MICROBIAL COMMUNITIES AND ECOSYSTEM RECOVERY OF GALVESTON BAY POST-HURRICANE HARVEY

An Undergraduate Research Scholars Thesis

by

JORDAN RONALD WALKER

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Jessica M. Labonté

May 2019

Major: Marine Biology

TABLE OF CONTENTS

ABSTRACT

Microbial Communities and Ecosystem Recovery of Galveston Bay Post-Hurricane Harvey

Jordan Ronald Walker
Department of Marine Biology
Texas A&M University

Research Advisor: Dr. Jessica M. Labonté
Department of Marine Biology
Texas A&M University

On August 25, 2017, Hurricane Harvey made landfall off the coast of Texas then stalled over the Houston area for five days, setting the record for the highest rainfall in the United States. The waters drained through Houston and into Galveston Bay bringing terrestrial, freshwater, and anthropogenic chemicals and microbes with them. Since microbial communities are constantly changing and adapting to the ecosystem around them this study aimed to characterize the impact of Hurricane Harvey on the microbial communities in the Galveston Bay using 16S ribosomal RNA (rRNA) genes and metagenomic data. Heterotrophs were most prevalent directly following the storm and, throughout the 5-week study period, the autotrophic microbes rebounded to pre-Hurricane Harvey levels. Preliminary data analysis has shown that metagenomic data can provide a glimpse into the metabolism of prevalent groups. Further metagenomic analysis will provide insight into the specific metabolic pathways driving the recovery of Galveston Bay's microbial community. The expectation of larger and more intense storms in North America will increase the amount of stormwater runoff per year in ecologically and economically important ecosystems. Understanding the effects of these storms and how microbial communities respond will help inform management and recovery efforts.

ACKNOWLEDGEMENTS

I would like to thank my principal investigator Dr. Jessica M. Labonté for not only her support and guidance but the opportunities she has allowed me to pursue in the amazing and fascinating world of marine microorganisms.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University at Galveston as rich, fun, and rewarding as it could have possibly been. Special thanks to my lab mates who have helped me develop and hone the skills required to complete this work. I also want to extend my gratitude to Texas Sea Grant and Aggies Commit to Excellence for their support to purchase lab equipment necessary to the research. I would like to specifically thank Dr. Orissa Moulton in her role with Aggies Commit to Excellence and guiding me through the process of applications as well as guiding me when I was searching for lab experiences.

Thanks to my fiancé, Brenda Guzman, for all her love and encouragement over the past two years as I have chased my dream of becoming a researcher. I would also like to thank my parents, Ron and DeAnn Walker, for all of their patience and support through everything.

CHAPTER I

INTRODUCTION

Bacteria in the ocean have been estimated at ~one million organisms per ml of seawater (Schmidt et al. 1998) while viral particles are estimated to be one to two orders of magnitude more abundant than the bacterial estimates (Williamson et al. 2014). Microbes make up >90% of the living biomass in the ocean (Suttle 2007) and influence global nutrient cycles, local fisheries, and coastal public health (Mallin et al. 2001; Williamson et al. 2014). The "microbial loop" was first described by Azam et al. (1983) which states that carbon and other nutrients are inefficiently introduced to the main food web by the loop of dissolved organic matter (DOM) between bacteria, flagellates, and microzooplankton. Estimates of viral populations and their ability to lyse bacteria and introduce additional DOM into the microbial loop have only inflated the effect of the microbial loop on the ocean's food web (Suttle 2007). As the understanding of the size and complexities of microbial communities in aquatic environments has grown, so has the understanding of their importance and contribution to global nutrient cycles.

Modern microbiology and its challenges

Microbes from ocean and coastal environments remain notoriously challenging to study due to the difficulty of cultivating the majority of environmental bacteria (Handelsman 2004). Uncultured microorganisms represent a majority of the microbial abundance in the world and the inability to culture these organisms has presented a substantial barrier to understanding their morphology, physiology, and ecology (Pace et al. 1986; Handelsman 2004). Estimates of the percentage of culturable bacteria in environmental samples has been shown to be as low as 0.1-1.0% of total bacterial abundance in various environments, including soil, fresh- and marine waters

(Rappé and Giovannoni 2003). Concerted efforts to characterize and monitor the world's aquatic microbial communities on a large scale began with major expeditions including Sorcerer II Global Ocean Survey in 2007 and the Tara Oceans expedition in 2017 (Armbrust and Palumbi 2015). Both expeditions were extremely successful in assessing the global microbial diversity using metagenomics – the direct extraction and sequencing of total DNA from an environment – as a means to study microorganisms without the use of cultivation.

