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Jessica Labonté, for all her guidance and support during the course of this research. I would also like to thank my committee members, Drs. Kaiser, Turner and Quigg for their time spent helping me with my thesis. Thank you to the Labonté Viral Ecology lab members specifically, Kate Campbell and Jordan Walker, for their patience in answering my many questions and allowing me to pinball ideas off of them!

Thank you to my mom, dad and sisters for their support during this process, I would not have been able to make it here without them. Mom and dad, your constant support throughout my education has meant the world to me, thank you for always pushing me! Thank you to my friends who stuck with me through this stressful process, I appreciate all of you!

Finally, thank you to my partner, Travis, for all of his support and help. Putting up with the late nights and early mornings, taking over puppy duty so I could write and focus, cooking or ordering so many meals to make sure I was eating and generally putting up with the level of stress that comes with finishing a master's thesis! I truly appreciate everything you have done for me.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Dr. Jessica Labonté [advisor] of the Department of Marine Biology, Dr. Karl Kaiser of the Department of Marine and Coastal Environmental Science, Dr. Jeffrey Turner of Texas A&M University- Corpus Christi Department of Marine Biology and Dr. Antonietta Quigg associated with the departments of Marine Biology and Oceanography.

The 16S rRNA gene data analyzed in Chapter III was provided by Dr. Antonietta Quigg and the Quigg Phytoplankton Dynamics Lab. This data was published by Dr. Jamie Steichen (Department of Marine Biology) and colleagues in April of 2020 in Frontiers in Marine Science.

Funding Sources

Graduate study was supported by the 2-Year Competitive Graduate Student Fellowship from Texas A&M University at Galveston. This research was funded by the National Science Foundation—RAPID Response (Award #1801367).

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

With ~$1x10^{29}$ bacterial and archaeal cells, the ocean, which covers 71% of Earth's surface, houses an incredible amount of microbial diversity (Flemming and Wuertz 2019). Prokaryotes and protists represent a large proportion of the biomass on Earth, 70 Gt C (1 Gt C = 10^{15}g C) and 4 Gt C, respectively, accounting for ~70% of the total biomass in the ocean (Bar-On et al 2018). The ocean contains ~$4x10^{30}$ viruses (Suttle 2005), which translates to about 0.2 Gt C (Bar-On et al 2018). Although they represent less biomass due to their smaller sizes, 0.2 Gt C (Bar-On et al 2018), viruses are on average 10-fold more abundant than their hosts in marine waters (Suttle 2005, Wigington et al 2015), making them the most abundant biological entities on Earth. For comparison, all the animals on our planet (molluscs, chordates, arthropods, etc.) represent 2 Gt C (Bar-On et al 2018). Microorganisms (or microbes, which include protists, bacteria, and viruses) fill diverse roles such as nutrient cycling, evolution and adaptation, and population control within the marine environment.

1.1 Role of microorganisms in nutrient cycling

1.1.1 Carbon cycling

The majority of the organic material in the ocean is in the dissolved phase, which accounts for the largest global reservoir of fixed carbon (Hansell and Carlson 2001). The carbon is fixed from the atmosphere by the primary producers via photosynthesis. Once the atmospheric carbon is fixed, it can be transported to upper trophic levels or enter the

dissolved phase as dissolved organic matter (DOM) or particulate organic matter (POM). DOM enters the ecosystem by sloppy feeding and excretion by zooplankton, viral lysis of cells, and direct exudation from phytoplankton (Anderson and Ducklow 2001). DOM can then be taken up by bacteria (and archaea) where it can then move up trophic levels, be released through lysis of virally-infected cells, or be exported to the deep ocean through the biological pump (Ducklow et al 2001). Half or more of the DOM in the marine food web passes through marine microorganisms (Azam et al 1983, Fuhrman 1999).

The biological pump is the combined biological processes that export organic matter to the deep ocean. In the surface ocean, primary producers use dissolved atmospheric carbon dioxide and convert it to organic carbon. The organic carbon is then converted back to carbon dioxide through food web processes, mainly from consumption by heterotrophic bacteria (bacteria that derive nutritional requirements from complex organic matter) and zooplankton, while the other fraction, the POM, is used by the traditional food web (e.g. plankton, fishes, mammals) or sinks to the deep ocean where it settles into sediment for long-term carbon storage (Figure 1) (Ducklow et al 2001).

The microbial loop describes the process of DOM being returned to higher trophic levels via bacterial biomass, instead of being exported to the deep ocean via the biological pump. Heterotrophic bacteria within the water column use DOM as their primary energy source. These bacteria utilize the DOM that is too small to be utilized by the traditional food chain, playing an important role in the microbial loop (Figure 1) (Azam et al 1983, Pomeroy et al 2007).

A study characterizing the composition of bacterial communities during spring phytoplankton blooms in the southern Indian Ocean found significant evidence that the available DOM shapes the bacterial community within the water column (Landa et al 2016). The increased supply of phytoplankton DOM, from the spring algal bloom, increased bacterial production, suggesting that the source of available DOM impacts the composition of the bacterial community within the marine environment. These findings show a connection between the biological pump and microbial loop; heterotrophic bacteria are utilizing DOM that would otherwise be lost to the deep ocean.

1.1.2 Nitrogen cycling

Microorganisms control remineralization of nitrogen, nitrification and nitrogen fixation within the ocean (Hutchins and Fu 2017). Rivers are responsible for inputting 40-66 Tg (teragram = megaton = 10^{12} g) of nitrogen per year to coastal ecosystems (Seitzinger et al 2005, Voss et al 2013). Other major sources of nitrogen to coastal ecosystems include sediment and vegetation, totaling about 15 Tg of nitrogen per year (Voss et al 2013). Nitrogen fluxes to the coastal ocean increase with higher precipitation and freshwater discharge rates, although these forms of nitrogen are not usable by the traditional food web (Howarth et al 2002, Howarth 2008). Of the nitrogen input into the coastal ecosystem through rivers, 40% is in the form of particulate nitrogen, which has to go through remineralization and nitrification by microorganisms in order to be used by the traditional food web as nitrate (Seitzinger et al 2005, Voss et al 2013).

Microorganisms are also responsible for fixation of atmospheric nitrogen in the ocean (Riemann et al 2010, Sohm et al 2011). Cyanobacteria control the fixation of

atmospheric nitrogen to ammonium and dissolved organic nitrogen in the ocean (Riemann et al 2010, Sohm et al 2011). The annual average rate of nitrogen fixation by the filamentous cyanobacteria *Trichodesmium* spp. is 250 μmol N m^{-2} d^{-1} in the North Atlantic Ocean (Capone et al 2005, Sohm et al 2011) and 30-120 μmol N m^{-2} d^{-1} in the North Pacific Ocean (Dore et al 2002, Sohm et al 2011). Nitrogen fixation is normally associated with warmer oligotrophic waters (Capone et al 1997), and warmer waters allow for high respiration rates (Stal 2009), lowered water temperatures, normally associated with hurricanes, would slow nitrogen fixation rates.

1.1.3 Sulfur cycling

Along with DOM and nitrogen cycles, the sulfur cycle, which is especially important within marine sediment, is also dependent upon microbial activities and microbial transformation of inorganic and organic sulfur (Jørgensen and Kasten 2006). The oceans are one of the largest sulfur pools on Earth, housing $1.3x10^9$ Tg of sulfate (Jørgensen and Kasten 2006). The main source of sulfur into the ocean is the weathering of rocks and sediment which is then cycled through marine microorganisms that are metabolically diverse (Sievert et al 2007). Within marine sediment the sulfur cycle is driven by anaerobic microorganisms reducing sulfate to sulfide. Biotic sulfur oxidation, done by microorganisms, in the surface sediment, oxidize sulfide far more quickly, orders of magnitude more, than abiotic chemical oxidation is capable of (Luther et al 2011, Thamdrup et al 1994). Sulfate reducing microorganisms are responsible for 29% of the remineralization of organic matter on the seafloor (Bowles et al 2014, Wasmund et al 2017). Due to the high rates of remineralization by sulfate reducing

microorganisms, sulfate reduction is the driver of the sulfur cycle in marine sediment (Bowles et al 2014, Wasmund et al 2017).

In the coastal ocean, phytoplankton convert methionine into dimethylsulfoniopropionate (DMSP), a highly soluble form of reduced sulfur, with each molecule of DMSP containing five carbon atoms (Gage et al 1997, Sievert et al 2007). Due to the five carbon atoms associated with DMSP, the production of DMSP is estimated to account for 3–10% of global marine primary productivity budgets (Kiene et al 2000, Sievert et al 2007). DMSP is also a source of sulfur, through assimilation, to bacteria in the surface water, specifically alpha- and betaproteobacteria (Malmstrom et al 2004, Sievert et al 2007). Due to the production of DMSP, marine phytoplankton play a crucial role in the global oceanic sulfur cycle (Alcolombri et al 2015, Sievert et al 2007)

1.1.4 Role of viruses in nutrient cycling

In the ocean, lytic viruses have impacts on biogeochemical cycles (Fuhrman 1999). The lytic life cycle starts when the virus attaches itself onto the host, injects its DNA, and then takes over the host's metabolism to support viral production. Viral production within the host ends with cell lysis and release of the viral progeny and the host cell's contents (Fuhrman 1999). Viral lysis is responsible for the viral shunt. The viral shunt, first described by Wilhelm and Suttle (1999), refers to the redirection of nutrients from higher trophic levels back to the microbial food web through viruses infecting and lysing microbial cells, which releases the host cells contents back into the total DOM pool (Figure 2) (Suttle 2007). The DOM released via lysis is then utilized by heterotrophic bacteria in the microbial loop (Figure 1), instead of being passed up to the

traditional food chain. Without the viral shunt, the host cells would be consumed by secondary consumers and the nutrients would continue up the food chain (Wilhelm and Suttle 1999).

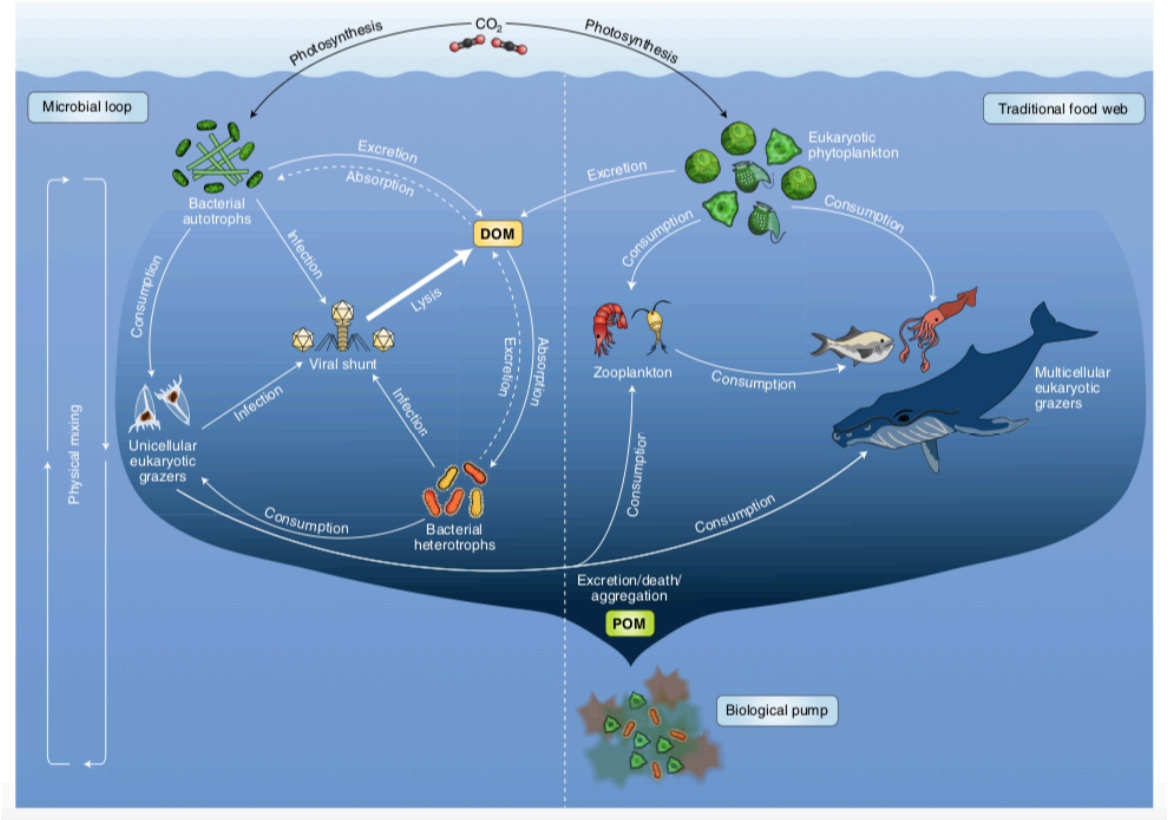

Figure 1 Graphic depiction of the microbial loop and viral shunt, showing sources of particulate organic matter (POM) and dissolved organic matter (DOM) along with export pathways. Image from (Breitbart et al 2018).

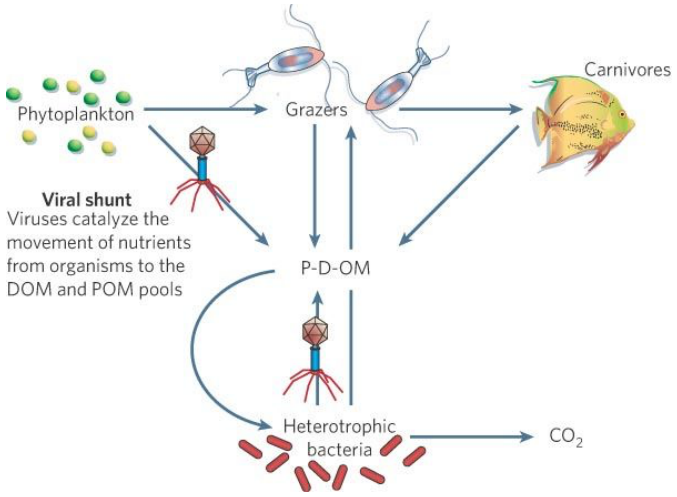

Figure 2 Graphic depiction of the viral shunt. POM stands for particulate organic matter and DOM stands for dissolved organic matter. Image from (Suttle, 2005).

1.2 Microbial evolution and adaptation

Viruses and their hosts are in a constant race to out evolve each other, therefore driving one another's adaptation and evolution. As the host cell develops defense mechanisms, the virus evolves mechanisms to circumvent the host's defenses to continue its lifecycle, which will force the host cell to evolve new defense mechanisms (Rohwer and Thurber 2009). This competition is known as the "Red Queen Hypothesis", coined by Leigh Van Valen (Van Valen 1974). Support for this hypothesis was established through the sequencing of the virus EhV86's genome infecting the coccolithophore *Emiliania huxleyi* (Bidle et al 2007). The genome of EhV86 encodes genes that are similar to those in ceramide production (ceramides are involved in triggering apoptosis). Apoptosis is a host cell defense mechanism used to decrease viral spread (Galluzzi et al 2008); however, EhV86 has developed mechanisms to use apoptosis to its advantage and increase replication (Bidle et al 2007).

Temperate viruses, or viruses that can undergo the lysogenic cycle, are important players of horizontal gene transfer (HGT). The lysogenic cycle occurs when the phage (a virus that infects bacteria) integrates its nucleic acids into the host genome. Nucleic acids are therefore passed down to other generations as the host divides. The lytic cycle can be induced when the host cell is subjected to an environmental stressor such as changes in abiotic factors like nutrient availability, temperature or pH (Howard-Varona et al 2017). Temperate phages can be involved in transduction, or virally mediated HGT, by moving host genes from one cell to another. HGT is one of the dominant ways to expand bacterial metabolic capabilities, which can aid organisms in creating condition-specific adaptations and increase functional diversity in their ever-changing environments (Pál et al 2005, Smets and Barkay 2005). HGT can occur in three ways: transformation (uptake of free-floating genetic material), conjugation (gene transfer from cell-to-cell contact) and transduction (virally mediated movement of nucleic material). It is suggested that there could be as many as 10^{24} (one septillion) genes transferred through transduction every year in the ocean, and this number might be far greater (Rohwer and Thurber 2009).

A bacterium that was isolated from an environment highly contaminated with crude oil provided a good example of HGT (Das et al 2015). Sequencing of the genome of the isolate *Pseudomonas aeruginosa* N002 found that the bacterium had acquired many genetic adaptations to survive in contaminated soil, and at least two of these adaptations were transferred via HGT. Two specific genes that were identified as being transferred by HGT encoded for organic solvent transporter systems, which aid in crude

oil degradation. This example shows that environmental stressors impact the genetic make-up of microorganisms within that environment (Das et al 2015).

Auxiliary metabolic genes (AMGs) are used by viruses, especially by phages, during viral infection to redirect the host energy and resources to support viral production (Breitbart et al 2007, Thompson et al 2011). As many as 243 different AMGs were found within the metagenomic datasets collected on the *Tara* Oceans and *Malaspina* research expeditions (Roux et al 2016). Diversity ranges from AMGs that encode for phosphate uptake (*phoH*) and stress (Crummett et al 2016, Goldsmith et al 2011), alteration of host carbon fluxes (Hurwitz et al 2013), nucleotide synthesis (Crummett et al 2016), to parts of the photosynthetic process (Crummett et al 2016, Mann et al 2003, Thompson et al 2011), to give a few examples.

An example of how AMGs can alter host metabolic pathways to support phage production was seen in a study on uncultured cyanophages. The study identified a Calvin cycle inhibitor gene, *cp12*, in 16 out of 17 uncultured cyanophage genomes sequenced (Thompson et al 2011). The *cp12* gene allows for support of the pentose phosphate pathway (PPP) in the host. It is suggested that when a cyanophage having the *cp12* gene infects a host cell, the Calvin cycle is down-regulated, which disassociates the light reaction from the Calvin cycle. By doing this, the light reaction and PPP occur together within the host cell (these are usually offset by ~12 hours). These two processes working simultaneously can fuel phage dNTP biosynthesis, increasing phage production and effectively altering the host's normal metabolic activity (Thompson et al 2011). AMGs, along with their function, are also transferred to the host organism (Tsiola et al 2020).]

1.3 Role of viruses in population control

Viruses play an important role in population control and diversity of prokaryotes within the ocean (Bouvier and Giorgio 2007, Thingstad and Lignell 1997). One density dependent model showed that viruses eliminate the most competitive hosts, which increases resource availability allowing for a more diverse community, a concept known as the "Kill the Winner" hypothesis (Thingstad and Lignell 1997). Support for "Kill the Winner" is best seen when algal blooms occur. One mesocosm study found that when a bloom of the marine algae *Emiliania huxleyi* occurred, there were high levels of virally mediated mortality (Bratbak et al 1993). Indeed, when the algal bloom collapsed it was found that about 3% of the algal cells were infected with intracellular viruses (Bratbak et al 1993). When a bloom collapses, resources become available to other species allowing for an increase in species diversity. It has been shown that viral lysis plays a significant role in algal bloom control, along with zooplankton grazing and bacterial interactions (Paerl and Otten 2013).

Another example of viral role in population control was seen with the algae *Phaeocystis pouchetii* and its virus PpV. When cultivated alone, the abundance of *P. pouchetii* reached a maximum of 1.1×10^6 cells ml^{-1} within two weeks. When PpV was added to the sample (with a virus-to-algae ratio of 0.5) the entire population of *P. pouchetii* was lysed in eight days (Haaber and Middelboe 2009). Decreasing the virus-to-algae ratio to 0.005, *P. pouchetii* abundances only reached a maximum of 3.9×10^5 cells ml^{-1} and the PpV virus lysed the entire population in 12 days. In competition experiments, it was shown that virally reducing *P. pouchetii* abundances stimulated

growth of *Rhodomonas salina,* the non-infected algae. Based on these findings, it is hypothesized that viruses play a role in succession and relative distribution and diversity of marine algae.

1.4 Microbial response to abiotic changes

Studies have found abiotic factors such as salinity and temperature are drivers of aquatic microbial community composition (Langenheder et al 2003, Langenheder and Ragnarsson 2007, Rubin and Leff 2007). One study isolated three abiotic factors (salinity, chlorophyll-a concentration and water color) that had a significant effect on the bacterial community composition within Baltic Sea rock pools. These abiotic factors explained 16.7 and 14.6% of the total variation in bacterial community composition between the rock pools. Salinity was the most significant factor shaping the bacterial community composition, accounting for 8.3% of variation (Langenheder and Ragnarsson 2007). Bacterial community composition is strongly correlated with differences in salinity, which is supported by research done in surface waters of the Baltic Sea where salinity was the major factor that structured bacterial community composition (Herlemann et al 2011).

Temperature has also been shown to be a dominant abiotic factor that influences bacterial community composition by altering abundances (Fuhrman et al 2008, Gilbert et al 2009, Rubin and Leff 2007). In an Ohio river, temperature was the abiotic environmental variable that showed the strongest correlation to bacterial abundance (Rubin and Leff 2007). These results were further supported by more recent studies done in the Western English Channel (Gilbert et al 2009) and at other locations across the

oceans such as North and South Pacific, and South Atlantic (Fuhrman et al 2008). There were highly significant relationships found between specific bacterial taxa (Actinobacteria, Bacteroidetes, Cyanobacteria and Alpha- Gammaproteobacterial) and how temperature changed their relative abundance (Gilbert et al 2009). The water temperature at the time of sampling and average sea-surface temperature positively correlated with community composition (Fuhrman et al 2008). Given the significant drop in salinity and temperature seen following Hurricane Harvey, these abiotic factors may have altered the microbial community composition and abundance in Galveston Bay, which is the focus of this thesis.

