



Article

Record-Breaking Rain Event Altered Estuarine Viral Assemblages

Alaina C. Woods , Jordan R. Walker, Cameron D. Jackson [†] and Jessica M. Labonté ^{*}

Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX 77553, USA; alaina.woods@maine.edu (A.C.W.); jronwalker@tamug.edu (J.R.W.); cjacks89@vols.utk.edu (C.D.J.)

^{*} Correspondence: labontej@tamug.edu; Tel.: +1-(409)-740-4921

[†] Current address: Department of Microbiology, University of Tennessee Knoxville, Knoxville, TN 37996, USA.

Abstract: Viruses are the dominant biological entity in the ocean, play a vital role in biogeochemical cycles, and provide their hosts with novel metabolic capabilities through auxiliary metabolic genes (AMGs). Hurricane Harvey was a category 4 hurricane that made landfall on the Texas coast in 2017 and lashed the Houston area with $1.4\text{--}1.7 \times 10^{10} \text{ m}^3$ of rainfall. In this paper, we aim to characterize how the changes in abiotic conditions brought by Hurricane Harvey altered the viral assemblages of Galveston Bay at the taxonomic level and determine how viral ecosystem functions were altered. Metagenomes of the viruses and their hosts were sequenced from a transect in Galveston Bay over the five weeks following the storm. Our results show that the viral assemblages of Galveston Bay dramatically changed following Hurricane Harvey's landfall. Of the abiotic parameters measured, salinity had the strongest effect on shaping the viral assemblages. In the five weeks following Hurricane Harvey, there was a steady increase of metabolic genes and putative viral infections. Our study provides the first in-depth look at how marine viral assemblages respond and recover from extreme rainfall events, which models predict will become more frequent and intense with climate change.



Citation: Woods, A.C.; Walker, J.R.; Jackson, C.D.; Labonté, J.M.

Record-Breaking Rain Event Altered Estuarine Viral Assemblages.

Microorganisms **2022**, *10*, 729.

<https://doi.org/10.3390/microorganisms10040729>

<https://doi.org/10.3390/microorganisms10040729>

Academic Editors: Kristina Mojica, Corina P.D. Brussaard and Thomas Mock

Received: 7 March 2022

Accepted: 25 March 2022

Published: 29 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: viral ecology; hurricane; metagenomics; auxiliary metabolic genes

1. Introduction

Viruses are the dominant biological entity in the ocean, with their numbers estimated to be as high as $\sim 4 \times 10^{30}$ [1]. They play many important roles in aquatic ecosystems, including top-down control, nutrient recycling, and increasing the fitness of the host in unfavorable conditions [2,3]. Viral infections cause the lysis of an estimated $\sim 10\text{--}20\%$ of the marine prokaryotic population daily [1,4,5], and the ecosystem-wide effects of viral lysis can alter geochemical cycles in the ocean by redirecting nutrients from higher trophic levels back to the microbial food web as dissolved organic matter, a process called the viral shunt [6–11]. Viruses are important in directing carbon and energy flux in marine ecosystems.

Auxiliary metabolic genes (AMGs) are viral encoded genes that typically originate from their hosts, with metabolic capabilities that can redirect energy and resources, improving viral production within the host [12]. In surface ocean environments, some cyanophages (viruses infecting Cyanobacteria) encode photosynthetic genes [13], and nearly 40% of sequenced phage genomes encode for the phosphate regulation uptake gene, *phoH* [14]. At least 31 additional AMGs have been identified in cyanophages, and phylogenetic analyses revealed that rare AMGs are more frequently gained or lost as a result of fluctuating selection pressures, while common AMGs are associated with stable selection pressures [15]. As the viral genomic diversity is revealed, more AMGs are discovered. For example, among the viral genomes assembled from the *Tara* Oceans and Malaspina research expeditions, a total of 243 virally-encoded AMGs were identified, comprising nearly 150 genes not previously known as viral AMGs [7]. With AMGs, viruses may directly manipulate the biogeochemical cycles such as those of sulfur, nitrogen, and phosphorus [7], although more studies are necessary to quantify the impact of viral AMGs within these cycles. The

presence of AMGs in viral genomes and potential transfer of AMGs between hosts and viruses or viruses and their hosts may provide greater fitness to the microbial communities in the face of unfavorable conditions [2], such as those brought by a large-scale rain event.

While it has been well documented that microbial communities respond to pulse disturbances such as hurricanes [16,17] and oil spills [18], there has been little work done on the impact of pulse disturbances and large rain events on viral assemblages. After Hurricane Sandy, Williamson et al. [19] evaluated the impact of stormwater runoff on the viral community composition in a freshwater system and found that the storm had negative impacts on viral species richness and viral abundances due to increased rainfall and bacterial abundance. In contrast, another study used metagenomics to study the impact of a large rain event on the viral assemblages of a retention pond and found that following the storm, stormwater runoff contributed to a sharp increase in viral richness and diversity [20]. In this study, we looked at the impact of a pulse disturbance, Hurricane Harvey, on the viral communities inhabiting the Galveston Bay estuary. Hurricane Harvey made landfall near Port Aransas, TX, on 25 August 2017. Galveston, TX, located ~280 km northeast of Port Aransas, did not experience hurricane-force winds but experienced a tropical storm with winds between 60–120 km/h and a storm surge of <1 m. Moreover, Hurricane Harvey stalled over Houston and Galveston Bay, causing a record-breaking $1.4\text{--}1.7 \times 10^{10} \text{ m}^3$ of rainfall to lash the area over the next four days (25–29 August 2017). Floodwater, stormwater, and rainwater all flushed into Galveston Bay, causing a drop in salinity, temperature, and dissolved oxygen [17,21,22] (Supplementary Table S1). Moreover, 9.86×10^7 metric tons of sediment were resuspended into Galveston Bay, causing an increase in dissolved organic matter, nutrients, and turbidity [21] (Supplementary Table S1).

To characterize the impact of Hurricane Harvey on the viral assemblages, we sequenced viral and microbial metagenomes from a transect in Galveston Bay for five weeks following the storm. We characterized how the change in environmental conditions changed the viral assemblages at the taxonomic level. We hypothesized there would be an increase in viruses associated with the terrestrial environment in Galveston Bay following Hurricane Harvey. We also identified how the storm impacted viral ecosystem functions through the identification of AMGs and estimation of putative viral infections. We hypothesized that viral function would change due to the changing environmental conditions [17,21,22] and host community [17], and there would be a decrease in putative viral infections due to the excess runoff [21], diluting viral concentrations and reducing virus–host contacts. Our results show that the viral assemblages changed after the storm and returned to estuarine viruses as the environmental conditions (i.e., salinity and temperature) returned to pre-storm levels. There was also a constant increase of metabolic genes and putative viral infections over the following five weeks, suggesting that the storm’s long-term impact on viruses resulted in accelerated geochemical cycles.

2. Materials and Methods

2.1. Sample Collection and Processing

Samples were collected from four stations in Galveston Bay after Hurricane Harvey left the Galveston Bay area on 4 September, 9 September, 16 September, and 28 September 2017 onboard the Texas A&M University at Galveston (TAMUG) vessel R/V Trident. A four-station transect (Stn1 (−94.976944, 29.6725), Stn4 (−94.897222, 29.535556), Stn7 (−94.825833, 29.4125), and Stn10 (−94.688611, 29.333611)) was sampled from the mouth of the San Jacinto River to the Gulf of Mexico (Figure 1). Sampling trips started in the morning and were always transiting from Stn1 to Stn10. Pre-Harvey samples were collected from TAMUG’s boat basin (−94.85, 29.32) on 31 July and 22 August 2017 prior to Hurricane Harvey’s landfall. The pre-Harvey samples are geographically closest to Stn7 but are not directly in Galveston Bay. The environmental conditions at the TAMUG Boat Basin are more closely related to those found at Stn10, at the mouth of the Gulf of Mexico (Supplementary Table S1). Due to the possible differences in environmental parameters, we will compare the post-Harvey samples to the pre-Harvey samples but refrain from

making assumptions regarding ecosystem recovery. For this study, 20–25 L of surface water was collected using a 5-gallon bucket attached to a rope. All samples were pre-filtered onboard immediately after sampling with a Nitex filter (30 μm) to remove small grazers and large particles. After pre-filtration, the samples were stored on the boat in the dark at room temperature for up to 4 h in HDPE carboys and brought to the laboratory, where they were stored at 4 $^{\circ}\text{C}$ until filtration. Filtration always occurred on the same day as sampling. Generally, each sample was filtered through a 142 mm glass fiber filter (Whatman GF/F with a 0.7 μm pore-size or Whatman GF/D with a 2.7 μm pore-size), followed by a 142 mm 0.22 μm pore-size polyvinylidene fluoride (PVDF) filter (Supplementary Table S2 details which filter was used for each sample). Due to the availability of supplies, for the sampling of 9 September, pre-filtration was performed with a 142 mm 0.45 μm filter and the virus concentrate was then filtered through 47 mm 0.22 μm PVDF filter. The volume of sample filtered ranged from 4 to 20 L, depending on time constraints, available material, and personnel (Supplementary Table S2). All GF and PVDF filters were stored at -20°C until further use. Filtered samples, which contained the viruses, were stored at 4 $^{\circ}\text{C}$ until virus concentration, for no more than 24 h. Viruses were concentrated using tangential flow filtration with a 30 kDa cut-off as previously described [23] (Supplementary Table S2). All virus concentrates were stored at 4 $^{\circ}\text{C}$ until DNA extractions.



