Identification of microbiological and water quality drivers of brown tide in Baffin Bay

Final Report

GLO Contract No. 20-034-000-B742
September 2022

Prepared by:

Paxton T. Bachand
Jordan R. Walker
Lydia Hayes
Wingman Lee
Jessica M. Labonté
Lin Zhang
Jeffrey W. Turner
Texas A&M University-Corpus Christi
6300 Ocean Dr., Unit 5858
Corpus Christi, Texas 78412
Phone: 361-825-6206
Email: jeffrey.turner@tamucc.edu

Submitted to:
Texas General Land Office
1700 Congress Ave.
Austin, TX 78701-1495

A report funded by a Texas Coastal MANAGEMENT Program Grant approved by the Texas Land Commissioner pursuant to National Oceanic and Atmospheric Administration Award No. NA19NOS4190106.

CONTENTS

EXECUTIVE SUMMARY

The pelagophyte *Aureoumbra lagunensis* forms environmentally disruptive algal blooms (EDABs), colloquially known as brown tides, in shallow estuaries around the world. In Texas, brown tides occur on a near annual basis in the Laguna Madre, with the longest lasting for eight years. The disruption of ecosystem processes, following a brown tide bloom, can cripple coastal economies that rely on tourism and coastal natural resources. The purpose of this study was to identify the drivers of a persistent brown tide bloom that occurs in the tidal segment of Los Olmos Creek in Baffin Bay. Drivers were identified through the completion of three tasks: 1) characterize the bacterial and archaeal communities associated with the bloom, 2) characterize the viral communities associated with the bloom, and 3) characterize the water quality parameters associated with the bloom. Additionally, to promote education and increase awareness, students at Kaufer High School in Riviera Texas were engaged through lecture- and field-based outreach events (Task 4).

To fulfill Task 1 (Characterize the bacterial and archaeal communities associated with the Los Olmos bloom), we used 16S rRNA metagenomics to characterize the prokaryotic communities associated with the brown tide bloom. Findings showed that brown tide bloom-associated microbial communities were unique in comparison to non-bloom communities. Additionally, several taxa were enriched by the bloom. Members of Puniceicoccaceae (1.10%), Terasakiellaceae (1.21%), Trueperaceae (1.66%), Izemoplasmataceae (1.65%), Crocinitomicaceae (4.00%), the PeM15 group of Actinobacteria (2.47%), and Phycisphaeraceae (1.10%) composed on average 13.1% of the bloom site community. In comparison, these taxa were less abundant in the non-bloom community, composing only ~1% of the community. These taxa likely play an important

functional role in the modulation of bloom dynamics. Further, the decreased microbial diversity recorded at the bloom site provides evidence of ecosystem disruption.

To fulfill Task 2 (Characterize the viral communities associated with the Los Olmos bloom), we used metagenomics to characterize the viral assemblages. We compared the viral community composition 1) to the microbial communities to identify potential interactions between viruses and their hosts, 2) to the environmental data to identify the potential drivers of viral community composition, and 3) to *A. lagunensis* abundance to better understand the relationships between the bloom-forming alga and viruses. Finding showed that *Phycodnaviridae* viruses, the family of viruses containing most known algal viruses, was underrepresented at the bloom site. This suggests that viruses capable of infecting *A. lagunensis* do not belong to the *Phycodnaviridae* family. We found that typically cosmopolitan viral families like *Siphoviridae*, *Podoviridae*, and *Myoviridae* made up the largest proportion of viruses at each site. In contrast, rarer viral families such as *Herelleviridae, Phycodnaviridae*, and *Autographiviridae* were more closely associated with shifts in the bacterial community. These results are consistent with previous studies that suggest that dominant viral taxa form stable, long-term relationships with their hosts while rarer taxa tend to boom and bust with their hosts.

To fulfill Task 3 (Characterize the water quality parameters associated with the Los Olmos bloom), we measured physical and chemical factors and tested for relationships with algal cell abundance. Non-bloom water samples were primarily characterized by higher nitrate, ammonia, and oxygen concentration, and higher transparency whereas bloom water samples were characterized by higher salinity, pH, water temperature, and algal cell counts. Additionally, the water quality parameters at the non-bloom site were generally more stable, likely due to the greater depth and stronger connectivity to the larger Baffin Bay system. Findings confirmed that

temperature and salinity are drivers of brown tide abundance but demonstrated that elevated temperature and salinity are not bloom requirements. In fact, the highest brown tide cell counts were observed during the winter.

To fulfill Task 4 (Promote public awareness through education and training at Riviera High School), students at the Kaufer Early College High School in Riviera, Texas were engaged in five education and training events. These events introduced students to the research problem, i.e., the recurrent and persistent brown tide blooms in Baffin Bay. Students were also introduced to broader global concerns about HABs and connections with long-term trends in eutrophication and climate change. Concepts were introduced through lecture and discussion while hands-on skill were cultivated during field sampling activities. We remain in communication with the school about continuing outreach events in connection with ongoing research that builds on this study.

In summary, this study expanded understanding of brown tide blooms through the characterization of microbiological and water quality drivers. A continuous bloom appears to be the natural state in Los Olmos Creek sans the disruption of the bloom by storm events that dramatically increase freshwater inflows. Further, the Los Olmos Creek bloom appears to be a reservoir for *A. lagunensis* that can spread to neighboring water segments across Baffin Bay when water quality parameters select for its proliferation. Increasing freshwater inflows to Los Olmos Creek stands out as a promising remediation strategy.

ACKNOWLEDGEMENTS

This study was funded by the Texas General Land Office Coastal Management Plan (GLO-CMP) contract #20-034-000-B742, the Texas Research and Development Fund (TRDF), and Texas Sea Grant (TSG). Computational data analysis was performed on TAMU-CC's high-performance computing cluster (funded in part by the National Science Foundation's CNS MRI Grant No. 1429518) and the advanced computing resources provided by Texas A&M University's High Performance Research Computing.

INTRODUCTION

The pelagophyte *Aureoumbra lagunensis* forms brown tide blooms in shallow estuaries around the world (Gobler and Sunda, 2012). In Texas, brown tides occur on a near annual basis in the Laguna Madre, with the longest lasting for eight years (Buskey *et al.*, 2001). Brown tide blooms impair ecosystems through the disruption of ecosystem processes (Gobler and Sunda, 2012). Namely, *A. lagunensis* blooms attenuate sunlight necessary for the productivity of submerged vegetative habitats such as seagrass beds (Onuf, 1996). In the Laguna Madre, brown tide blooms have been implicated in the decreased abundance of seagrasses (Onuf, 1996). This loss of seagrass habitat then affects higher trophic levels as ecologically and economically important fisheries depend on seagrasses for nursery and foraging habitat (Sheridan and Minello, 2003) The disruption of seagrass habitat can also impact larger scale processes, such as nutrient cycling through decreased oxygen levels, decreased sediment stabilization, and decreased carbon sequestration (Orth *et al.*, 2006).

The disruption of ecosystem processes, following a brown tide bloom, can cripple coastal economies that rely on tourism and coastal natural resources. In Texas, in 2014, the shrimp, oyster, blue crab, snapper, and drum landings were valued at $278.4 million (Kildow *et al.*, 2016). Two of Texas' largest recreational fisheries, red drum and spotted seatrout, were estimated to be worth $350 million and $220 million, respectively (Vega *et al.*, 2011). A bloom-related decrease in seagrass habitat and water quality would inevitably decrease landings and increase seafood costs (Hoagland *et al.*, 2002). Along the Atlantic seaboard, from 1985 to 1987, a brown tide formed by *Aureococcus* species was responsible for mass mortality and recruitment failure in populations of bay scallops causing the collapse of a multi-million-dollar industry (Kraeuter *et al.*, 2008).

Similarly, in the late 1980s through early 1990s, brown tide blooms were responsible for the near collapse of *Mercinia mercinia* clam landings.

Physical factors, such as temperature, salinity, and nutrient concentrations can aid in initiation and maintenance of brown tide blooms (Bricker *et al.*, 2008; Gobler and Sunda, 2012). The average surface water temperatures in the Laguna Madre range from 15 to 30˚C between May and October and culture-based studies have determined that growth is optimal at 20 to 25˚C (Buskey *et al.*, 1998). Salinities in the Laguna Madre can reach 70 (Buskey *et al.*, 1998), and *A. lagunensis* has been observed to proliferate in hypersaline conditions, although hypersalinity is not a strict growth requirement (Cira and Wetz*.,* 2019). In addition, high levels or organic nutrients have been linked to bloom formation, as blooms have been seen to form in highly eutrophic systems (Wetz *et al.,* 2017). As a result, the physical factors that select for *A. lagunensis* growth, such as high concentrations of organic nutrients, temperatures, and hypersalinity, establish the Laguna Madre as an ideal system for brown tide blooms (Buskey *et al.*, 1998).

The limited flushing of the lagoon is a compounding factor that also contributes to bloom initiation and maintenance. Residence times for the Laguna Madre have been estimated to range from 300 days to several years (Buskey *et al.*, 1998). High residence times allow for the accumulation of large amounts of algal biomass, further exacerbating low nutrient availability (Gobler and Sunda, 2012). High residence times and rapid evaporation rates, coupled with decreased freshwater inflows, also contribute to hypersalinity, further selecting for the growth of halotolerant algal species. As a result, the lagoon is regarded as a negative or inverse estuary, and the persistence of this inverse condition is thought to have favored the maintenance of the aforementioned eight-year bloom event (Buskey *et al.*, 1998).

Depressed grazing activity has been implicated as a contributor to brown tide maintenance (Buskey *et al.*, 1997). Grazing may decrease due to a bloom-associated deterioration of environmental parameters (e.g., depleted food sources, increased salinity, predation) (Buskey *et al.*, 1998). Additionally, the production of a protective and recalcitrant sheath of extracellular polymeric substances (EPS) was shown to decrease algal grazing (Liu and Buskey, 2000). The EPS layer was subsequently shown to protect cells during zooplankton digestion (Bersano *et al.*, 2002). This depressed grazing activity is thought to be yet another reason for long-term bloom maintenance (Gobler and Sunda, 2012).

A critical gap in our understanding of brown tide blooms concerns the influence of co-occurring microorganisms: bacteria, archaea, and viruses. Numerous studies have reported that similar types of bacteria are associated with algal blooms (e.g., Alphaproteobacteria, Gammaproteobacteria, Actinobacteria, and Bacteroidetes) (Buchan *et al.*, 2014; Tan *et al.*, 2015) and additional studies have shown that changes in microbial community structure tracked with specific bloom stages (i.e., initiation, maintenance, termination) (Zhou *et al.*, 2018). Algicidal bacteria negatively affect algal growth and are thought to influence bloom dynamics including termination and Frazier et al. showed that several genera e.g., *Bacillus, Halomonas, Marinobacter, Croceibacter* were algicidal against *Aureococcus anophagefferens* brown tides (Frazier *et al.*, 2007).

Viruses control algal blooms, harmful or otherwise, by infecting and killing the host, causing increased production of exopolymeric substances, or inducing spore formation (Short, 2012; Nissimov *et al.*, 2018; Pelusi *et al.*, 2021). However, viruses and algae are known to follow complex evolutionary trajectories where the outcome can be a host community with strong resistance to viral infection and a viral population favoring lysogeny or longer lytic cycles (Sanda

et al., 2022). Additionally, physiochemical properties of the environment can dampen viral infection rates including but not limited to salinity, temperature, and light intensity (Mojica and Brussaard, 2014). Although previous studies have investigated viruses associated with various algal species (reviewed by Short *et al.*, 2020), no viruses capable of infecting *A. lagunensis* have been described.

While much is known about the physical factors associated with brown tides, the potentially complex dynamics between *A. lagunensis* brown tides and co-occurring microorganisms are unknown. Microorganisms could play a heretofore unknown role in the modulation of bloom dynamics. This critical knowledge gap was addressed through the completion of three tasks: 1) characterize the bacterial and archaeal communities associated with the bloom, 2) characterize the viral communities associated with the bloom, and 3) characterize the water quality parameters associated with the bloom. Additionally, to promote education and increase awareness, students at Kaufer High School in Riviera Texas were engaged through lecture- and field-based outreach (Task 4).

METHODS

Sampling collection. Water samples (N=12) were collected from Los Olmos Creek (27°16'23.62" N, 97°48'08.01" W) and Riviera Beach (27°17'00.09" N, 97°39'52.80" W) in the Baffin Bay watershed (Figure 1). Sampling events occurred 01/15/2020, 02/24/2020, 05/18/2020, 06/15/2020, 09/13/2020, 10/16/2020, 12/18/2020, 01/22/2021, 02/26/2021, 03/22/2021, 05/25/2022, and 06/18/2021. Los Olmos Creek is a shallow, typically hypersaline tributary that flows into the Laguna Salada (Tunnell and Judd, 2002) and a well-documented reservoir for persistent brown tide blooms (Cira and Wetz, 2019). Riviera Beach is a public recreational area at the confluence of the Laguna Salada and Baffin Bay 13.7 km from the Los Olmos Creek study site. The Los Olmos Creek site served as the experimental "bloom site" while the Riviera Beach site served as the experimental "non-bloom" or control site.

Water quality parameters. Water temperature (˚C), pH, dissolved oxygen (%), and salinity were recorded with a YSI 556 Multi Probe System (YSI Incorporated, Yellow Springs, OH, United States). Wind speed (mph) and air temperature (°C) were measured using a Kestrel 3000 Wind Meter (Kestrel Meters, Boothwyn, PA, United States). For identification and quantification of different nitrogen species, concentrations of NO_3^-, NO_2^-, and NH_4^+ were measured using a SEAL AQ300 Discrete Analyzer (AQ300 methods EPA-148-D, EPA-115-D, and EPA-126-D, respectively).