Molecular approaches have provided a path to which microbes can be studied without having to use cultivation. The first molecular studies of environmental microbes were completed in 1986 by Pace and colleagues when they sequenced 5S and 16S ribosomal RNA (rRNA) (Pace et al. 1986). This and many modern studies employ the use of shotgun sequencing, in which DNA or RNA is broken into short segments which are then sequenced independently. The sequenced segments are then aggregated into libraries of reads and assembled into partial or complete genomes using computers (Segata et al. 2013). The gathering of large amounts of data including DNA, RNA, and proteins required that computational methods be developed to analyze the vast amounts of data being collected. Bioinformatics, or the science of utilizing computers to retrieve, manipulate, distribute, and analyze information related to biological macromolecules, became necessary in order to overcome the challenges involved with understanding the data collected from genomic and metagenomic studies. The advent of the internet transformed this field by allowing for the aggregation of biological databases across the world so that sequences, models, and biological knowledge could be shared and compared to all known data (Tyson et al. 2004).

Stormwater impacts on microbial communities

Hurricane Harvey made landfall off the southern Texas coast as a Category 4 hurricane on August 25, 2017. The storm stalled over the Texas coast and spent the next five days meandering

from Corpus Christi to Beaumont. Rainfall amounts for those five days in the Houston area were at least 20 inches and in some areas as high as 48.2 inches (Liberto 2017). The rain filled reservoirs past capacity which when drained flowed through the city's bayous causing additional floods over the next ten days. With much of Houston being situated in Galveston Bay Watershed, the resulting stormwater runoff from the hurricane and release of the reservoirs meant that Galveston Bay would see a constant influx of huge amounts of cold stormwater containing anthropogenic and soil contaminants. These conditions provided a unique opportunity to study the effects of large-scale terrestrial and anthropogenic stormwater runoff on resident microbial populations of a coastal/estuarine ecosystem that is directly influenced by the fourth largest urban area in the United States.

Coastal urban areas have the potential to affect the resident microbial communities as terrestrial and anthropogenic sources of microbes and nutrients are introduced. Mallin et al. (2001) were able to use historical data relating to urbanization, farmland, and shellfish harvest closures, due to high fecal coliform bacteria levels, to show a relationship between population, farmland, impervious surfaces, and higher concentration of fecal coliform bacteria. A study on large rainfall events was conducted by Williamson et al. (2014) where they were able to identify connections between bacteria concentrations, changes in water characteristics, and viral concentrations; However, the study was completed in freshwater and neither study attempted to identify or characterize the microbes. This project will use metagenomics to identify how resident microbial communities responded to the introduction of transient microbes and pollution from a large-scale rain event.

The objective of this project is to identify the microbial populations and their metabolic potential present before and after Hurricane Harvey, and how that community changed over a five-

week period from September 4 to September 28, 2017 using metagenomic techniques. It is hypothesized that anthropogenic, freshwater, and soil bacteria will dominate the microbials populations directly following the storm, and that as the water warms and the salinity returns to that of an estuary, the population will shift to one containing mostly coastal marine microbial species.

CHAPTER II

METHODS

Sampling, collection, and filtration

Sampling of seawater in Galveston Bay Post- Hurricane Harvey occurred during 5 cruises with the R/V Trident on 09/06, 09/09, 09/16, and 09/28 of 2017. Samples were taken on transect from the San Jacinto River to the Gulf of Mexico. For this study, only Stations 1, 4, 7, and 10 were sampled, as indicated in Figure 1. Two control samples were collected from the Texas A&M University at Galveston's boat basin prior to Hurricane Harvey on 07/ 31 and 08/ 22, 2017.

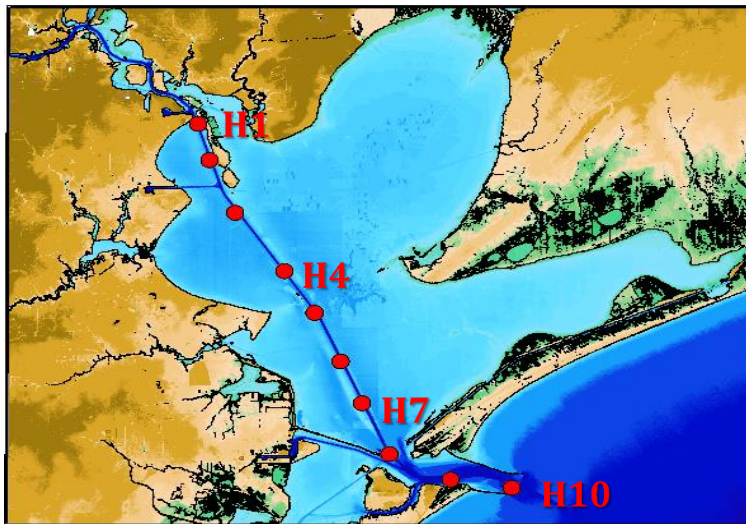

Figure 1. Map of the transect sampled in Galveston Bay from Station H1 (mouth of the San Jacinto River) to Station H10 (Gulf of Mexico).