1.5 Microbial response and resilience to pulse disturbance

In ecology, a pulse disturbance is described as an intense, short term disturbance that increases and decreases in severity in a short period of time (Lake 2000). Microbial communities have a unique ability to adapt to pulse disturbances, such as oil spills, like Deep Water Horizon (Mason et al 2012), and hurricanes (Amaral-Zettler et al 2008). A main factor aiding in microbial success during pulse disturbances is the versatility in utilizing different carbon acquisition pathways and energy sources (i.e., photoautotrophy and heterotrophy) (Eiler 2006).

Shifts in microbial community composition were observed as a result of the 2010 Deepwater Horizon oil spill (Mason et al 2012) and following Hurricane Katrina's landfall in 2005 (Amaral-Zettler et al 2008). In April 2010, the Deepwater Horizon offshore drilling platform exploded and sunk in the Gulf of Mexico (GOM). From April to July there were over ~4.9 million barrels of oil released into the GOM. The release of

oil into the GOM altered the structure and function of the microbial community. Analysis of the PCR 16S ribosomal RNA (rRNA) gene showed that the microbial community changed and displayed an increase in hydrocarbon-degrading bacteria within the oil plume. Due to the increase in hydrocarbon-degrading bacteria within the oil plume, the microbial community diversity was lower within the oil plume than outside of the plume (Mason et al 2012). Analysis of 16S rRNA gene sequences showed that a single operational taxonomic unit (OTU), identified as *Oceanospirillales,* comprised almost 90% of the total microbial community within the oil plume. However, outside of the oil plume, *Oceanospirillales* only comprised ~3% of the total microbial community. Metagenomic and metatranscriptomics of samples within the oil plume also revealed higher relative abundance of hydrocarbon degradation genes, showing that the microbial community within the oil plume had a different genetic makeup when compared to the control samples (Mason et al 2012).

Hurricane Katrina was a category 4 hurricane that made landfall on August 29th, 2005. Hurricane Rita was a category 3 hurricane that made landfall on September 24th, 2005. Both hurricanes hit the United States Gulf Coast and resulted in massive flooding in New Orleans, LA, changing the composition of the microbial community (Amaral-Zettler et al 2008). Samples collected from flood waters, canals and two lakes (Lake Pontchartrain and Lake Charles) showed 2,139 separate bacterial OTUs, with 69 of them related to potentially pathogenic species. The most abundant OTUs found in floodwater samples were from *Rhodobacter* and *Allochromatium*, two genera that are usually found in wastewater and highly polluted systems, not in aquatic ecosystems. The data

suggested that the change in composition is a result of nutrient availability in the ecosystem following the mass flooding event (Amaral-Zettler et al 2008).

Few studies have assessed the impact pulse disturbances have on marine viral communities. One study by Williamson et al (2014) looked at the impact of stormwater runoff on viral community composition in freshwater systems. Viral community composition was explored at Grim Dell retention pond before, during and after Hurricane Sandy hit Virginia on October 26th – November 1st, 2012. Hurricane Sandy averaged about 2 mm of rain per hour for 36 hours, with a peak of 28 mm in 6 hours at 72 hours. This rainfall lowered ammonium, phosphate, nitrate and nitrite concentrations within the pond. Over the six day observational period, viral abundance ranged from 1.4 to 3.6 x10^6 ml^{-1} and was consistently higher than bacterial abundance, by a factor of 10. A multiple regression analysis showed that conductivity, pH and bacterial abundance had a significant impact ($p<0.001$) on the variation in viral abundance. Viral species richness decreased following the hurricane, an effect of increased rainfall, and then increased as the ecosystem recovered. These findings show that extreme events have an impact on the aquatic viral community (Williamson et al 2014). With Hurricane Harvey being a 1-in-1000 year flooding event (Blake and Zelinsky 2018), it is necessary to determine the impact on the viral community in Galveston Bay.

Community resilience is the rate at which the community returns to an original composition following the disturbance (Allison and Martiny 2008, Shade et al 2012a). Community resilience stems from the idea that there is a strong "core" group of microorganisms that are resilient, and the community retains similar composition and

function (Needham et al 2013). Over a 78 day sampling period off the coast of California, the microbial community changed very little and function did not appear to be impacted by daily or seasonal changes (Needham et al 2013). Using mixing as a pulse disturbance, the microbial communities in the epilimnion and hypolimnion layers returned to pre-disturbance conditions in seven and eleven days, respectively (Shade et al 2012b). These studies support that microbial communities tend to return to pre-disturbance conditions. However, quantifying the impact on microbial communities and the time it takes to return to pre-disturbance conditions could be used as another measure of pulse disturbance intensity.

1.6 Hurricane Harvey

Hurricane Harvey was a category 4 hurricane that hit San Jose Island, TX on August 25th, 2017. Hurricane Harvey brought high winds to San Jose Island and mass amounts of rain to the Houston area. Over the course of five days, one station, about 40 km west of Houston, reported a rainfall accumulation of 1,318 mm (Oldenborgh et al 2017). For comparison, NOAA National Weather Service reports the annual amount of rainfall in Houston, over the last 30 years, to average 1,264.16 mm per year. Estimations predict that the city of Houston sank into the Earth by ~2 cm in areas where rainfall totals peaked due to weight of water (Milliner et al 2018). The influx of water into Galveston Bay following Hurricane Harvey brought excess nutrients, a persistent sediment plume (Figure 3), and drops in salinity, temperature (Du et al 2019), and dissolved oxygen (Figure 4). The sediment plume brough 9.86 x 10^7 metric tons of

sediment into Galveston Bay, which is equivalent to the 18 year average of sediment load (Du et al 2019).

Models are predicting rainfall rates associated with tropical cyclones will increase by 10-15% in the coming years, along with the intensity of these storms (Knutson et al 2015). With rainfall rates and intensities of tropical cyclones predicted to increase, it is important to understand how organisms are impacted following a pulse disturbance, and further our understanding of their role in ecosystem recovery. It is unknown how the ecosystem changes brought on by Hurricane Harvey impacted the marine microbial community structure and function, or the role that viruses played in ecosystem recovery within Galveston Bay.

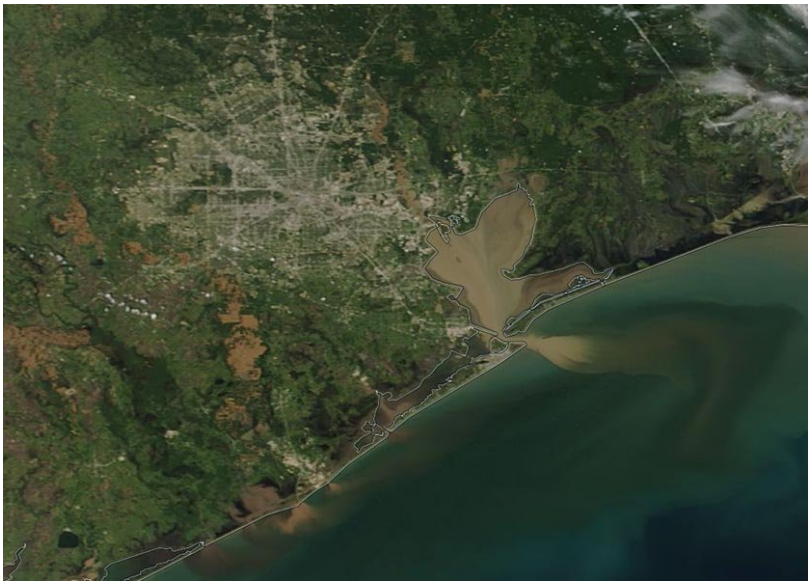

Figure 3 Aerial photo showing the sediment plumes in Galveston Bay and the Gulf of Mexico after Hurricane Harvey. Plumes were caused by flood waters draining into Galveston Bay and the Gulf. The image was taken on August 31st, 2017 from NASA's Terra satellite. Image credit: NASA.

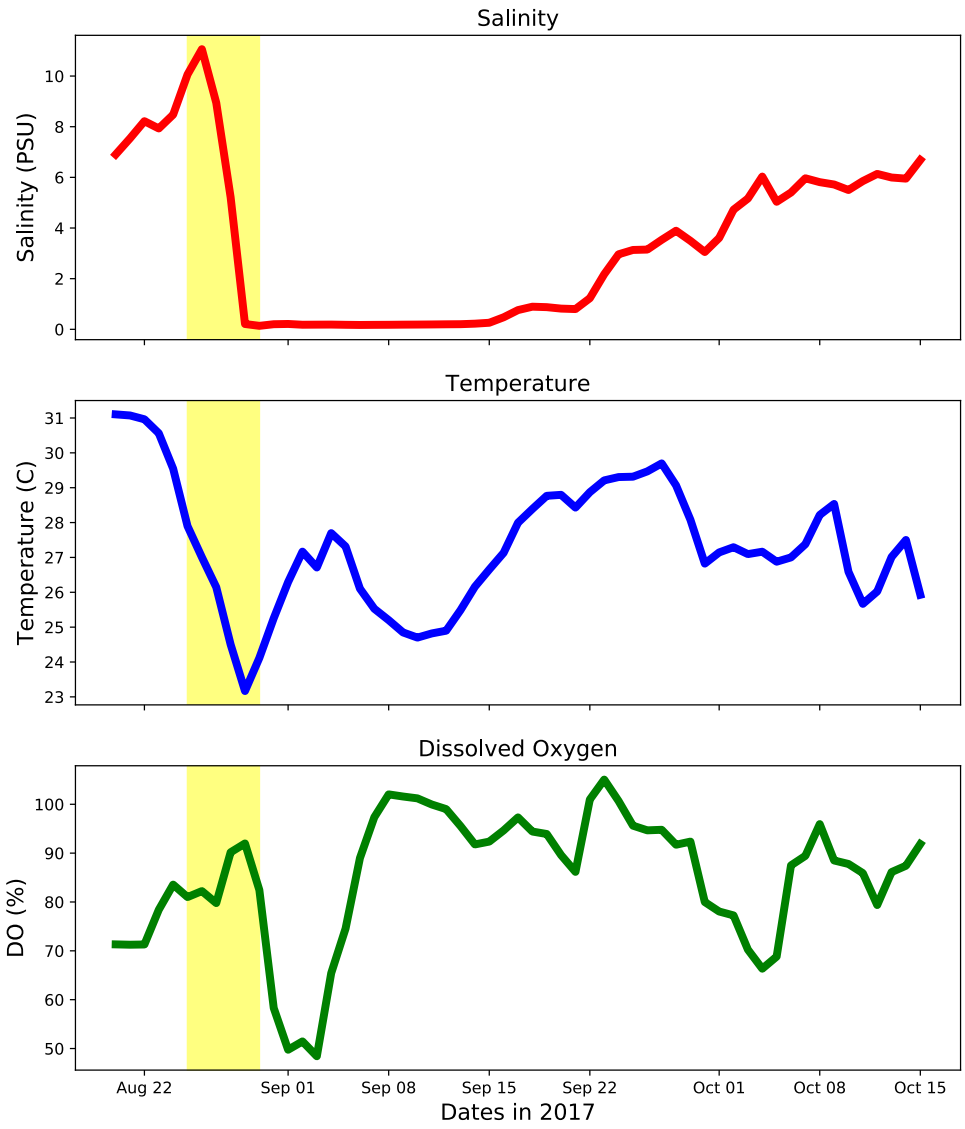

Figure 4 Graphs showing the drop-in salinity, temperature and dissolved oxygen from August 20th – October 15th, 2017. The shaded area represents August 26th-30th, 2017 when heavy precipitation from Hurricane Harvey occurred. All data was collected from the Trinity Bay station operated by Texas Water Development Board. Data was downloaded from Water Data for Texas (waterdatafortexas.org).

CHAPTER II

OBJECTIVES

The main goal of this study is to characterize how Hurricane Harvey, which was a pulse disturbance, impacted the microbial communities in Galveston Bay. This goal was accomplished through the completion of the following three objectives:

2.1 Objective 1

Describe the changes in the microbial community composition in Galveston Bay following Hurricane Harvey.

Hypothesis: Since the microbial community structure is directly impacted by the environment, the lower salinity and influx of terrestrial sediment, in the form of $9.86x10^7$ metric tons of sediment (Du et al 2019), brought by Hurricane Harvey *would cause there to be an increase in bacterial sequences associated with sediment and terrestrial environments within Galveston Bay* due to the freshwater influx. *There would be a decrease in genes associated with photosynthetic processes,* due to photosynthetic bacteria being flushed out of the bay and high turbidity due to sediment load because of the unfavorable conditions created in Galveston Bay. *There would be an increase in genes associated with nitrogen metabolisms* due to fresh water loading into Galveston Bay. Finally, *there would be an increase in genes associated with sulfur metabolisms* due to the influx of sediment within the bay.

2.2 Objective 2

Identify how Hurricane Harvey impacted microbial metabolisms, especially the metabolisms associated with the carbon, nitrogen and sulfur cycles.

Hypothesis: Hurricane Harvey brought a sudden change to the marine environment in Galveston Bay by dropping temperature, and introducing $9.86x10^7$ metric tons of sediment (Du et al 2019). Due to these sudden changes, *there would be a decrease in genes associated with photosynthetic processes, due to photosynthetic bacteria being flushed out of the bay and high turbidity due to sediment load. There would be an increase in genes associated with nitrogen metabolisms due to fresh water loading into Galveston Bay. Finally, there would be an increase in genes associated the sulfur metabolisms due to the influx of sediment within the bay.*

2.3 Objective 3

Demonstrate the role of viruses in ecosystem recovery in Galveston Bay by showing how the viral community adapts and changes with host abundances following Hurricane Harvey.

Hypothesis: Based on the Kill-the-Winner hypothesis, *the viral community would change following their host community, therefore play a role in population control. There would also be an increase in AMGs associated with nitrogen and sulfur metabolisms within the viral community because of the influx of freshwater and sediment into Galveston Bay.*

CHAPTER III

HURRICANE HARVEY IMPACT TO BACTERIAL COMMUNITY

3.1 Introduction

Prior research has shown that hurricanes have altered marine microbial communities with the changes being attributed to the availability of nutrients such as ammonium, nitrate and nitrite (Amaral-Zettler et al 2008). Impacts of hurricanes on coastal bays include decreased temperature, salinity, and dissolved oxygen, and increased sediment loading (Figure 4) (Balthis et al 2006, Du et al 2019, Park et al 2007). However, microbial communities are highly resilient to changes due to the versatility in utilizing different carbon acquisition pathways and energy sources (Eiler 2006). Here, we aimed to determine the changes in the microbial community composition following Hurricane Harvey and identify how Hurricane Harvey impacted key microbial functions involved in nitrogen and sulfur metabolisms and carbon acquisition pathways.

Hurricane Harvey made landfall in Port Aransas on the 25th of August 2017. While Galveston did not experience hurricane-force winds, but rather a record-breaking amount of rainfall that then flushed into Galveston Bay, 1.4-1.7 x 10^{10} m^3. Hurricane Harvey not only brought rainfall accumulation, but also drops in salinity, temperature, and dissolved oxygen, and 9.86 x 10^7 metric tons of sediment into Galveston Bay (Figure 4) (Du et al 2019). A previous study showed that Hurricane Harvey changed the microbial community composition from a marine to a terrestrial community (Steichen et

al 2020). Here, we will complement other studies performed on the impact of Hurricane Harvey by looking at the changes in diversity and genomic potential.

We characterized the changes Hurricane Harvey brought to the microbial community at the taxonomic and genomic level. To accomplish this, we sampled a transect for five consecutive weeks in Galveston Bay for metagenomics. We hypothesized 1) there would be an increase in bacterial sequences associated with sediment and terrestrial environments within Galveston Bay due to the freshwater influx; 2) there would be a decrease in genes associated with photosynthetic processes, due to photosynthetic bacteria being flushed out of the bay and high turbidity due to sediment load because of the unfavorable conditions created in Galveston Bay; 3) there would be an increase in genes associated with nitrogen metabolisms due to fresh water loading into Galveston Bay and; 4) there would be an increase in genes associated with sulfur metabolisms due to the influx of sediment within the bay. We observed a large decrease in genes involved in primary production as runoff flowed out of Galveston Bay and an increase in genes associated with metabolisms and bacterial taxa associated with the terrestrial and sediment environments as sediment was resuspended and flood waters flowed into Galveston Bay.

3.2 Material and Methods

Note: Methods in sections 3.2.1 and 3.2.2 were performed by members of the Labonté Viral Ecology Lab prior to my arrival at Texas A&M University at Galveston.

3.2.1 Sample collection and processing

Samples were collected from four stations in Galveston Bay, after Hurricane Harvey's landfall, on September 4th (Samp1), 9th (Samp2), 16th (Samp3), 21st (Samp4) and 28th (Samp5) of 2017. A transect was sampled from the mouth of the San Jacinto River to the Gulf of Mexico (Figure 5 & Table 1). Pre-Harvey samples were collected from Texas A&M University at Galveston's boat basin on July 31st and August 22nd, of 2017 prior to Hurricane Harvey's land fall. The volume of samples collected ranged from 4 to 20 L, depending on time constraints, available material, and personnel (Table 7 in Appendix A).

All samples were pre-filtered immediately after sampling with a Nitex filter (30 µm) to remove small grazers and large particles. The total volume of each sample varied depending on time and manpower constraints for each sampling day. After filtration the samples were stored on the boat in the dark and brought to the laboratory for further filtration. Generally, each sample was filtered through a glass fiber filter (GF-F with a 0.7 µm pore-size or GF/D with a 2.7 µm pore-size), followed by a 0.22 µm pore-size polyvinylidene fluoride (PVDF) filter. Due to the availability of supplies, for the sampling of 09/09, prefiltration was performed with a 0.45 µm and the virus concentrate was the filtered through 0.22 µm PVDF filter. All GF and PVDF filters were stored at -20°C until further use (Table 7 in Appendix A).

3.2.2 DNA extractions, PCR and sequencing

Metagenomic DNA extraction and sequencing

Total DNA was extracted from the filters using a standard phenol chloroform extraction protocol (Green and Sambrook 2017). The lysis buffer recipe used was 400 mM NaCl, 750 mM sucrose, 20 mM EDTA, 50 mM Tris-HCl with a pH of 9.0. Between 0.5–1 μg of DNA per sample was sequenced using Illumina HiSeq 4000 with 150 bp paired-end sequencing technologies at the Texas A&M Genomics and Bioinformatics facility (College Station, Texas, United States).

16S rRNA DNA extraction, PCR and sequencing

This section was done in collaboration with the Quigg Phytoplankton Dynamics Lab at Texas A&M University at Galveston. Two hundred mL of water were filtered through 0.2μm polyethylsulfone membrane filter and stored at -80ºC until nucleic acid was extracted. DNA was extracted using the MO Bio PowerSoil DNA Isolation Kit (cat. No. 128888-50) (Steichen et al 2020). PCR was used to amplify the 16S rRNA gene from prokaryotes. The DNA polymerase used was Promega GoTaq Flexi DNA polymerase and PCR primers used were 16S-515F (5'-GTG YCA GCM GCC GCG GTA A-3') and 16S-806R (5'-CCG YCA ATT YMT TTR AGT TT-3') (10 μM each) (Parada et al 2016, Steichen et al 2020). The following cycling parameters were used: 95ºC for 3 minutes, 30 cycles at 95ºC for 45 seconds, 50ºC for 60 seconds, and 72ºC for 90 seconds, the final step was 72ºC for 10 minutes (Parada et al 2016, Steichen et al 2020). PCR reactions were carried out in triplicate to maximize the overall yield, then pooled, and purified with UltraClean PCR Clean-Up Kit (MoBio Laboratories; Carlsbad,

CA, United States). Sequencing was then completed at the Georgia Genomics Facility (Athens, GA, United States) with MiSeq sequencing technology (v2 chemistry, 2 x 250 bp) (Steichen et al 2020).