Bloom characterization. Water transparency was measured using a transparency tube fixed with a Secchi disk (cm) (Carolina Biological Supply Company, Burlington, NC, United States). Water/bloom samples (50 mL) were preserved with Lugol's solution (1% final concentration, v/v) and cells having a morphology consistent with *A. lagunensis* were quantified by directly counting five random cells of the center field with a Leica DM750 fluorescence

microscope (Wetzlar, Germany) using a 0.1 mL hemacytometer at 400x magnification (Hall *et al*., 2018). The presence and dominance of *A. lagunensis* was confirmed by 18S rRNA gene sequencing analysis as detailed below.

DNA isolation. For the characterization of the microbial communities, water samples (100 mL) were filtered through low-protein binding 0.22 µm polyethersulfone (PES) filters (MilliporeSigma, Burlington, MA, United States). Total genomic DNA was isolated from the filters using a DNeasy PowerSoil kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. The DNA quantity (ng/µl) and quality (A260/A280 and A260/A230 absorbance ratios) were measured using a biospectrophotometer (Bio-Rad, Hercules, CA, United States) and the DNA was stored at -20˚C until sequencing.

For the characterization of the viral communities, approximately 10 L of water was pre-filtered using Nitex filters (30 µm) and stored in 24 L acid-washed, field-rinsed carboys. Samples were then sequentially filtered through 142 mm GF (0.7 or 2.7 µm) and 142 mm 0.22 PVDF filters. The filtrate was concentrated using the iron flocculation method (John *et al.*, 2011; Poulos *et al.*, 2018). Briefly, the filtrate was treated with 100 µL of 10g/L iron chloride stock solution, vigorously shook, and allowed to incubate for 1 hour. The samples were then filtered through 142 mm 0.8 µm hydrophilic polyethersulfone filters. The filters were retained in 50 mL falcon tubes on ice during transit to Texas A&M University at Galveston where they were stored at 4˚C . Filters were treated with an oxalic acid buffer to remove the iron and resuspend viral particles and allowed to incubate over night at 4˚C. The following day the buffer containing viral particles was passed through a 0.2 µm syringe filter and placed in a sterile 15 mL falcon tube and stored at 4˚C. Virus concentrates were further concentrated to a final volume of < 1 ml using 30 kDa molecular weight cutoff 50 mL centrifugal filtration units (Amicon Ultra-15, Millipore). Wizard® Plus Miniprep

DNA Purification System (Promega, Madison, WI, United States) was used to extract DNA from the virus concentrates using an adapted protocol (Poulos *et al.*, 2018). DNA extracts were further purified using magnetic beads (Sera-mag SpeedBeads, Cytiva, Marlborough, MA, United States) (Faircloth and Glenn, 2011). Final DNA extracts were stored at -20˚C until sequencing.

16S and 18s rRNA gene sequencing. The 515f (5' – GTG YCA GCM GCC GCG GTA A – 3') and 806r (5' – GGA CTA CNV GGG TWT CTA AT – 3') primers were used to amplify the V4 region of the 16S rRNA gene (Walters *et al*., 2015). The euk1391F (5' – GTA CAC ACC GCC CGT C – 3') and EukB-Rev (5' –TGA TCC TTC TGC AGG TTC ACC TAC – 3') primers were used to amplify the V9 region of the 18S rRNA gene (Caporaso *et al*., 2012). Both were amplified using a HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA, United States). For the 16S rRNA gene, the following cycling conditions were used for amplification: 94˚C for 3 minutes, 30 cycles of 94˚C for 30 seconds, 53˚C for 40 seconds, and 72˚C for one minute, followed by a 5-minute hold at 72˚C. For the 18S rRNA gene, the following cycling conditions were used for amplification: 94˚C for 3 minutes, 35 cycles of 94˚C for 45 seconds, 57˚C for 60 seconds, and 72˚C for 90 seconds, followed by a 10-minute hold at 72˚C. Amplicons were visualized in a 2% agarose gel, pooled and purified with calibrated Ampure XP beads (Beckman Coulter, Indianapolis, IN, United States), and sequencing was performed on an Illumina MiSeq platform using paired-end chemistry (2 x 300 bp) at Molecular Research L.P. (Shallowater, TX, United States).

Sequence reads were processed using a combination of QIIME version 1.9 (Caporaso *et al*., 2010) and QIIME2 version 2018.11 (Bolyen *et al*., 2019). Barcodes were extracted from the paired-end reads using the 'extract_barcodes.py' tool in QIIME. The following steps were performed within QIIME2. Reads were demultiplexed and denoised with DADA2 (Callahan *et al*., 2016), resulting in amplicon sequence variants (ASVs). Trim lengths of 242 bp were used on both

the forward and reverse reads. In addition to denoising the data, DADA2 filtered sequences for quality, removed chimeric sequences, and merged paired-end reads. A phylogenetic tree was then generated using the 'q2-phylogeny' pipeline with default settings, which was used to calculate phylogeny-based diversity metrics. Taxonomy was assigned using a Naïve Bayes classifier trained on the SILVA v. 132 99% OTUs database (Quast *et al*., 2013), including only the 250 bases from the V4 region bound by the 515F/806R primer pair. Reads mapped to chloroplast and mitochondrial sequences were filtered out from the sequence variants table using the 'filter_taxa' function. Faith's phylogenetic distance (PD) and the Shannon-Weiner diversity index were calculated for all samples using the 'qiime diversity alpha-phylogenetic' function. Data were then imported into phyloseq (McMurdie and Holmes, 2013) using the 'import_biom' and 'import_qiime_sample_data' functions and merged into a phyloseq object. All samples maintained high read counts and therefore did not require proportional transformation to a normalized read count. Beta diversity was analyzed using weighted UniFrac (Lozupone *et al*., 2011) distances calculated in phyloseq. From these distances, a principal coordinate analysis (PCoA) was calculated and plotted. Permutational multivariate analysis of variance (PERMANOVA) was used to test for significant differences between communities using the 'vegan' (Oksanen *et al*., 2019) and 'pairwiseAdonis' (Arbizu, 2017) packages in R version 3.5.2 (R Core Team, 2017). To test if variance was due to dispersion of variability among groups, permutational analyses of dispersion (PERMDISP) were conducted for all significant PERMANOVA outcomes with the 'vegan' package in R. Community differences were further confirmed via linear discriminant analysis (LDA) effect size (LEfSe) using the LEfSe tool on the Galaxy server (https://huttenhower.sph.harvard.edu/galaxy/). Genera that made up more than 0.10% of the

communities' relative abundance were analyzed using default settings with the significance threshold set to a p-value of < 0.01.

Viral metagenomic sequencing. DNA extractions from the viral communities were shotgun sequenced by the Texas A&M Genomics and Bioinformatics facility using Illumina NovaSeq 6000 SP (150 bp paired-ends) technology. Raw sequencing files were quality controlled using BBtools (Bushnell *et al.*, 2017). Quality control consisted of predicting sequence adapters, merging reads, trimming sequencing adapters and artifacts, removing low quality reads, masking low complexity regions, and removing common laboratory contaminants. Reads were assembled using the *de novo* sequence assembler MEGAHIT using both the merged and unmerged quality-controlled sequences (Li *et al.*, 2015). Coding sequences were predicted using Prodigal and the subsequent coding regions were annotated using DIAMOND against the blast-nr database. The annotated genes were then uploaded to MEGAN6 Ultimate in order to visualize and extract sequence counts for statistical analysis.

Statistical analysis. ANOVA was performed to determine differences in the physical parameters between sites, and Kendall's tau correlation was used in addition to principal component analysis (PCA) to characterize site conditions. Generalized linear models (GLMs) were generated for bloom and non-bloom sites to identify drivers of algal cell count concentrations. GLMs were generated in R (version 3.6.1) and RStudio (version 1.2.1335) using the glm function. The variables in each model were tested for multicollinearity using the VIF function (car package; Fox and Weisberg, 2012) to ensure VIF < 10. Marginal R^2 values were determined with the r.squaredGLMM function (MuMIn package; Barton and Barton, 2021), and the models were tested for significance against null models with all variables removed, using the ANOVA function. Metagenomic sequence data was always normalized to the total number of coding sequences found

within each metagenome. Stacked bar histograms of the viral metagenomes were constructed in R (2015). Diversity, richness, and evenness indices were calculated using vegan and fossil in R (Vavrek, 2011). NMDS plots of the viral metagenomic and abiotic conditions were constructed and analyzed using analysis of similarity and procrustes tests in vegan in R (Oksanen *et al.*, 2020). PERMANOVA analysis was done to determine if samples significantly grouped by sample or site using the adonis2 function in vegan (Oksanen *et al.*, 2020).

Outreach. Students enrolled in the Ocean Science course at the Kaufer Early College High School (Riviera, TX) were engaged through education and training events in coordination with two Kaufer instructors: Shelby Szteiter (2020-2021) and Selina Ramirez Leos (2022-2023). These events were designed to introduce students to the research problem, i.e., the recurrent and persistent brown tide blooms in Baffin Bay. Students were also introduced to broader global concerns about harmful algal blooms and connections with long-term trends in eutrophication and climate change. Concepts were introduced through lecture and discussion while hands-on skills were cultivated during field sampling activities.

Data availability. Data were transferred to the TGLO Box.com folder. The raw 16S bacterial sequence data can be found in directory 092821PBillcus515F-298329269 – 16S, the raw 18S eukaryotic sequence data can be found in directory 092821PBeuk1391F-298329270 – 18S, and the raw viral sequence data can be found in directory 21411Lbn_[Control or Bloom]-[Samp #]-[Bac or V].tar.gz. An Excel spreadsheet descriptive of environmental variables was also uploaded to the folder (filename TGLO Metadata Spreadsheet.xlsx).

RESULTS AND DISCUSSION

Task 1: Characterize the bacterial and archaeal communities associated with the Los Olmos bloom. The bloom and non-bloom sites microbial communities maintained relatively high alpha diversities. Figure 3 shows bar-plots of the Faith's PD (A) and Shannon-Weiner (SW) diversity values (B) observed throughout the study. As reported by Faith's PD (Figure 3A), the non-bloom site maintained significantly higher alpha diversity values when compared to the bloom communities (p-value < 0.05), although the differences were not significant when SW values were compared. Differences in diversity may be explained by the system being impacted negatively by the bloom (i.e., an ecological disturbance) and subsequently harboring fewer taxa, or 2) that more phylogenetically similar taxa would report lower PD values than phylogenetically dissimilar taxa, such as in the non-bloom site. It should be noted that this continuous bloom, if acting as an ecological disturbance, could result in reduced resilience of an ecosystem's taxa, decreasing stability within the system and resulting in a lower relative diversity.

Microbial community compositions were investigated using PCoA, as shown in Figure 4, with a clear distinction between bloom site and non-bloom site communities (PERMANOVA; adj. p-value < 0.05). In terms of abundance, Cyanophyceae, Actinobacteria, and Rhodothermia were the three most commonly observed classes at either site whereas at the non-bloom site, Cyanophyceae, Alphaproteobacteria, and Bacteroidia (shown in Figure 5). Cyanophyceae (family Cyanobiaceae) was the most abundant taxa found at both sites, with 17.25% at the bloom site and 15.04% at the non-bloom. Actinobacteria (family Nitriliruptoraceae) and Rhodothermia (family Balneolaceae) were the second and third most abundant taxa found among the bloom site's community, composing 11.94 and 10.20%, respectively. Alphaproteobacteria (SAR-11 clade) and Bacteroidia (family Saprospiraceae) were the second and third most abundant taxa found among

the non-bloom site's community, composing 13.33 and 8.10%, respectively. Gammaproteobacteria (Gammaproteobacteria *incertae sedis*; 6.83%) and Alphaproteobacteria (family *Rhodobacteraceae*; 6.31%) were observed to be the fourth and fifth most abundant bloom community taxa, respectively. Alternatively, the Acidimicrobiia (Actinomarinaceae; 6.65%) and Rhodothermia (Balneolaceae; 6.23%) were the fourth and fifth most abundant classes observed in the non-bloom's community. Regardless of similarities, beta diversity values (i.e., weighted UniFrac) highlight clear distinctions between both sites' communities as visualized through PCoA (Figure 4). Additionally, the non-bloom community was composed disproportionately of one microbial group (Alphaproteobacteria; SAR-11 and AEGEAN-169).

Importantly, several bacterial classes were detected in higher abundance (i.e., enriched) within the bloom site community. Further, the higher abundance of those classes coincided with higher concentrations of *A. lagunensis* cells. Throughout the study, members of Puniceicoccaceae (1.10%), Terasakiellaceae (1.21%), Trueperaceae (1.66%), Izemoplasmataceae (1.65%), Crocinitomicaceae (4.00%), the PeM15 group of Actinobacteria (2.47%), and Phycisphaeraceae (1.10%) composed on average 13.1% of the bloom sites community. In comparison, these taxa were less abundant in the non-bloom community, composing only ~1% of the community. Additionally, these taxa, along with the most abundant classes described above, represented over 51% of the *A. lagunensis* microbiome community throughout the study period. By comparison, few bacterial taxa were enriched in the non-bloom site: marine groups SAR86 and 324 (1.70 and 2.00%), Flavobacteriaceae (1.85%), and the Alphaproteobacteria AEGEAN-169 marine group (3.41%). Similarly, to further investigate and confirm differences in the bloom and non-bloom community, linear discriminant analysis (LDA) effect size (LEfSe) was implemented. Results showed that several taxa (e.g., Planctomycetacia, Phycisphaerae, Bacilli, Desulfuromonadia,

Deinococci, Bdellovibrionia, Rhodothermia, Kapabacteria, Bacteroidia, Actinobacteria, and various Gamma- and Alphaproteobacteria) were enriched in the bloom community (p-value $<$ 0.01) (Figure 6). In contrast, fewer taxa (e.g., marine Alpha- and Gammaproteobacteria and Acidimicrobiia) were enriched in the non-bloom community.