Figure 1. 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 basin, where pre-Harvey samples were taken on 31 July and 22 August 2017.

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 (Supplementary Table S2). Before extraction, virus concentrates were further concentrated to a final volume of $\sim 500\ \mu\text{L}$ using 30 kDa Amicon Centrifugal Filters (Millipore, Burlington, MA, USA). Viral DNA was extracted using the DNeasy PowerSoil DNA extraction kit (Qiagen, Germantown, MD, USA). Since the yield was too low for metagenomic sequencing, the DNA was amplified using the Repli-g DNA Mini kit (Qiagen) according to the manufacturer's recommendations. The amplified DNA was further purified with the QIAamp DNA purification kit (Qiagen) following the cleanup of genomic DNA protocol. Total DNA from the putative host fraction was extracted from the GF/F, GF/D, and 0.22 μm filters to the corresponding volume of 4 L using a standard phenol-chloroform extraction protocol [24], where the lysis buffer consisted of 400 mM NaCl, 750 mM sucrose, 20 mM EDTA, and 50 mM Tris-HCl, pH = 9.0. All DNA extracts were quantified with a QuBit fluorometer [with the HS dsDNA Assay kit (Invitrogen, Waltham, MA, USA)] and stored at -20°C until further use. Between 1–10 μg of DNA from each sample 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, TX, USA).

2.3. Metagenomic Data Analysis

BBtools software suite (version 38.31; available at <https://sourceforge.net/projects/bbtools/> (accessed on 1 June 2019)) was used for quality control of the metagenomic raw reads [25], as follows. BBduk was run twice, once to remove contaminants (kmers between 11–23 bp for common adapter artifacts) and once to trim reads that contained adapter sequences and had low quality (Phred score $Q < 10$). BBmerge was used to merge the forward and reverse paired-end reads using the default settings [25]. BBmask was used to soft-mask human, cat, dog, and mouse contamination. MEGAHIT (version 1.2.8) was then used for de novo assembly into contigs [26]. VirSorter2 (version 2.2.2) was then used to identify putative viral contigs ($<0.22 \mu\text{m}$ fraction) [27,28]. All identified contigs were kept for further analyses, and these contigs were confirmed to be host-free as they all returned no SSU rRNA sequences when screened with PhyloFlash (version 2.0) [29]. Gene prediction was made with Prodigal (version 2.6.3) [30]. Genes and proteins were annotated with DIAMOND (version 0.9.26), utilizing the “sensitive” setting and NCBI protein non-redundant database (created on 23 August 2019) [31]. Viral assemblage analysis and diversity measures were calculated in MEGAN [32] with the R package vegan [33]. vCONTACT2 [34] was used to compare viral contigs from Stn7 and construct a network visualized with Cytoscape [35]. We used the number of viral contigs identified in the microbial fraction ($>0.22 \mu\text{m}$) as a proxy to estimate the frequency of viral infections. To do so, VirSorter was used to identify putative viral contigs in the $>0.22 \mu\text{m}$ fraction metagenomes. For putative lytic infections, categories 1 and 2 were kept, while for the putative prophages, categories 4 and 5 were kept.

Environmental variables were normalized and put on the same unitless scale utilizing a z-score transformation. Viral families were identified in MEGAN and square root transformed before being plotted on the non-metric multidimensional scaling plot (NMDS). The NMDS was plotted with the R packages vegan and ggplot2 [36]. Due to the lack of complete metadata for the pre-Harvey samples, the NMDS only includes the post-storm samples. Diversity analysis (Shannon index, Pielou index, and ACE index) of viral metagenomes was done in R-Studio version 4.0.3 with packages from fossil version 0.4.0 [37], vegan version 2.5.7 [33], cluster version 2.1.2 [38], ggplots version 3.1.1 [39], pvclust version 2.2.0 [40], metagenomeSeq version 1.32.0 [41], and RCurl version 1.98.1.5 [42]. Analysis was done, and figures were made following an R script previously published [43].

2.4. Identification of Putative AMGs

VIBRANT (version 1.2.0) was used to identify and annotate viral contigs in microbial ($>0.22 \mu\text{m}$) and viral ($<0.22 \mu\text{m}$) metagenomes utilizing the -virome argument and default parameters (1000 bp sequence input and 4 ORFs per scaffold) and to find potential metabolic genes, or AMGs [44]. AMGs are identified in VIBRANT by assigning functional genes to KEGG metabolic pathway maps and then further sorting into broader KEGG annotation categories. In this study, we kept only metabolic genes involved in energy and carbohydrate metabolisms that have been identified as AMGs in other studies. AMG counts were normalized to one million proteins, and heat maps were built in R Studio with ggplot2 [36].

2.5. Viral Abundance

Viral-like particles were enumerated from all the Stn7 samples and the 22 August pre-Harvey sample using epifluorescence microscopy as described in [45]. Briefly, 100 μL of virus concentrate was fixed with 37% paraformaldehyde before filtration on a 0.02 μm Anodiscs (Whatman, Marlborough, MA, USA). Filters were stained with 1:400 SYBR Green I, and 10% (*w/v*) *p*-phenylenediamine dihydrochloride was used as an anti-fade. At least 20 fields of view were counted, and the average was used to calculate the concentration per ml of virus concentrate. Viral abundance in the original sample was estimated based on the concentration factor. It is worth noting that the concentrations are only estimates because there may be loss during tangential flow filtration virus concentration [46,47], and we did not have preserved samples from the original samplings.

2.6. Phylogenetic Analysis

All gokushovirus genomes available in NCBI and viral contigs sharing similarity with gokushoviruses (Clusters 2 and 3) were annotated using RAST [48]. Genes annotated as “Phage major capsid protein #Fam0039” and “Phage major capsid protein” were translated into proteins and aligned using MAFFT with the E-INS-I algorithm [49,50]. Sequences containing the N-terminal half (as in [51]) were kept for further analysis. Maximum likelihood phylogenetic analysis was performed in MEGA11 [52] with the LG + I + F model as determined within MEGA11. Trees were viewed in FigTree (<https://github.com/rambaut/figtree/> (accessed on 16 November 2021)).

2.7. Data Availability

All metagenomes are publicly available in the MG-RAST metagenomics analysis server [53] and NCBI SRA Archive (accession numbers listed in Supplementary Table S3).

3. Results and Discussion

Models predict that rainfall rates associated with tropical cyclones will increase by 10–15% in the coming years, and these cyclones will intensify [54]. With viruses playing a vital role in marine geochemical cycles by recycling organic nutrients via the lysis of plankton in pelagic [9,10,55] and coastal environments [56,57], it is important to understand how viruses are impacted by such disturbances and the consequences of that impact. Two other studies evaluated the impacts of heavy rainfall on viral communities and showed that stormwater runoff can significantly change the composition of the viral assemblages in freshwater retention ponds [19,20]. In this study, we provide one of the first looks at how heavy rainfall events can impact marine viral assemblages, both at the taxonomic and functional level, in estuarine environments. Utilizing Hurricane Harvey as a pulse disturbance, we can evaluate how the viruses in other coastal bays similar in structure to Galveston Bay (i.e., riverine input, hurricane-prone, coastal locations), such as those located along the Gulf of Mexico coast, could be impacted by large rain events.

3.1. Hurricane Harvey Changed the Viral Community Diversity and Composition

We sequenced viral metagenomes (<0.22 μm fraction) from four samplings over the five weeks following Hurricane Harvey along a four-station transect across Galveston Bay, from the San Jacinto River to the mouth of the Gulf of Mexico (Figure 1 and Supplementary Table S2). We also sequenced the viral metagenomes from two pre-Harvey samples located nearby Stn7. On average, each metagenome consisted of ~37 million reads, which resulted in an average of ~169 thousand contigs with an N50 length of 944 bp and ~38 thousand proteins (Supplementary Table S3).

Previous studies have shown that Hurricane Harvey dramatically changed Galveston Bay. The heavy rainfall, stormwater runoff, and sediment resuspension contributed to decreasing salinity, pH, and temperature and increasing total resuspended sediment, turbidity, and total nitrogen in the system [17,21,22]. A non-metric multidimensional scaling (NMDS) analysis revealed clear temporal differences between samples taken at each sampling date (Figure 2) as stations grouped together by sampling date (Figure 2). We identified correlations between environmental variables and the viral composition to determine which factor was potentially driving the changes. Salinity was the environmental driver with the strongest impact ($p < 0.001$) on the viral assemblages, followed by temperature, turbidity (Secchi depth), and pH ($p < 0.05$) (Figure 2). Total suspended sediment (TSS), total nitrogen, and dissolved oxygen had no significant impact on the viral assemblages ($p > 0.1$) (Figure 2). The diversity of viruses in the ecosystem decreased after the rain event, which is the opposite of what was observed for the microbial community (Walker et al., unpublished) [17].