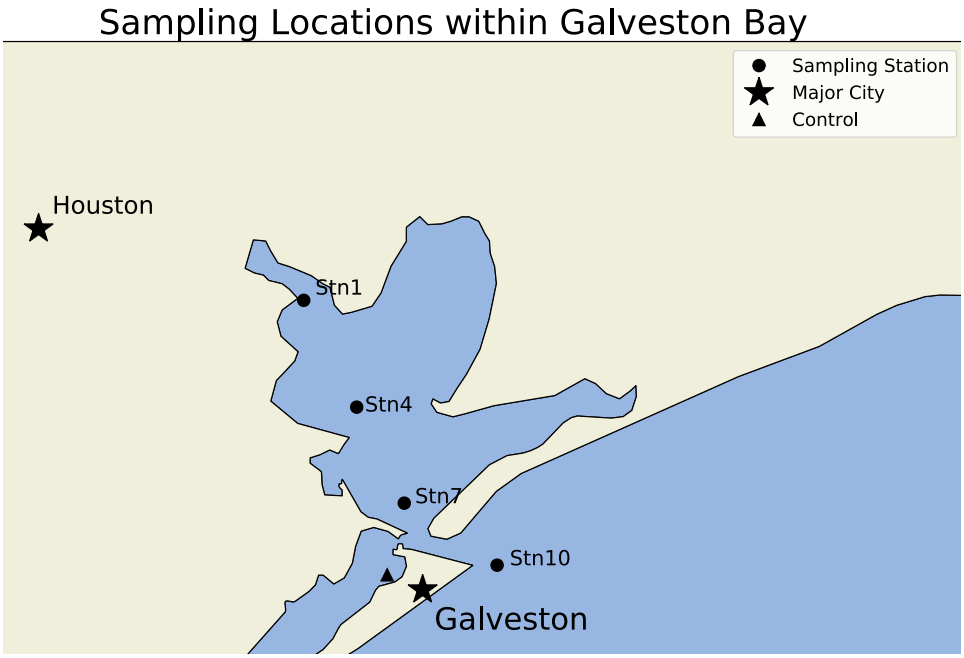

Figure 5 Sampling map of Galveston Bay, Texas (United States) showing the sampling stations used during each of the sampling efforts. The triangle represents the Texas A&M University at Galveston boat Pre-Harvey samples taken on 07/31/2017 and 08/22/2017.

Table 1 Coordinates of the four sampling stations and pre-Harvey samples.

Station ID	Coordinates
Stn1	-94.976944, 29.6725
Stn4	-94.897222, 29.535556
Stn7	-94.825833, 29.4125
Stn10	-94.688611, 29.333611
Pre-Harvey	-94.85, 29.32

3.2.3 16S rRNA gene analysis

Analysis of 16S rRNA gene was performed with *mothur* (version 132) following the standard operating procedure (SOP) MiSeq protocol accessed in July of 2019 (https://www.mothur.org/wiki/MiSeq_SOP) (Kozich et al 2013, Schloss et al 2009). That included quality control, merging duplicated sequences, alignment utilizing the Silva non redundant database version 132 file, clustering utilizing the cluster.split command, and then an ASV (Amplicon Sequence Variant) list was generated using the classify.otu command with the label "asv". Taxonomy and diversity measures were visualized (maps, graphs, etc.) with R-Studio and python 3 (Figure 6) (RStudio Team 2016). 16S rRNA sequences are publicly available through the Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC) at https://data.gulfresearchinitiative.org/ (doi: 10.7266/PGC99C7D) (Steichen et al 2020).

3.2.4 Metagenomic analysis

BBtools software suite (https://sourceforge.net/projects/bbtools/) was used for quality control of the metagenomic raw reads. The commands for BBtools were built following a tutorial from the USDA ARS Microbiome Workshop in 2017 titled Metagenomics tutorial part 1: Quality control, assembly and mapping (https://usda-ars-gbru.github.io/Microbiome- workshop/tutorials/metagenomics/). BBduk (version 38.31) was used to filter out contaminated reads, including kmers of 31 bp in length, and dropping the extra base that is occasionally added by Illumina sequencing. BBduk was also used in conjunction with the Phred algorithm for more accurate adapter trimming and removal of kmers between 11 and 23 bp in length – a common length for adapter

artifacts. BBmerge (version 38.31) was used to merge the forward and reverse paired-end reads using default settings (Bushnell et al 2017). BBmask (version 38.31) was used to soft-mask human, cat, dog and mouse contamination. MEGAHIT (version 1.2.8) was then used for the *de novo* assembly into contigs (Li et al 2015). Gene prediction was completed with Prodigal (version 2.6.3) (Hyatt et al 2010). Genes and proteins were annotated with DIAMOND (version 0.9.26), utilizing the fast setting and NCBI protein non-redundant database (March 19th, 2019) (Buchfink et al 2015). PhyloFlash was used to extract and identify the 16S rRNA genes from the microbial metagenomes (Gruber-Vodicka et al 2019). All metagenomes will be uploaded into the MG-RAST metagenomics analysis server (Meyer et al 2008) for public access. Community analysis and figures were done in MEGAN (Huson et al 2007) and the R package vegan (Figure 7) (Oksanen et al 2019). The Bray-Curtis similarity matrix and non-metric multidimensional scaling (NMDS) plot was run using vegan (Oksanen et al 2019) in R-Studio (RStudio-Team 2016).

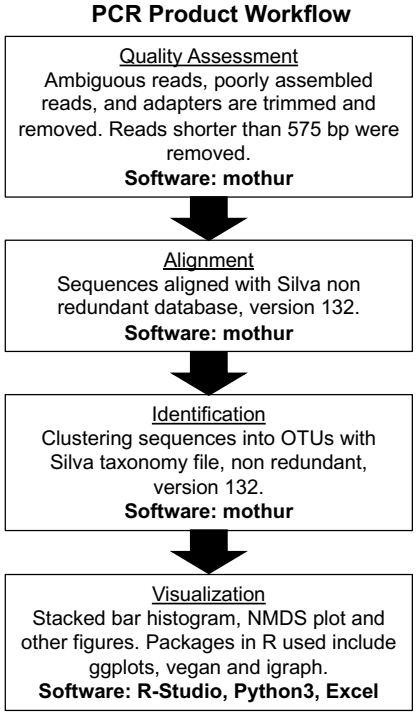

Figure 6 Bioinformatic pipeline for 16S rRNA gene analysis from PCR product data.

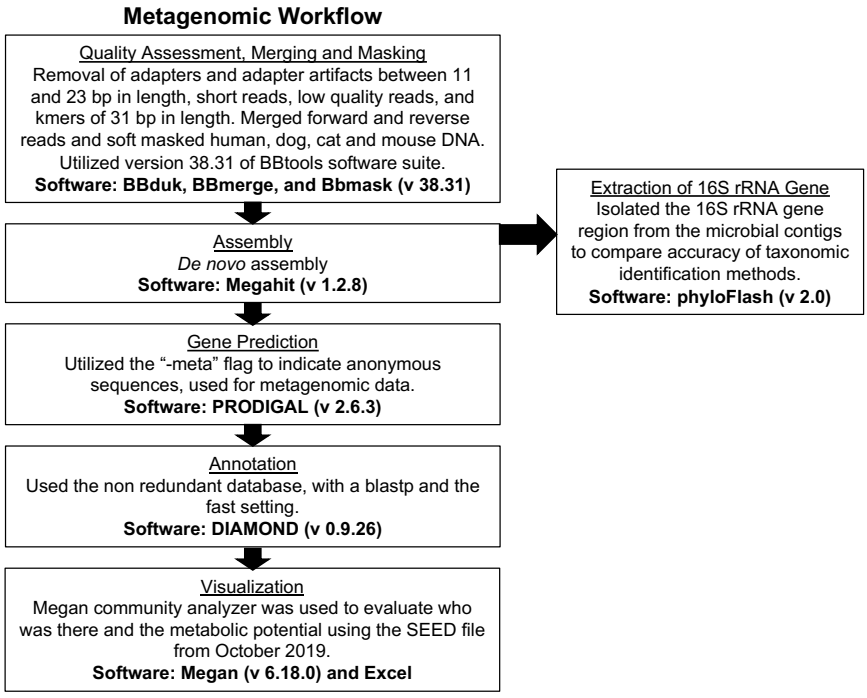

Figure 7 Bioinformatic pipeline for microbial metagenomes and viral DNA.

3.4 Results

For metagenomic data, we sampled four times over the course of the six weeks following Hurricane Harvey at four sampling stations arranged in a transect from the San Jacinto River to the mouth of the Gulf of Mexico (Figure5 and Table 2). We also collected two pre-Harvey samples taken from TAMUG's boat basin (Table 2). Sampling efforts resulted in 18,788,809 contigs, with 12,977,903,603 bp assembled and 26,740,520 proteins recovered in the 18 microbial metagenomes (Table 8 in Appendix A). On average, each sample had 360,497,322 bp, 521,911 contigs, and 742,792 proteins. The PCR product data had four sampling stations at five time points with two pre-Harvey samples. This resulted in 12,341,512 sequences prior to quality control

measures; following quality control there was 9,069,244 sequences. The average N50 value was 753.

3.4.1 Hurricane Harvey changed the microbial community composition

16S rRNA gene sequences from the metagenomic data and PCR amplicons both revealed that the relative abundance of common marine bacteria was greatly reduced in Samp1 when compared to pre-Harvey samples, and was replaced by terrestrial bacteria (Figure 8). Proteobacteria was the most abundant phylum, representing on average 37.1% of the total community. Cyanobacteria showed the greatest change in relative abundance from pre-Harvey to the Samp1, relative abundances went from 18.0% of the community pre-Harvey to 10.4% of the community during Samp1 (Figure 8B). The Verrucomicrobia phylum increased in relative abundance following Hurricane Harvey's landfall, going from a relative abundance of 3.5% prior to Hurricane Harvey to 6.2% during Samp1. The class Gammaproteobacteria increased in relative abundance following Hurricane Harvey's landfall, with average relative abundance going from 17.1% pre-Harvey to 19.4% during the first sampling period; by Samp5, the community had returned to 18.4% (Figure 8B). Bacteroidia decreased in relative abundance later in the sampling period showing the highest relative abundances during Samp1, averaging 9.2% of the community and dropping to 7.9% of the total community by Samp5 (Figure 8B), which this is lower than the pre-Harvey samples that were 9.3%. Actinobacteria represented 1.0% of the community prior to Hurricane Harvey's landfall; however, during Samp1, this class represented 4.8% of the total community (Figure 8B). Acidimicrobiia had a relative abundance of 4.7% prior to Hurricane Harvey's landfall.

Following landfall Acidimicrobiia dropped in relative abundance and had a relative abundance of 2.9%. Levels of Acidimicrobiia had almost returned to pre-Harvey conditions by Samp5 with a relative abundance of 3.2% (Figure 8B). Gammaproteobacteria had a relative abundance of 17.1% of the total community prior to Hurricane Harvey's landfall. Following landfall Gammaproteobacteria relative abundances increased to 19.4% of the total community. Levels of Gammaproteobacteria had not completely recovered by Samp5, still representing 18.4% of the total community (Figure 8B).

We evaluated the microbial diversity from each metagenome using a Shannon index. Samp1_Stn4 and Samp1_Stn7 had the highest microbial richness and diversity, 3.128 and 3.121 respectively. While pre-Harvey samples had lower microbial richness and diversity, 2.707 and 2.655 respectively (Table 3).

A non-metric multidimensional scaling plot was used to group the samples and determine if samples cluster based on time post-Harvey or geographic distribution in the Bay. There was a clear difference in groupings in samples taken prior to Hurricane Harvey's landfall and subsequent sampling dates (Figure 9). Knowing that abiotic factors have significant impacts on shaping microbial communities, we wanted to see what abiotic factors controlled the community after Hurricane Harvey. Five environmental variables had significant impacts on the microbial community: salinity, secchi depth, total phosphorus, total nitrogen and pH. Total nitrogen and total phosphorus had significant impacts on the community ($p=0.05$) while pH and Secchi

depth had a stronger significant impact on the microbial community ($p=0.005$) and salinity had the strongest significant impact ($p=0.001$) (Figure 9).

Utilizing a Bray-Curtis Dissimilarity matrix, we compared the two microbial community datasets, the PCR amplicons and metagenomic datasets. All of the metagenomic samples clustered together in groupings distinct from the PCR amplicons, and all values produced by the matrix were >0.7, confirming the two methods produced very different communities (Figure 10). For the metagenomic dataset, all but one Samp5 sampling time points grouped together, closely to the two pre-Harvey samples. For the PCR product samples, Samp1, Samp2, and Samp3 grouped together while samplings Samp4 and Samp5 grouped with the pre-Harvey samples (Figure 10). In the PCR product dataset, Samp1_Stn10 grouped more closely with Samp4, Samp5 and the pre-Harvey samples than the other Samp1 samples.

Table 2 Shannon diversity indices from the PCR product data at the class level. The darker the green the higher the community richness and evenness.

Dates in 2017	Station	Shannon Index
	Stn4	3.148
	Stn7	3.121
Sep. 4th	Stn10	2.954
	Stn1	3.081
	Stn4	3.064
	Stn7	3.013
Sep. 9th	Stn10	2.773
	Stn1	2.889
	Stn4	2.702
	Stn7	2.608
Sep. 16th	Stn10	2.718
	Stn1	3.101
	Stn4	3.028
	Stn7	2.781
Sep. 21st	Stn10	2.872
	Stn1	2.883
	Stn4	2.643
	Stn7	2.837
Sep. 28th	Stn10	2.712
July 31st	Pre-Harvey1	2.707
August 22nd	Pre-Harvey2	2.656

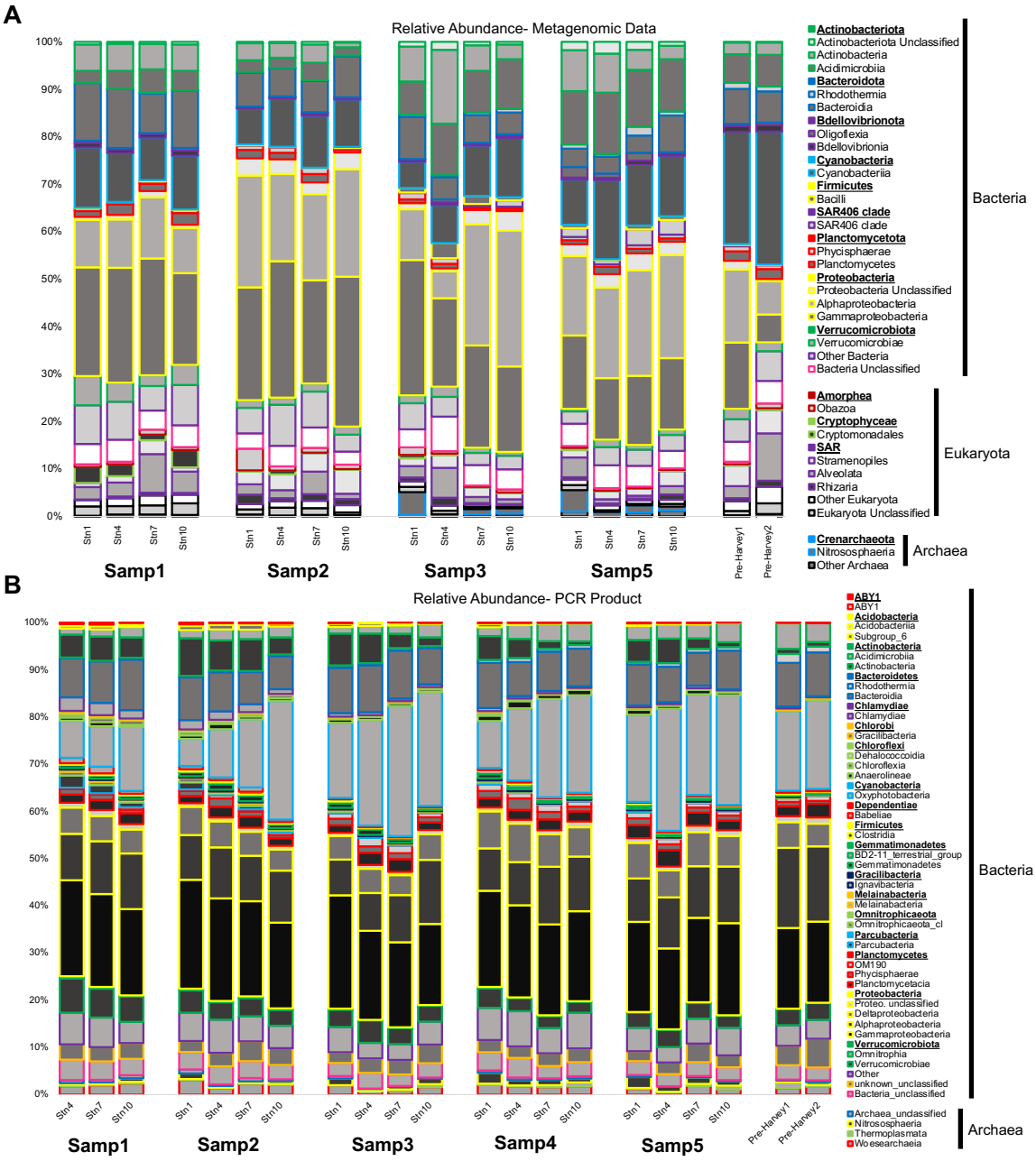

Figure 8 Stacked bar histogram showing the relative abundance of bacterial ASVs at the class level of the microbial community identified by A) 16S and 18S rRNA extracted from the metagenomic data and B) 16S rRNA gene PCR amplicons data. Pre-Harvey1 and Pre-Harvey2 are samples taken before Hurricane Harvey on 7/31/17 and 8/22/17.

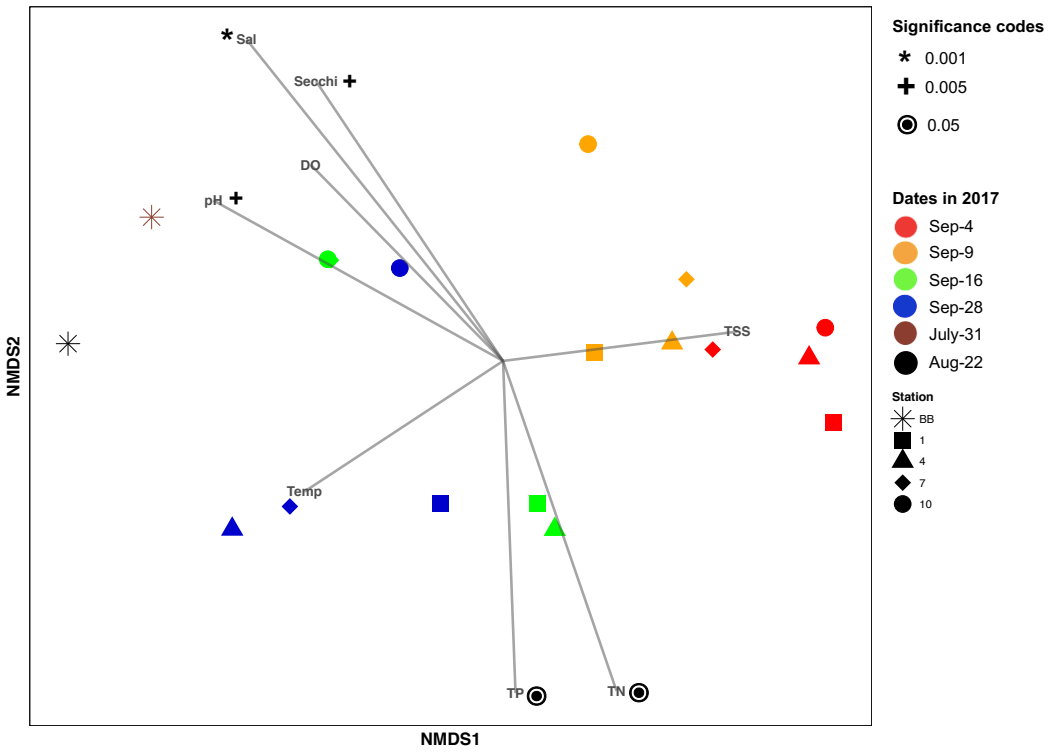

Figure 9 NMDS of the metagenomic community composition showing the correlation between temperature (Temp; ^{o}C), salinity (Sal), pH, dissolved oxygen (DO;%), total nitrogen (TN; μM^{1}), total phosphorus (TP; μM), Secchi (m) and total suspended sediment (TSS; g mL^{-1}). All data for environmental variables were taken from (Steichen et al 2020). Colors represent sampling time points in 2017 and symbols represent sampling stations. NMDS was constructed using metagenomic ASV abundances at the phylum level using a square root transformation and a Bray-Curtis similarity matrix, with a stress level of 0.084.

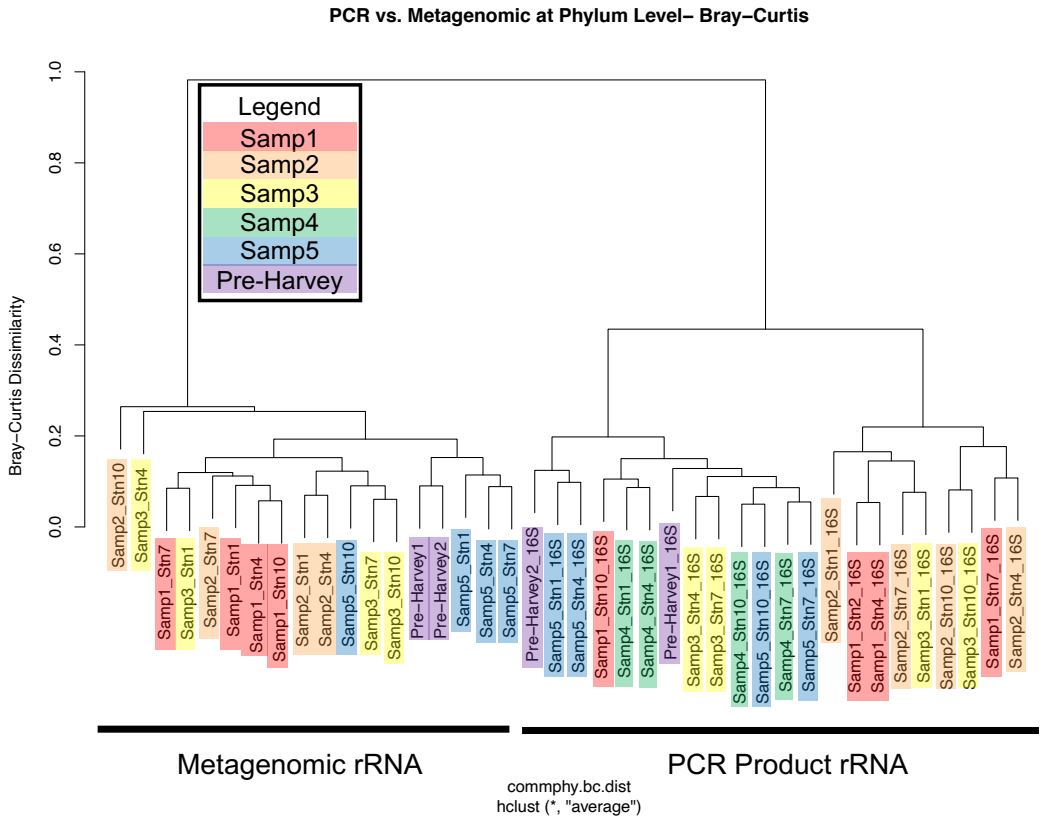

Figure 10 Bray-Curtis Dissimilarity matrix, represented in a dendrogram, showing a clear separation in the 16S rRNA gene sequences from metagenomes and PCR amplicons.