In general, Alphaproteobacteria, commonly reported as constituents of HAB microbiomes, were less enriched in the bloom site. For example, HAB associated Alphaproteobacteria such as AEGEAN-169 and SAR-116 were only detected at the non-bloom site, making up 19.42% of the total community. Although the amount of Alphaproteobacteria was disproportionately high in the non-bloom compared to the bloom community, some Alphaproteobacteria, Rhodobacterales in particular, made up 8.41% of the bloom community. Interestingly, this observation could be explained by Rhodobacterales' ability to produce vitamin B_{12} as *A. lagunensis* is a B-vitamin auxotroph and requires a mutualistic relationship for its supply (Gómez-Consarnau *et al*., 2018). *A. lagunensis* also requires B_6 and Verrucomicrobiae (1.90%) could have supplied the micronutrient, having been previously observed to supply disproportionate amounts relative to their abundance in other environments (Gómez-Consarnau *et al*., 2018). These findings suggest that B-vitamin relationships could be a major factor affecting bloom dynamics, although additional research is needed to further explore this hypothesis.

Cyanobacterial taxa were common within the bloom community. Members of Oxyphotobacteria (Cyanobacteria) composed 8.52% of the bloom microbiome. Previous research has shown that Synechococcus are correlated with *A. lagunensis* blooms, observed to co-occur as well as maintain co-dominance within systems (Hall *et al*., 2018); however, while Synechococcus was observed in both the bloom and control, the results presented here do not support the co-dominance hypothesis as the taxa was maintained in similar abundances (~2.50%) at both sites.

Rather, our results show that Cyanobacteria of the order Nostocales, especially members of the family Microcystaceae, were enriched in comparison to non-bloom samples, composing 4.50% of the bloom community on average. Cyanobacterial taxa are thought to affect bloom dynamics through providing reduced forms of nitrogen as well as contributing to the organic carbon pool via carbon fixation. The co-occurrence of cyanobacteria may also simply be explained by the taxa's proclivity for organic nutrients, given Los Olmos Creek is highly eutrophic. However, it appears Cyanobacteria are correlated with *A. lagunensis* blooms regardless of whether increased algal abundances, eutrophication, or both are a prerequisite for this group's succession.

The majority of Gammaproteobacteria identified within the bloom community were also observed in the non-bloom community. However, several taxa, such as the Chromatiaceae (2.11%) and Alcanivoracaceae (1.73%), were enriched within the bloom. The Chromatiaceae (2.09%) are a group of phototrophic purple sulfur bacteria capable of organic and inorganic sulfur reduction (Frigaard and Dahl, 2008). As *A. lagunensis* is known to produce high concentrations of dimethylsulfoniopropionate (DMSP), it follows that a member of the microbial community may be responsible in supplying sulfate to the bloom (Buskey and Hyatt, 1995). Curiously, while not necessarily indicative of an algal bloom, Alcanivoracaceae are typically associated with oil contamination, potentially a result of the bloom's location relative to that of the high-traffic state highway. Future research to investigate the potential role of oil contamination in *A. lagunensis* bloom systems and their microbiomes could be warranted.

Cyclobacteriaceae and Saprospiraceae were the only members of the Bacteroidetes enriched within the bloom, making up ~11% of the 17.84% total detected. Bacteroidetes is known to contain members commonly in league with various algal species (reviewed by Amin *et al.*, 2012). Not surprisingly, Bacteroidetes have been shown to produce enzymes capable of degrading

laminarin, a storage form of glucose commonly associated with brown algae, pointing to a possible ecological function for Cyclobacteriaceae and Saprospiraceae in this system. Saprospiraceae could also serve a predatory role within the bloom, as the family is an important degrader of complex carbon sources and has been observed to exhibit algicidal behavior (Furusawa *et al*., 2003). Additionally, both families have been associated with activated wastewater sludge (Ren *et al.*, 2016), indicating that eutrophication and pollution may be a co-factor affecting their abundances.

Rhodothermia were heavily enriched within the bloom site, with all the group's abundance supplied by the family Balneolaceae. Composing an average of 7.06%, this taxon was one of the single most abundant individual bacterial groups observed. Members of this family are heavily associated with moderate salinity tolerance. Although elevated salinity is not a prerequisite for bloom development, findings suggest that Balneolaceae could be normal constituents of *A. lagunensis* blooms. In contrast, several taxa were present in much smaller concentrations. Of the remaining taxa enriched within the bloom (e.g., Phycisphaerae (2.02%), Planctomycetacia (3.12%), Kiritimatiellae (1.41%), Mollicutes (4.50%), Nitriliruptoria (4.84%), and Deltaproteobacteria (Bdellovibrioaceae, 1.58%), their presence may suggest potentially relevant functions within the bloom such as organics degraders and algal pathogens/parasites and even small variations in these low abundance organisms could be biologically important.

Task 2: Characterize the viral communities associated with the Los Olmos bloom. To characterize the viral assemblages at the bloom and non-bloom control sites, we extracted and sequenced total viral DNA (<0.22 μm fraction) on all eight samples collected for viral analysis. The original plan was to use PCR amplification of genes specific to viral families, i.e., including the DNA polymerase of algal viruses from the *Phycodnaviridae* family (Chen *et al.*, 1996), the major capsid protein of T-4 like phages (family *Myoviridae*) (Filée *et al.*, 2005), and the DNA

polymerase of phages from the *Podoviridae* family (Labonté *et al.*, 2009). Unfortunately, none of the primer sets tested resulted in amplification, suggesting that the viruses present in Los Olmos Creek and the Baffin Bay watershed are different and distinct from marine and freshwater viruses identified in other environments.

To circumvent the PCR amplification problems, we used metagenomics for all samples, allowing for a characterization of all the viruses present in the samples. Within the viral metagenomes, between 17-32% of the sequences were identified as viral (Figure 7A), highlighting the lack of viral isolates and sequences representative of the studied environments in the databases. Interestingly, the relative abundances of sequences identified as viral was lowest during the winter months in both the bloom site and the control site, when *A. lagunensis* abundances were at their highest. At the taxonomic level, the dominant viral families were *Siphoviridae* (long non-contractile tailed phages), *Myoviridae* (contractile tailed phages), and *Podoviridae* (short non-contractile tailed phages) (Figure 7B). At the bloom site, the family *Autographiviridae* (short non-contractile tailed phages, formerly part of the *Podoviridae* family) were virtually absent in the cold winter months of 2021, which also coincide with the highest abundance of *A. lagunensis*. Those results suggest that the viruses present in the winter of 2021 may be infecting hosts that are tolerant to cold, and/or hosts that are associated with a bloom of *A. lagunensis* (Figure 7B).

There are currently no known viral isolates for *A. lagunensis*. Most known algal viruses available in culture belong to the family *Phycodnaviridae* [from the phylum *Nucleocytoviricota* (NCLDV)] (Coy *et al.*, 2018), yet very few sequences sharing similarity to *Phycodnaviridae*, or even NCLDVs, were found at the bloom site, where higher concentrations of *A. lagunensis* are found, suggesting that viruses infecting *A. lagunensis* are likely not NCLDV. NCLDV sequences were on average five times more prevalent in the samples at the control site compared to those

collected at the bloom site. Possible explanations for this could be that (1) the conditions such as higher salinity are negatively impacting virulence (Mojica & Brussaard, 2014), (2) that large numbers of *A. lagunensis* offer protection from viral infection through the production of exopolymeric substances, or (3) that the NCLDVs seen at the control site are infecting some other algae more prevalent in typical marine waters. Future studies should focus on the isolation of novel viruses of *A. lagunensis* and further our understanding of how *A. lagunensis* interacts with viruses and escapes epidemic level viral lysis.

A large fraction of the sequences shared similarity with bacterial sequences, such as bacteria from the classes Alphaproteobacteria, Betaproteobacteria, and Actinomycetia (Figure 7A), which indicates a bias within the databases for more studied microorganisms and environments rather than contamination of the metagenomes by microbial sequence. These bacterial groups are likely hosts to these viruses, but it is impossible to confirm at this time. Deeper analyses of the genomic content (e.g., similar tetranucleotide frequencies) and gene markers and signatures (e.g., auxiliary metabolic genes shared by viruses and their hosts) may inform about virus–host interactions by allowing the linkage of viruses to their putative hosts. Moreover, more analyses are needed to confirm that the sequences are viral and devoid of microbial contamination.

As expected, the abundance of *A. lagunensis* was higher in the bloom site than in the non-bloom control site. The abundance of *A. lagunensis* increased in the winter of 2021 as the concentration was higher in the months of January and March, declining to normal levels in June (Figure 8A). The bloom may have been caused by the increase of nitrate (Figure 8C). The increase in dissolved oxygen observed during those same months may have been caused by the increase of *A. lagunensis* via photosynthesis (Figure 8C). Interestingly, the viral diversity was generally

slightly lower at the bloom-site than at the control site, except during the months when *A. lagunensis* was blooming (Fig. 8B).

While the bloom and non-bloom sites were similar at the taxonomic level (Figure 9), a non-metric multidimensional scaling (NMDS) analysis revealed that the two sites were distinct as none of the samples overlapped (Figure 9). We performed an analysis of similarity on the samples grouped by site and by season that confirmed samples group significantly by site rather than by season ($\alpha < 0.05$). Additionally, we performed a procrustes analysis between the NMDS plot of the community data and the environmental parameters and found a significant correlation between the two ($\alpha < 0.05$), indicating a correlation between the environmental parameters and community data. We used the bioenv() function from vegan to find the optimal set of environmental parameters that have the maximum rank correlation with the community dissimilarity. The results of the bioenv() function showed that temperature, dissolved oxygen, and salinity were optimal and correlated at an R^2 of 0.74. Temperature and salinity extremes have been implicated in lowering the abundance of bacteria competing with *A. lagunensis* allowing the brown algae to bloom (Buskey et al, 1998).

To determine the viral sequences most contributing to the dissimilarity in the community data when grouped by location we performed a SIMPER analysis. The SIMPER analysis indicated that *Myoviridae*, *Siphoviridae*, *Podoviridae*, *Phycodnaviridae*, *Autographiviridae*, and *Herelleviridae* were the most influential viral families. SIMPER analysis is biased towards groups that constitute larger portions of the community, so we performed Kruskal Wallis tests to confirm that the contribution was from differences between sites and not just overwhelming numbers. Kruskal Wallis tests showed that *Podoviridae*, *Phycodnaviridae* and *Herelleviridae* were the only viral families contributing due to differences between sites.

We used Spearman rank-base correlations to determine positive and negative correlations between viruses and the microbial sequences (Figure 10). A strong negative correlation could mean that a specific virus infected a specific species to the point where the species is barely detected, following the Kill-The-Winner Hypothesis (Thingstad, 2000). On the other hand, a strong positive correlation could mean that the viruses tend to be abundant when their host species is abundant, such is the case of *Pelagibacter* (SAR11), one of the most abundant species in the ocean, and its phages (Zhao *et al.*, 2013). Interestingly, the most abundant viral families in our data and in the literature (*Myoviridae*, *Podoviridae*, and *Siphoviridae*) had relatively lower numbers of correlated bacteria compared to rarer viral families like *Herelleviridae* (known to infect Firmicutes), *Phycodnaviridae*, and *Autographiviridae* (Figure 10). It has been posited that rarer viral taxa tend to replicate faster with larger burst sizes in response to fluctuations in the rarer microbial hosts they infect (Suttle, 2007). Alternatively, cosmopolitan viruses will have longer replications periods and lower burst sizes forming long-term stable relationships with their abundant hosts (Suttle, 2007). The higher degree of correlations between rarer viral families and a larger number of less abundant bacterial classes could be explained by the fact that shifts in the viral families and their hosts would be acute and pronounced when conditions change. Moreover, the more abundant bacterial groups and viruses are more temporally stable and resistant to fluctuations in abiotic conditions.

We searched for auxiliary metabolic genes (AMGs) in the viral metagenomes. 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 (Rosenwasser *et al.*, 2016). Preliminary analysis revealed that during the bloom of *A. lagunensis* that occurred during the winter months, the number of genes related to the carbohydrate, amino acid, nucleotide, and

methane metabolisms increased (Figure 11). Similarly, these peaks also occurred during the times where genes assigned to a viral class were least prevalent (Figures 7 and 11). The co-occurrence of these three factors makes it plausible that the bloom altered the microbial populations and in turn changed the relationships between viruses and host bacteria within the community.

Task 3: Characterize the water quality parameters associated with the Los Olmos bloom. Hypersalinity was regular throughout the investigation, especially within the Los Olmos bloom site regardless of bloom condition. At the beginning of sampling, salinity at Los Olmos Creek was 59.81, reaching a maximum of 82.19 half-way through the sampling period and a minimum of 2.57 by the end of sampling. *A. lagunensis* cell concentrations reached a maximum of 2.22 x 10^6 cells/mL during this period of elevated salinity, relative to that of other salinities experienced, and then subsequently some of their lowest concentrations due to a storm event flushing the system with large amounts of fresh water. Hypersalinity was normal within the non-bloom site as well, with salinities reaching 49.78. The non-bloom site also saw relatively low salinities near the end of sampling, in comparison, again due to the large amounts of storm-associated freshwater inflow. At Los Olmos Creek, water temperature ranged from 32.52-12.10˚C whereas at the non-bloom site, water temperature ranged from 28.93-14.60˚C. Dissolved oxygen (DO) at both sites typically ranged from supersaturated to near saturated, but oxygen levels approaching hypoxia were also observed at the Los Olmos bloom site. On average, Los Olmos Creek's DO (72.43%) was lower than the non-bloom's DO (97.70%). The lowest DO experienced at the bloom site was 11.00% whereas the lowest experienced at the non-bloom site was 65.00%. In general, the bloom site experienced greater fluctuations and extremes in water quality parameters and these extremes coincided with higher agal abundances.