The drastic environmental changes caused by Hurricane Harvey changed the viral richness and diversity in the samples (Table 1 and Supplementary Figure S1). The Shannon diversity index calculations revealed that 4 September had the lowest richness (ACE

index) and evenness (Pielou index) among the post-storm samples (Table 1). This suggests an uneven distribution of the taxa with a few dominant members. Viral diversity and richness then gradually increased throughout the study (Table 1). One hypothesis for the lower diversity is that the record amount of rainfall possibly diluted the viruses. To test whether the change in richness and evenness was the result of a potential dilution of the viruses following the rain, we counted the number of virus-like particles in each virus concentrate and estimated viral abundance in the original sample based on the concentration factor. If we assume that each sample had similar viral recovery and decay prior to counting, we observe fewer viruses after the storm, which could be due to the large input of rainwater that diluted the viruses (Supplementary Figure S2). Another explanation for the lower abundance of viral particles after the storm is that the inflow of water and new viruses reduced the virus–host contact rate, reducing viral infections and, therefore, viral production. It is also possible that viral concentrations were reduced due to changes in environmental parameters such as the change in temperature, presence of pollutants, or grazing that may have increased viral degradation [58].

Table 1. Diversity (Shannon), evenness (Pielou), and richness (ACE) indices of the viral metagenomes at the family level.

Dates in 2017	Station	Shannon Index	Pielou Index	ACE Index
31 July	TAMUG Boat Basin	2.01	0.59	33.73
22 August	TAMUG Boat Basin	1.57	0.45	36.91
4 September	Stn1	1.93	0.55	34.45
	Stn4	1.73	0.49	36.01
	Stn7	1.72	0.49	35.13
	Stn10	1.70	0.47	37.08
9 September	Stn1	1.86	0.51	38.47
	Stn4	2.16	0.60	37.69
	Stn7	1.92	0.54	37.44
	Stn10	2.07	0.58	36.43
16 September	Stn1	2.08	0.58	36.21
	Stn4	2.22	0.62	37.46
	Stn7	2.09	0.58	37.42
	Stn10	2.14	0.60	35.00
28 September	Stn1	2.23	0.62	38.25
	Stn4	2.08	0.58	40.54
	Stn7	1.96	0.55	40.75
	Stn10	1.85	0.53	40.51

The viral assemblages' taxonomic composition at the family level revealed drastic, short-term changes after Hurricane Harvey. Viral contigs were first compared and grouped into clusters (Figure 3). Most clusters had no similarity to reference genomes unless otherwise stated. Five of the 20 largest clusters, Clusters 15, 8, 2 (related to *Gokushovirinae*), 6, and 12, had no similarity with the viruses found prior to Hurricane Harvey (Figure 3), suggesting that the storm brought novel viruses into the system. Three viral groups, Clusters 11, 18 (related to *Myoviridae* Rhodothermus phage RM378), and 10, present before the storm, were absent on 4 September, and their abundances gradually increased as the environmental conditions returned to estuarine conditions similar to pre-Harvey (Figure 3). The group with the most contigs, Cluster 1 (sharing similarities to many Cyanophages, *Pelagibacter* phages, and other *Caudovirales* infecting marine microorganisms), was predominantly composed of contigs from 28 September and pre-Harvey metagenomes. The re-emergence of Cluster 1 on 28 September suggests the viral community was returning to that of an estuarine environment.

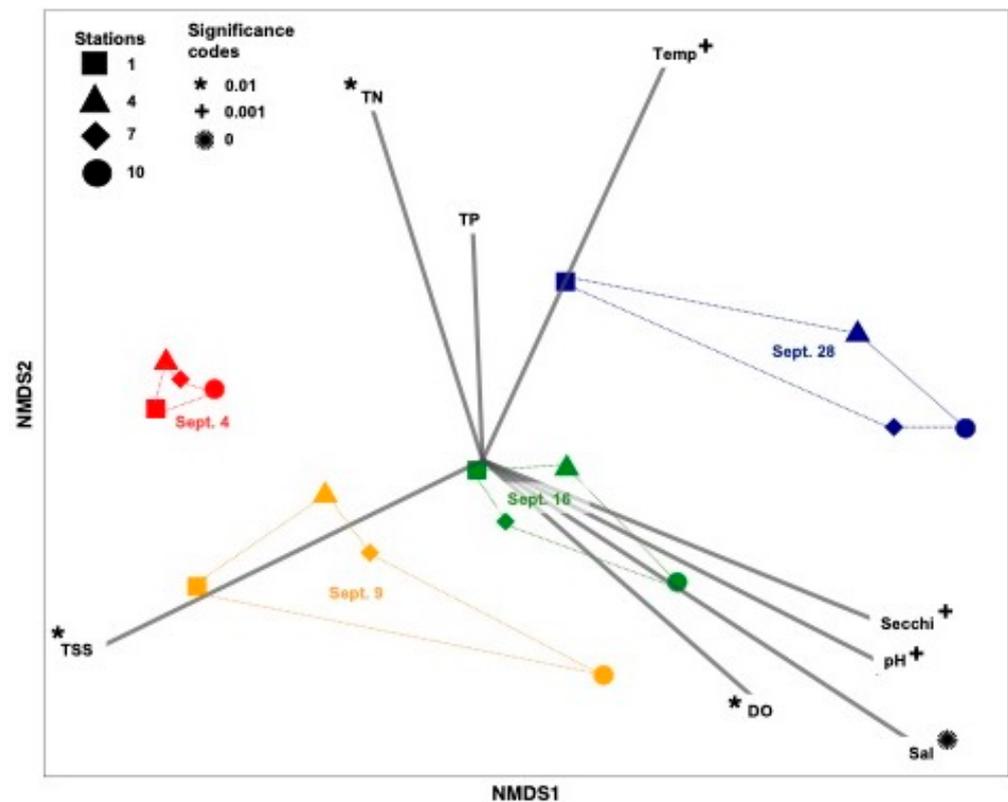


Figure 2. NMDS of the viral community composition showing the correlation between temperature (Temp; °C), salinity (Sal), pH, dissolved oxygen (DO; %), total nitrogen (TN; mM), total phosphorus (TP; mM), Secchi (m), and total suspended sediment (TSS; g mL⁻¹). All data for environmental variables were taken from [17]. 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.

We then compared the viral assemblages to the NCBI nr database to assign taxonomy. Taxonomic classification of the viral proteins also revealed a change in the viral composition immediately after the storm (Figure 4). The family *Siphoviridae* was the most abundant viral family in July, prior to Hurricane Harvey’s landfall, with a relative abundance of 26% of the total viral assemblages. Relative abundances of *Siphoviridae* fell to an average relative abundance of 0.54% on 4 September (Figures 4 and 5). Relative abundances of the family *Microviridae* increased almost three-fold following Hurricane Harvey, from an average relative abundance of 14.13% of the total viral assemblage pre-Hurricane Harvey to 45.16% on 4 September (Figure 4). Following Hurricane Harvey’s landfall, proteins representing the viral families *Siphoviridae*, *Podoviridae*, and *Myoviridae* almost disappeared from Galveston Bay but returned by 28 September. There was also an influx of proteins associated with freshwater invertebrate viruses grouped into the category Lake Sarah Associated Circular Viruses (Figure 5), a group of freshwater viruses assembled from metagenomes recovered from a New Zealand lake, Lake Sarah [59]. These freshwater viruses were very low in abundance in the viral assemblages prior to Hurricane Harvey (0.32% of the total assemblage); however, on 4 September, the proteins from this group comprised 4.40% of the total viral assemblage. The average relative abundance of proteins associated with these freshwater viruses had receded to 0.58% of the total viral assemblage by 28 September (Figure 5).