3.4.2 Hurricane Harvey changed the microbial genomic composition

We evaluated the genomic potential of the microbial community following Hurricane Harvey. Genes associated with common photosynthetic processes (photosystems I and II and chlorophyll biosynthesis) decreased after Hurricane Harvey's landfall (Figure 11). Following the decrease of photosynthetic genes there was an increase in genes associated with respiration during Samp1 (Figure 11). By Samp5, the incidence of genes associated with photosynthetic and respiration processes almost recovered to pre-Hurricane Harvey levels.

There was an increase in the abundance of inorganic sulfur assimilation genes during Samp1. The abundance then decreased over time to return to pre-Hurricane Harvey levels by Samp4 (Figure 11). The abundance of genes associated with sulfur oxidation also increased from the pre-Harvey samples and peaked during Samp3. There was a decrease in the abundance of dimethylsulfoniopropionate (DMSP) mineralization genes in Samp1 (Figure 11). DMSP remineralization genes did not return to pre-Hurricane Harvey levels by Samp5. There was a decrease in ammonia assimilation gene abundance. Gene abundances of ammonia assimilation genes dropped from the controls to Samp1. Gene abundances of ammonia assimilation genes almost recovered to pre-Hurricane Harvey levels by Samp5. Genes associated with nitrate and nitrite ammonification increased from the pre-Harvey samples to Samp1. Gene abundances of nitrate and nitrite ammonification genes had almost returned to pre-Harvey levels by Samp5 (Figure 11).

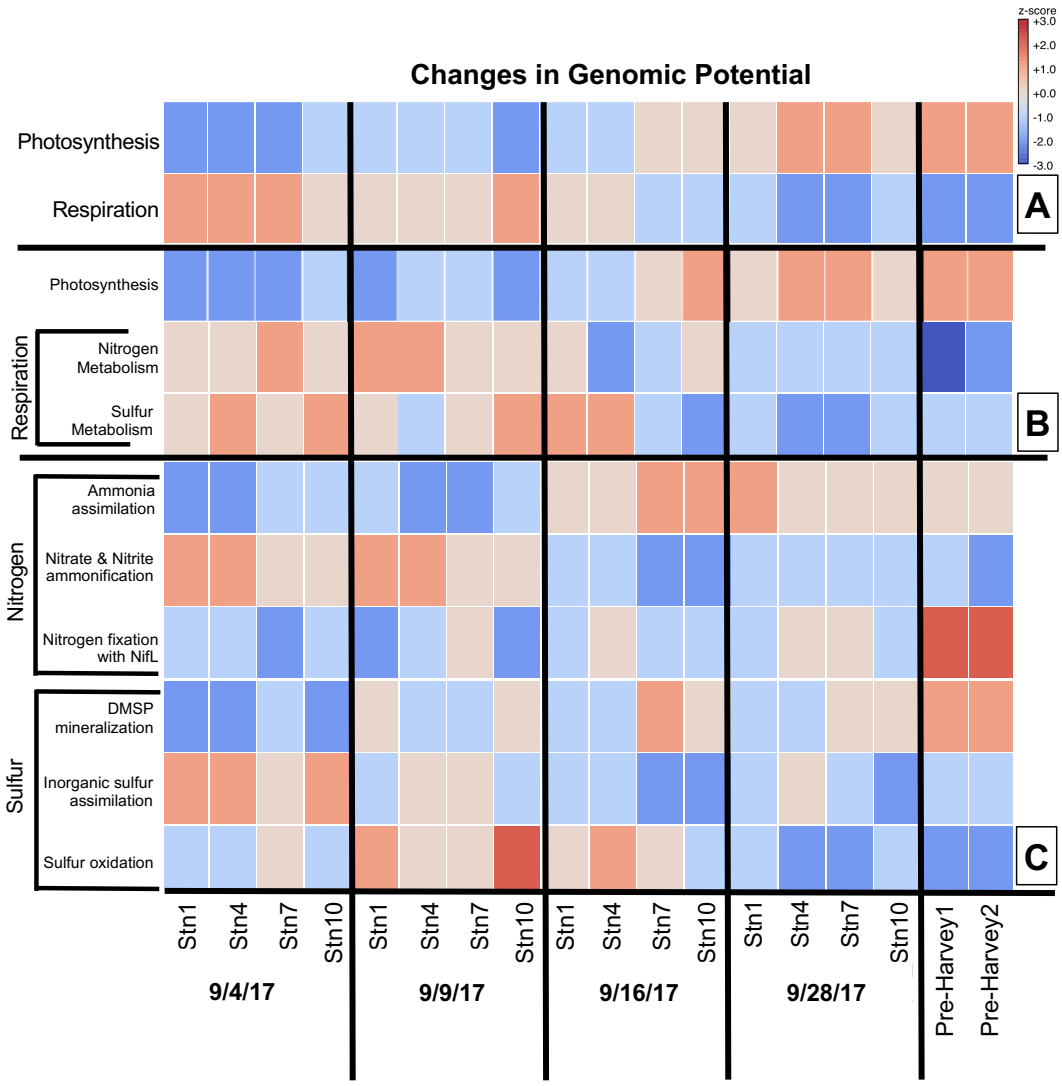

Figure 11 Heat map showing the z-score for genes associated with photosynthetic processes and respiration (A); genes associated with photosynthetic processes, nitrogen and sulfur metabolisms (B); and a further breakdown of genes associated with pathways specific to nitrogen and sulfur metabolisms. The z-score is calculated separately for each part of the figure (A, B & C) and is based on averages from that section only, darker red (3.0) there are above average reads and darker blue (-3.0) there are below average reads. Samples are broken down by date with two pre-Harvey samples at the end taken 7/31/17 and 8/22/17.

3.5 Discussion

In ecology, a pulse disturbance is described as an intense, short term disturbance that increases and decreases in severity in a short period of time (Lake 2000). Based on this definition, Hurricane Harvey is considered a pulse disturbance. Hurricane Harvey caused an intense, short-term disturbance in Galveston Bay with a sediment loading of 9.86×10^7 metric tons of sediment brought into the system, along with a drop in salinity due to the record-breaking amount of rainfall, as well as decreased temperature and dissolved oxygen (Du et al 2019, Steichen et al 2020). Changes brought to the ecosystem from the pulse disturbance impacted the microbial community by flushing out dominant marine members, such as Cyanobacteria and Acidimicrobiia, which were replaced with terrestrially derived microorganisms such as Actinobacteria and Verrucomicrobiota (Figure 8).

After the pulse disturbance, there were many bacterial phyla of terrestrial origin in Galveston Bay. We chose to focus our discussion on two of the more abundant phyla: Actinobacteria and Verrucomicrobiota. The majority of the Actinobacteria phylum are a common soil-dwelling, terrestrial microorganisms (Barka et al 2016, Mayfield et al 1972). Within the phylum Actinobacteria, we recovered two classes: Acidimicrobiia and Actinobacteria. Acidimicrobiia is an abundant class of Actinobacteria that is found in a variety of aquatic habitats (Hu et al 2018). However, Actinobacteria is one of the top three dominant classes found within soil samples taken from cultivated, forested and pastured environments (Shange et al 2012). Within the phylum Verrucomicrobiota, we recovered two separate classes of bacteria after the pulse disturbance: Verrucomicrobiae

and Omnitrophia. Verrucomicrobiae is an abundant bacterial class in soil samples, specifically soil samples from grassland and prairie environments. On average, 35% of all bacterial sequences taken from soil samples in grassland and prairie environments belonged to the class Verrucomicrobiae (Bergmann et al 2011). There was also an increase in Gammaproteobacteria, which are known to occur in marine sediment (Gribben et al 2017). An increase in these classes within the microbial community can be attributed to the resuspension of marine sediment in the water column, high sediment loading and flux of freshwater into the bay after the pulse disturbance, which brought an influx of marine sediment and terrestrial soil bacteria. These results are supported by Steichen et al (2020), who found similar conclusions of Hurricane Harvey's impact to Galveston Bay.

Common marine bacteria, such as Cyanobacteria, were removed from the ecosystem after the pulse disturbance. Possible explanations for the removal of Cyanobacteria include the drop in salinity and high sediment loading. The Cyanobacteria commonly found in Galveston Bay belong to the Synechococcales order, which are not well adapted to strive in freshwater environments (Khatoon et al 2010, Taylor et al 2001). The ecosystem became a freshwater environment after the pulse disturbance with salinities in the range of 0.1-0.4 PSU (Figure 4), which most likely inhibited the growth of these Cyanobacteria. The removal of Cyanobacteria could also be attributed to the high sediment loading. Following this pulse disturbance there was high turbidity due to sediment loading, which could account for the decrease in Cyanobacteria following the pulse disturbance.

There was a shift in the genomic potential in the microbial community from a community dominated by primary producers (dominated by Cyanobacteria) to a community dominated by heterotrophs following the pulse disturbance. This shift can be seen with an increase in genes associated with respiration and a decrease in genes associated with photosynthesis (Figure 11A). Respiration is a metabolic process associated with environments such as terrestrial soil and is performed by heterotrophic bacteria (Bond-Lamberty et al 2018). Respiration in terrestrial soil originates from microorganism's decomposition of organic matter within the soil (Stotzky 1965). Microorganisms are one of the main controllers of respiration rates within soil, specifically Alphaproteobacteria and Bacteroidetes (Fierer et al 2007, Liu et al 2018). The increase in respiration genomic potential is likely due to terrestrial sediment loading in the ecosystem or a shift to a heterotrophic community that more quickly consumes dissolved organic carbon (DOC).

There was an increase in genomic potential associated with nitrogen and sulfur metabolisms following the pulse disturbance (Figure 11B). Half of the nitrogen input into the coastal ocean is via riverine input, while atmospheric deposition of nitrogen impacts the open ocean more (Voss et al 2013). The main source of sulfur to the oceanic ecosystem comes from the weathering and leaching of rocks and sediment (Sievert et al 2007). The pulse disturbance brought an increase in freshwater input from three major rivers along with a massive sediment load (Du et al 2019), which would increase the nitrogen and sulfur input into the ecosystem thus increasing the need for nitrogen and sulfur metabolisms (Figure 11B).

Certain pathways within the sulfur and nitrogen metabolisms did not follow the overall trend of increasing after the pulse disturbance. Pathways that did not follow the trend are associated with Cyanobacteria and still decreased such as ammonia assimilation and DMSP mineralization. The decrease in ammonia assimilation genes further supports the loss of phytoplankton after the pulse disturbance as ammonia assimilation genes are found within phytoplankton (Capone et al 2008, Glibert et al 2016, Klawonn et al 2019). The removal of photosynthetic microbes from the bay is not only supported by the decrease in genes associated with photosynthetic pathways (Figure 11 A & B) but is also supported by a decrease in genes associated with DMSP mineralization. DMSP is a metabolite primarily produced by marine phytoplankton (Keller et al 1999), with a decrease in phytoplankton within the system there would be less DMSP to mineralize thus causing abundances of genes associated with DMSP mineralization to drop.

Two genes that did not follow the overall trends of sulfur and nitrogen metabolisms were decreases in genes associated with nitrogen fixation with *nifL* and sulfur oxidation, which had a delay in increased abundance (Figure 11C). The *nif* genes are found in most bacteria, specifically Proteobacteria and are nitrogen sensitive . Nitrogen fixation with the *nifL* gene is done when there is an ample amount of reduced nitrogen or oxygen present in the environment (Dixon 1998, O'Carroll and Dos Santos 2011). A decrease in nitrogen fixation with *nifL* genes could mean a lack of usable nitrogen and oxygen within the ecosystem following the pulse disturbance. Although we saw an increase in nitrogen to the bay via riverine input, this nitrogen is most likely input in the unusable

form of nitrate (NO_3^-) or dissolved organic nitrogen (DON) (Voss et al 2013). Genes associated with sulfur oxidation increased during Samp2 and Samp3 (Figure 11C). Since sulfur oxidizers are commonly found within marine sediment (Baker et al 2015), these findings further support that resuspension of marine sediment in the water column following a pulse disturbance.

The incidence of genes associated with nitrate and nitrite ammonification and inorganic sulfur assimilation increased, not following the overall trends of sulfur and nitrogen metabolisms (Figure 11C). Riverine input of nitrogen into the oceanic ecosystem is in the form of nitrate ions (NO_3^-) and dissolved organic nitrogen (Voss et al 2013); however, this is not a usable form of nitrogen in the oceanic ecosystem. Nitrate must undergo nitrate ammonification, via microorganisms, in order to become the usable form of ammonium (NH_4^+) (Voss et al 2013). Due to the increase of nitrate in the ecosystem due to riverine input there was an increase in nitrate ammonification genes to get usable forms of nitrogen (Figure 11C). After the pulse disturbance, we also saw an increase in genes associated with inorganic sulfur assimilation. This increase can be explained by the increase of sulfur in the system that needed to be assimilated by microorganisms. Sulfur is input into the ocean via weathering and leaching of rocks as sulfate (SO_4^{2-}), which needs to be assimilated by microorganisms (Sievert et al 2007).

As seen in this study, the microbial community was impacted at the taxonomic and genomic level by this pulse disturbance. When comparing to the pre-Harvey samples, we find that the community had not fully recovered by our last sampling time point (Sep. 28th). Recovery at the taxonomic and genomic level in Galveston Bay, from a pulse

disturbance of this intensity it is hypothesized to take longer than a month after the event. Further research should be done to determine recovery lengths in other coastal bays and on metabolic changes as these could lead to long term imbalances in the biogeochemical cycles in the coastal ocean. Although the community had not fully recovered by the last sampling point, we did see a resemblance of the community return by Samp2 and Samp3. That recovery consisted of a gradual return of common marine bacteria returning such as Cyanobacteria and Acidimicrobiia. These common marine bacteria decreased in abundance after the pulse disturbance, but the communities were resilient and recovery started as soon as by Samp2, only two weeks after the pulse disturbance.

3.6 Conclusions

Utilizing Hurricane Harvey as a pulse disturbance, we can evaluate how other coastal bays similar in structure to Galveston Bay (i.e. riverine input, hurricane prone, coastal location), could be impacted from large rain events. Bacteria commonly found in the marine ecosystem, such as Cyanobacteria and Acidimicrobiia, were replaced with common terrestrial bacteria, such as Actinobacteria (class) and Verrucomicrobiota (Figure 8B). We also observed the genomic potential of the microbial community shift from a photosynthetic community to a heterotrophic community (Figure 11). With rainfall rates and intensities of tropical cyclones predicted to increase in the coming years (Knutson et al 2015), it is important to understand how microorganisms, which play viral roles in oceanic nutrient cycling, are impacted following pulse disturbances and further understand how the microbial ecosystem recovers.

CHAPTER IV

HURRICANE HARVEY IMPACTS TO THE VIRAL COMMUNITY

4.1 Introduction

With an estimated ~$4x10^{30}$ viruses, viruses are the dominant biological entity within the ocean (Suttle 2005). Viruses play many important roles in the ecosystem, including top-down control, nutrient recycling, aiding in host metabolisms and increasing fitness of the host in unfavorable conditions (Tsiola et al 2020, Warwick-Dugdale et al 2019). While it has been well documented how the host community responds to pulse disturbances such as hurricanes (Amaral-Zettler et al 2008, Mason et al 2012, Steichen et al 2020) and oil spills (Mason et al 2012), there has been little work done on the impact of disturbances on the viral communities.

Viruses play key roles as agents of mortality as ~20% of the prokaryotic population in the ocean is lysed daily by viruses (Fuhrman 1999, Suttle 2007). The effects of viral lysis can range from altering chemical cycles in the ocean via the viral shunt (Suttle 2007, Wilhelm and Suttle 1999). The viral shunt refers to the redirection of nutrients from higher trophic levels back to the microbial food web through viruses infecting and lysing microbial cells, which releases the host cells contents back into the total dissolved organic matter (DOM) pool (Suttle 2007). Without the viral shunt, the host cells would be consumed by secondary consumers and the nutrients would continue up the food chain (Wilhelm and Suttle 1999).

Viruses can alter their host metabolisms, in ways such as altering phosphate uptake (*phoH*) and stress (Crummett et al 2016, Goldsmith et al 2011), alteration of host carbon fluxes (Hurwitz et al 2013), to parts of the photosynthetic process (Crummett et al 2016, Mann et al 2003, Thompson et al 2011) and through auxiliary metabolic genes (AMGs) (Breitbart et al 2007, Thompson et al 2011). During infection, phages use AMGs to redirect the host energy and resources to support viral production (Breitbart et al 2007, Thompson et al 2011). AMGs play a significant role in marine ecosystem functioning and the function of the AMG can be transferred to the host and integrated into the host genome via viruses (Tsiola et al 2020). The transfer of AMGs from virus to host allow the host to reproduce in unfavorable conditions (Tsiola et al 2020), such as the conditions a pulse disturbances would bring.

Very few studies have investigated the impact of abiotic conditions in aquatic environments. Williamson et al (2014) evaluated the impact of stormwater runoff on viral community composition in a freshwater system and found Hurricane Sandy, a pulse disturbance and large rain event, had negative impacts on viral species richness and viral abundances, due to increased rainfall and bacterial abundance. The impacts of pulse disturbances on marine viral communities has yet to be evaluated.

Hurricane Harvey made landfall in Port Aransas on the 25th of August 2017. While Galveston did not experience hurricane-force winds, there was a record-breaking 1.4-1.7 x 10^{10} m^3 of rainfall that that lashed on the area and flushed into Galveston Bay. Hurricane Harvey not only brought rainfall accumulation, but also drops in salinity, temperature and dissolved oxygen and 9.86 x 10^7 metric tons of sediment into Galveston

Bay (Figure 4) (Du et al 2019). Utilizing Hurricane Harvey as a pulse disturbance and large rain event, we aimed to understand the impact large scale rain events have on coastal bay marine ecosystems.

We characterized the changes Hurricane Harvey brought to the viral community at the taxonomic level and determined the potential role AMGs played in ecosystem recovery. To accomplish this, we sequenced host and viral metagenomes from a 4-station transect in Galveston Bay for five weeks following Hurricane Harvey. To discover the impact of Hurricane Harvey on the viral diversity and determine the role viruses play in ecosystem recovery, we hypothesized that 1) the viral community would change following their host community and aid in population control of the microbial community in Galveston Bay; and 2) there would also be an increase in AMGs associated with nitrogen and sulfur metabolisms within the viral community due to the influx of host metabolisms associated with these cycles.

4.2 Methods

Note: Methods in sections 4.2.1 and 4.2.2 were performed by members of the Labonté Viral Ecology Lab prior to my arrival at Texas A&M University at Galveston.

4.2.1 Sample collection and processing

Samples were collected from four stations in Galveston Bay, after Hurricane Harvey's landfall, on September 4th (Samp1), 9th (Samp2), 16th (Samp3), and 28th (Samp5) of 2017. A transect was sampled from the mouth of the San Jacinto River to the Gulf of Mexico (Figure 5 & Table 1). Pre-Harvey samples were collected from Texas A&M University at Galveston's boat basin on July 31st and August 22nd, of 2017 prior to

Hurricane Harvey's land fall. The volume of samples collected ranged from 4 to 20 L, depending on time constraints, available material and personnel (Table 7 in Appendix A).

All samples were pre-filtered immediately after sampling with a Nitex filter (30 µm) to remove small grazers and large particles. The total volume of each sample varied depending on time and manpower constraints for each sampling day. After filtration the samples were stored on the boat in the dark and brought to the laboratory for further filtration. Generally, each sample was filtered through a glass fiber filter (GF-F with a 0.7 µm pore-size or GF/D with a 2.7 µm pore-size), followed by a 0.22 µm pore-size polyvinylidene fluoride (PVDF) filter. Due to the availability of supplies, for the sampling of 09/09, prefiltration was performed with a 0.45 µm and the virus concentrate was the filtered through 0.22 µm PVDF filter. All GF and PVDF filters were stored at -20°C until further use (Table 7 in Appendix A). Viruses, which are in the remaining aqueous fraction, were concentrated using tangential flow filtration with a 30 kDa cut-off as previously described (Suttle et al 1991).