After six sampling events (approx. 10 months), a persistent and dense bloom had formed in Los Olmos Creek, known to maintain persistent and reoccurring blooms. The bloom site maintained an average of 4.76 x 10^5 cells/mL in comparison to the non-bloom site in the bay, which maintained modest densities (3.04 x 10^3 cells/mL on average). Once reformed, the continuous bloom at Los Olmos Creek was highly elevated in comparison to the non-bloom site. At its reformation, the bloom's concentration at Los Olmos Creek was 1.72 x 10^6 cells/mL, increasing to a maximum concentration of over two million cells/mL (2.22 x 10^6). In contrast, the concentration at the non-bloom site was relatively stable throughout the study, ranging from a minimum concentration of 0 cells/mL to 2.00 x 10^5 cells/mL. This concentration gradient suggests that Los Olmos Creek is a reservoir for the spread of *A. lagunensis* to other segments of the bay.

Figure 2 depicts a PCA of water quality parameters of the bloom and non-bloom sites. 31.32% variation among samples can be explained through axis PC1, and 19.82% though PC2. Non-bloom samples were mainly characterized by higher nitrates, ammonia, oxygen, and transparency whereas bloom samples were characterized by higher salinity, pH, water temperature, and algal cell counts. The water quality parameters at the non-bloom site were generally more "stable", likely due to the greater depth and connection to the larger bay system. In a separate study at Los Olmos Creek, organic nutrient concentrations at the bloom site were shown to be highly elevated compared to the non-bloom waters: DOC and DON concentrations at the bloom site were observed to be on average 3,529.25 µmol/L and 293.93 µmol/L (respectively), nearly four times higher than observed at the non-bloom site (unpublished data).

Large storm events, especially in systems with high residence times, are an important factor in algal bloom reduction and termination (Buskey *et al.*, 2001). The lack of flushing, low freshwater inflow, high residence time, and high salinity of the Laguna Madre/Baffin Bay system

allow this water segment to support long-lasting, persistent blooms (Buskey *et al.*, 2001; Cira and Wetz, 2019). However, no bloom was detected until half way through our sampling effort, despite the creek maintaining hypersalinity well before reformation. Regardless, when the bloom did reform, salinity reached its maximum (82.19) accompanied by a dense bloom. Direct correlations between *A. lagunensis* and elevated salinity have been reported by several studies, as this relationship has been thought to increase *A. lagunensis*' competitive advantage with other phytoplankton species and decrease grazing pressure (Buskey *et al.*, 1998, 2001). Although hypersalinity is not a strict growth requirement, it does appear to be a co-factor that supports dense bloom formation and persistence (Hall *et al.*, 2018).

Nutrient sources in the watershed, such as wastewater treatment plant (WWTP) outfalls, septic effluent, groundwater discharge, agricultural run-off, and fecal waste from a large Mexican Free-tailed Bat colony (identified in this study), may all be supporting the continuous bloom. Eutrophication is evident, as organic nitrogen and carbon for the site have been observed to be highly elevated in the Laguna Salada (Wetz *et al.*, 2017). As observed in other brown tide blooms, it appears that organic nutrient eutrophication likely plays a major role in bloom development and persistence. Indeed, the Laguna Salada segment of Baffin Bay, where Los Olmos Creek is situated, maintains a consistently high abundance of *A. lagunensis* in addition to elevated DON and DOC concentrations (Cira and Wetz, 2019). Flushing in the Laguna Salada, or a lack thereof, is thought to cause the buildup of organic nutrients, and may explain how this continuous bloom can maintain such elevated algal and nutrient concentrations. In a separate study at Los Olmos Creek, nearly 70% of the variation between algal cell counts in the bloom site could be explained by DOC, DON, and DO, whereas the same relationship was not observed within the non-bloom site (unpublished data).

Task 4: Promote public awareness through education and training at the Kaufer Early College High School in Riviera, Texas. Students enrolled in the Ocean Science course at the Kaufer Early College High School were engaged through five education and training events. These outreach events spanned two years and were coordinated with two Kaufer instructors: Shelby Szteiter (2020-2021) and Selina Ramirez Leos (2022-2023). The first outreach event occurred 02/03/2020 and was attended by Jeffrey Turner, Paxton Bachand (graduate student assigned to the project), Jay Tarkington (Outreach Director, TAMU-CC Center for Coastal Studies), and Edgar DeLaGarza (Digital Imaging Specialist, TAMU-CC Marketing and Communications). This was an introductory trip to detail the research problem i.e., the recurrent and persistent brown tide blooms in Baffin Bay. As an ice-breaker, students were asked to participate in the 'draw a scientist activity' (https://www.nsta.org/draw-scientist). This activity was used to assess preconceptions of scientists including gender bias, dispel stereotypes, and spark discussion. Students were also given a peer-reviewed article on brown tides as a reading assignment: The decline and recovery of a persistent Texas brown tide algal bloom in the Laguna Madre (Buskey *et al.*, 2001). Photographs taken during this event can be found at https://photos.tamucc.edu/Events/Events-By-Year/2020/020320-Riviera-High-School-Visit/ (see Appendix A).

The second outreach event occurred 02/19/202 and was attended by Jeffrey Turner, Paxton Bachand, Robert Duke (Research Assistant, TAMU-CC Center for Coastal Studies), and Jay Tarkington. During this event, students were trained, in a classroom setting, how to collect water samples and measure various water quality parameters. This training included the use of a portable water quality sonde for measuring temperature, salinity, conductivity, pH, and dissolved oxygen. TAMU-CC gifted the students university T-shirts and the class was provided a transparency tube

and refractometer for future field sampling activities. This training was reinforced in a hands-on field sampling activity at Los Olmos Creek on 02/24/2020 that was attended by Jeffrey Turner, Paxton Bachand, Jordan Walker, and Edgar DeLaGarza. During this third outreach event, students visited the study site, collected water samples, and recorded water quality parameters. Photographs taken during this event can be found at https://photos.tamucc.edu/Events/Events-By-Year/2020/022420-Riviera-High-School-Water-Sampling/ (see Appendix B). A data collection sheet was provided to record various environmental parameters: water and air temperature, salinity, dissolved oxygen, pH, transparency, nitrate, phosphate, wind intensity, antecedent rainfall, and current weather conditions (see Appendix C).

The intent was to replicate hands-on sampling events on a monthly basis; however, the Kaufer High School cancelled in-person classes in March 2020, due to the COVID-19 pandemic. Classes were also held virtually through Fall 2020 with restrictions and disruptions continuing through Spring 2021. In-person outreach events resumed in October 2021 and were coordinated with the new Ocean Science instructor Selina Ramirez Leos. The fourth event occurred 10/01/2021 and was attend by Jeffrey Turner. During this event, a new cohort of students were introduced to the research problem i.e., the recurrent and persistent brown tide blooms in Baffin Bay. To help students prepare for the visit, they were assigned a documentary on HABs: Troubled Waters (https://calusawaterkeeper.org/troubled-waters/). Student knowledge of HABs was assessed during a discussion about the documentary. Concerns in the documentary were then made relevant to the students by making connections between the documentary and HABs occurring in Baffin Bay. A follow up outreach event occurred 10/22/2021 and was attended by Jeffrey Turner and Paxton Bachand. During this event, Paxton Bachand shared a PowerPoint presentation of the

preliminary results of his brown tide research in Baffin Bay. He then facilitated a discussion about his results and the important of the research in the Baffin Bay watershed.

CONCLUSIONS

Algae blooms have become a progressively widespread issue over the past half-century, having their increased occurrence and range attributed to coastal eutrophication and climate change (Bricker *et al.*, 2008). Increased coastal urbanization and warming waters will continue to exacerbate the formation and proliferation of algae blooms, highlighting the need for a better baseline understanding. Brown tides caused by *A. lagunensis* occur regularly in the northwestern Gulf of Mexico (Cira and Wetz, 2019) and more recently in the Indian River Lagoon (Florida) and Guatanamo Bay (Cuba) (Gobler *et al.*, 2013; Hall *et al.*, 2018). The reoccurrence of persistent blooms in Baffin Bay makes the bay a natural laboratory for studying brown tide dynamics. This study followed the formation of a dense brown tide and characterized the water quality parameters and microorganisms (i.e., prokaryotes and viruses) associated with the bloom. Although previous studies have reported relationships between brown tide blooms and physical and chemical parameters (e.g., temperature, salinity, organic nutrient concentrations), this study expands our knowledge of brown tides. In particular, this was the first study to show that brown tide bloom-associated microbial communities are unique in comparison to non-bloom communities. Additionally, this was the first study to propose that guano from a nearby free-tailed bat colony could be contributing to the nutrient pollution driving the bloom. Ongoing research is focused on 1) further defining the roles these microorganisms play in modulating bloom dynamics, 2) identifying virus–host interactions associated with the bloom, and 3) investigating potential connections between the bloom and the nutrient-rich guano deposited in Los Olmos Creek by free-tailed bats.

REFERENCES

Amin, S. A., Parker, M. S., and Armbrust, E. V. (2012). Interactions between diatoms and bacteria. *Microbiol Mol Biol Rev* 76(3): 667–684.

Arbizu, P.M. (2017). pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.0, 1.

Barton, K. and Barton, M.K. (2021). Package 'MuMin'. Version 1 (18): 439.

Bersano, J.G.F., Buskey, E.J., and Villareal, T.A. (2002) Viability of the Texas brown tide alga, *Aureoumbra lagunensis*, in fecal pellets of the copepod *Acartia tonsa*. *Plankt Biol Ecol* 49: 88–92.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, and G.A., Alexander *et al.* (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 378(37): 852–857.

Bricker, S.B., Longstaff, B., Dennison, W., Jones, A., Boicourt, K., Wicks, C., and Woerner, J. (2008). Effects of nutrient enrichment in the nation's estuaries: A decade of change. *Harmful Algae* 8: 21–32.

Buchan, A., LeCleir, G.R., Gulvik, C.A., and González, J.M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol* 12: 686–698.

Bushnell, B., Rood, J., and Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. *PLoS One*: e0185056.

Buskey, E. J., and Hyatt, C. J. (1995). Effects of the Texas (USA) 'brown tide' alga on planktonic grazers. *MEPS* 126: 285–292.

Buskey, E.J., Montagna, P.A., Amos, A.F., and Whitledge, T.E. (1997). Disruption of grazer

populations as a contributing factor to the initiation of the Texas brown tide algal bloom. *Limnol Oceanogr* 42: 1215–1222.

Buskey, E.J., Wysor, B., and Hyatt, C. (1998). The role of hypersalinity in the persistence of the Texas 'brown tide' in the Laguna Madre. *J Plankton Res* 20: 1553–1565.

Buskey, E.J., Liu, H., Collumb, C., and Bersano, J.G.F. (2001). The decline and recovery of a persistent Texas brown tide algal bloom in the Laguna Madre (Texas, USA). *Estuaries* 24: 337–346.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 137(13): 581–583.

Caporaso, J. Gregory, *et al.*, (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7(5): 335-336.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., and Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6: 1621–1624.

Chen, F., Suttle, C.A., and Short. S.M. (1996). Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes." *Appl Environ Microbiol* 62(8): 2869–74.

Cira, E.K., and Wetz, M.S. (2019). Spatial-temporal distribution of *Aureoumbra lagunensis* ("brown tide") in Baffin Bay, Texas. *Harmful Algae* 89: 101669.

Coy S.R., Gann E.R., Pound H.L., Short S.M., and Wilhelm S.W. (2018). Viruses of eukaryotic algae: diversity, methods for detection, and future directions. *Viruses*. 10(9): 487.

Faircloth, B., and Glenn, T. (2011). *Serapure*. Ethanomics. https://ethanomics.files.wordpress.com/2012/08/serapure_v2-2.pdf

Filée, J., Tétart, F., Suttle, C.A., and Krisch, H.M. (2005). Marine T4-yype bacteriophages, a ubiquitous component of the dark matter of the biosphere. *PNAS* 102(35): 12471–76.

Frazier, A.D., Rowe, J.M., Rentz, C.A., Gobler, C.J., and Wilhelm, S.W. (2007). Bacterial lysis of *Aueococcus anophagefferens* CCMP 1784 (Pelagophyceae). *Phycol Soc Am* 43: 461–465.

Frigaard, N. U., and Dahl, C. (2008). Sulfur metabolism in phototrophic sulfur bacteria. *Adv Microbial Phys* 54: 103–200.

Furusawa, G., Yoshikawa, T., Yasuda, A., and Sakata, T. (2003). Algicidal activity and gliding motility of *Saprospira sp.* SS98-5. *Can J Microbiol* 49(2): 92–100.

Gobler, C.J. and Sunda, W.G. (2012). Ecosystem disruptive algal blooms of the brown tide species, *Aureococcus anophagefferens* and *Aureoumbra lagunensis*. *Harmful Algae* 14: 36–45.