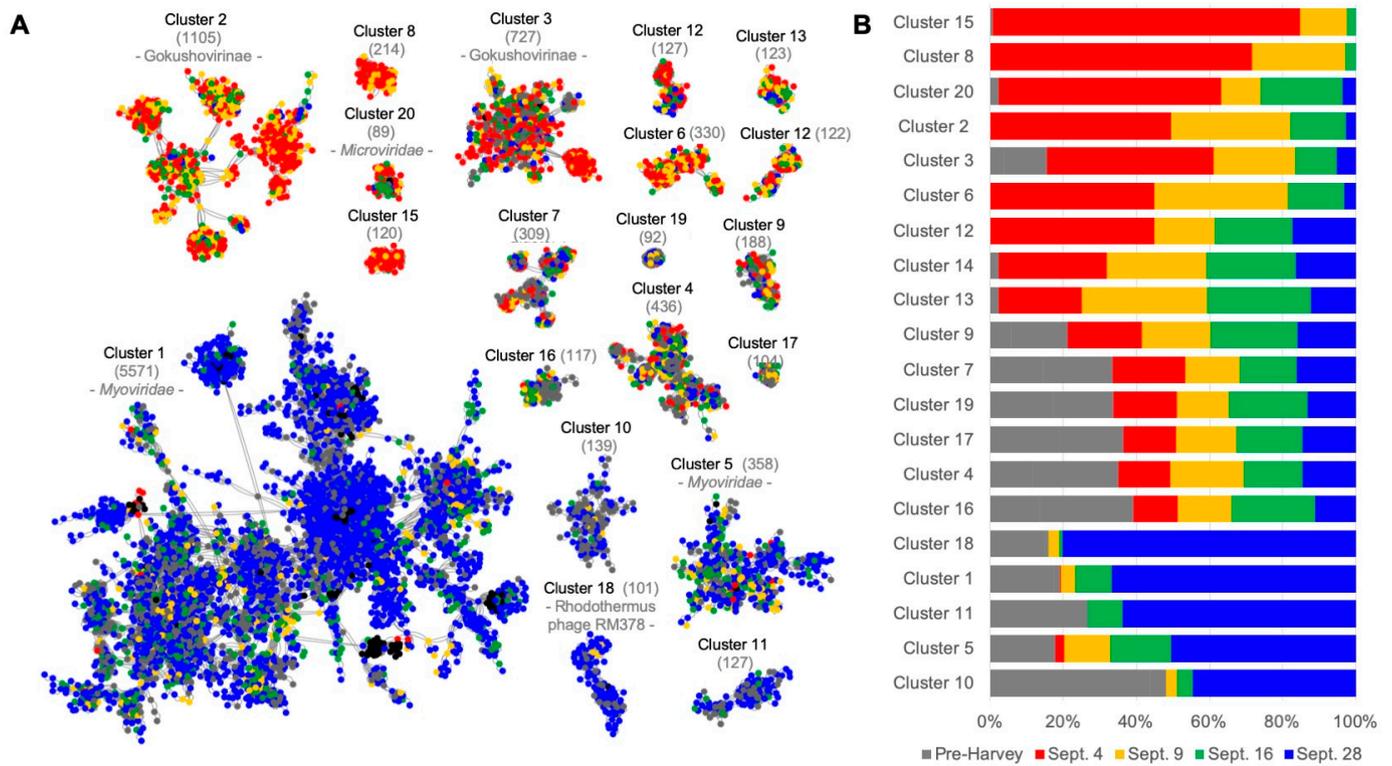


Figure 3. Comparison of the viral contigs to each other at Stn7. (A) Gene-sharing network of viral contigs generated with vConTACT v2.0 and visualized using an edge-weighted spring-embedded algorithm that locates more similar contigs closer to each other. Nodes represent contigs from 4 September (red), 9 September (yellow), 16 September (green), 28 September (blue), and pre-Harvey (grey). Clusters were named based on the number of contigs in each, 1 being the cluster with the most contigs. The number of nodes in each cluster is indicated in parentheses, and the identification of the reference genomes is indicated when available. (B) Relative abundance of the number of contigs from each sampling date in each identified cluster, ordered by the abundance of 4 September contigs (decreasing) and 28 September contigs (increasing).

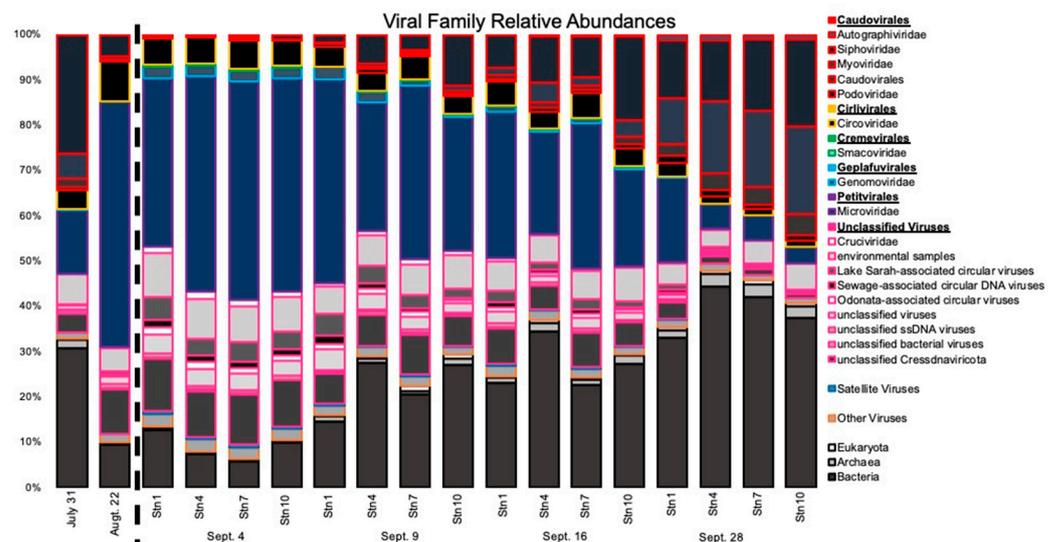


Figure 4. Stacked bar histogram showing the relative abundance of viral families identified by metagenomic viral DNA samples. Sampling points are broken up into blocks separated by white space. The dotted line separates pre- and post-Harvey samples.

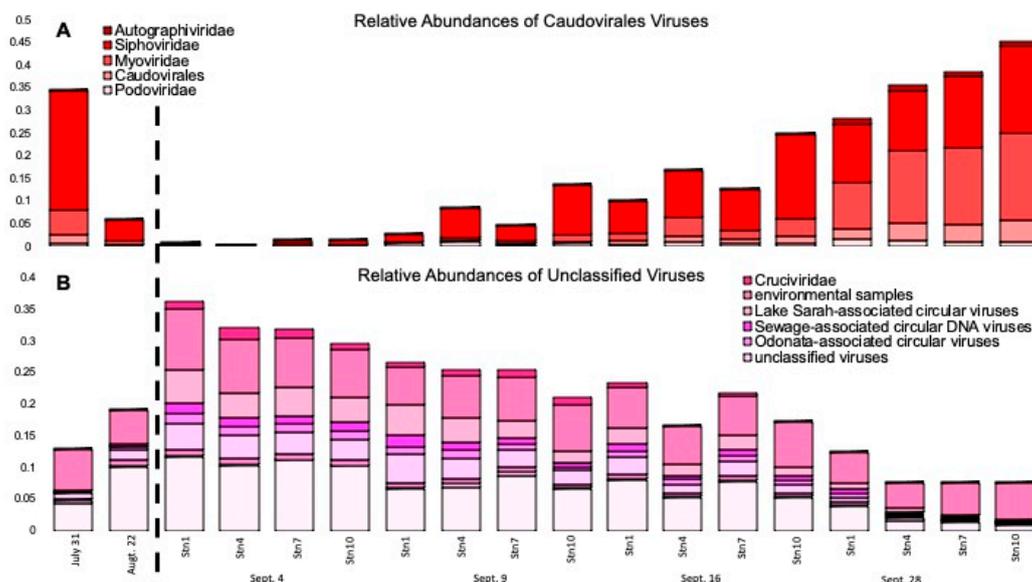


Figure 5. Stacked bar histogram showing the relative abundance of the viral family (A) Caudovirales and (B) Unclassified viruses. Both families were identified by metagenomic viral DNA samples. The dotted line separates pre- and post-Harvey samples.

The decrease in the relative abundance of *Podoviridae* and *Myoviridae* and the dominance of a single family, the *Microviridae*, could be responsible for the changes in richness and diversity. *Podoviridae* and *Myoviridae* are commonly the most abundant viral families in coastal environments [8]. They were also the most abundant families prior to the storm and in our last sampling when environmental conditions were back to estuarine conditions. Most known Cyanophages belong to the *Podoviridae* or *Myoviridae* families [60,61]. Besides Cyanobacteria, *Podoviridae* are also known to infect common marine bacteria such as *Pelagibacter* [62]. These findings are comparable to how stormwater runoff rapidly changes freshwater viral communities and is a driver in community composition [19].

While *Microviridae*, especially *Gokushovirinae* (small circular genomes of <5.5 kb), are present in our Galveston Bay samples, their relative abundances may not be accurate due to the polymerase used for whole genome amplification that has a template bias for small circular genomes, resulting in a higher number of reads associated with these genomes [63–66]. Since we had to resort to whole genome amplification to get enough DNA for metagenomic sequencing, it is possible that members of the *Microviridae* family were over-amplified. It has been well documented that gokushoviruses are widely distributed throughout the marine environment [67,68]. The widespread nature of these viruses in the marine environment makes it difficult to attribute the increase in gokushoviruses after Hurricane Harvey to the storm. We performed a phylogenetic analysis of the major capsid protein gene to look at the genetic relatedness of the gokushoviruses from this study and other environments and clinical isolates (Supplementary Figure S3). Our analysis shows that the gokushoviruses from Galveston Bay are distinct from the isolates, which are phages of *Chlamydia* and *E.coli*, and distinct from most gokushoviruses isolated from animal and human guts. One cluster contained sequences observed only after Hurricane Harvey hit, suggesting that these are of terrestrial and freshwater origin and were brought into the system by the storm. Our results confirm that gokushoviruses are diverse and ubiquitous but occupy specialized niches [67].