3.2.2 DNA extractions and sequencing

Viral DNA was extracted from the virus concentrates corresponding to an initial volume of 4 L estimated from the concentrated final volume and concentration factor (Table 7 in Appendix A). Before extraction, virus concentrates were further concentrated to a final volume of ~ 500 µl using 30 kDa Amicon Centrifugal Filters (Millipore). Viral DNA was extracted using the Qiagen DNeasy PowerSoil DNA extraction kit (Qiagen). Because the yield was too low for metagenomic sequencing, the DNA was amplified using the Repli-g DNA Mini kit according to the manufacturer's recommendations,

which provided enough material for metagenomic sequencing. DNA extracts were quantified with QuBit and stored at -20˚C until further use. Between 1–10 µg of DNA were sent for sequencing using Illumina HiSeq 4000 with 150 bp paired-end sequencing technologies at the Texas A&M Genomics and Bioinformatics facility (College Station, Texas, United States).

4.2.1 Metagenomic data analysis

BBtools software suite (https://sourceforge.net/projects/bbtools/) was used for quality control of the metagenomic raw reads. The commands for BBtools were built following a tutorial from the USDA ARS Microbiome Workshop in 2017 titled Metagenomics tutorial part 1: Quality control, assembly and mapping (https://usda-ars-gbru.github.io/Microbiome- workshop/tutorials/metagenomics/). BBduk (version 38.31) was used to filter out contaminated reads, including kmers of 31 bp in length and dropping the extra base that is occasionally added by Illumina sequencing. BBduk was also used for the Phred algorithm for more accurate adapter trimming and removal of kmers between 11 and 23 bp in length- a common length for adapter artifacts. BBmerge (version 38.31) was used to merge the forward and reverse paired-end reads using default settings (Bushnell et al 2017). BBmask (version 38.31) was used to soft-mask human, cat, dog and mouse contamination. MEGAHIT (version 1.2.8) was then used for the *de novo* assembly into contigs (Li et al 2015). Gene prediction was done with Prodigal (version 2.6.3) (Hyatt et al 2010). Genes and proteins were annotated with DIAMOND (version 0.9.26), utilizing the fast setting and NCBI protein non-redundant database (March 19th, 2019) (Buchfink et al 2015). All metagenomes were uploaded into

the MG-RAST metagenomics analysis server (Meyer et al 2008) for public access (supplementary material Table 1 for accession numbers). Community analysis, diversity measures and figures were done in MEGAN (Huson et al 2007) and the R package vegan (Figure 12) (Oksanen et al 2019).

Binning of metagenomes

Taking the BBmasked, merged and unmerged files, a large co-assembly of all the samples was done using MEGAHIT (version 1.2.8) with a 2.5 kb limit flag (Li et al 2015). Annotations and taxonomic classification of contigs was done using Prokka (version 1.14.5) and Kraken 2 (version 2.0.8) and the NCBI protein non-redundant database (March 19th, 2019) (Seemann 2014, Wood et al 2019). Sequence coverage of all contigs was performed using BWA (version 0.7.17) to map the reads using the BWA-MEM setting (Li and Durbin 2010, Li 2013) followed by Samtools (version 1.10) to convert SAM files to BAM files and filter out unmapped reads (Ramirez-Gonzalez et al 2012). MetaBAT (version 2.12.1) was used to calculate contig depths and MetaBAT-2 (version 2.15) was used to generate species bins (Kang et al 2015, Kang et al 2019). Finally, CheckM (version 1.1.3) lineage workflow was used to assess bin quality (Figure 7) (Parks et al 2015).

AMG identification and matching to host bins to viral contigs

VIBRANT (version 1.2.0) was used to identify viral scaffolds within the host metagenomes and annotate all viral contigs in host and viral metagenomes utilizing the -virome argument and find AMGs (Kieft et al 2020). VirHostMatcher (version 1.0) was used to link viral scaffolds from VIBRANT, greater than 5,000 bp in length, to the host

species bins (Ahlgren et al 2017). Viral scaffolds were identified using VirSorter (version 1.0.3), then linked to the most likely host bin in VirHostMatcher (Ahlgren et al 2017, Roux et al 2015). Once the host bin was identified all depths of that host bin family were added together to get accurate family depth counts across the samples. The host bin summed depths were then plotted against the viral scaffold depths. All depths were taken from the MetaBAT output (version 2.12.1) (Kang et al 2015).

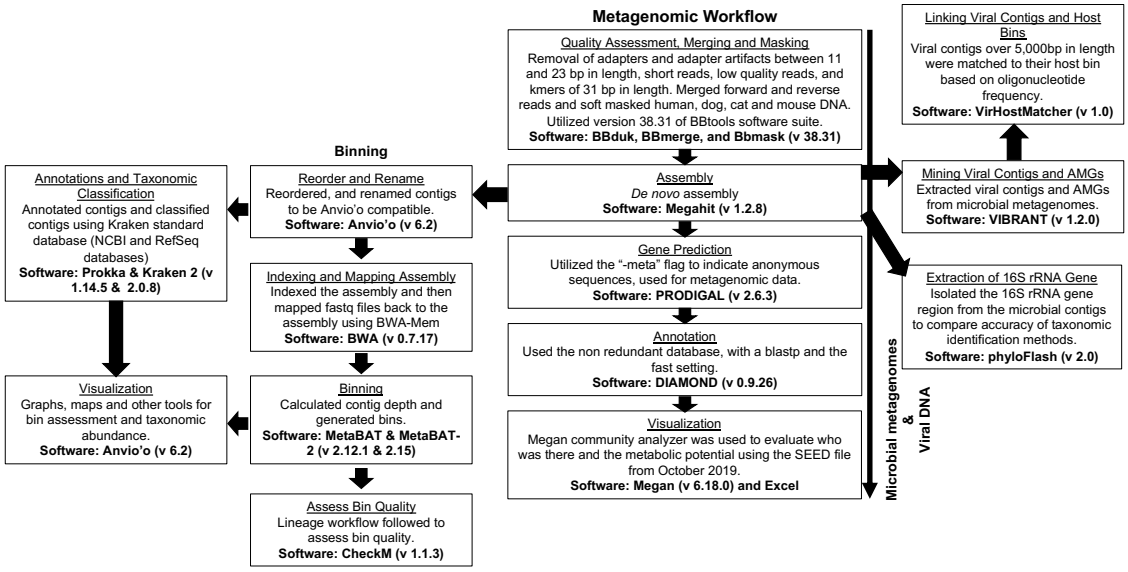

Figure 12 Bioinformatic pipeline for microbial metagenomes and viral DNA.

4.2.2 Viral production experiments

In order to estimate viral production, we used the dilution technique of Matteson et al (2010). Briefly, 300-1000 mL of sample were filtered through a 0.2 µm pore size Sterivex filter to retain the prokaryotes, which were then resuspend in virus-free water (water filtered using a 30 kDa cutoff to remove all viral particles) from the same sample as the initial sample volume. Triplicates of 100 mL were incubated at *in situ* temperature

in the dark. Two-milliliter subsamples were collected at 0, 2.5, 5, 7.5 and 12 h and fixed with 0.2 μm filtered paraformaldehyde and stored at -80ºC until further use.

Epifluorescence microscopy slides were prepared with SYBR Green I as described by Patel et al (2007). Briefly, samples were subsequentially filtered through 0.2 and 0.02 μm Anodisc filters to count prokaryotes and viruses, respectively. One mL of the sample was filtered with 1 mL of filtered, autoclaved milliQ. The Anodisc filters were dried for four minutes, stained with 1:400 solution of SYBR Green I for 15-20 minutes and dried for an additional four minutes. Each filter was then placed on a slide and preserved with Vectashield (Vector Laboratories). Slides were then stored at -20ºC until they were counted with a Leica DM2000 LED Microscope with Leica DFC295 digital camera. Twenty fields of view were counted and averaged per slide. From these averages bacteria and virus concentrations were calculated using the following equation:

$$\frac{RSF \times X \times 100}{V}$$

Where RSF is the grid reticle scaling factor (calculated for our microscope viewing area), X is the average number of viruses per field of view and V is the volume of sea water sampled.

Viral production was determined from the production of new viral particles after the dilution of the initial viral abundance. Viral production rates were calculated from first-order regressions of viral abundance versus time after correcting for the loss of the bacterial hosts between the experimental samples and the natural seawater community.

4.3 Results

4.3.1 Hurricane Harvey impacted viral community diversity

The viral and host metagenomic dataset had four sampling points, at four sampling stations with two pre-Harvey samples (Table 8 in Appendix A). Sampling efforts resulted in 21,833,421 contigs, with 15,339,749,905 bp assembled and 31,690,471 proteins recovered in the metagenomes (Table 8 in Appendix A).

We evaluated the viral diversity from each metagenome using a Shannon index diversity measure. The Samp1 sampling had the lowest of all the samplings in richness and diversity (Table 3). Diversity of the viral community peaked during the Samp2 sampling at Stn1 and Stn4, before decreasing again (Table 3). The richness and evenness of the viral community was the inverse of the microbial community (Table 3). While the host diversity peaked during Samp1, the viral diversity peaked during Samp2, when host diversity was already returning to pre-Hurricane Harvey conditions (Table 3).

Investigating viral community taxonomic composition at the family level revealed drastic changes from pre-Hurricane Harvey samples (pre-Harvey1 and pre-Harvey2) to the first sampling time point. The family *Siphoviridae* was the most abundant viral family prior to Hurricane Harvey's landfall, with a relative abundance of ~25.2% of the total viral community. Relative abundances of *Siphoviridae* fell to an average relative abundance of 0.80% in Samp1 (Figure 13). Relative abundances of the family *Microviridae* increased almost three-fold from pre-Hurricane Harvey levels. *Microviridae* relative abundances increased from an average of 21.23% of the total viral community pre-Hurricane Harvey, to an average relative abundance of 60.43% in

Samp1. Relative abundances of *Microviridae* fell to below pre-Hurricane Harvey conditions in Samp5, averaging 14.33% of the total viral community (Figure 13). However, the abundances of *Microviridae* may not be accurate due to the small nature of the genomes (small circular genomes < 5.5 kb (Roux et al 2012)) and amplification of viral DNA using Repli-g, which uses the phi29 DNA polymerase that uses the rolling-circle mechanism to get a high enough concentration of DNA for sequencing.

Following Hurricane Harvey's landfall, the viral families *Podoviridae* and *Myoviridae* almost disappeared from Galveston Bay. These two families returned and then exceeded, pre-Hurricane Harvey levels by Samp5. The viral families *Podoviridae* and *Myoviridae* were replaced in the community by high levels of the viral family *Microviridae* (Figure 13). There was also an influx of freshwater invertebrate viruses that were all grouped into the category Lake Sarah Associated Viruses, a group of freshwater viruses assembled from metagenomes recovered from a New Zealand lake, Lake Sarah (Dayaram et al 2016). These freshwater viruses were not present in the viral community prior to Hurricane Harvey; however, at Samp1 this group comprised 3.12% of the total viral community. Relative abundances of freshwater viruses had almost returned to pre-Hurricane Harvey levels by Samp5, averaging 0.11% of the total viral community (Figure 13).

Like the host communities (Figure 9), viral communities are impacted by abiotic environmental variables. There is a clear difference in groupings in samples taken prior to Hurricane Harvey's landfall and subsequent sampling dates (Figure 14). All stations are grouping together by sampling date, for example all four Samp1 stations group

together (Figure 14). Salinity, Secchi depths, total phosphorus, total nitrogen and pH had significant impacts on the community ($p<0.05$) (Figure 14). Salinity and pH both decreased after Hurricane Harvey's landfall, while total suspended sediment (TSS) was higher during Samp1 when compared to Samp5 (Steichen et al 2020).

Table 3 Shannon diversity indices from the viral (green) and microbial (blue) metagenomes at the family and class level, respectively. The darker the color the higher the richness and evenness of the community.

Dates in 2017	Station	Shannon Index- Viral Comm.	Shannon Index- Microbial Comm.
	Stn1	1.590304467	NA
	Stn4	1.422408073	3.148
	Stn7	1.436567098	3.121
Sep. 4th	Stn10	1.538954139	2.954
	Stn1	2.026885971	3.081
	Stn4	2.047100514	3.064
	Stn7	1.640636886	3.013
Sep. 9th	Stn10	1.786637221	2.773
	Stn1	1.942369194	2.889
	Stn4	1.981107394	2.702
	Stn7	1.995486205	2.608
Seo. 16th	Stn10	1.927588721	2.718
	Stn1	1.94158621	2.883
	Stn4	1.799051695	2.643
	Stn7	1.685317539	2.837
Sep. 28th	Stn10	1.617820151	2.712
July 31st	Pre-Harvey1	1.828214349	2.707
Aug. 22nd	Pre-Harvey2	1.939305548	2.656

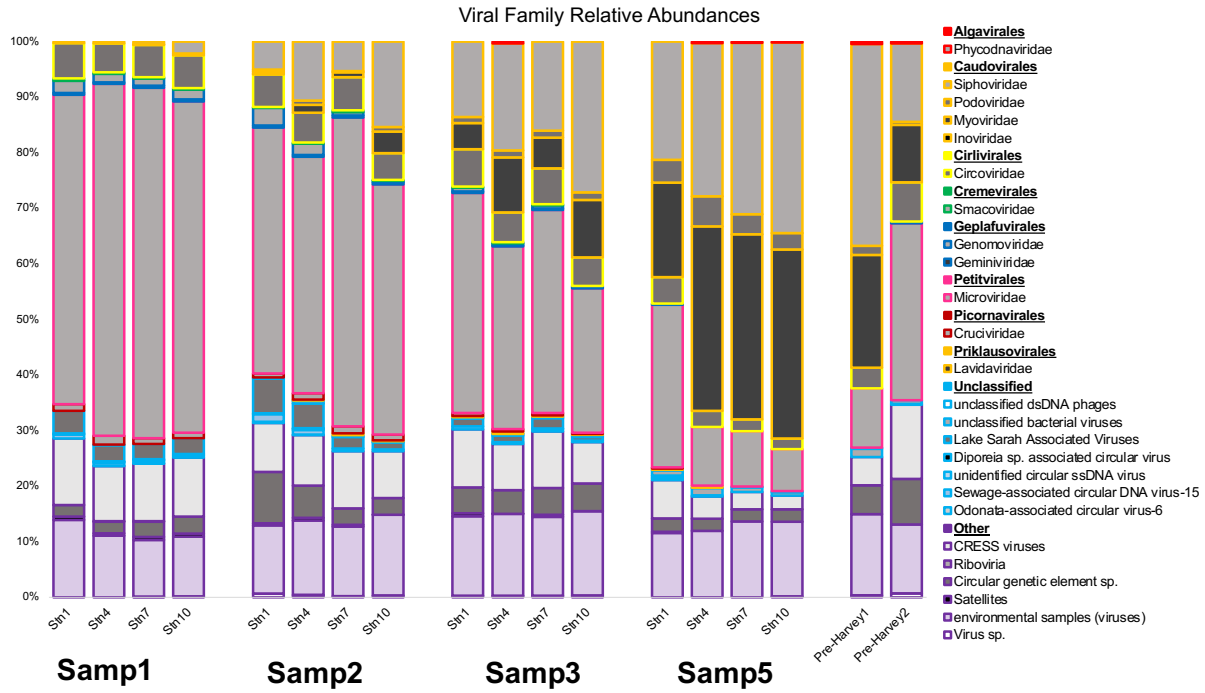

Figure 13 Stacked bar histogram showing the relative abundance of viral families identified by metagenomic viral DNA samples. Sampling points are broken up into blocks separated by white space. Pre-Harvey1 and Pre-Harvey2 are samples taken before Hurricane Harvey on 7/31/17 and 8/22/17.

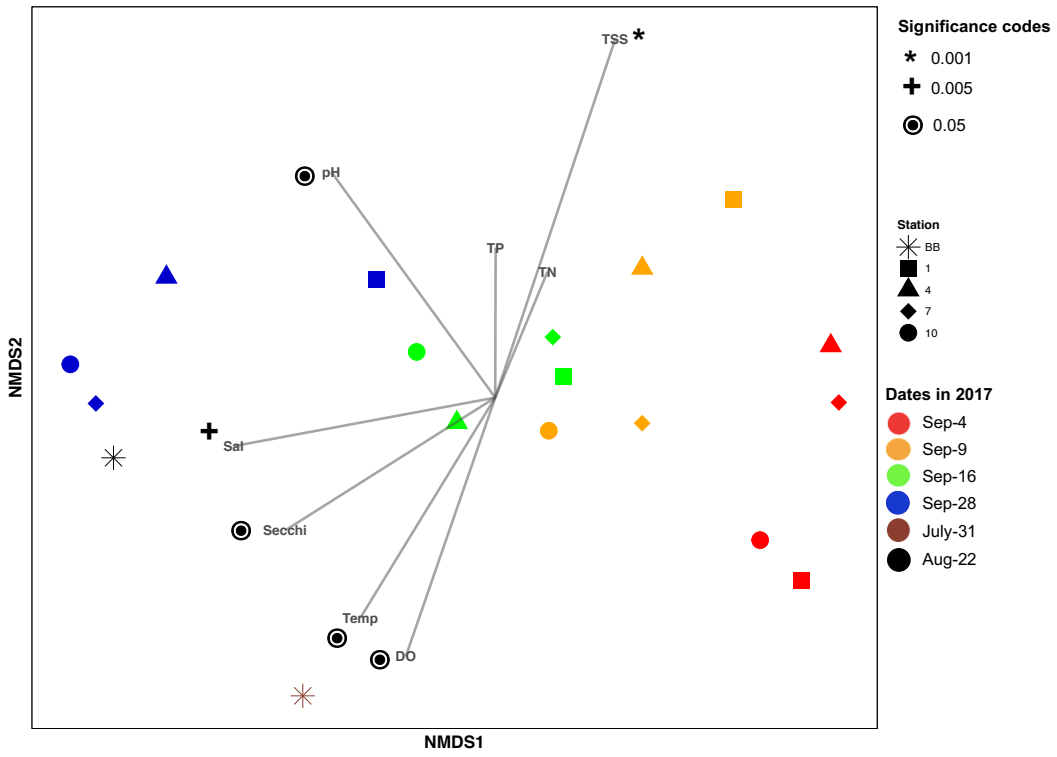

Figure 14 NMDS of the viral community composition showing the correlation between temperature (Temp; ºC), salinity (Sal), pH, dissolved oxygen (DO;%), total nitrogen (TN; µM[1]), total phosphorus (TP; µM), Secchi (m) and total suspended sediment (TSS; g mL^{-1}). All data for environmental variables were taken from (Steichen et al 2020). A "*" represents a significance level of p=0.01. Colors represent sampling time points in 2017 and symbols represent sampling stations. NMDS was constructed using viral metagenomic abundances at the family level using a square root transformation and a Bray-Curtis similarity matrix, with a stress level of 0.0510.

4.3.2 Trends in viral and host abundances following Hurricane Harvey

When host abundances increased, viral abundances decreased, therefore viruses are reversed from host abundances in the environment (Figure 17). We calculated viral and bacterial cell production using viral production experiments on samples Samp2, Samp3 and Samp5, along with tracking a select few complete circular viral genomes mined from the microbial metagenomes and linking them to their host bins and tracking

abundances through our sampling time points. Viral production was an order of magnitude lower in Samp2 compared to Samp3 and Samp5 (Figure 15). The low levels of viral production in Samp2 was due to a significant decrease in viral abundance in the system, limiting the encounter rate (Figure 15, Table 4). The most representative results of viral production can be seen in Samp3 at Stn10, where bacterial cell abundances decreased due to viral lysis and viral particle abundances peaked due to viral release following lysis (Figure 15). The highest rate of viral production almost always occurred between 2.5 and 5 hours of incubation (Figure 15).

Linking viruses to their microbial hosts can show changes in the marine community as well as impact of microbial relationships and the environment. In one sample, we had 300 separate viral contigs and 995 separate host bins (Figure 16). There are clear groupings (yellow blocks) of viruses that are matching to microbial host bins, but no real distinct patterns are emerging at the family or higher taxonomic levels (Figure 16). We selected three viral contigs and tracked the contigs following their most closely related host and the host with the best score was selected (Table 5). Viruses selected needed to be above 2,000 bp in length to have high confidence in the matching to the host. There were no viruses over 2,000 bp in length in Samp1 or Samp2 as a result of the freshwater load that flushed viruses from the system (Table 4), therefore the selected viruses are from Pre-Harvey samples, Samp3 and Samp5. All viruses that were selected followed reverse trends of the host population, when the host abundances peaked, viral abundances dipped (Figure 17). Relative abundances of viruses selected stayed below host abundances (Figure 17). Viruses following reverse trends of the host

population can also be seen during the viral production experiments. When the host population dipped, the viral population peaked (Figure 15).