Gobler, C. J., Koch, F., Kang, Y., Berry, D. L., Tang, Y. Z., Lasi, M., and Miller, J. D. (2013). Expansion of harmful brown tides caused by the pelagophyte, *Aureoumbra lagunensis* DeYoe et Stockwell, to the US east coast. *Harmful Algae 27*: 29–41.

Gómez-Consarnau, L., Sachdeva, R., Gifford, S. M., Cutter, L. S., Fuhrman, J. A., Sañudo-Wilhelmy, S. A., and Moran, M. A. (2018). Mosaic patterns of B-vitamin synthesis and utilization in a natural marine microbial community. *Environ Microbiol* 20(8): 2809–2823.

Hall, N.S., Litaker, R.W., Kenworthy, W.J., Vandersea, M.W., Sunda, W.G., Reid, J.P., Slone, D.H., and Butler, S. (2018). Consortial brown tide − picocyanobacteria blooms in Guantánamo Bay, Cuba. *Harmful Algae* 73: 30–43.

Hoagland, P., Anderson, D., Kaoru, Y., and White, A.W. (2002). The economic effects of harmful algal blooms in the United States: estimates, assessment issues, and information needs. *Estuaries* 25: 819–837.

John, S. G., Mendez, C. B., Deng, L., Poulos, B., Kauffman, A. K. M., Kern, S., Brum, J., Polz, M. F., Boyle, E. A., and Sullivan, M. B. (2011). A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ Microbiol Rep*, 3(2): 195–202.

Kildow, J., Colgan, C.S., Farnum, M.G., Johnson, P., and Scorse, J. (2016). State of the U.S. Ocean and Coastal Economies 2016 Update.

Kraeuter, J.N., Klinck, J.M., Powell, E.N., Hofmann, E.E., Buckner, S.C., Grizzle, R.E., and Bricelj, V.M. (2008). Effects of the fishery on the northern quahog (hard clam, *Mercenaria mercenaria* L.) population in Great South Bay, New York: A Modeling Study. *J Shellfish Res* 27: 653–666.

Labonté, J.M., Reid, K.E., and Suttle, C.A. (2009). Phylogenetic analysis indicates evolutionary diversity and environmental segregation of marine Podovirus DNA polymerase gene sequences. *Appl Environ Microbiol* 75(11): 3634–40.

Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10), 1674-1676.

Liu, H. and Buskey, E.J. (2000). The exopolymer secretions (EPS) layer surrounding *Aureoumbra lagunensis* cells affects growth, grazing, and behavior of protozoa. *Limnol Oceanogr* 45: 1187–1191.

Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J., and Knight, R. (2010). UniFrac: an

effective distance metric for microbial community comparison. ISME J. 52(5): 169–172.

McMurdie, P.J., and Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8: e61217.

Mojica, K. D. A., and Brussaard, C. P. D. (2014). Factors affecting virus dynamics and microbial host–virus interactions in marine environments. *FEMS Microbiol Ecol* 89(3): 495–515.

Nissimov, J. I., Vandzura, R., Johns, C. T., Natale, F., Haramaty, L., and Bidle, K. D. (2018). Dynamics of transparent exopolymer particle production and aggregation during viral infection of the coccolithophore, *Emiliania huxleyi*. *Environ Microbiol* 20(8): 2880–2897.

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., and Stevens, M.H.H. (2019). Vegan: community ecology package (version 2.5-6). *The Comprehensive R Archive Network*.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., and Solymos, P., *et al.* (2020). *vegan: Community Ecology Package. R package version 2.5-6.*

Onuf, C.P. (1996). Seagrass responses to long-term light reduction by brown tide in upper Laguna Madre, Texas: distribution and biomass patterns. *Mar Ecol Prog Ser* 138: 219–231.

Orth, R.J., Carruthers, T.J.B., Dennison, W.C., Duarte, C.M., Fourqurean, J.W., and Heck, K.L., *et al.* (2006). A global crisis for seagrass ecosystems. *Bioscience* 56: 987–996.

Pelusi, A., de Luca, P., Manfellotto, F., Thamatrakoln, K., Bidle, K. D., and Montresor, M. (2021). Virus-induced spore formation as a defense mechanism in marine diatoms. *New Phytologist*, *229*(4): 2251–2259.

Poulos, B. T., John, S. G., and Sullivan, M. B. (2018). Iron chloride flocculation of

bacteriophages from seawater. In *Bacteriophages* (pp. 49–57). Springer.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41: D590.

Tan, S., Zhou, J., Zhu, X., Yu, S., Zhan, W., Wang, B., and Cai, Z. (2015). An association network analysis among microeukaryotes and bacterioplankton reveals algal bloom dynamics. *J Phycol* 51: 120–132.

R Development Core Team. (2015). R Internals. In *R Development Core Team* (R Foundati, Vol. 1). https://doi.org/3-900051-14-3

Ren, G., Xu, X., Qu, J., Zhu, L., and Wang, T. (2016). Evaluation of microbial population dynamics in the co-composting of cow manure and rice straw using high throughput sequencing analysis. *World J Microbiol Biotechnol* 32(6): 1–11.

Rosenwasser, S., Ziv, C., Creveld, S.G. van, and Vardi, A. Virocell (2016). Metabolic innovations during host virus interactions in the ocean. *Trends Microbiol. 24*: 821–832

Sheridan, P. and Minello, T.J. (2003). Nekton use of different habitat types in seagrass beds of lower Laguna Madre, Texas. *Bull Mar Sci* 72: 37–61.

Short, S. M. (2012). The ecology of viruses that infect eukaryotic algae. *Environ Microbiol 14*(9): 2253–2271.

Short, S.M., Staniewski, M.A., Chaban, Y.V., Long, A.M., and Wang, D. (2020). Diversity of viruses infecting eukaryotic algae. *Curr Issues Mol Biol* 39: 29-62.

Vega, R.R., Neill, W.H., Gold, J.R., and Ray, M.S. (2011). Enhancement of Texas Sciaenids (Red Drum and Spotted Seatrout). In *Interactions of Fisheries and Fishing Communities Related to Aquaculture Proceedings of the Thirty-eighth U.S.-Japan Aquaculture Panel*

Symposium. pp. 1–12.

Suttle, C. (2007). Marine viruses — major players in the global ecosystem. *Nat Rev Microbiol* 5: 801–812.

Thingstad, T. F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr* 45(6): 1320–28.

Tunnell, J.W., and Judd, F.W. (2002). The Laguna Madre of Texas and Tamaulipas, 1st ed. ed. Texas A & M University Press, College Station.

Vavrek M.J. (2011). fossil: palaeoecological and palaeogeographical analysis tools. *Palaeontologia Electronica* 14(1), 1T. R package version 0.4.0.

Wetz, M.S., Hutchinson, E.A., Lunetta, R.S., Paerl, H.W., and Taylor, J.C. (2011). Severe droughts reduce estuarine primary productivity with cascading effects on higher trophic levels. *Limnol Oceanogr* 56: 627–638.

Wetz, M.S., Cira, E.K., Sterba-Boatwright, B., Montagna, P.A., Palmer, T.A., and Hayes, K.C. (2017). Exceptionally high organic nitrogen concentrations in a semi-arid South Texas estuary susceptible to brown tide blooms. Estuar Coast Shelf Sci 188: 27–37.

Zhao, Y., Temperton, B., Thrash, J.C., Schwalbach, M.S., Vergin, K.L., Landry, Z.C., Ellisman, M., Deerinck, T., Sullivan, M.B., and Giovannoni, S.J. (2013). Abundant SAR11 viruses in the ocean. *Nature* 494 (7437): 357–60.

Zhou, J., Richlen, M.L., Sehein, T.R., Kulis, D.M., Anderson, D.M., and Cai, Z. (2018). Microbial community structure and associations during a marine dinoflagellate bloom. *Front Microbiol* 9: 1–21.

FIGURES

Figure 1. Map showing locations of Los Olmos Creek (bloom) and Riviera Beach (non-bloom) study sites in Baffin Bay, Texas, United States. Figure rendered with GIS. Map credit Md Mahabubur Rahman.

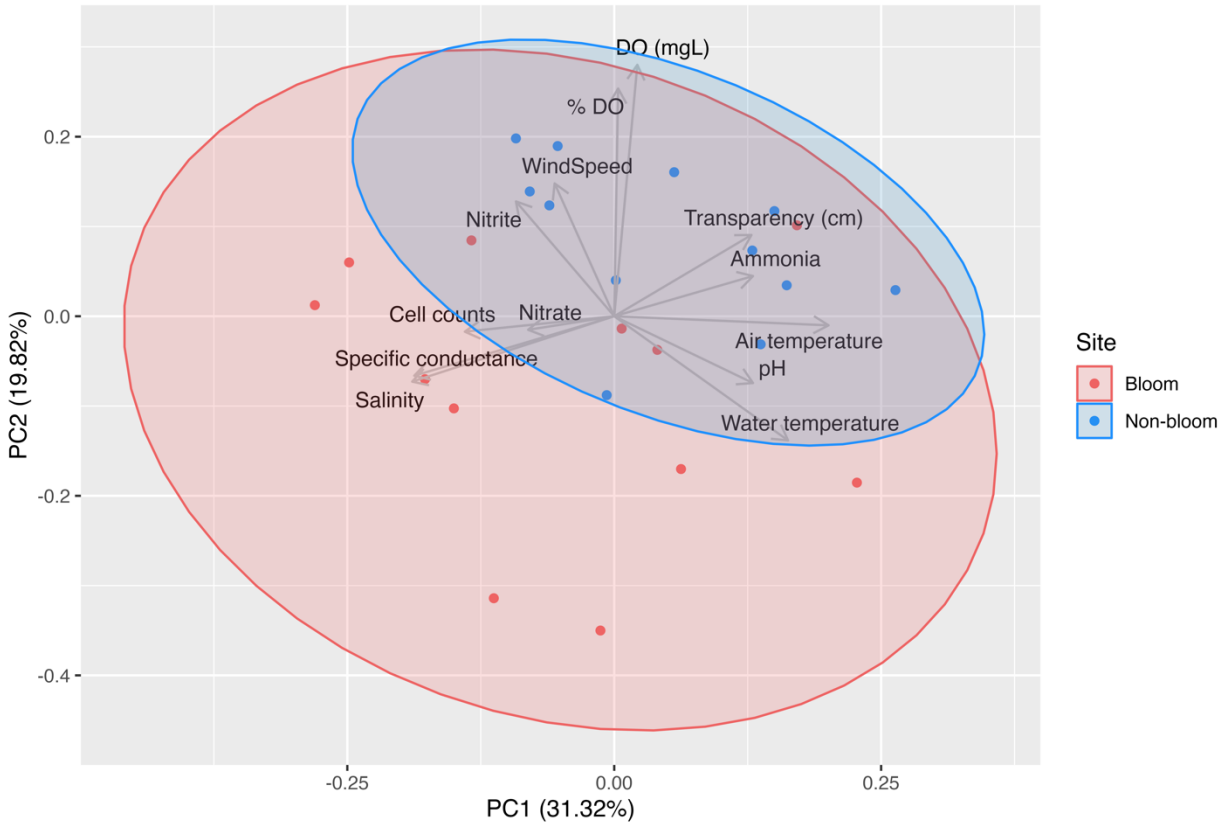


Figure 2. Principle component analysis (PCA) of water quality parameters of the bloom and non-bloom sites.

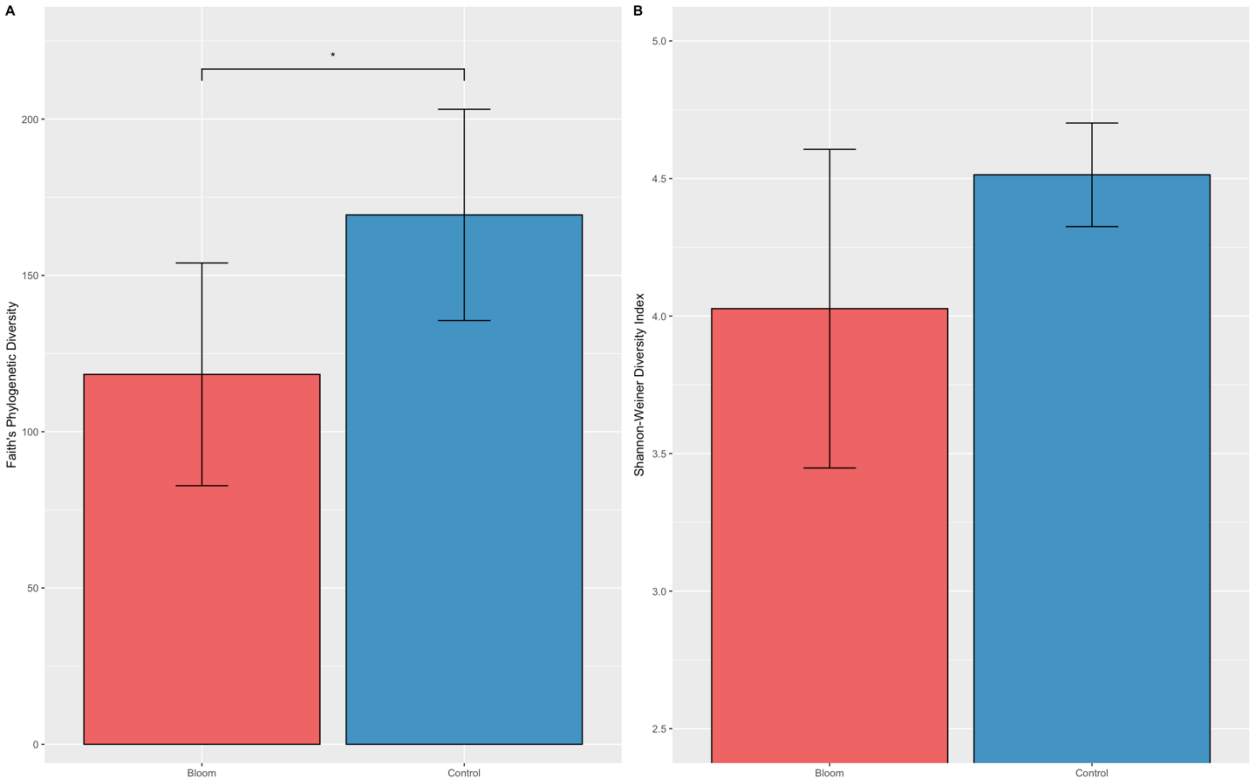


Figure 3. Subset bar-plots of Faith's phylogenetic diversity (A) and Shannon-Weiner diversity (B) values for bloom and non-bloom sites.