3.2. The Number of Viruses Identified in the Microbial Fraction Escalated throughout the Weeks Following Hurricane Harvey

We identified the viral contigs from lytic infections and prophages in the microbial metagenomes (>0.22 μm ; Table 2). The lowest number of viral contigs identified in the

microbial metagenomes were observed prior to and immediately after Hurricane Harvey. Over time, the abundance of lytic viruses increased, and by 16 September, they were 3.6× more abundant than on 4 September or prior to the storm. The presence of viral contigs in the microbial fraction can indicate ongoing infections [69], ingested viruses [70], or viruses that attached to cells, particles, or filters during the filtration process. Here we used the abundance of viral contigs from lytic infections as a proxy to estimate the relative abundance of ongoing viral infections. To do so, we assume that active viruses should be more abundant than particle-associated viruses due to the high burst size of marine heterotrophic bacterial infections in marine environments (average burst size of ~24–34 viruses [71]). Moreover, there were more resuspended particles in the earlier samplings after the storm, yet the number of identified viral contigs was lower. The low incidence of lytic infections following the storm is likely the result of the heavy rainfall that reduced the virus–host encounter rate and created a disconnection between viruses and their hosts. After Hurricane Harvey, with the virus–host contact rate reduced, the novel microbial hosts introduced into the system from stormwater runoff and sediment were able to thrive. Once host abundances became high enough to increase the contact rate, viral infections increased, behaving in a “Kill-The-Winner” fashion [72]. An increase in lytic infections will likely increase the viral shunt, accelerating nutrient recycling [8,11]. It has been documented that Hurricane Harvey altered carbon exchange in the marine environment [73]. The contribution of viruses to this increase in the carbon cycle may be greater than previously thought and may have long-term effects on the ecosystem.

Table 2. Total number of lytic viruses and prophages, normalized to 10,000 contigs, recovered from microbial metagenomes.

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

The abundance of identified prophages also increased after Hurricane Harvey. The increase started on 4 September for Stns 7 and 10, increased leading to 9 September, and peaked on 16 September. Only the freshwater stations had prophage-associated sequences by 28 September. No metabolic genes were identified within the prophages, which is probably due to the low numbers of identified prophages (536 total contigs) compared to the total number of contigs in the viral metagenomes (160,999 contigs). However, an increased incidence of prophages could be linked to the increase in observed metabolic genes. Indeed, viruses have been shown to act as a great source of available functions that can be transferred to their hosts via transduction [2].

We taxonomically classified the viral proteins identified within the microbial metagenomes (Supplementary Figures S4 and S5). For the lytic viruses, a majority of the proteins were classified as cellular organisms or as members of the order Caudovirales (Supplementary Figure S4). Members of the order Caudovirales contain some of the most abundant viruses identified in the ocean [8,74] and infected cells [75]. The taxonomic distribution differed greatly from the viral proteins identified in the viral metagenomes, which was most likely due to the small number of contigs identified and the use of whole-genome amplification on the viral DNA, which skews viral concentrations. However, the identification of single-stranded DNA viruses in the microbial metagenomes, which have not been subjected to whole genome amplification, confirms that the sequences identified in the viral metagenomes are not artifacts. As for the prophages, after the storm, most of the prophages carried a majority of genes related to their bacterial hosts (i.e., Verrucomicrobia, Comamonadaceae, Rhodospirillaceae, and Rhodobacteraceae), which were related to the terrestrial and freshwater microbes brought in by the flood and storm waters [17]. In the later samplings, the prophages carried, almost exclusively, genes related to the order Caudovirales.

3.3. Hurricane Harvey Increased the Abundance of Virally Encoded Metabolic Genes Involved in Major Biogeochemical Cycles

To determine the potential impact of viral infections on the geochemical cycles of Galveston Bay, we searched for metabolic genes in the viral (<0.22 μm) and microbial (>0.8–0.22 μm) metagenomes. The density of metabolic genes was more than one order of magnitude more in viral metagenomes than in microbial metagenomes. While the metabolic gene concentration was higher in viruses, their detection was delayed by one sampling for each station (e.g., metabolic genes involved in energy metabolisms were identified first on 16 September and peaked on 28 September in the microbial fraction, but only on 28 September for viruses). Metabolic genes associated with the metabolisms of cofactors and vitamins, amino acids, and carbohydrates were the most abundant (Figure 6). The abundance of metabolic genes associated with energy metabolism follows a similar trend to the ones found in the microbial fraction. The metabolic genes involved in energy and carbohydrates were barely detectable prior to Hurricane Harvey, then increased over time to display the highest abundances in the later samples, five weeks following the storm (A in Figure 6). While the microbial fraction displayed the highest concentrations of metabolic genes on 16 September at Stn4 and Stn7, and on 28 September, at all of the stations, viruses were higher only on 28 September, suggesting that the viral cycle is delayed compared to the microbial fraction or their putative host population. This trend is most obvious with carbohydrate metabolism, which includes pathways associated with glycolysis, pentose phosphate, and amino and nucleotide sugar metabolism (Supplementary Figure S6). Metabolic genes associated with energy metabolism, such as photosynthesis, methane, and sulfur, dropped after Hurricane Harvey's landfall (A and B in Figure 6). Metabolic genes associated with sulfur metabolism were the first to appear consistently in the viral metagenomes on 9 September at Stn10 (B in Figure 6), which is congruent with the sulfur metabolic genes observed in the microbial metagenomes in all stations on 9 September (A in Figure 6).

Viral AMGs reflecting host metabolisms and adaptations have previously been documented in different datasets, including in the Pacific Ocean Virome database, where it was shown that viral AMGs are depth-stratified based on the host distribution [76]. A similar trend where viruses follow the cellular population distribution can also be observed in our dataset, especially when looking at the genes associated with photosynthesis (A and B in Figure 6); therefore, we hypothesize that our metabolic genes are, in fact, auxiliary metabolic genes (AMGs) although we are unable to confirm. Photosynthetic microbes, such as Cyanobacteria, were depleted from Galveston Bay following Hurricane Harvey [17]. Viral metabolic genes associated with the photosynthetic pathway were detected only once the microbial community, mainly consisting of Cyanobacteria, had recovered on and after

16 September [17] (B in Figure 6). The loss of primary producers is also corroborated with a decrease in surface chlorophyll *a* following Hurricane Harvey in other coastal Texas bays [22].

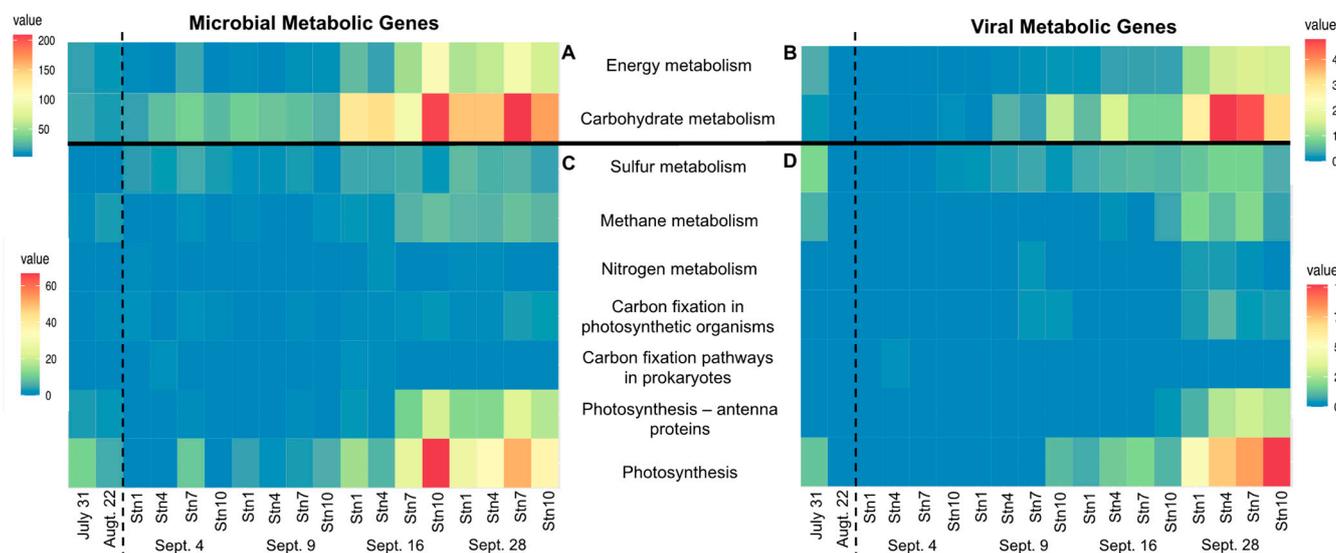


Figure 6. Heat map showing the density of metabolic genes, normalized to one million proteins, recovered from microbial metagenomes (A, C) and viral metagenomes (B, D). A and B show a breakdown of metabolic genes associated with all metabolism pathways in microbial and viral metagenomes, respectively. C and D show a breakdown of metabolic genes associated with energy metabolisms in microbial and viral metagenomes, respectively. In the viral metagenomes on 4 September at Stn1 and Stn7 and on 22 August, no metabolic genes were found. The dotted line separates pre- and post-Harvey samples.