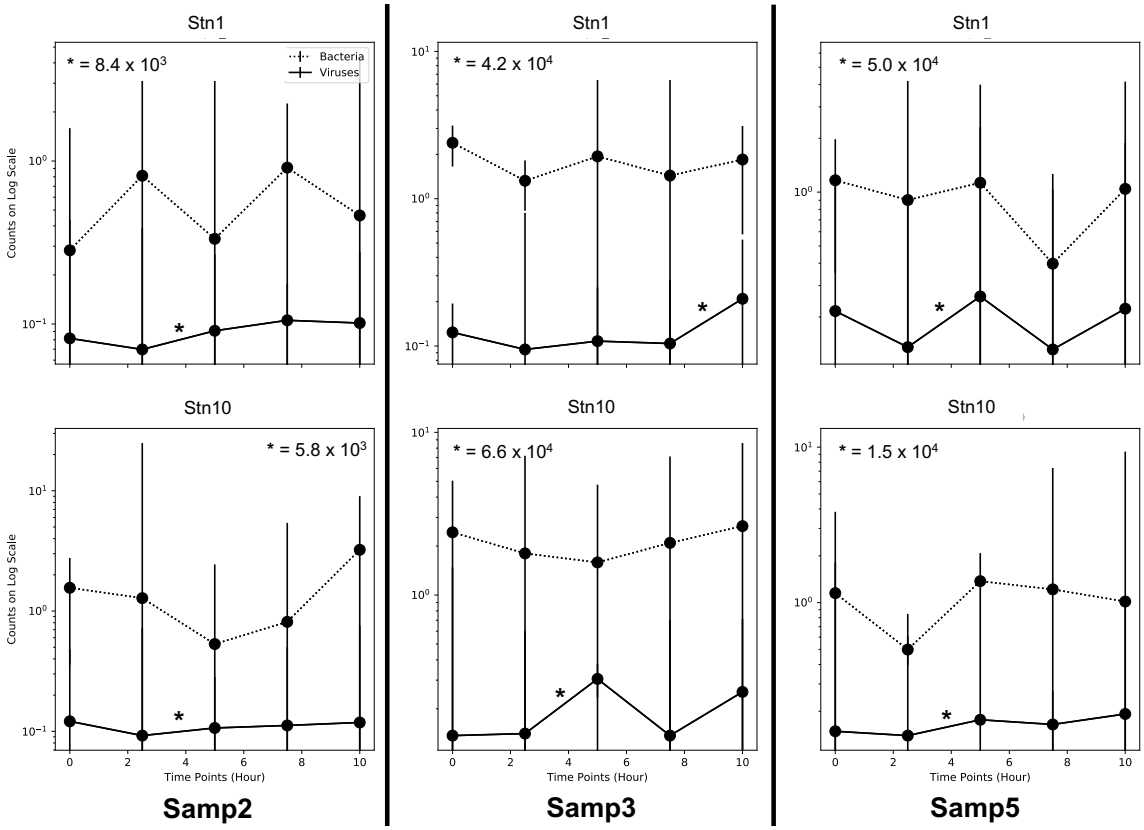

Figure 15 Viral production experiments for three samples at two stations following Hurricane Harvey at five separate hour points (0, 2.5, 5, 7.5 and 10). These counts are on a log scale with standard deviation bars. Stars represent the highest level of viral production and are placed at the time point the highest level of viral production occurred.

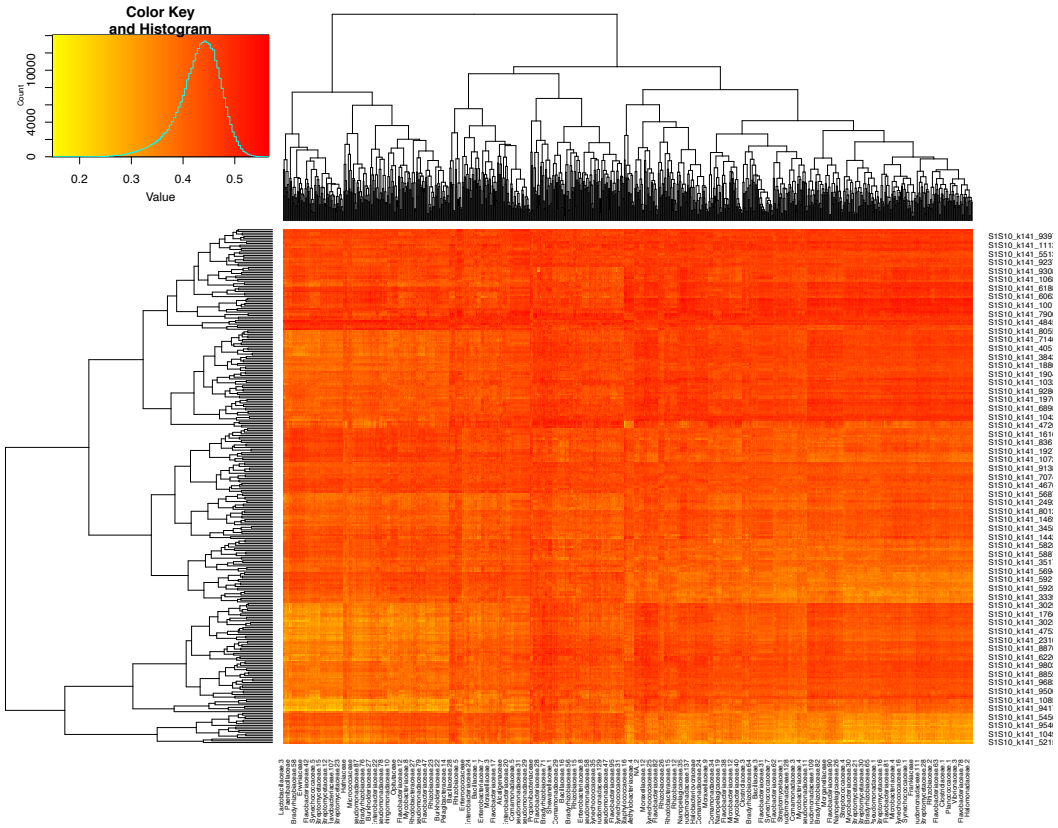

Figure 16 Heatmap showing groupings of most likely hosts of different viral contigs. This is a sampling from Samp1_Stn10 to showcase the number of different microbial bins and viral scaffolds we are trying to link to each other. The lower the value (yellow) the more likely that viral scaffold infects the respective microbial bin.

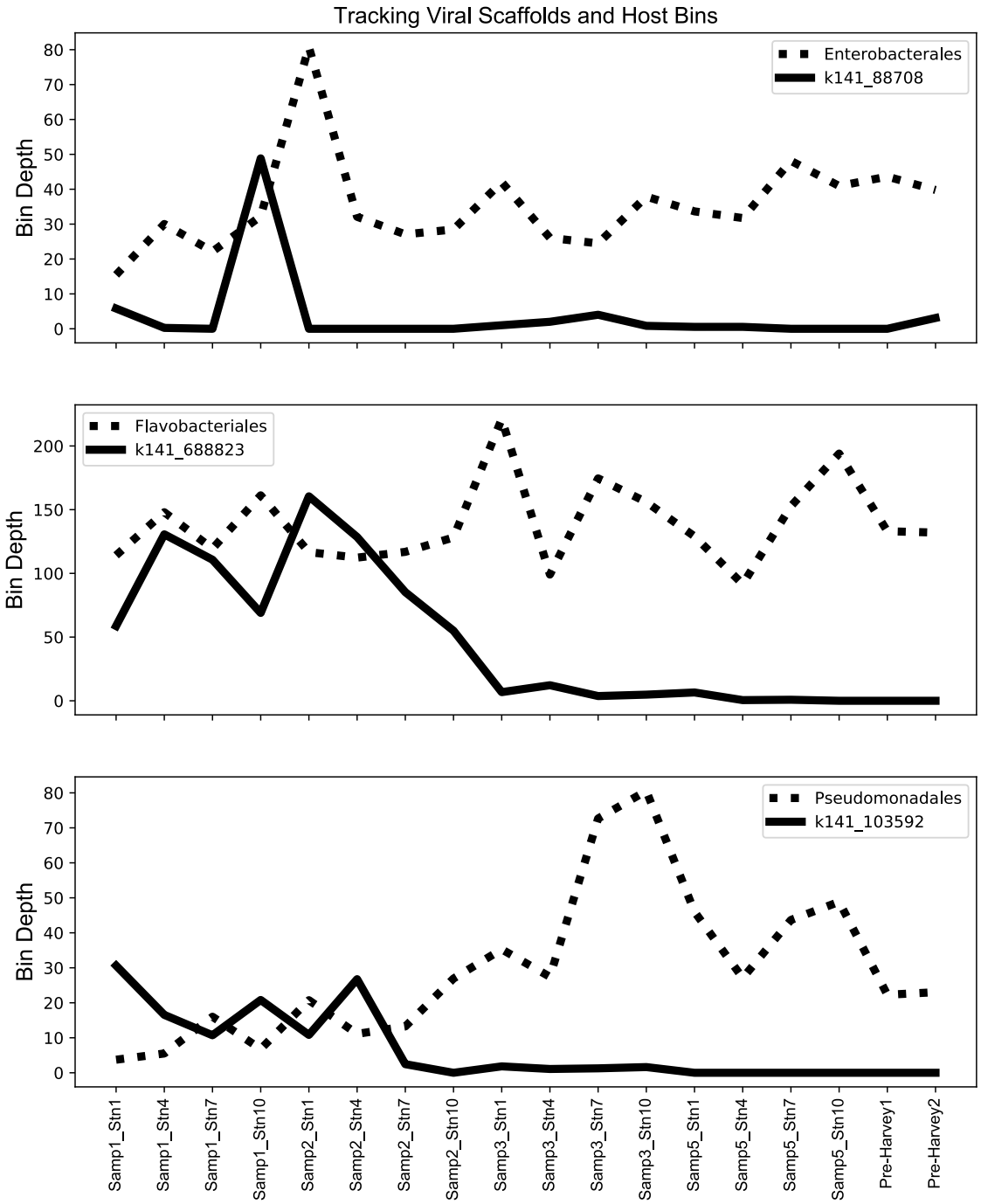

Figure 17 Line graph showing the abundance of three separate marine viruses and the predicted host bins. The viral scaffold was identified off the lowest E-value obtained from VirSorter and was matched to the host bin (bacterial) with the lowest score obtained from VirHostMatcher. Pre-Harvey1 and Pre-Harvey2 are samples taken before Hurricane Harvey on 7/31/17 and 8/22/17.

Table 4 Total number of lytic viruses and prophages, normalized to 10,000 contigs, recovered from microbial metagenomes.

Dates in 2017	Station	Lytic viruses	Prophages
	Stn1	28.7	0.0
	Stn4	44.6	0.0
	Stn7	69.2	0.5
Sep. 4th	Stn10	47.8	3.0
	Stn1	58.6	0.6
	Stn4	61.3	1.9
	Stn7	62.0	1.6
Sep. 9th	Stn10	37.5	1.6
	Stn1	147.1	1.1
	Stn4	152.7	5.0
	Stn7	116.8	2.6
Sep. 16th	Stn10	199.6	0.6
	Stn1	184.8	2.2
	Stn4	165.3	4.6
	Stn7	199.5	0.0
Sep. 28th	Stn10	144.0	0.2
July 31st	Pre-Harvey1	39.5	0.0
August 22nd	Pre-Harvey2	28.5	0.0

Table 5 Summary of complete viral genomes that were linked and tracked with the host bin.

Viral Contig	Closest relative	E-value	Host Bin Best Match (Order)	Score ($d2^*$ w/ kmer size of 6)
k141_88708	Burkholderia phage KS9	$6.5x10^{-125}$	Proteobacteria; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae	0.408073
k141_688823	Nonlabens phage	$9.5x10^{-91}$	Bacteroidetes; Falvobacteriia; Flavobacteriales; Flavobacteriaceae	0.28062
k141_103592	Synechococcus phage	$0x10^{0}$	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae	0.257876

4.3.3 Abundances of viral AMGs following Hurricane Harvey

AMGs associated with metabolism of cofactors and vitamins, amino acid metabolism and carbohydrate metabolisms were the most abundant AMGs recovered in the genomes (Figure 18 A & C). Further breaking down AMGs associated with energy metabolism reveals trends that support those of the host metabolisms. AMGs associated with energy metabolisms, such as photosynthesis, dropped after Hurricane Harvey's landfall, while sulfur and nitrogen metabolism AMGs increased (Figure 18 A). Looking at a breakdown of energy metabolism, prior to Hurricane Harvey photosynthetic AMGs dominated the Galveston Bay ecosystem (Figure 18 B & D). After Hurricane Harvey's landfall, we saw a sharp increase in AMGs associated with sulfur metabolism (Figure 18 B & D). After Hurricane Harvey, there was also an increase in AMGs associated with methane and nitrogen metabolisms, although these trends were only seen in two stations during Smap1 (Figure 18D). The viral DNA was almost free from bacterial contamination (Table 3).

Data from the vDNA (Viral DNA) dataset was shown to not be contaminated by bacterial DNA. PhyloFlash was used to get percentages of 16S rRNA genes that mapped to viral metagenomes (Gruber-Vodicka et al 2019) (Table 6). For comparison, in the microbial metagenomes there was approximately a tenth of reads mapped more per metagenome (~0.05% mapped per genome) (Table 6).

Table 6 Percent of reads mapped, from viral metagenomes and microbial metagenomes, to the 16S rRNA gene.

	16S rRNA gene reads mapped (%)		
Sample ID	Viral DNA	Microbial DNA	
		0.22 Filtration	GFF Filtration
Samp1_Stn1	0.005	0.016	0.022
Samp1_Stn4	0.002	0.016	0.022
Samp1_Stn7	0.001	0.017	0.020
Samp1_Stn10	0.005	0.016	0.021
Samp2_Stn1	0.000	0.021	0.021
Samp2_Stn4	0.000	0.021	0.022
Samp2_Stn7	0.003	0.019	0.018
Samp2_Stn10	0.003	0.024	0.022
Samp3_Stn1	0.002	0.017	0.017
Samp3_Stn4	0.002	0.016	0.017
Samp3_Stn7	0.001	0.023	0.019
Samp3_Stn10	0.001	0.024	0.018
Samp5_Stn1	0.002	0.022	0.019
Samp5_Stn4	0.002	0.020	0.021
Samp5_Stn7	0.002	0.023	0.019
Samp5_Stn10	0.001	0.025	0.019
Pre-Harvey1	0.002	0.022	0.019
Pre-Harvey2	0.001	0.017	0.017

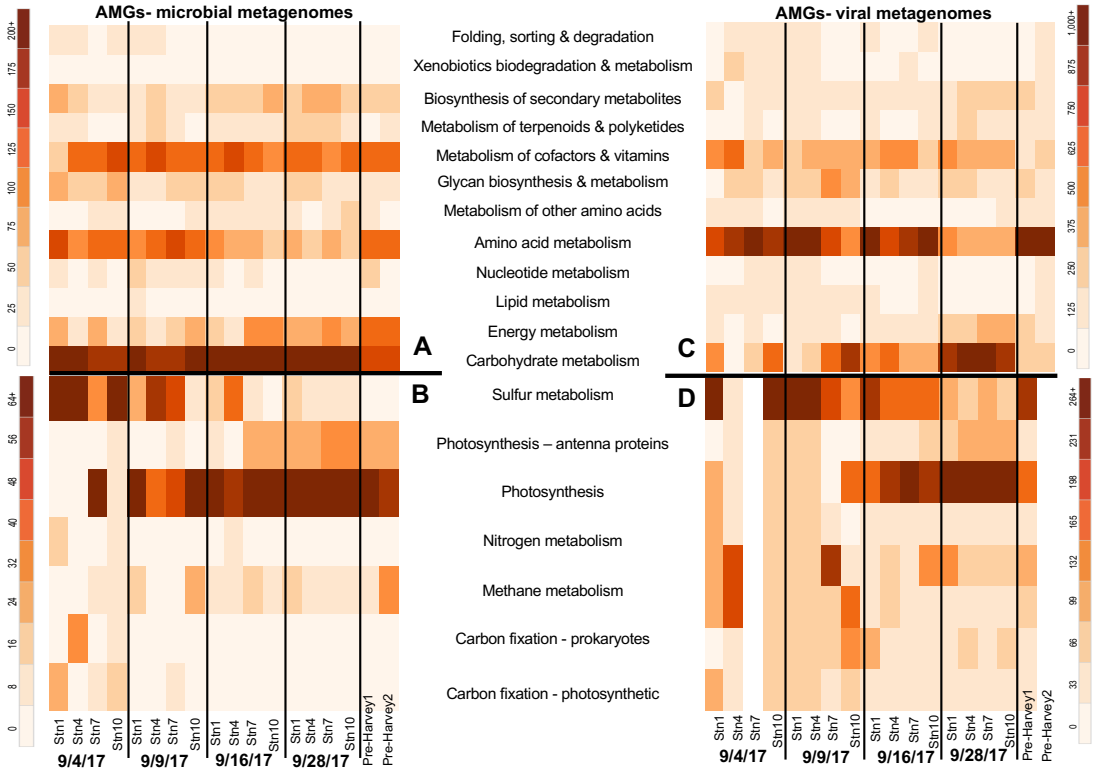

Figure 18 A four part heat map showing the density of auxiliary metabolic genes (AMGs) recovered from microbial metagenomes (A & B) and viral metagenomes (C & D). B and D show a breakdown of AMGs associated with energy metabolisms in microbial and viral metagenomes, respectively. In the viral metagenomes on 9/4/17 Stn7 and Pre-Harvey2, no AMGs were found within the dataset. Pre-Harvey1 and Pre-Harvey2 are samples taken before Hurricane Harvey on 7/31/17 and 8/22/17.

4.4 Discussion

Models are predicting that rainfall rates associated with tropical cyclones will increase by 10-15% in the coming years and these cyclones will intensify (Knutson et al 2015). It is important to understand how viruses, which recycle as much as one quarter of the organic carbon in the ocean (Wilhelm and Suttle 1999), are impacted from extreme rainfall events and how viruses aid in marine ecosystem recovery. To the best of our knowledge, only one other study evaluated impacts of heavy rainfall on viral communities and focused on evaluating freshwater viral community composition (Williamson et al 2014). Here, we provide one of the first looks at heavy rainfall impacts on marine viral communities in a coastal bay. Utilizing Hurricane Harvey as a pulse disturbance, we can evaluate how other coastal bays similar in structure to Galveston Bay (i.e., riverine input, hurricane prone, coastal location) could be impacted from large rain events. Hurricane Harvey removed common marine viruses of the *Caudovirales* order, such as *Podoviridae*, *Siphoviridae* and *Myoviridae*, while ssDNA viruses, especially *Microviridae* and freshwater viruses, such as Lake Sarah Associated Viruses, increased in abundance.

The diversity of viruses in the ecosystem decreased after the rain event. There was a decrease in viral community diversity, compared to the host community which increased in diversity during Samp1, as the viral diversity peaked during Samp3 and Samp5, while host diversity peaked during Samp1. The drop in viral diversity after Hurricane Harvey could be attributed to the dilution effect, as the heavy rainfall flushed viruses out of the system (Table 3).

The decrease in diversity could be also be due to two common marine viral families being removed in the marine ecosystem, *Podoviridae* and *Myoviridae*. Cyanophages (viruses that infect Cyanobacteria) are most commonly in the *Podoviridae* or *Myoviridae* families (Ignacio-Espinoza and Sullivan 2012, Sullivan et al 2005). Besides Cyanobacteria, *Podoviridae* are also known to infect common marine bacteria such as SAR11 (*Pelagibacterales*) in the class Alphaproteobacteria (Zhao et al 2019).

We did see a large increase in the *Microviridae* family. While the presence of Microviridae is true, their abundance may be biased by our methodology. Indeed, it is possible that members of the *Microviridae* family were over amplified due to having to use whole genome amplification to get enough material for DNA sequencing. It is possible that the rolling-circle DNA polymerase used to amplify the DNA over amplified *Microviridae* viruses, due to nature of genomes of these viruses, which have very small ssDNA circular genomes (< 5 kb). The only genus of *Microviridae* recovered were unclassified *Gokushoviruses.* It has been well document that *Gokushoviruses* are widely distributed throughout the marine environment (Labonté and Suttle 2013, Tucker et al 2011). Due to the widespread nature of *Gokushoviruses* in the marine environment, it is difficult to attribute the increase in *Gokushoviruses* after Hurricane Harvey to the storm. The rationale for this is all other common marine bacteria and viruses were flushed out of the marine ecosystem following the rain event. Therefore the leading hypothesis explaining such a dominance of *Microviridae* viruses is that while present, they were overamplified.

Viruses follow trends of their hosts in the marine environment. Utilizing viral production experiments, we found that viral production rates were an order of magnitude lower in Samp2, when compared to Samp3 and Samp5 (Figure 15). The low level of viral production found in Samp2 was most likely due to the fact that there were few viruses left in the ecosystem following the large rain event (Table 4). Further support of viral abundances being opposite of host abundances can be seen when tracking selected viruses and matching them to the host bin (Figure 17). All three viral scaffolds tracked were found to infect two of the eight globally abundant microbial groups Gammaproteobacteria and Bacteroidetes (the other six groups include: Actinobacteria, Alphaproteobacteria, Deltaproteobacteria, Cyanobacteria, Deferribacteres and Euryarchaeota) (Roux et al 2016). Roux et al (2016) found similar trends in their dataset as the four most abundant viral clusters isolated, infected seven of the eight globally abundant microbial groups. Even though viral densities were low following the large rain event, viruses were still in the system infecting hosts and keeping one host from dominating the system.