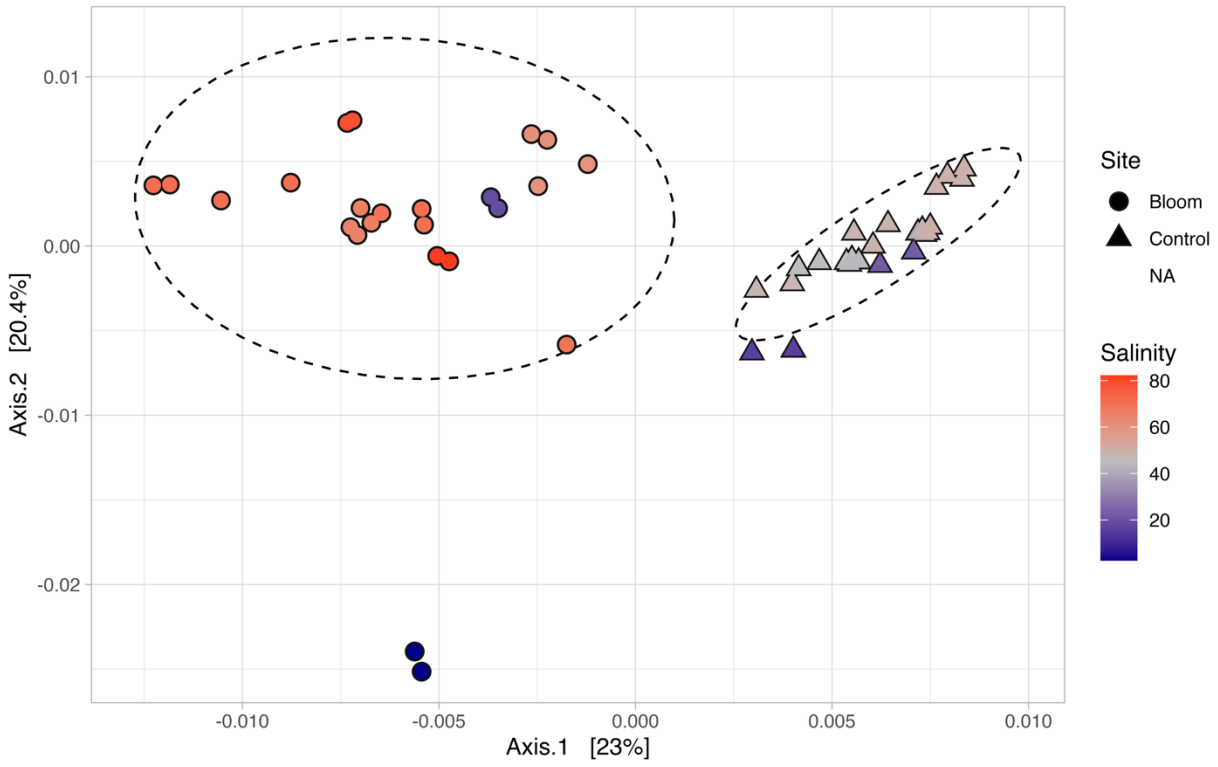


Figure 4. Principal coordinate analysis (PCoA) of beta diversity values from the bloom and non-bloom sites using weighted UniFrac distance values. Communities in the bloom and non-bloom samples were compositionally distinct (p-value < 0.05). Site "NA" represents non-template control.

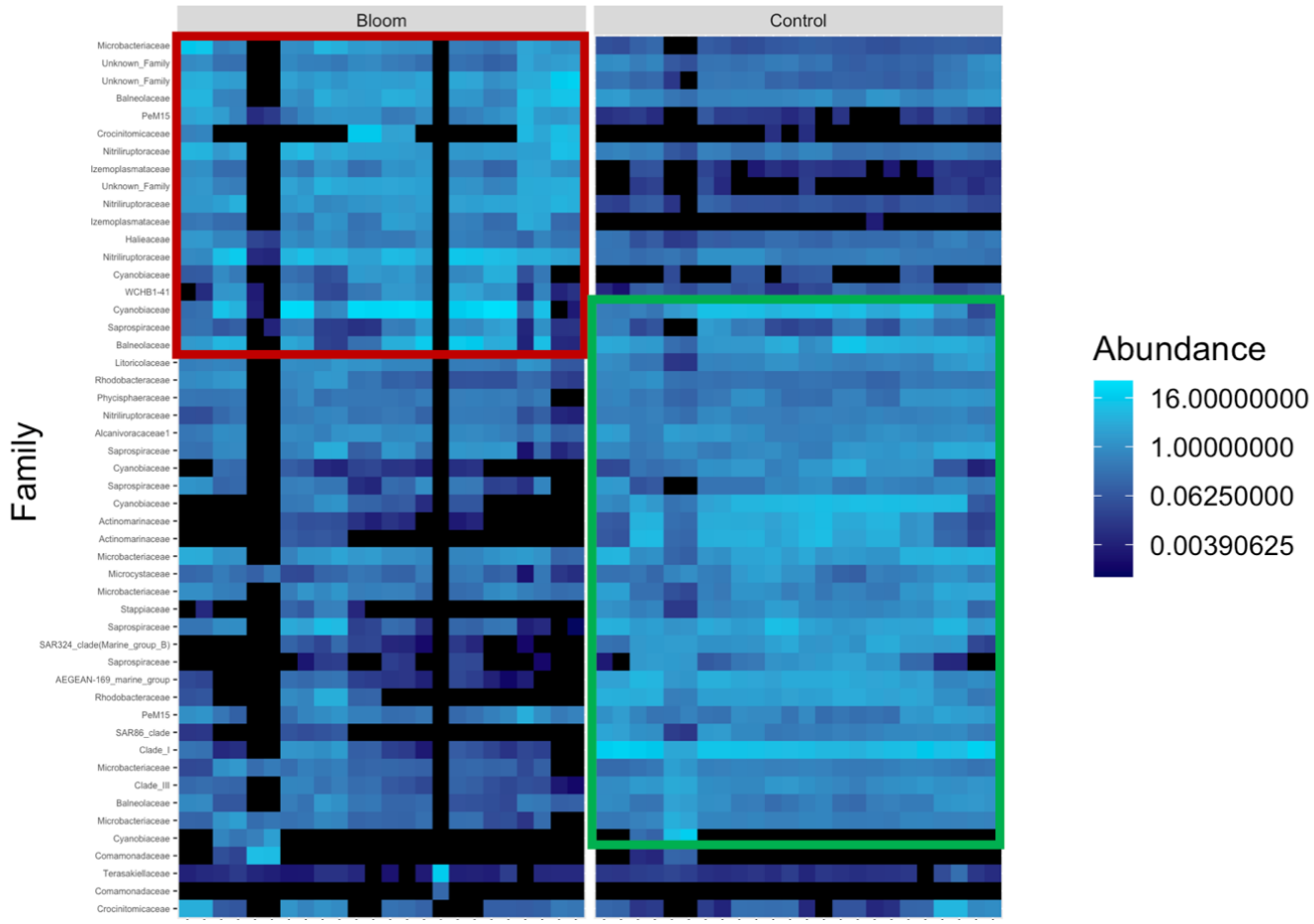


Figure 5. Heatmap of the 50 most abundant taxa amplicon sequence variants (ASVs) observed at the bloom (left) and non-bloom (right) sites.

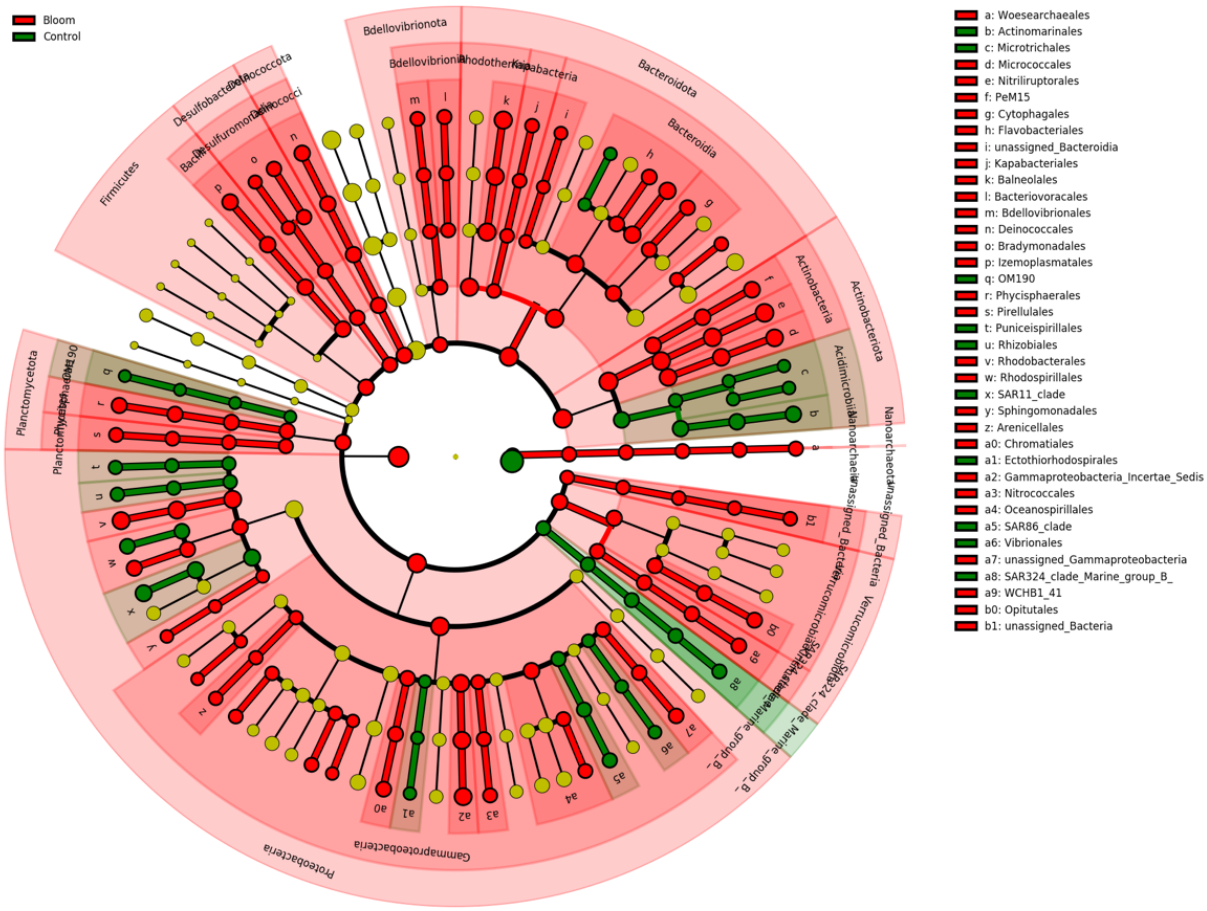


Figure 6. Linear discriminant analysis (LDA) effect size (LEfSe) cladogram depicting enriched taxa enriched of the bloom (red) and non-bloom control (green) control site samples (p-value < 0.01).