There was a slight increase in metabolic genes associated with sulfur metabolisms following Hurricane Harvey (B in Figure 6). Genes in this pathway are similar in function to *cysP*, which binds to thiosulfate and aids in transmembrane transport [77]. This increase in metabolic genes associated with sulfur metabolisms is hypothesized to be linked to the sediment resuspension seen in Galveston Bay following Hurricane Harvey. The resuspension of sediment is confirmed by the highest concentrations of total suspended sediment occurring during time points following Hurricane Harvey's landfall [17] and documented heavy sediment loading [21]. A significant part of the microbial community within marine sediment utilizes sulfur compounds as an energy source [78]. Metabolic genes associated with sulfur metabolisms have previously been observed in other environments, including the epipelagic zone [7] and the deep-sea [79], suggesting that viruses may play an important role in sulfur cycling.

Within the identified pathways associated with carbohydrate metabolism, five had obvious changes following Hurricane Harvey's landfall: the pentose phosphate pathway, pentose and glucuronate interconversions, galactose metabolism, fructose and mannose metabolisms, and amino and nucleotide sugar metabolisms (Supplementary Figure S6). Genes represented in these metabolism pathways include *pgi*, *galM*, and *mak*, which are involved in catalytic activity; *kduI*, which substitutes the normal hexuronate degrading enzyme under osmotic stress; and *nagZ*, which is involved in peptidoglycan recycling [80]. Hurricane Harvey floodwaters caused a large influx of dissolved organic carbon (DOC) within Galveston Bay [17,73]. The influx of DOC to the bay from the floodwaters could be the cause of the increase in metabolic genes associated with carbohydrate metabolisms (Supplementary Figure S6). Previous work done in the Arctic Ocean found that a decrease in total dissolved carbohydrates accounted for half of the DOC decrease seen within the upper 200 m of the ocean [81], linking carbohydrates and DOC. Similarly, an increase in

DOC in Galveston Bay could cause an increase in carbohydrates, which would account for the increase in heterotrophs and microorganisms able to degrade carbohydrates. Since virally encoded metabolic genes are typically also found within their hosts [3], the infection of these microorganisms could lead to an increase in viruses carrying metabolic genes from carbohydrate degradation pathways. Moreover, Hurricane Harvey caused an acceleration of the DOC cycle in Galveston Bay via higher rates of mineralization of terrigenous DOC, which was linked to the active microbial community [73]. Our results show that this acceleration could be exacerbated by viruses, as is shown by the increase in viral metabolic genes associated with carbohydrate metabolisms, which are orders of magnitude higher post-Hurricane Harvey compared to the pre-storm samples (Supplementary Figure S6).

4. Conclusions

It is projected that rainfall rates associated with tropical cyclones will increase in the coming years due to climate changes [54], making these intense rainfall events more common. It is important to understand how viral assemblages, which are linked to their hosts, respond to intense rain events. We provided the first in-depth look at marine viral assemblages' impact and recovery from intense rainfall. Immediately after the storm, viral communities were altered by the introduction of novel taxa, a decreased viral abundance, and a decreased frequency of viral infections, all of which resulted in lower viral diversity. There was an increase in AMGs associated with carbon, nitrogen, and sulfur metabolisms that may have aided the host community in the unfavorable environmental conditions brought by the storm; highlighting how viruses can play an important role in ecosystem recovery from large scale rain events or other pulse disturbances. After five weeks, there was greater representation of virus-associated sequences in the microbial fraction than prior to or just after the storm, suggesting that the impact of the storm on the geochemical cycles in the long term may be more important than previously thought. Future studies should quantify how the viral load is impacted by rainfall to further understand the impact of storms on viruses and their role in ecosystem recovery.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/microorganisms10040729/s1>, Supplementary figure and Tables (docx); Table S1: Metadata from Steichen et al. showing a drop in salinity, temperature, and dissolved oxygen following Hurricane Harvey; Table S2: Types of filters used and volume filtered for each sample; Table S3: Reads and contigs statistics and accession numbers for each sequenced metagenome. Figure S1: Estimates of viral diversity (Shannon index), evenness (Pielou index) and richness (ACE index) from the viral metagenomes of (A) pre-Harvey samples (July 31 & Aug. 22) compared to the first sampling (Sept. 4) and (B) Sept. 4 sampling, compared to the last sampling effort (Sept. 28); Figure S2: Virus-like particles abundances (with standard error bars) in samples from Stn7 and the Aug. 22 pre-storm sample. Counts were performed on the viral concentrates and the total for the original samples are estimated based on the concentration factor, assuming a 100% concentration efficiency.; Figure S3: Unrooted phylogenetic analysis (maximum likelihood; LG + I + F model; 100 bootstrap replicates) of the gokushovirus major capsid protein (Sept. 4 in red, Sept. 9 in orange, Sept. 16 in green, and Sept. 28 in blue, pre-Hurricane Harvey in grey, and isolates in black). Bootstrap support is represented by a gradient from white (0% support) to black (100% support). The large black clade solely contains sequences from isolates from animal guts. The highlighted grey clade contains sequences from post-Hurricane Harvey only.; Figure S4: Stacked bar histogram showing the relative abundance of the lytic viruses identified in the host metagenomes. The dotted line separates Pre- and Post-Harvey samples. Figure S5: Stacked bar histogram showing the relative abundance of the prophages identified in the host metagenomes. The dotted line separates Pre- and Post-Harvey samples.; Figure S6: Heat map showing the density of auxiliary metabolic genes (AMGs) related to carbohydrate metabolisms, normalized to one million proteins, recovered from microbial metagenomes (A) and viral metagenomes (B). July 31 and Aug. 22 were sampled prior to Hurricane Harvey's landfall.

Author Contributions: Conceptualization, J.M.L.; methodology, C.D.J., J.R.W., A.C.W. and J.M.L.; bioinformatics analyses, A.C.W. and J.R.W.; formal analysis, A.C.W. and J.M.L.; data curation, A.C.W.; writing—original draft preparation, A.C.W.; writing—review and editing, A.C.W. and J.M.L.; visualization, A.C.W. and J.M.L.; supervision, J.M.L.; project administration and funding acquisition, J.M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation—RAPID Response (Award #1801367).

Data Availability Statement: All metagenomes are publicly available in the MG-RAST metagenomics analysis server [53] and NCBI SRA Archive (accession numbers listed in Supplementary Table S3).

Acknowledgments: We would like to thank Quigg, Kaiser, Hala, and Hui at Texas A&M University at Galveston (TAMUG) for assisting in cruise plans and organization, sample collection, and processing. Many thanks to the crew of the TAMUG R/V Trident and to everyone from TAMUG that came along (with very short notice) to assist in sample collection, sample processing, and data analysis. A special thank you to Jed Fuhrman for graciously sending us filters when none could be ordered after the storm. Thank you to all the members of the Labonté Viral Ecology Lab that assisted with filtration and virus concentration. Environmental data were provided by Quigg and collected via funds from Texas Parks and Wildlife and Environmental TPWD and the U.S. Environmental Protection Agency by the Phytoplankton Dynamics Lab.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Suttle, C.A. Viruses in the sea. *Nature* **2005**, *437*, 356–361. [[CrossRef](#)] [[PubMed](#)]
2. Tsiola, A.; Fodelianakis, S.; Karakassis, I.; Simboura, N.; Da, D.; Tsapakis, M. Viral Metagenomic Content Reflects Seawater Ecological Quality in the Coastal Zone. *Viruses* **2020**, *12*, 806. [[CrossRef](#)] [[PubMed](#)]
3. Warwick-Dugdale, J.; Buchholz, H.H.; Allen, M.J.; Temperton, B. Host-hijacking and planktonic piracy: How phages command the microbial high seas. *Virol. J.* **2019**, *16*, 15. [[CrossRef](#)]
4. Mojica, K.D.A.; Huisman, J.; Wilhelm, S.W.; Brussaard, C.P.D. Latitudinal variation in virus-induced mortality of phytoplankton across the North Atlantic Ocean. *ISME J.* **2016**, *10*, 500–513. [[CrossRef](#)]
5. Suttle, C.A. The Significance of Viruses to Mortality in Aquatic Microbial Communities. *Microbiol. Ecol.* **1994**, *28*, 237–243. [[CrossRef](#)]
6. Jover, L.F.; Effler, T.C.; Buchan, A.; Wilhelm, S.W.; Weitz, J.S. The elemental composition of virus particles: Implications for marine biogeochemical cycles. *Nat. Rev. Microbiol.* **2014**, *12*, 519–528. [[CrossRef](#)]
7. Roux, S.; Brum, J.R.; Dutilh, B.E.; Sunagawa, S.; Duhaime, M.B.; Loy, A.; Poulos, B.T.; Solonenko, N.; Lara, E.; Poulain, J.; et al. Ecogenomics and biogeochemical impacts of uncultivated globally abundant ocean viruses. *bioRxiv* **2016**, *537*, 053090. [[CrossRef](#)]
8. Suttle, C.A. Marine viruses—Major players in the global ecosystem. *Nat. Rev. Microbiol.* **2007**, *5*, 801–812. [[CrossRef](#)]
9. Wilhelm, S.W.; Suttle, C.A. Viruses and Nutrient Cycles in the Sea. *Bioscience* **1999**, *49*, 781–788. [[CrossRef](#)]
10. Weitz, J.S.; Stock, C.A.; Wilhelm, S.W.; Bourouiba, L.; Coleman, M.L.; Buchan, A.; Follows, M.J.; Fuhrman, J.A.; Jover, L.F.; Lennon, J.T.; et al. A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *ISME J.* **2015**, *9*, 1352–1364. [[CrossRef](#)]
11. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. *Nature* **1999**, *399*, 541–548. [[CrossRef](#)] [[PubMed](#)]
12. Rosenwasser, S.; Ziv, C.; van Creveld, S.G.; Vardi, A. Virocell Metabolism: Metabolic Innovations During Host-Virus Interactions in the Ocean. *Trends Microbiol.* **2016**, *24*, 821–832. [[CrossRef](#)] [[PubMed](#)]
13. Sullivan, M.B.; Lindell, D.; Lee, J.A.; Thompson, L.R.; Bielawski, J.P.; Chisholm, S.W. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol.* **2006**, *4*, 1344–1357. [[CrossRef](#)] [[PubMed](#)]
14. Goldsmith, D.B.; Crosti, G.; Dwivedi, B.; McDaniel, L.D.; Varsani, A.; Suttle, C.A.; Weinbauer, M.G.; Sandaa, R.A.; Breitbart, M. Development of phoH as a novel signature gene for assessing marine phage diversity. *Appl. Environ. Microbiol.* **2011**, *77*, 7730–7739. [[CrossRef](#)]
15. Crummett, L.T.; Puxty, R.J.; Weihe, C.; Marston, M.F.; Martiny, J.B.H. The genomic content and context of auxiliary metabolic genes in marine cyanomyoviruses. *Virology* **2016**, *499*, 219–229. [[CrossRef](#)]
16. Amaral-Zettler, L.A.; Rocca, J.D.; Lamontagne, M.G.; Dennett, M.R.; Gast, R.J. Changes in microbial community structure in the wake of Hurricanes Katrina and Rita. *Environ. Sci. Technol.* **2008**, *42*, 9072–9078. [[CrossRef](#)]
17. Steichen, J.L.; Labonté, J.M.; Windham, R.; Hala, D.; Kaiser, K.; Setta, S.; Faulkner, P.C.; Bacosa, H.; Yan, G.; Kamalanathan, M.; et al. Microbial, Physical, and Chemical Changes in Galveston Bay Following an Extreme Flooding Event, Hurricane Harvey. *Front. Mar. Sci.* **2020**, *7*. [[CrossRef](#)]