Viral communities were impacted by abiotic environmental factors (Figure 14). However, viruses still played an important role in ecosystem recovery. AMGs carried in the viral metagenome (virome) can impact ecosystem functioning by transferring the function of the AMG to the host which would allow the host to reproduce in unfavorable conditions (Tsiola et al 2020). AMGs putatively used in ecosystem recovery are seen after this rain event. The marine ecosystem was most likely loaded with excess nitrogen and sulfur after the rain event. There was an increase in freshwater input from three

major rivers into Galveston Bay along with a massive sediment load (Du et al 2019), which would increase the nitrogen and sulfur input. We observed an increase in AMGs associated with nitrogen, sulfur and methane energy metabolisms after the rain event in both viral and microbial metagenomes (Figure 18 B & D). The recovery of AMGs associated with sulfur and nitrogen metabolisms has previously occurred in another study that isolated AMGs from epipelagic viruses that have accepted roles in nitrogen and sulfur cycling (Roux et al 2016).

The viral community had not completely recovered by our last sampling time point on September 28th, 2017, when compared to the pre-Harvey samples. If the viral community had returned to pre-Harvey conditions, the Samp5 samples would have grouped with the pre-Harvey samples on the nMDS plot (Figure 14). However, the Samp5 samples did not group closely to pre-Harvey samples (Figure 14). The changes to the community can also be seen in the shear abundance of viruses present. By Samp5, there were over three-fold more viruses present in the community in Samp5 when compared to the pre-Harvey samples (Table 4). Future studies should continue sampling efforts past a month after the rain event to accurately estimate the time it takes for the viral community to return to pre-rain event conditions. Impacts to the viral community impact the rate of viral lysis in the ecosystem. Since the viral shunt is dependent on lysis events, a decrease in viral lysis within the marine ecosystem throws off the carbon cycle in the ocean. It has already been documented that Hurricane Harvey altered carbon exchange in the marine environment (Yan et al 2020). The impacts seen on carbon exchange following Hurricane Harvey could be to the decrease in viral lysis in the

system following Hurricane Harvey and then an accelerated carbon exchange due to a higher incidence of lytic viruses five weeks after Hurricane Harvey.

4.5 Conclusions

It is important to understand how marine viral communities respond to intense rain events, as described here. It is projected that rainfall rates associated with tropical cyclones will only increase in the coming years, due to climate warming (Knutson et al 2015), making these intense rainfall events only more common. The viral community plays an important role in host community recovery through AMGs and host community population control, as seen here. There were higher abundances of AMGs associated with nitrogen and sulfur metabolisms, which aid hosts in adapting and reproducing in unhospitable environmental conditions. Viral predation is important in the marine microbial community in maintaining population levels. This is seen here by not one single host taking over the microbial community. Here, we provide the first in depth look at marine viral communities' impact and recovery from intense rainfall. Viral communities are altered, at the taxonomic level, from intense rainfall events. The viral communities' abundances and production levels were impacted as well. There were notably less viruses in the marine ecosystem at the beginning of our sampling, which resulted in lower viral diversity (Table 3 & 4). There was an increase in AMGs associated with nitrogen, sulfur and methane metabolisms that aided the host community in reproduction in unfavorable environmental conditions, meaning viruses play a highly important role in ecosystem recovery from large scale rain events. Research on viral impacts of large scale rain events should continue to quantify how viral load is impacted

at a varying degree of rainfall rates and to further understand viral role in ecosystem recovery, as this study did not sample long enough post rain event to show the community returning to pre-event conditions. With the likelihood of large rain events occurring more frequently in the coming years, it is imperative to understand how these events impact the viral community and the long term impacts on coastal biogeochemical cycles due to the lowered abundances of viruses in the system and lower rates of viral production and cell lysis.

CHAPTER V

CONCLUSIONS

Hurricane Harvey was a category 4 hurricane that hit San Jose Island, TX on August 25th, 2017. Hurricane Harvey caused an intense, short-term disturbance in Galveston Bay in regards to sediment loading to the system, along with increased fresh water loads and drops in salinity, temperature and dissolved oxygen (Du et al 2019, Steichen et al 2020). We have characterized how this 1-in-1000 year flooding event impacted the marine microbial communities in Galveston Bay through three objectives: 1) describe the changes in the microbial community composition relative abundance in Galveston Bay following Hurricane Harvey; 2) identify how Hurricane Harvey impacted microbial metabolisms, especially the metabolisms associated with the carbon, nitrogen and sulfur cycles; and 3) demonstrate the role of viruses in ecosystem recovery in Galveston Bay by showing how the viral community adapts and changes with host abundances following Hurricane Harvey (Figure 19).

5.1 Hurricane Harvey increased bacterial diversity and altered microbial genomic potential in Galveston Bay

There was an increase in bacteria known to dominate marine sediment, Gammaproteobacteria, in the water column microbial communities. There was also an increase in terrestrially derived bacteria, such as Actinobacteria and Verrucomicrobiae, present in Galveston Bay's microbial community following Hurricane Harvey's landfall. The increase in sediment and terrestrially derived bacteria in Galveston Bay can be

attributed to the resuspension of marine sediments in the water column and influx of freshwater and sediment to the marine ecosystem.

There was a change in metagenomic potential in Galveston Bay. This was shown through answering three hypotheses: 1) there will be a decrease in genes associated with photosynthetic processes; 2) there will be an increase in genes associated with nitrogen metabolism; and 3) there will be an increase in genes associated with sulfur metabolisms. Our findings supported all three of these hypotheses. There was a decrease in genes associated with photosynthesis, this is due to the decrease of Cyanobacteria from Galveston Bay because of the low salinity and high turbidity seen in the bay following Hurricane Harvey. We also saw an increase in genes associated with nitrogen and sulfur metabolisms due to the resuspension of marine sediments and influx of freshwater into Galveston Bay.

5.2 There was a decrease in viral production and an increase in viral AMGs associated with sulfur and nitrogen metabolisms

Recovery of the microbial communities in Galveston Bay was aided by the viral community. This was seen through support for our two hypotheses: 1) the viral community will change following their host community to aid in population control of the microbial community; and 2) there will be an increase in AMGs associated with nitrogen and sulfur metabolisms. We saw an increase in *Microviridae* and terrestrially associated viruses and a decrease in common marine viruses, such as *Podoviridae* and *Myoviridae*. There was also an increase in AMGs associated with nitrogen and sulfur metabolisms in Galveston Bay.

This study was the first to look at how Hurricane Harvey impacted microbial genomic potential in Galveston Bay and the first to examine how a large-scale rain event impacted marine viral communities. With the intensity of tropical cyclones and rainfall rates associated with tropical cyclones predicted to increase in the coming years (Knutson et al 2015), it is important to have an understanding of how microorganisms, who are at the base of the food web, control nutrient cycling and have the second largest biomass of organisms in the ocean, respond and recover to massive rainfall events.

5.3 How does this research impact the rest of the coastal ecosystem?

Within our sampling period, the bacterial and viral communities had not fully recovered in Galveston Bay by September 28th, 2017, just over a month after Hurricane Harvey made landfall. The lack of full recovery of these communities have lasting impacts on the coastal biogeochemical cycles. The decrease in photosynthetic organisms results in a loss of primary productivity, which translates to a loss of oxygen (Chavez et al 2010). Respiration will occur until the system is depleted of oxygen and then it becomes an anoxic environment. Once the ecosystem becomes anoxic there is a whole new community of microorganism that thrive, creating a long term imbalance. The photosynthetic microorganisms were also replaced by a heterotrophic bacterial community, which would accelerate the carbon cycle, as seen in Yan et al (2020), creating more of an imbalance in the carbon cycle.

Further imbalance in nutrient cycling can be seen from the marine sediment resuspension in the water column. Sulfur is normally stored in marine sediment. However, the sulfur stored in the marine sediment and its microbial community, was

resuspended into the water column, where it normally does not occur. As for the excess nitrogen in the system from storm water runoff, it was most likely depleted more quickly in the system by organisms that use ammonia and convert it into ammonium. Future studies should use transcriptomics or stable isotope probing to evaluate the rates at which these nutrient cycles were impacted from large scale rain events, like Hurricane Harvey.

The decrease of viral abundances (Table 4) and lowered viral production rates, in turn show a decrease in cell lysis, which decreases the amount of dissolved organic matter (DOM) and dissolved organic carbon (DOC) in the marine system. These forms are the most labile in the marine environment and are easily usable for heterotrophic bacteria, which connect microorganisms to the traditional food-web via unicellular eukaryotic grazers. However, in the long-term, we see viral lysis accelerated during Samp3. This acceleration of viral lysis has long term effects on the carbon cycle for a still unknown amount of time. Future studies should determine the time period in which viral lysis returns to pre-disturbance conditions to further understand the impacts large scale rain events have on the coastal carbon cycle. The role of the diverse microorganisms in coastal ecosystems is significant for our planet's balanced ecosystem. It is crucial to understand how these organisms respond to large scale rain events to better understand the impacts such events have on coastal biogeochemical cycles.

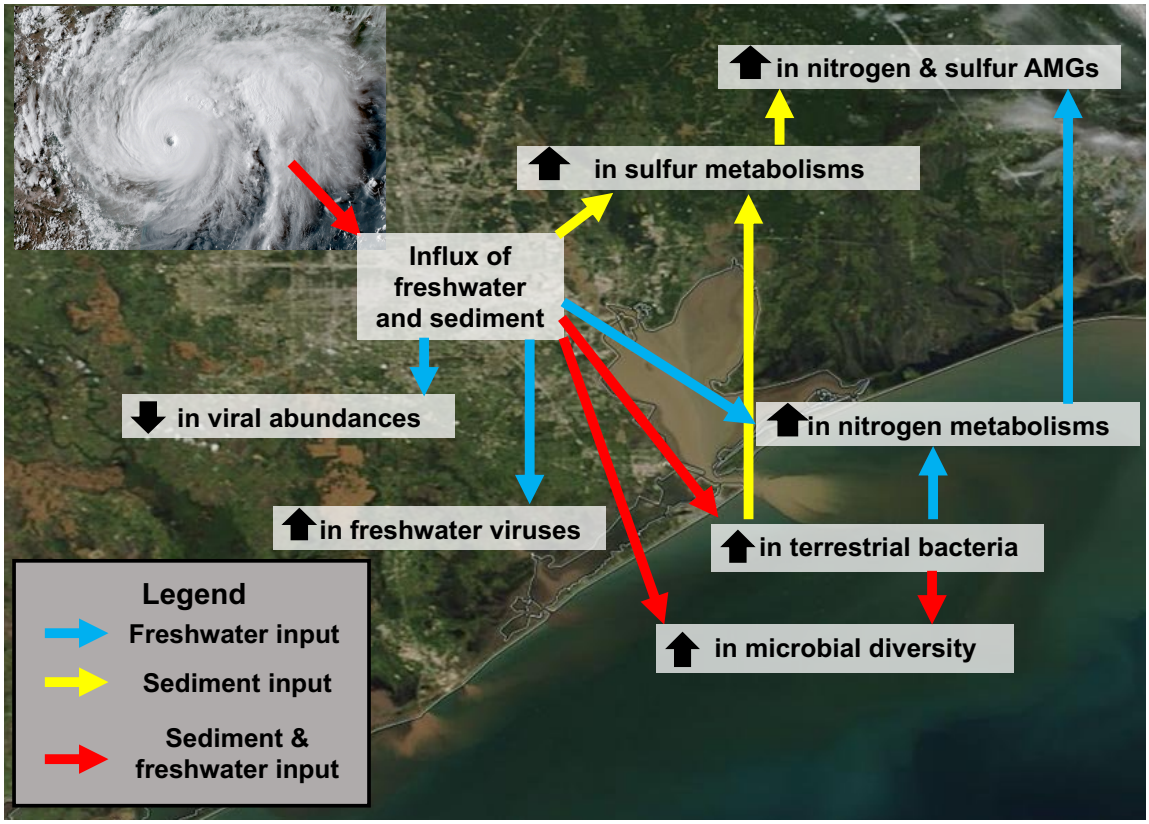

Figure 19 Summary figure detailing findings of this research project. Both images, Hurricane Harvey cyclone and Galveston Bay sediment plume, are from NASA satellites.

REFERENCES

Ahlgren NA, Ren J, Lu YY, Fuhrman JA, Sun F (2017). Alignment-free oligonucleotide frequency dissimilarity measure improves prediction of hosts from metagenomically-derived viral sequences. *Nucleic acids research* **45:** 39-53.

Alcolombri U, Ben-Dor S, Feldmesser E, Levin Y, Tawfik DS, Vardi A (2015). Identification of the algal dimethyl sulfide–releasing enzyme: A missing link in the marine sulfur cycle. *Science* **348:** 1466-1469.

Allison SD, Martiny JB (2008). Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences* **105:** 11512-11519.

Amaral-Zettler LA, Rocca JD, Lamontagne MG, Dennett MR, Gast RJ (2008). Changes in microbial community structure in the wake of Hurricanes Katrina and Rita. *Environmental science & technology* **42:** 9072-9078.

Anderson TR, Ducklow HW (2001). Microbial loop carbon cycling in ocean environments studied using a simple steady-state model. *Aquatic Microbial Ecology* **26:** 37-49.

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983). The ecological role of water-column microbes in the sea. *Marine Ecology* **10:** 257-263.

Balthis WL, Hyland JL, Bearden DW (2006). Ecosystem responses to extreme natural events: impacts of three sequential hurricanes in fall 1999 on sediment quality and condition of benthic fauna in the Neuse River Estuary, North Carolina. *Environmental monitoring and assessment* **119:** 367-389.

Bar-On YM, Phillips R, Milo R (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences* **115:** 6506-6511.

Bidle KD, Haramaty L, Ramos E, Barcelos J, Paul F (2007). Viral activation and recruitment of metacaspases in the unicellular coccolithophore, Emiliania huxleyi. *Proceedings of the National Academy of Sciences* **104:** 6049-6054.

Blake ES, Zelinsky DA (2018). National Hurricane Center Tropical Cyclone Report Hurricane Harvey. National Oceanic and Atmospheric Administration.

Bouvier T, Giorgio PAd (2007). Key role of selective viral-induced mortality in determining marine bacterial community composition. *Environmental Microbiology* **9:** 287-297.

Bowles MW, Mogollón JM, Kasten S, Zabel M, Hinrichs K-U (2014). Global rates of marine sulfate reduction and implications for sub–sea-floor metabolic activities. *Science* **344:** 889-891.

Bratbak G, Egge JK, Heldal M (1993). Viral mortality of the marine alga Emiliania huxleyi (Haptophyceae) and termination of algal blooms. *Marine Ecology Progress Series* **93:** 39-48.

Breitbart M, Thompson LR, Suttle CA, Sullivan MB (2007). Exploring the vast diversity of marine viruses. *Oceanography* **20:** 135-139.

Breitbart M, Bonnain C, Malki K, Sawaya NA (2018). Phage puppet masters of the marine microbial realm. *Nature Microbiology* **3:** 754-766.

Capone DG, Zehr JP, Paerl HW, Bergman B, Carpenter EJ (1997). Trichodesmium, a globally significant marine cyanobacterium. *Science* **276:** 1221-1229.

Capone DG, Burns JA, Montoya JP, Subramaniam A, Mahaffey C, Gunderson T *et al* (2005). Nitrogen fixation by Trichodesmium spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Global Biogeochemical Cycles* **19**.

Chavez FP, Messié M, Pennington JT (2010). Marine primary production in relation to climate variability and change.

Crummett LT, Puxty RJ, Weihe C, Marston MF, Martiny JB (2016). The genomic content and context of auxiliary metabolic genes in marine cyanomyoviruses. *Virology* **499:** 219-229.

Das D, Baruah R, Roy AS, Singh AK, Boruah HPD, Kalita J *et al* (2015). Complete genome sequence analysis of Pseudomonas aeruginosa N002 reveals its genetic adaptation for crude oil degradation. *Genomics* **105:** 182-190.

Dayaram A, Galatowitsch ML, Argüello-Astorga GR, van Bysterveldt K, Kraberger S, Stainton D *et al* (2016). Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. *Infection, Genetics and Evolution* **39:** 304-316.

Dore JE, Brum JR, Tupas LM, Karl DM (2002). Seasonal and interannual variability in sources of nitrogen supporting export in the oligotrophic subtropical North Pacific Ocean. *Limnology and Oceanography* **47:** 1595-1607.

Du J, Park K, Dellapenna TM, Clay JM (2019). Dramatic hydrodynamic and sedimentary responses in Galveston Bay and adjacent inner shelf to Hurricane Harvey. *Science of the Total Environment* **653:** 554-564.

Ducklow HW, Steinberg DK, Buesseler KO (2001). Upper ocean carbon export and the biological pump. *Oceanography* **12:** 50-58.

Eiler A (2006). Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: implications and consequences. *Appl Environ Microbiol* **72:** 7431-7437.

Flemming H-C, Wuertz S (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nature Reviews Microbiology* **17:** 247-260.

Fuhrman JA (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* **399:** 541-548.

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL *et al* (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences* **105:** 7774-7778.

Gage DA, Rhodes D, Nolte KD, Hicks WA, Leustek T, Cooper AJ *et al* (1997). A new route for synthesis of dimethylsulphoniopropionate in marine algae. *Nature* **387:** 891-894.

Galluzzi L, Brenner C, Morselli E, Touat Z, Kroemer G (2008). Viral control of mitochondrial apoptosis. *PLoS pathogens* **4:** e1000018.

Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T *et al* (2009). The seasonal structure of microbial communities in the Western English Channel. *Environmental microbiology* **11:** 3132-3139.

Goldsmith DB, Crosti G, Dwivedi B, McDaniel LD, Varsani A, Suttle CA *et al* (2011). Development of phoH as a novel signature gene for assessing marine phage diversity. *Appl Environ Microbiol* **77:** 7730-7739.

Gribben PE, Nielsen S, Seymour JR, Bradley DJ, West MN, Thomas T (2017). Microbial communities in marine sediments modify success of an invasive macrophyte. *Scientific reports* **7:** 1-8.

Haaber J, Middelboe M (2009). Viral lysis of Phaeocystis pouchetii: implications for algal population dynamics and heterotrophic C, N and P cycling. *The ISME journal* **3:** 430.

Hansell DA, Carlson CA (2001). Marine dissolved organic matter and the carbon cycle. *Oceanography* **14:** 41-49.

Herlemann DP, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME journal* **5:** 1571.

Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB (2017). Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *The ISME journal* **11:** 1511-1520.

Howarth RW, Sharpley A, Walker D (2002). Sources of nutrient pollution to coastal waters in the United States: Implications for achieving coastal water quality goals. *Estuaries* **25:** 656-676.

Howarth RW (2008). Coastal nitrogen pollution: a review of sources and trends globally and regionally. *Harmful algae* **8:** 14-20.

Hurwitz BL, Hallam SJ, Sullivan MB (2013). Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome biology* **14:** R123.

Huson DH, Auch AF, Qi J, Schuster SC (2007). MEGAN analysis of metagenomic data. *Genome research* **17:** 377-386.

Hutchins DA, Fu F (2017). Microorganisms and ocean global change. *Nature microbiology* **2:** 17058.

Ignacio-Espinoza JC, Sullivan MB (2012). Phylogenomics of T4 cyanophages: lateral gene transfer in the 'core'and origins of host genes. *Environmental microbiology* **14:** 2113-2126.

Kang DD, Froula J, Egan R, Wang Z (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3:** e1165.

Kang DD, Li F, Kirton E, Thomas A, Egan R, An H *et al* (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7:** e7359.

Kieft K, Zhou Z, Anantharaman K (2020). VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* **8:** 1-23.

Kiene RP, Linn LJ, Bruton JA (2000). New and important roles for DMSP in marine microbial communities. *Journal of Sea Research* **43:** 209-224.

Knutson TR, Sirutis JJ, Zhao M, Tuleya RE, Bender M, Vecchi GA *et al* (2015). Global projections of intense tropical cyclone activity for the late twenty-first century from dynamical downscaling of CMIP5/RCP4. 5 scenarios. *Journal of Climate* **28:** 7203-7224.

Labonté JM, Suttle CA (2013). Metagenomic and whole-genome analysis reveals new lineages of gokushoviruses and biogeographic separation in the sea. *Frontiers in microbiology* **4:** 404.

Lake PS (2000). Disturbance, patchiness, and diversity in streams. *Journal of the north american Benthological society* **19:** 573-592.

Landa M, Blain S, Christaki U, Monchy S, Obernosterer I (2016). Shifts in bacterial community composition associated with increased carbon cycling in a mosaic of phytoplankton blooms. *The ISME journal* **10:** 39.

Langenheder S, Kisand V, Wikner J, Tranvik LJ (2003). Salinity as a structuring factor for the composition and performance of bacterioplankton degrading riverine DOC. *FEMS Microbiology Ecology* **45:** 189-202.

Langenheder S, Ragnarsson H (2007). The role of environmental and spatial factors for the composition of aquatic bacterial communities. *Ecology* **88:** 2154-2161.

Li H, Durbin R (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26:** 589-595.

Li H (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:13033997*.