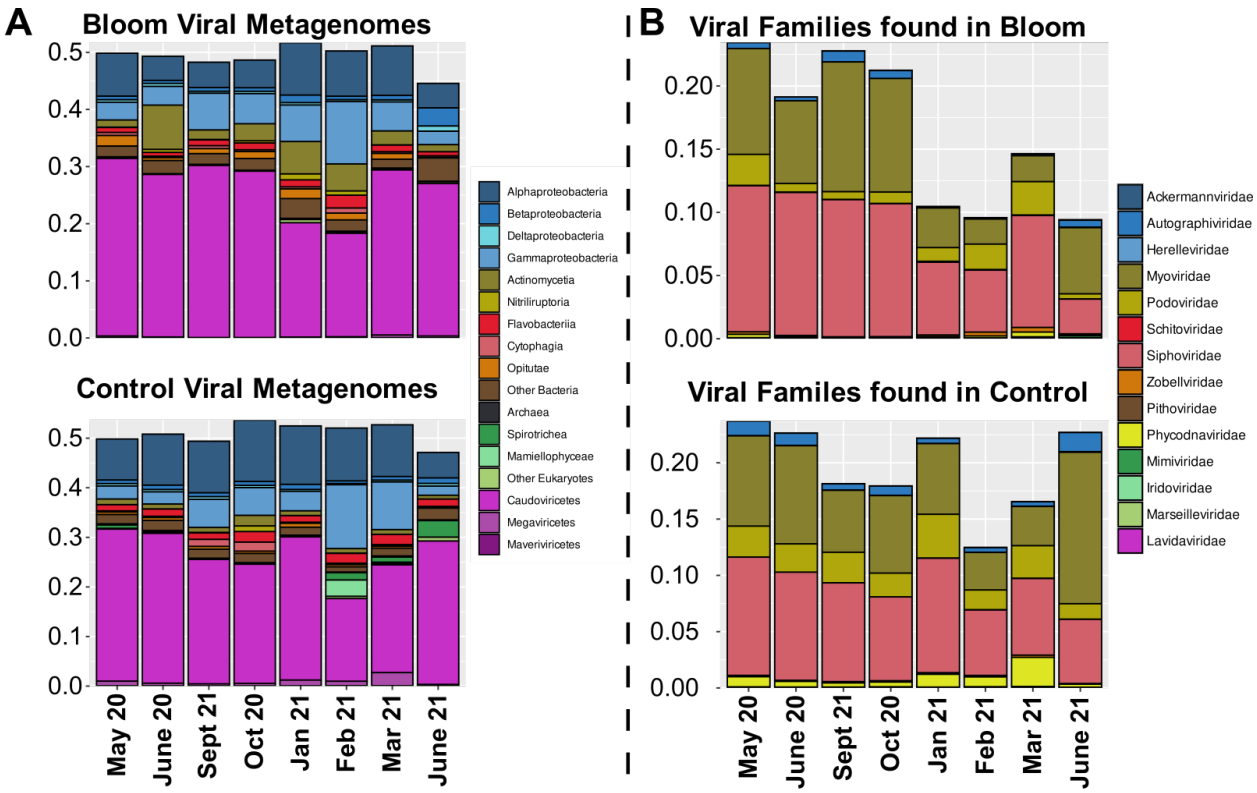


Figure 7. A) Stacked bar histograms of the Class level taxonomic assignments of identifiable sequences in the viral metagenomes. B) Stacked bar histograms of the sequences which were classifiable as viral at the Family level in the viral metagenomes.

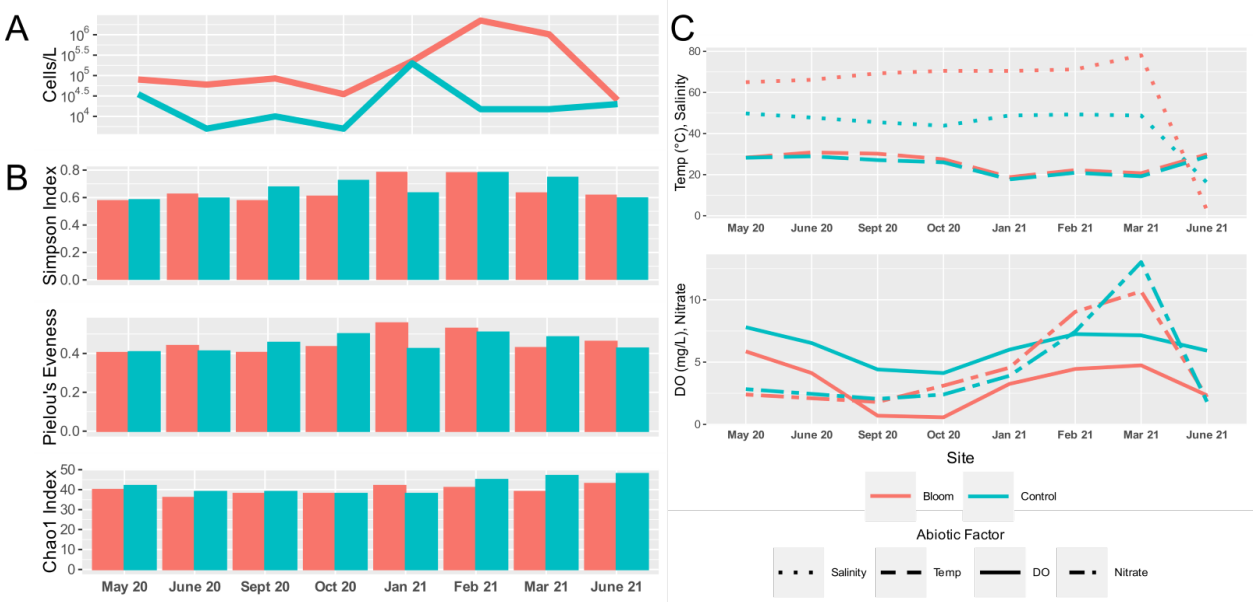


Figure 8. A) Fluorescent cell counts of samples taken at the bloom and control sites. B) Diversity, evenness, and richness indices of the viral metagenomes at the bloom and control sites. C) Abiotic factors that were considered significant at an $\alpha < 0.05$ according to the NMDS.

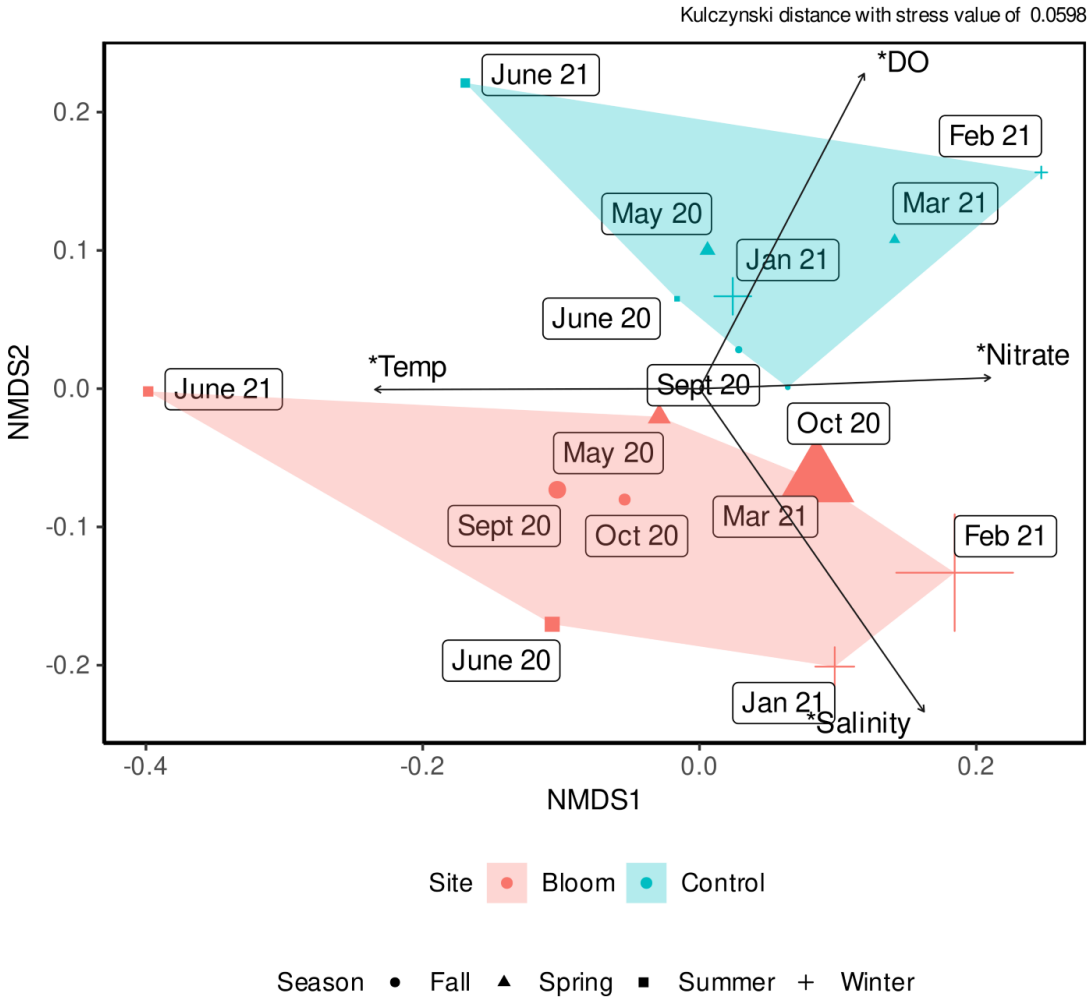


Figure 9. NMDS plot of the taxonomic assignments of the viral metagenomes. The size of the points on the plot are proportional to the number of cells counted in each sample. The vectors indicate abiotic factors that were significantly correlated to the ordination plot at $\alpha < 0.1$, those significant at an $\alpha < 0.05$ are indicated with *.

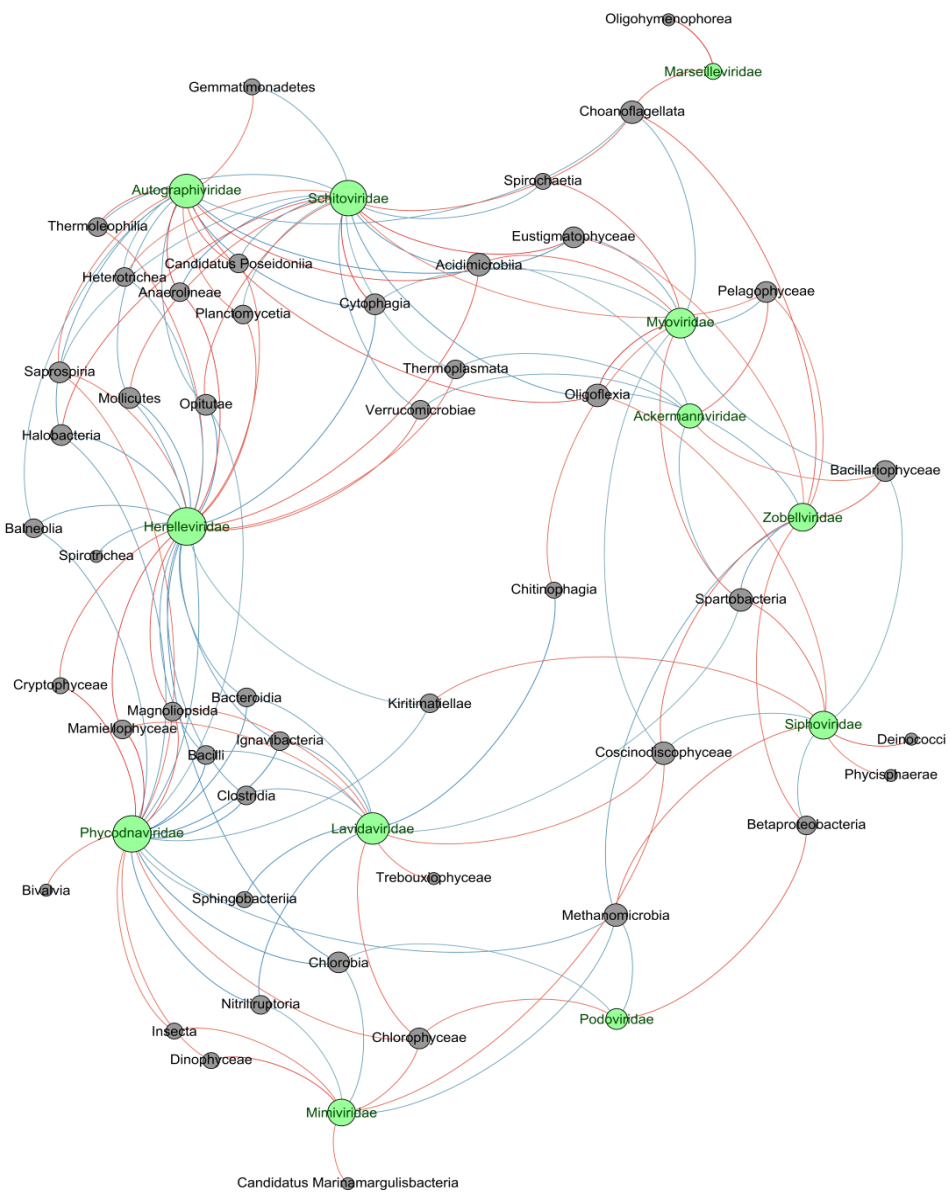


Figure 10. Network analysis of viruses that co-occurred (red) or were negatively related (blue) to the classes within the bacterial metagenomes. Green dots indicate viral families and gray dots indicate bacterial Classes. The size of the dot is relative to the number of connections that bacterial class or viral family contains. The network is based on a Spearman Rank-Order Correlations with $\alpha < 0.10$.