18. Mason, O.U.; Hazen, T.C.; Borglin, S.; Chain, P.S.G.; Dubinsky, E.A.; Fortney, J.L.; Han, J.; Holman, H.Y.N.; Hultman, J.; Lamendella, R.; et al. Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. *ISME J.* **2012**, *6*, 1715–1727. [[CrossRef](#)]
19. Williamson, K.E.; Harris, J.V.; Green, J.C.; Rahman, F.; Chambers, R.M. Stormwater runoff drives viral community composition changes in inland freshwaters. *Front. Microbiol.* **2014**, *5*, 105. [[CrossRef](#)]
20. Green, J.C.; Rahman, F.; Saxton, M.A.; Williamson, K.E. Quantifying aquatic viral community change associated with stormwater runoff in a wet retention pond using metagenomic time series data. *Aquat. Microb. Ecol.* **2018**, *81*, 19–35. [[CrossRef](#)]
21. Du, J.; Park, K.; Dellapenna, T.M.; Clay, J.M. Dramatic hydrodynamic and sedimentary responses in Galveston Bay and adjacent inner shelf to Hurricane Harvey. *Sci. Total Environ.* **2019**, *653*, 554–564. [[CrossRef](#)] [[PubMed](#)]
22. Walker, L.M.; Montagna, P.A.; Hu, X.; Wetz, M.S. Timescales and Magnitude of Water Quality Change in Three Texas Estuaries Induced by Passage of Hurricane Harvey. *Estuaries Coasts* **2021**, *44*, 960–971. [[CrossRef](#)]
23. Suttle, C.A.; Chan, A.M.; Cottrell, M.T. Use of ultrafiltration to isolate viruses from seawater which are pathogens of marine phytoplankton. *Appl. Environ. Microbiol.* **1991**, *57*, 721–726. [[CrossRef](#)] [[PubMed](#)]
24. Green, M.R.; Sambrook, J. Isolation of high-molecular-weight DNA using organic solvents. *Cold Spring Harb. Protoc.* **2017**, *2017*, 356–359. [[CrossRef](#)]
25. Bushnell, B.; Rood, J.; Singer, E. BBMerge—Accurate paired shotgun read merging via overlap. *PLoS ONE* **2017**, *12*, e0185056. [[CrossRef](#)] [[PubMed](#)]
26. Li, D.; Liu, C.M.; Luo, R.; Sadakane, K.; Lam, T.W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **2015**, *31*, 1674–1676. [[CrossRef](#)] [[PubMed](#)]
27. Roux, S.; Enault, F.; Hurwitz, B.L.; Sullivan, M.B. VirSorter: Mining viral signal from microbial genomic data. *PeerJ* **2015**, *3*, e985. [[CrossRef](#)]
28. Guo, J.; Bolduc, B.; Zayed, A.A.; Varsani, A.; Dominguez-huerta, G.; Delmont, T.O.; Pratama, A.A.; Gazitúa, M.C.; Vik, D.; Sullivan, M.B.; et al. VirSorter2: A multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **2021**, *9*, 37. [[CrossRef](#)]
29. Gruber-Vodicka, H.R.; Seah, B.K.B.; Pruesse, E. phyloFlash: Rapid Small-Subunit rRNA Profiling and Targeted Assembly from Metagenomes. *mSystems* **2020**, *5*. [[CrossRef](#)]
30. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **2010**, *11*, 119. [[CrossRef](#)]
31. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2014**, *12*, 59–60. [[CrossRef](#)] [[PubMed](#)]
32. Huson, D.H.; Beier, S.; Flade, I.; Górska, A.; El-Hadidi, M.; Mitra, S.; Ruscheweyh, H.-J.; Tappu, R. MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* **2016**, *12*, e1004957. [[CrossRef](#)] [[PubMed](#)]
33. Oksanen, J.; Blanchet, F.G.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P.R.; O’Hara, R.B.; Simpson, G.L.; Solymos, P.; et al. vegan: Community Ecology Package 2020. Available online: <https://cran.r-project.org/web/packages/vegan/index.html> (accessed on 6 March 2022).
34. Bolduc, B.; Jang, H.B.; Doucier, G.; You, Z.Q.; Roux, S.; Sullivan, M.B. vConTACT: An iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. *PeerJ* **2017**, *2017*, e3243. [[CrossRef](#)]
35. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)]
36. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; ISBN 978-3-319-24277-4.
37. Vavrek, M.J. fossil: Palaeoecological and palaeogeographical analysis tools. *Palaeontol. Electron.* **2011**, *14*. Available online: https://www.uv.es/pe/2011_1/238/238.pdf (accessed on 6 March 2022).
38. Maechler, M.; Rousseeuw, P.; Struyf, A.; Hubert, M.; Hornik, K. cluster: Cluster Analysis Basics and Extensions 2021. Available online: <https://cran.r-project.org/web/packages/cluster/citation.html> (accessed on 6 March 2022).
39. Warnes, G.R.; Bolker, B.; Bonebakker, L.; Gentleman, R.; Liaw, W.H.A.; Lumley, T.; Maechler, M.; Magnusson, A.; Moeller, S.; Schwartz, M.; et al. *Package “ggplots”: Various R Programming Tools for Plotting Data*, R Packag. version 2.17.0. 2016; pp. 1–68.
40. Suzuki, R.; Terada, Y.; Shimodaira, H. Package ‘pvclust’. 2019.
41. Paulson, J.N.; Colin Stine, O.; Bravo, H.C.; Pop, M. Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* **2013**, *10*, 1200–1202. [[CrossRef](#)] [[PubMed](#)]
42. Lang, D.T. *The CRAN team RCurl: General Network (HTTP/FTP/...) Client Interface for R*. R package version 1.95-4.7. 2015.
43. Vázquez-Castellanos, J.F. Diversity analysis in viral metagenomes. *Methods Mol. Biol.* **2018**, *1838*, 203–230. [[CrossRef](#)] [[PubMed](#)]
44. Kieft, K.; Zhou, Z.; Anantharaman, K. VIBRANT: Automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* **2020**, *8*, 90. [[CrossRef](#)]
45. Patel, A.; Noble, R.T.; Steele, J.A.; Schwabach, M.S.; Hewson, I.; Fuhrman, J.A. Virus and prokaryote enumeration from planktonic aquatic environments by epifluorescence microscopy with SYBR Green I. *Nat. Protoc.* **2007**, *2*, 269–276. [[CrossRef](#)]
46. Cai, L.; Yang, Y.; Jiao, N.; Zhang, R. Evaluation of tangential flow filtration for the concentration and separation of bacteria and viruses in contrasting marine environments. *PLoS ONE* **2015**, *10*, e0136741. [[CrossRef](#)]