Luther GW, Findlay AJ, MacDonald DJ, Owings SM, Hanson TE, Beinart RA *et al* (2011). Thermodynamics and kinetics of sulfide oxidation by oxygen: a look at inorganically controlled reactions and biologically mediated processes in the environment. *Frontiers in microbiology* **2:** 62.

Malmstrom RR, Kiene RP, Kirchman DL (2004). Identification and enumeration of bacteria assimilating dimethylsulfoniopropionate (DMSP) in the North Atlantic and Gulf of Mexico. *Limnology and oceanography* **49:** 597-606.

Mann NH, Cook A, Millard A, Bailey S, Clokie M (2003). Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* **424:** 741.

Mason OU, Hazen TC, Borglin S, Chain P, Dubinsky EA, Fortney JL *et al* (2012). Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. *ISME* **6:** 1715-1727.

Matteson AR, Budinoff CR, Campbell CE, Buchan A, Wilhelm SW (2010). Estimating virus production rates in aquatic systems. *JoVE (Journal of Visualized Experiments)***:** e2196.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al* (2008). The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics* **9:** 386.

Milliner C, Materna K, Bürgmann R, Fu Y, Moore AW, Bekaert D *et al* (2018). Tracking the weight of Hurricane Harvey's stormwater using GPS data. *Science advances* **4:** eaau2477.

Needham DM, Chow C-ET, Cram JA, Sachdeva R, Parada A, Fuhrman JA (2013). Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *The ISME journal* **7:** 1274-1285.

Oksanen J, Blanchet FG, Friendly M, Roeland Kindt, Legendre P, McGlinn D *et al* (2019). vegan: Community Ecology Package, 2.5-6 edn.

Oldenborgh GJv, Wiel Kvd, Sebastian A, Singh R, Arrighi J, Otto F *et al* (2017). Attribution of extreme rainfall from Hurricane Harvey, August 2017. *Environmental Research Letters* **12**.

Paerl HW, Otten TG (2013). Harmful cyanobacterial blooms: causes, consequences, and controls. *Microbial ecology* **65:** 995-1010.

Pál C, Papp B, Lercher MJ (2005). Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nature Genetics* **37:** 1372-1375.

Park K, Valentine JF, Sklenar S, Weis KR, Dardeau MR (2007). The effects of Hurricane Ivan in the inner part of Mobile Bay, Alabama. *Journal of Coastal Research* **23:** 1332-1336.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* **25:** 1043-1055.

Patel A, Noble RT, Steele JA, Schwalbach MS, Hewson I, Fuhrman JA (2007). Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. *Nature protocols* **2:** 269.

Pomeroy LR, Williams PJ, Azam F, Hobbie JE (2007). The microbial loop. *Oceanography* **20:** 28-33.

Ramirez-Gonzalez RH, Bonnal R, Caccamo M, MacLean D (2012). Bio-samtools: Ruby bindings for SAMtools, a library for accessing BAM files containing high-throughput sequence alignments. *Source code for biology and medicine* **7:** 6.

Riemann L, Farnelid H, Steward GF (2010). Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters. *Aquatic Microbial Ecology* **61:** 235-247.

Rohwer F, Thurber RV (2009). Viruses manipulate the marine environment. *Nature* **459:** 207-212.

Roux S, Krupovic M, Poulet A, Debroas D, Enault F (2012). Evolution and diversity of the Microviridae viral family through a collection of 81 new complete genomes assembled from virome reads. *PloS one* **7:** e40418.

Roux S, Enault F, Hurwitz BL, Sullivan MB (2015). VirSorter: mining viral signal from microbial genomic data. *PeerJ* **3:** e985.

Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A *et al* (2016). Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537:** 689.

Rubin MA, Leff LG (2007). Nutrients and other abiotic factors affecting bacterial communities in an Ohio River (USA). *Microbial ecology* **54:** 374-383.

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

Seitzinger SP, Harrison JA, Dumont E, Beusen AH, Bouwman A (2005). Sources and delivery of carbon, nitrogen, and phosphorus to the coastal zone: An overview of Global Nutrient Export from Watersheds (NEWS) models and their application. *Global Biogeochemical Cycles* **19**.

Shade A, Peter H, Allison SD, Baho D, Berga M, Bürgmann H *et al* (2012a). Fundamentals of microbial community resistance and resilience. *Frontiers in microbiology* **3:** 417.

Shade A, Read JS, Youngblut ND, Fierer N, Knight R, Kratz TK *et al* (2012b). Lake microbial communities are resilient after a whole-ecosystem disturbance. *The ISME journal* **6:** 2153-2167.

Sievert SM, Kiene RP, Schulz-Vogt HN (2007). The sulfur cycle. *Oceanography* **20:** 117-123.

Smets BF, Barkay T (2005). Horizontal gene transfer: Perspectives at a crossroads of scientific disciplines. *Nature Reviews: Microbiology* **3:** 675-678.

Sohm JA, Webb EA, Capone DG (2011). Emerging patterns of marine nitrogen fixation. *Nature Reviews Microbiology* **9:** 499-508.

Stal LJ (2009). Is the distribution of nitrogen-fixing cyanobacteria in the oceans related to temperature? *Environmental microbiology* **11:** 1632-1645.

Steichen JL, Labonté JM, Windham R, Hala D, Kaiser K, Setta S *et al* (2020). Microbial, physical, and chemical changes in Galveston Bay following an extreme flooding event, Hurricane Harvey. *Frontiers in Marine Science* **7:** 186.

Sullivan MB, Coleman ML, Weigele P, Rohwer F, Chisholm SW (2005). Three Prochlorococcus cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* **3:** e144.

Suttle C (2007). Marine viruses—major players in the global ecosystem. *Nature Reviews Microbiology* **5:** 801-812.

Suttle CA, Chan AM, Cottrell MT (1991). Use of ultrafiltration to isolate viruses from seawater which are pathogens of marine phytoplankton. *Applied and environmental microbiology* **57:** 721-726.

Suttle CA (2005). Viruses in the sea. *Nature* **437:** 356.

Thamdrup B, Fossing H, Jørgensen BB (1994). Manganese, iron and sulfur cycling in a coastal marine sediment, Aarhus Bay, Denmark. *Geochimica et Cosmochimica Acta* **58:** 5115-5129.

Thingstad TF, Lignell R (1997). Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquatic Microbial Ecology* **13:** 19-27.

Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J *et al* (2011). Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proceedings of the National Academy of Sciences* **108:** E757-E764.

Tsiola A, Michoud G, Fodelianakis S, Karakassis I, Kotoulas G, Pavlidou A *et al* (2020). Viral Metagenomic Content Reflects Seawater Ecological Quality in the Coastal Zone. *Viruses* **12:** 806.

Tucker KP, Parsons R, Symonds EM, Breitbart M (2011). Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. *The ISME journal* **5:** 822-830.

Van Valen L (1974). Molecular evolution as predicted by natural selection. *Journal of molecular evolution* **3:** 89-101.

Voss M, Bange HW, Dippner JW, Middelburg JJ, Montoya JP, Ward B (2013). The marine nitrogen cycle: recent discoveries, uncertainties and the potential relevance of climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368:** 20130121.

Warwick-Dugdale J, Buchholz HH, Allen MJ, Temperton B (2019). Host-hijacking and planktonic piracy: how phages command the microbial high seas. *Virology journal* **16:** 15.

Wasmund K, Mußmann M, Loy A (2017). The life sulfuric: microbial ecology of sulfur cycling in marine sediments. *Environmental microbiology reports* **9:** 323-344.

Wigington CH, Sonderegger DL, Brussaard CP, Buchan A, Finke JF, Fuhrman J *et al* (2015). Re-examining the relationship between virus and microbial cell abundances in the global oceans. *BioRxiv***:** 025544.

Wilhelm S, Suttle C (1999). Viruses and Nutrient Cycles in the Sea. *BioScience* **49:** 781-788.

Williamson KE, Harris JV, Green JC, Rahman F, Chambers RM (2014). Stormwater runoff drives viral community composition changes in inland freshwaters. *Frontiers in microbiology* **5:** 105.

Wood DE, Lu J, Langmead B (2019). Improved metagenomic analysis with Kraken 2. *Genome biology* **20:** 257.

Yan G, Labonte JM, Quigg A, Kaiser K (2020). Hurricanes accelerate dissolved organic carbon cycling in coastal ecosystems. *Frontiers in Marine Science*.

Zhao Y, Qin F, Zhang R, Giovannoni SJ, Zhang Z, Sun J *et al* (2019). Pelagiphages in the Podoviridae family integrate into host genomes. *Environmental microbiology* **21:** 1989-2001.

APPENDIX A

SUMMARY TABLES

Table 7 All sampling dates and stations, with sampling coordinates, the total volume filtered (in L), filter sizes used, final viral concentration (VC) volumes in mL and the VC concentration factor.

SampID	Sampling Date	Coordinates	Volume Filtered (L)	Filter Sizes	Final VC Volume (mL)	Concentration Factor
Samp1_Stn1	9/4/17	- 94.976944, 29.6725	4	0.22 & GFD	114	35.1
Samp1_Stn4	9/4/17	- 94.897222, 29.535556	4	0.22 & GFD	108	37
Samp1_Stn7	9/4/17	- 94.825833, 29.4125	4	0.22 & GFD	69	58
Samp1_Stn10	9/4/17	- 94.688611, 29.333611	4	0.22 & GFD	64	62.5
Samp2_Stn1	9/9/17	- 94.976944, 29.6725	10	0.45 & GFD	65	153.8
Samp2_Stn4	9/9/17	- 94.897222, 29.535556	10	0.45 & GFD	149	67.1
Samp2_Stn7	9/9/17	- 94.825833, 29.4125	10	0.45 & GFD	382	26.2
Samp2_Stn10	9/9/17	- 94.688611, 29.333611	10	0.45 & GFD	171	58.5
Samp3_Stn1	9/16/17	- 94.976944, 29.6725	17	0.22 & GFF	115	147.8
Samp3_Stn4	9/16/17	- 94.897222, 29.535556	17.5	0.22 & GFF	170	102.9
Samp3_Stn7	9/16/17	- 94.825833, 29.4125	20	0.22 & GFF	188	106.4
Samp3_Stn10	9/16/17	- 94.688611, 29.333611	10	0.22 & GFF	172	58.1
Samp4_Stn1	9/21/17	- 94.976944, 29.6725				
Samp4_Stn4	9/21/17	- 94.897222, 29.535556				
Samp4_Stn7	9/21/17	- 94.825833, 29.4125				
Samp4_Stn10	9/21/17	- 94.688611, 29.333611				
Samp5_Stn1	9/28/17	- 94.976944, 29.6725	20	0.22 & GFF	195	102.6
Samp5_Stn4	9/28/17	- 94.897222, 29.535556	10	0.22 & GFF	196	51
Samp5_Stn7	9/28/17	- 94.825833, 29.4125	15.75	0.22 & GFF	195	80.8
Samp5_Stn10	9/28/17	- 94.688611, 29.333611	19.5	0.22 & GFF	205	95.1

Table 7 continued.

SampID	Sampling Date	Coordinates	Volume Filtered (L)	Filter Sizes	Final VC Volume (mL)	Concentration Factor
Pre-Harvey1	7/31/17	- 94.85, 29.32		0.22 & GFD		
Pre-Harvey2	8/22/17	- 94.85, 29.32	21	0.2 & 5µm	60	350

Table 8 Assembly metrics from all samplings and stations broke up by filter size. Total number contigs, total bp assembled, max contig length, average contig length and N50 are from *de novo* assembly done in Megahit. Number of proteins was taken from Prodigal.

Samp ID	Filter Size	Total # Contigs	Total bp Assembled	Max Contig Length	Average Contig Length	N50	# Proteins
Samp1_Stn1	0.22	584,688	380,807,570	181,496	651	668	828,435
Samp1_Stn4	0.22	540,483	366,600,185	132,707	678	714	783,902
Samp1_Stn7	0.22	538,660	360,809,803	89,548	670	704	790,577
Samp1_Stn10	0.22	547,564	371,660,270	166,982	679	709	792,777
Samp2_Stn1	0.45	574,772	418,240,335	81,556	728	814	864,246
Samp2_Stn4	0.45	577,526	417,442,972	543,787	723	794	857,422
Samp2_Stn7	0.45	520,220	357,986,102	80,516	688	731	763,602
Samp2_Stn10	0.45	403,307	304,838,186	70,494	756	872	606,950
Samp3_Stn1	0.22	444,073	322,839,415	99,903	727	809	684,673
Samp3_Stn4	0.22	490,412	381,694,878	121,364	778	920	766,353
Samp3_Stn7	0.22	515,190	398,991,752	238,672	774	909	795,285
Samp3_Stn10	0.22	559,582	406,062,959	77,453	726	820	843,242
Samp5_Stn1	0.22	597,030	436,773,626	140,955	732	837	893,175
Samp5_Stn4	0.22	361,311	258,458,740	130,635	715	802	551,864
Samp5_Stn7	0.22	470,957	352,276,317	173,418	748	859	730,096
Samp5_Stn10	0.22	358,917	269,820,294	85,257	752	864	555,612
Pre-Harvey1	0.22	494,328	364,865,701	55,138	738	825	735,220
Pre-Harvey2	0.2	632,933	449,485,501	90,292	710	774	919,021
Samp1_Stn1	GFD	481,796	295,863,296	322,552	614	624	624,959
Samp1_Stn4	GFD	464,833	270,922,896	76,749	583	585	551,911
Samp1_Stn7	GFD	442,933	252,155,932	39,597	569	566	556,460
Samp1_Stn10	GFD	538,739	322,983,257	77,883	600	598	684,011
Samp2_Stn1	GFD	504,945	310,570,014	52,301	615	624	651,039
Samp2_Stn4	GFD	523,651	347,345,466	146,628	663	690	704,808
Samp2_Stn7	GFD	645,360	408,224,781	60,818	633	644	827,671
Samp2_Stn10	GFD	490,896	322,138,454	94,643	656	692	631,363
Samp3_Stn1	GFF	635,532	487,604,658	86,627	767	892	937,146
Samp3_Stn4	GFF	496,819	345,771,611	81,711	696	760	694,278
Samp3_Stn7	GFF	531,449	351,935,643	149,754	662	697	758,212
Samp3_Stn10	GFF	529,522	367,369,414	137,899	694	746	768,822
Samp5_Stn1	GFF	472,772	321,751,838	138,312	681	726	669,782
Samp5_Stn4	GFF	510,601	374,058,254	78,298	733	832	740,499
Samp5_Stn7	GFF	467,757	356,100,990	106,610	761	891	693,354
Samp5_Stn10	GFF	667,378	459,600,971	104,332	689	741	936,465
Pre-Harvey1	GFD	652,489	438,889,672	89,952	673	717	876,799
Pre-Harvey2	5µm	519,384	324,961,850	89,811	626	649	670,489

Table 8 continued.

Samp ID	Filter Size	Total # Contigs	Total bp Assembled	Max Contig Length	Average Contig Length	N50	# Proteins
Samp1_Stn1	vDNA	271,790	205,446,170	85,738	756	906	411,618
Samp1_Stn4	vDNA	209,809	151,437,740	40,358	722	858	316,498
Samp1_Stn7	vDNA	174,248	125,811,208	85,134	722	886	266,602
Samp1_Stn10	vDNA	176,236	157,750,747	383,657	895	1,175	297,413
Samp2_Stn1	vDNA	100,974	86,737,853	35,815	859	1,124	186,745
Samp2_Stn4	vDNA	158,742	113,784,826	65,587	717	863	258,937
Samp2_Stn7	vDNA	207,712	173,280,926	165,803	834	1,055	324,513
Samp2_Stn10	vDNA	192,152	168,472,086	67,562	877	1,151	313,657
Samp3_Stn1	vDNA	129,970	93,914,138	43,808	723	862	204,693
Samp3_Stn4	vDNA	176,761	125,664,202	46,485	711	807	280,476
Samp3_Stn7	vDNA	113,412	80,206,427	75,669	707	835	181,506
Samp3_Stn10	vDNA	124,924	90,746,825	71,043	726	845	204,072
Samp5_Stn1	vDNA	170,599	123,884,304	117,201	726	840	271,212
Samp5_Stn4	vDNA	183,330	153,554,927	220,073	838	1,045	327,881
Samp5_Stn7	vDNA	287,885	223,537,903	88,960	776	912	468,393
Samp5_Stn10	vDNA	182,326	157,841,925	113,869	866	1,120	336,615
Pre-Harvey1	vDNA	131,913	92,464,154	61,354	701	789	214,746
Pre-Harvey2	vDNA	51,829	37,309,941	15,582	720	918	84,374

Table 9 Quality control metrics from all samplings and stations broke up by filter size. All metrics were pulled from the BBtools suite prior to *de novo* assembly.

Samp ID	Filter Size	# reads	# reads post-QC	% merged	% masked (merged)
Samp1_Stn1	0.22	32,165,848	32,035,118	19.31	0.02
Samp1_Stn4	0.22	34,420,532	34,274,700	14.65	0.02
Samp1_Stn7	0.22	31,429,186	31,317,318	17.26	0.04
Samp1_Stn10	0.22	33,958,200	33,835,650	17.11	0.04
Samp2_Stn1	0.45	36,157,034	36,001,282	28.26	0.03
Samp2_Stn4	0.45	35,438,314	35,333,472	18.92	0.06
Samp2_Stn7	0.45	29,597,168	29,525,854	22.20	0.05
Samp2_Stn10	0.45	23,472,810	23,415,436	33.08	0.07
Samp3_Stn1	0.22	27,420,504	27,345,366	23.00	0.03
Samp3_Stn4	0.22	32,124,260	32,040,026	23.07	0.04
Samp3_Stn7	0.22	34,428,046	34,308,366	16.29	0.04
Samp3_Stn10	0.22	36,536,080	36,368,464	25.25	0.03
Samp5_Stn1	0.22	47,263,132	47,081,744	26.42	0.08
Samp5_Stn4	0.22	32,890,468	32,773,616	25.46	0.03
Samp5_Stn7	0.22	32,195,330	32,098,010	30.09	0.03
Samp5_Stn10	0.22	23,965,818	23,842,924	21.85	0.02
Pre-Harvey1	0.22	27,887,626	27,791,264	18.22	0.03
Pre-Harvey2	0.2	32,160,120	32,058,328	25.52	0.03
Samp1_Stn1	GFD	27,011,362	26,930,666	17.69	0.12
Samp1_Stn4	GFD	25,941,884	25,835,122	19.94	0.22
Samp1_Stn7	GFD	26,548,726	26,487,858	26.68	0.47
Samp1_Stn10	GFD	29,471,906	29,355,146	21.41	0.22
Samp2_Stn1	GFD	34,538,226	34,415,282	17.80	0.33
Samp2_Stn4	GFD	33,717,552	33,609,022	17.14	0.36
Samp2_Stn7	GFD	33,263,434	33,139,334	21.47	0.22
Samp2_Stn10	GFD	30,689,328	30,440,254	19.44	0.61
Samp3_Stn1	GFF	40,496,102	40,394,972	14.81	0.16
Samp3_Stn4	GFF	25,489,008	25,415,396	22.29	0.17
Samp3_Stn7	GFF	27,893,754	27,822,518	24.10	0.18
Samp3_Stn10	GFF	30,862,948	30,771,102	20.39	0.28
Samp5_Stn1	GFF	30,082,442	29,971,260	20.83	0.37
Samp5_Stn4	GFF	31,759,374	31,611,686	18.51	0.20
Samp5_Stn7	GFF	31,518,434	31,396,792	17.60	0.19
Samp5_Stn10	GFF	35,135,656	35,004,760	28.31	0.21
Pre-Harvey1	GFD	39,073,962	38,934,916	15.04	0.27
Pre-Harvey2	5µm	34,793,716	34,652,358	16.97	0.40

Table 9 continued.

Samp ID	Filter Size	# reads	# reads post-QC	% merged	% masked (merged)
Samp1_Stn1	vDNA	44,390,824	44,179,728	16.35	0.01
Samp1_Stn4	vDNA	36,243,410	36,082,230	15.39	0.01
Samp1_Stn7	vDNA	37,319,380	37,166,234	16.74	0.01
Samp1_Stn10	vDNA	35,043,570	34,878,432	12.86	0.01
Samp2_Stn1	vDNA	35,331,176	35,213,542	20.00	0.01
Samp2_Stn4	vDNA	31,865,640	31,694,256	19.33	0.02
Samp2_Stn7	vDNA	33,065,372	32,759,936	20.07	0.31
Samp2_Stn10	vDNA	37,041,430	36,900,658	11.56	0.35
Samp3_Stn1	vDNA	35,334,440	35,026,594	19.66	0.02
Samp3_Stn4	vDNA	46,440,864	46,251,562	12.61	0.06
Samp3_Stn7	vDNA	40,128,170	39,959,278	13	0.03
Samp3_Stn10	vDNA	31,271,224	31,143,614	12.45	0.02
Samp5_Stn1	vDNA	36,275,350	36,125,418	22.51	0.05
Samp5_Stn4	vDNA	35,707,596	35,564,428	12	0.01
Samp5_Stn7	vDNA	41,270,820	41,148,976	23.34	0.13
Samp5_Stn10	vDNA	37,979,588	37,823,114	12.38	0.01
Pre-Harvey1	vDNA	39,187,644	39,039,398	14.68	0.00
Pre-Harvey2	vDNA	40,898,748	40,740,756	14.43	0.00