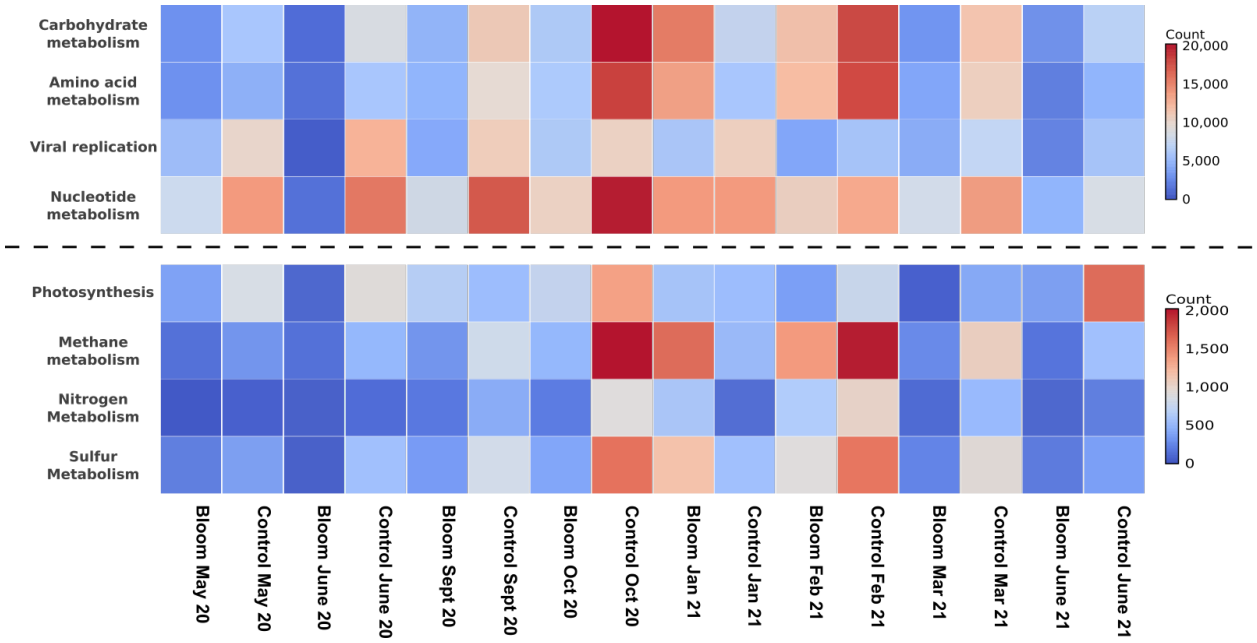
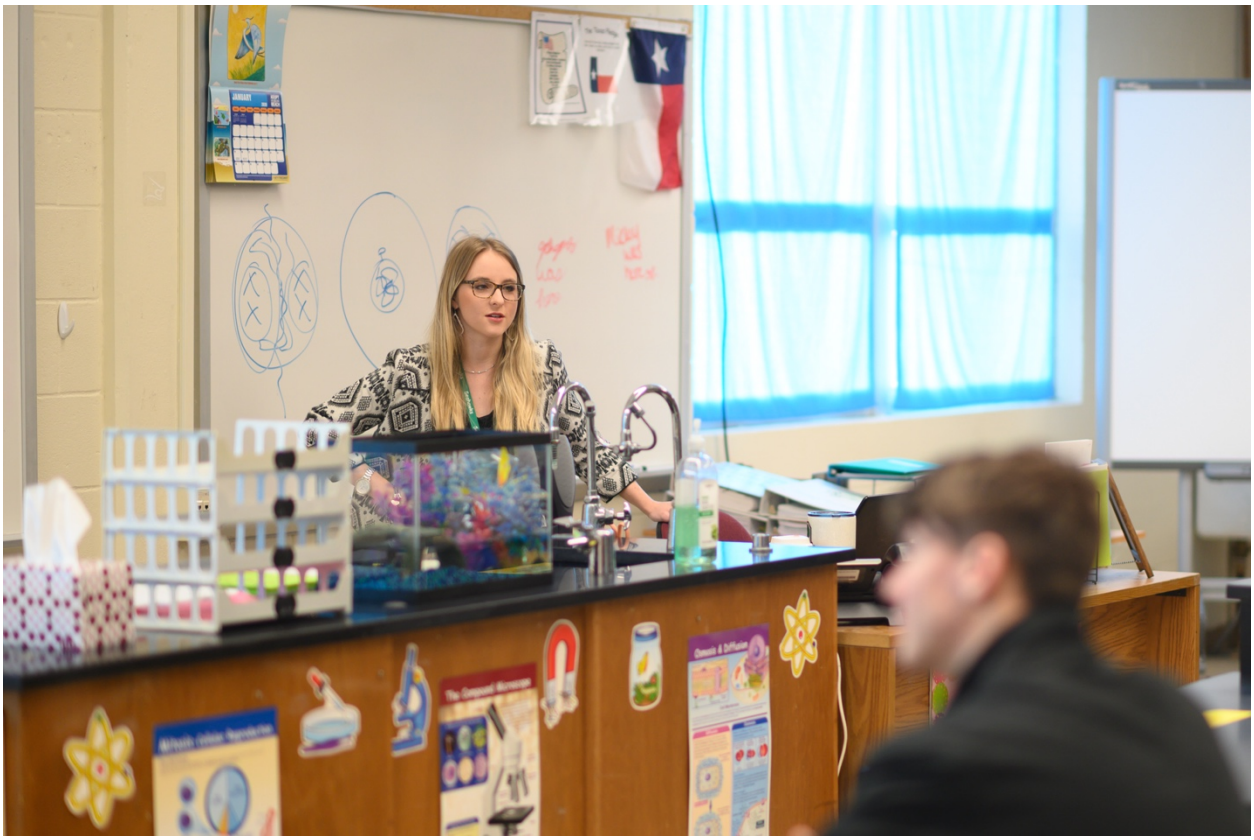


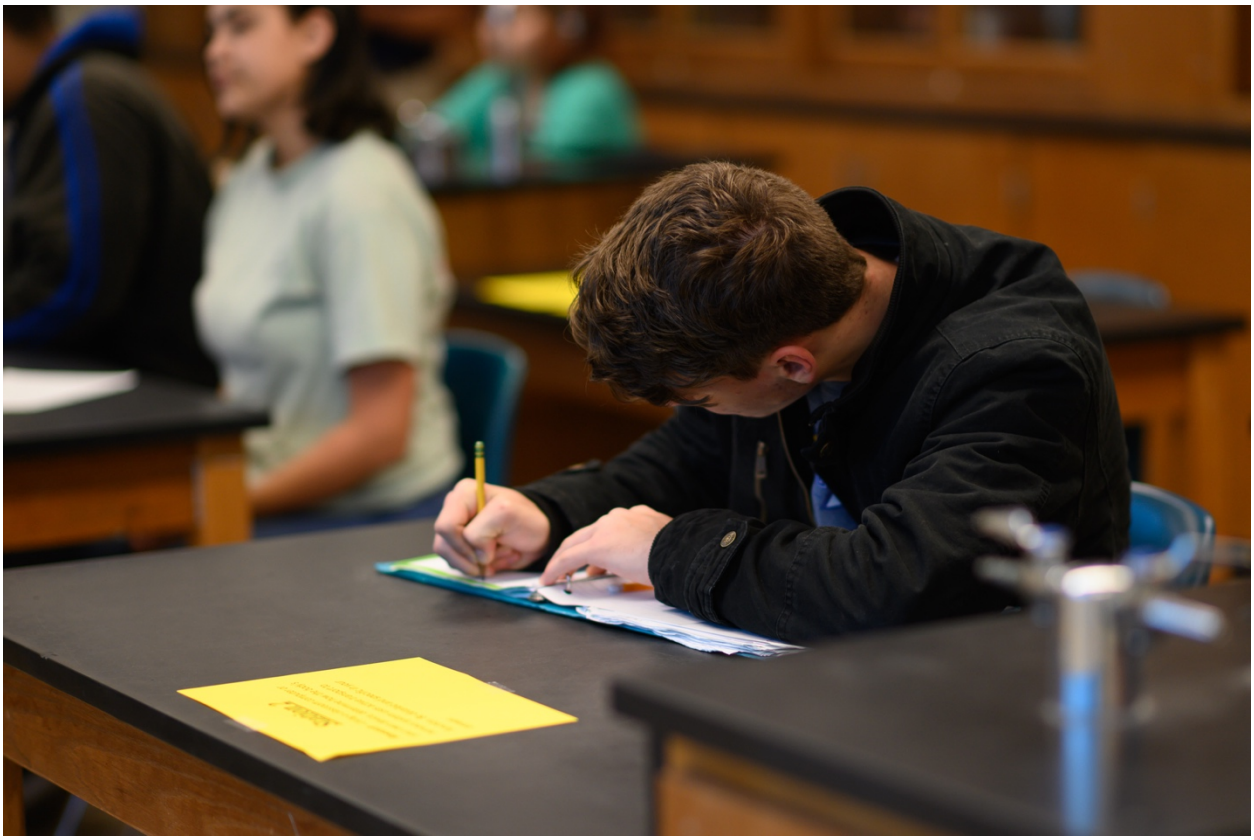
Figure 11. Heatmap of the sequences within the viral metagenomes assigned specific metabolic capabilities typical of viruses. The separation of the plots is to indicate that the two plots are on separate scales.

APPENDIX A: Photographs of 02/03/2020 outreach event at Kaufer High School




SCIENTIFIC METHOD
1. STATE THE QUESTION
2. COLLECT INFORMATION
3. FORM A HYPOTHESIS
4. TEST THE HYPOTHESIS


The Compound Microscope
Osmosis & Diffusion




RIVIERA
FIGHTING
SEAHAWKS
INDEPENDENT SCHOOL DISTRICT
Happy National
Counseling Week!
Sponsored By:
RIVIERA TELEPHONE
RTC
COMPANY, INC.


Rest Station
Use this time wisely.
Complete any station work
not yet finalized.


RIVIERA★
SEAHAWKS

Appendix B. Photographs of 02/24/2020 outreach event at Kaufer High School


SLANDER
RIDE


ISLANDER
PRIDE


Seahawks
Rivera ISD


DO IT
RIGHT
LOWE'S


DO IT
RIGHT
LOWE'S

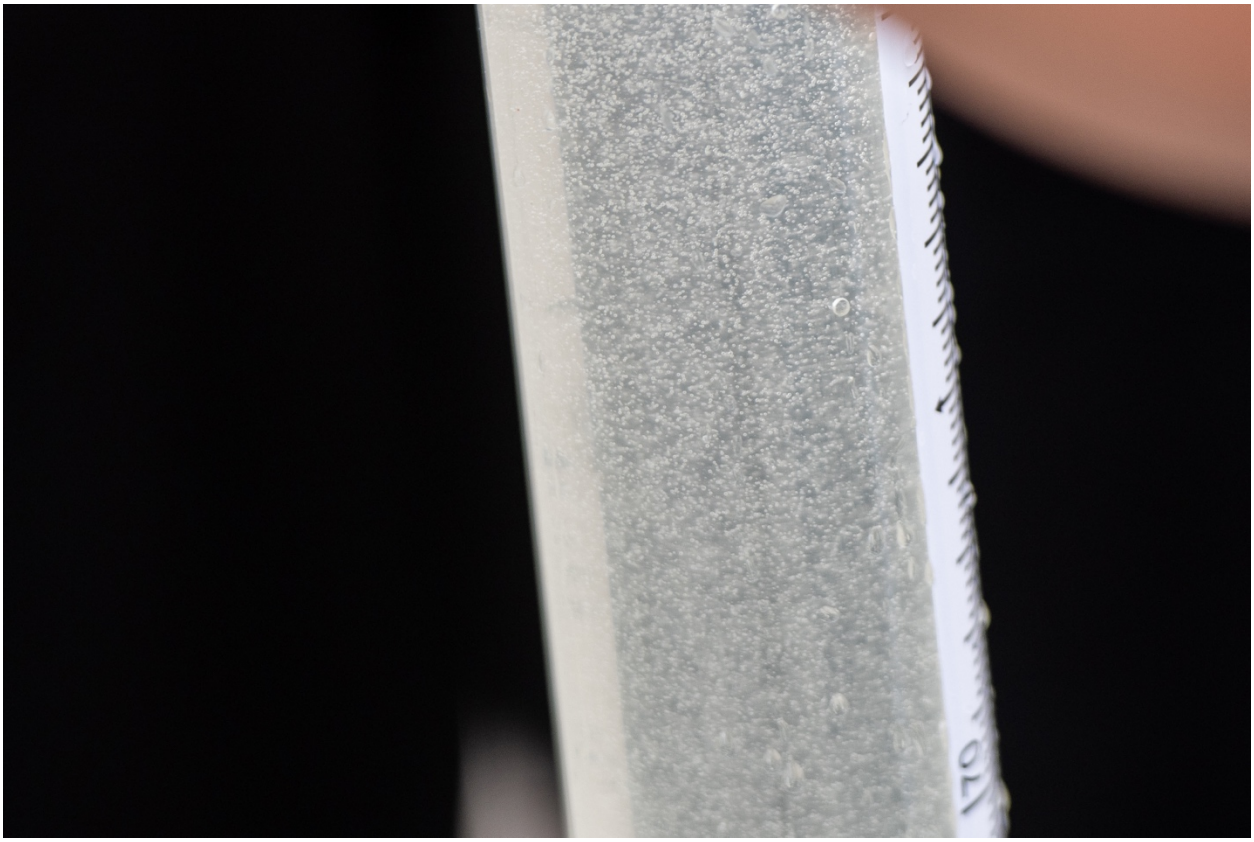

TEXAS A&M-CORPUS CHRIS
ISLANDER
PRIDE


ISLANDER
PRIDE


TEXAS A&M-CORPUS CHRISTI
ISLANDER
PRIDE


ISLANDER
PRIDE


ESTUARY
&MARINE
Monitoring




UA
UNDER ARMOUR


ISLANDER
PRIDE


pH
coliform
negative
positive
1 ppm
2 ppm
4 ppm
turbidity
0 JTU
40 JTU


turbidity
test fill line
kit
Code 5911


ÍSLAND


ISLANDER
PRIDE
LOWE'S


LANDER
PRIDE

APPENDIX C. Education and training materials developed for outreach events.

Date: _____

Time: _____

Sampling location: _____

Environmental parameters	**Value**	**Parameter code**
Water temperature (°C)		
Salinity (ppt)		Avg of 3 measures
Dissolved oxygen (% sat)		
pH		
Transparency (cm)		Avg of 3 measures
Nitrate (ppm)		
Phosphate (ppm)		
Wind intensity (mph)		
Air temperature (°C)		
Present weather		Clear, Cloudy, Overcast, Rain
Days since last rainfall		
Rainfall (inches past 1 day)		
Rainfall (inches past 7 days)		
Water color		Clear, Green, Brown, Other
Water odor		Sewage, Fishy, None, Other
Water surface		Calm, Ripples, Waves
Tidal stage		Low, Falling, Slack, Rising, High