47. John, S.G.; Mendez, C.B.; Deng, L.; Poulos, B.; Kauffman, A.K.M.; Kern, S.; Brum, J.; Polz, M.F.; Boyle, E.A.; Sullivan, M.B. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ. Microbiol. Rep.* **2011**, *3*, 195–202. [[CrossRef](#)] [[PubMed](#)]
48. Aziz, R.K.; Bartels, D.; Best, A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid annotations using subsystems technology. *BMC Genom.* **2008**, *9*, 75. [[CrossRef](#)] [[PubMed](#)]
49. Katoh, K. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
50. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
51. Labonté, J.M.; Hallam, S.J.; Suttle, C.A. Previously unknown evolutionary groups dominate the ssDNA gokushoviruses in oxic and anoxic waters of a coastal marine environment. *Front. Microbiol.* **2015**, *6*, 315. [[CrossRef](#)]
52. Kumar, S.; Stecher, G.; Li, M.; Nnyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
53. Keegan, K.P.; Glass, E.M.; Meyer, F. *MG-RAST, a Metagenomics Service for Analysis of Microbial Community Structure and Function BT—Microbial Environmental Genomics (MEG)*; Martin, F., Uroz, S., Eds.; Springer: New York, NY, USA, 2016; pp. 207–233. ISBN 978-1-4939-3369-3.
54. Knutson, T.R.; Sirutis, J.J.; Zhao, M.; Tuleya, R.E.; Bender, M.; Vecchi, G.A.; Villarini, G.; Chavas, D. Global projections of intense tropical cyclone activity for the late twenty-first century from dynamical downscaling of CMIP5/RCP4.5 scenarios. *J. Clim.* **2015**, *28*, 7203–7224. [[CrossRef](#)]
55. Weitz, J.S.; Wilhelm, S.W. Ocean viruses and their effects on microbial communities and biogeochemical cycles. *F1000 Biol. Rep.* **2012**, *4*, 17. [[CrossRef](#)]
56. Winget, D.M.; Wommack, K.E. Diel and daily fluctuations in virioplankton production in coastal ecosystems. *Environ. Microbiol.* **2009**, *11*, 2904–2914. [[CrossRef](#)]
57. Wilhelm, S.W.; Brigden, S.M.; Suttle, C.A. A dilution technique for the direct measurement of viral production: A comparison in stratified and tidally mixed coastal waters. *Microb. Ecol.* **2002**, *43*, 168–173. [[CrossRef](#)]
58. Mojica, K.D.A.; Brussaard, C.P.D. Factors affecting virus dynamics and microbial host-virus interactions in marine environments. *FEMS Microbiol. Ecol.* **2014**, *89*, 495–515. [[CrossRef](#)] [[PubMed](#)]
59. Dayaram, A.; Galatowitsch, M.L.; Argüello-Astorga, G.R.; van Bysterveldt, K.; Kraberger, S.; Stainton, D.; Harding, J.S.; Roumagnac, P.; Martin, D.P.; Lefeuvre, P.; et al. Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. *Infect. Genet. Evol.* **2016**, *39*, 304–316. [[CrossRef](#)] [[PubMed](#)]
60. Ignacio-Espinoza, J.C.; Sullivan, M.B. Phylogenomics of T4 cyanophages: Lateral gene transfer in the “core” and origins of host genes. *Environ. Microbiol.* **2012**, *14*, 2113–2126. [[CrossRef](#)]
61. Sullivan, M.B.; Coleman, M.L.; Weigele, P.; Rohwer, F.; Chisholm, S.W. Three Prochlorococcus cyanophage genomes: Signature features and ecological interpretations. *PLoS Biol.* **2005**, *3*, 0790–0806. [[CrossRef](#)] [[PubMed](#)]
62. Zhao, Y.; Temperton, B.; Thrash, J.C.; Schwabach, M.S.; Vergin, K.L.; Landry, Z.C.; Ellisman, M.; Deerinck, T.; Sullivan, M.B.; Giovannoni, S.J. Abundant SAR11 viruses in the ocean. *Nature* **2013**, *494*, 357–360. [[CrossRef](#)]
63. Roux, S.; Enault, F.; Robin, A.; Ravet, V.; Personnic, S.; Theil, S.; Colombet, J.; Sime-Ngando, T.; Debroas, D. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS ONE* **2012**, *7*, e33641. [[CrossRef](#)]
64. Labonté, J.M.; Suttle, C.A. Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J.* **2013**, *7*, 2169–2177. [[CrossRef](#)]
65. Angly, F.E.; Felts, B.; Breitbart, M.; Salamon, P.; Edwards, R.A.; Carlson, C.; Chan, A.M.; Haynes, M.; Kelley, S.; Liu, H.; et al. The marine viromes of four oceanic regions. *PLoS Biol.* **2006**, *4*, 2121–2131. [[CrossRef](#)]
66. Marine, R.; McCarren, C.; Vorrasane, V.; Nasko, D.; Crowgey, E.; Polson, S.W.; Wommack, K.E. Caught in the middle with multiple displacement amplification: The myth of pooling for avoiding multiple displacement amplification bias in a metagenome. *Microbiome* **2014**, *2*, 3. [[CrossRef](#)]
67. Labonté, J.M.; Suttle, C.A. Metagenomic and whole-genome analysis reveals new lineages of gokushoviruses and biogeographic separation in the sea. *Front. Microbiol.* **2013**, *4*, 404. [[CrossRef](#)]
68. Tucker, K.P.; Parsons, R.; Symonds, E.M.; Breitbart, M. Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. *ISME J.* **2011**, *5*, 822–830. [[CrossRef](#)]
69. Williamson, S.J.; Rusch, D.B.; Yooshef, S.; Halpern, A.L.; Heidelberg, K.B.; Glass, J.I.; Pfannkoch, C.A.; Fadrosch, D.; Miller, C.S.; Sutton, G.; et al. The sorcerer II global ocean sampling expedition: Metagenomic characterization of viruses within aquatic microbial samples. *PLoS ONE* **2008**, *3*, e1456. [[CrossRef](#)] [[PubMed](#)]
70. Brown, J.M.; Labonté, J.M.; Brown, J.; Record, N.R.; Poulton, N.J.; Sieracki, M.E.; Logares, R.; Stepanauskas, R. Single Cell Genomics Reveals Viruses Consumed by Marine Protists. *Front. Microbiol.* **2020**, *11*, 524828. [[CrossRef](#)]
71. Parada, V.; Herndl, G.J.; Weinbauer, M.G. Viral burst size of heterotrophic prokaryotes in aquatic systems. *J. Mar. Biol. Assoc. U. K.* **2006**, *86*, 613–621. [[CrossRef](#)]
72. Thingstad, T.F. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol. Oceanogr.* **2000**, *45*, 1320–1328. [[CrossRef](#)]

73. Yan, G.; Labonté, J.M.; Quigg, A.; Kaiser, K. Hurricanes Accelerate Dissolved Organic Carbon Cycling in Coastal Ecosystems. *Front. Mar. Sci.* **2020**, *7*. [[CrossRef](#)]
74. Gregory, A.C.; Zayed, A.A.; Sunagawa, S.; Wincker, P.; Sullivan, M.B.; Ferland, J.; Kandels, S.; Liu, Y.; Marec, C.; Vik, D. Marine DNA Viral Macro- and Microdiversity from Pole to Pole Article Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell* **2019**, *177*, 1109–1123. [[CrossRef](#)]
75. Labonté, J.M.; Swan, B.K.; Poulos, B.; Luo, H.; Koren, S.; Hallam, S.J.; Sullivan, M.B.; Woyke, T.; Eric Wommack, K.; Stepanauskas, R. Single-cell genomics-based analysis of virus-host interactions in marine surface bacterioplankton. *ISME J.* **2015**, *9*, 2386–2399. [[CrossRef](#)]
76. Hurwitz, B.L.; Brum, J.R.; Sullivan, M.B. Depth-stratified functional and taxonomic niche specialization in the “core” and “flexible” Pacific Ocean Virome. *ISME J.* **2015**, *9*, 472–484. [[CrossRef](#)] [[PubMed](#)]
77. Hryniewicz, M.; Sirko, A.; Palucha, A.; Bock, A.; Hulanicka, D. Sulfate and thiosulfate transport in *Escherichia coli* K-12: Identificaton of a gene encoding a novel protein involved in thiosulfate binding. *J. Bacteriol.* **1990**, *172*, 3358–3366. [[CrossRef](#)] [[PubMed](#)]
78. Wasmund, K.; Mußmann, M.; Loy, A. The life sulfuric: Microbial ecology of sulfur cycling in marine sediments. *Environ. Microbiol. Rep.* **2017**, *9*, 323–344. [[CrossRef](#)] [[PubMed](#)]
79. Anantharaman, K.; Duhaime, M.B.; Breier, J.A.; Wendt, K.A.; Toner, B.M.; Dick, G.J. Sulfur Oxidation Genes in Diverse Deep-Sea Viruses. *Science* **2014**, *344*, 757–760. [[CrossRef](#)] [[PubMed](#)]
80. Blattner, F.R.; Plunkett, G., 3rd; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; et al. The complete genome sequence of *Escherichia coli* K-12. *Science* **1997**, *277*, 1453–1462. [[CrossRef](#)] [[PubMed](#)]
81. Wang, D.; Henrichs, S.M.; Guo, L. Distributions of nutrients, dissolved organic carbon and carbohydrates in the western Arctic Ocean. *Cont. Shelf Res.* **2006**, *26*, 1654–1667. [[CrossRef](#)]