All samples were pre-filtered immediately after sampling with a nitex filter (30 µm) to remove small grazers and large particles. The total volume of each sample varied depending on time and manpower constraints for each sampling day. After filtration the samples were stored on the boat in the dark and brought to the laboratory for further filtration. Generally, each sample was filtered through a glass fiber filter (GF-F with a 0.7 µm pore-size or GF/D with a 2.7 µm pore-

size), followed by a 0.22 µm pore-size polyvinylidene fluoride (PVDF) filter. Due to the availability of supplies, for the sampling of 09/09, prefiltration was performed with a 0.45 µm and the virus concentrate was the filtered through 0.22 µm PVDF filter. All GF and PVDF filters were stored at -20°C until further use.

Evaluation of the microbial diversity using PCR of marker genes

To assess present microbial diversity for all samples collected , we used Illumina-based sequencing of 16S rRNA genes (Caporaso et al. 2012), with the primers 515F 5'-GTGYCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACNVGGGTWTCTAAT-3') from Parada et al. (2016), as detailed in the Earth Microbiome Project (http://press.igsb.anl.gov/earthmicrobiome/protocols-and-standards/16s/). A total of ~100,000 reads were produced for each amplicon. Analysis of the PCR amplicons included quality control of raw data, clustering into operational taxonomy units (OTUs), taxonomic assignment, and diversity calculations using the software package mothur (Schloss et al. 2009).

DNA extraction from filters

DNA from small eukaryotes and prokaryotes was then extracted from the GF and PVDF filters via a phenol chloroform method. Filters were aseptically cut to represent a volume of ~ 3L of initial sample water, i.e. if 4 L of water was filtered then ¾ of the filtered would be removed for extraction. All of the pieces for one filter portion were placed in individual 50 mL falcon tubes. Some of the GFF and GFD filters had to be done in multiple tubes due to the size of the filters, in general only ¼ to ½ of a GFF or GFD filter was used in each tube. 10 mL of lysis buffer (120 mM NaCl, 225 mM sucrose, 6 mM EDTA, and 15 mM Tris HCl, pH=9) was added, followed by 100 µL of lysozyme (100 mg/mL). The solutions were vortexed and securely placed in an incubator at 37˚C at 350 rpm for 30 minutes. The tubes were removed and 30 µL of stock proteinase K (20

mg/mL) was added followed by 525 µL of the 10% SDS. The solutions were mixed and securely placed in an incubator at 50 ˚C at 350 rpm overnight. A volume of 6.25 mL of saturated phenol (pH of 8) was added to each tube. The tubes were capped, vortexed, and incubated at room temperature for 5 minutes. A volume of 6.25 mL of chloroform:isoamyl alcohol 24:1 was added and the tubes were capped, vortexed, and incubated at room temperature for 5 minutes. Tubes were centrifuged at 3,220 g for 10 minutes. The top aqueous layer was removed using a pipette and placed in a clean falcon tube. A 2X volume of ice-cold ethanol was added to the aqueous solution, then 0.1X volume of 10 M ammonium acetate (pH=5.25) was added. The solutions were gently mixed by inversion and stored at -20 ˚C for at least one hour and up to two days. The DNA was pelleted by centrifugation at 20,446 g for 35 minutes at 4 ˚C. The supernatant was carefully removed to prevent the pellet from detaching, washed with 1 mL of ice-cold ethanol, then centrifuged for 10 minutes at 20,446 g and 4 ˚C. The ethanol was removed, and the samples were left to dry for approximately 2-3 hours. Once dry, 200 µL of Tris EDTA buffer was added. To remove residual phenol and chloroform, the QIAamp DNA mini and Blood mini kit was used following the DNA purification from Blood and Body Fluids (Spin Protocol) procedures. All DNA samples were stored at -20 ˚C until further use.

Sequencing and metagenomic analysis

DNA samples were sequenced using Illumina HiSeq chemistry (2x150bp) at the Texas A&M Genomics & Bioinformatics facility in College Station, TX. BBDuk was used to remove any adaptor sequences or DNA identified as sequencing artifact contamination from the raw reads. The reads were then merged together using BBMerge (Khan et al. 2012). BBMask was then used to determine if there were any sequences from species that often contaminate samples including humans, mice, and dogs. The sequences from each Sample date were combined and assembled

into contigs using MEGAHIT (Li et al. 2015) on the High Performance Research Computers at Texas A&M University, the pre-Harvey samples were aggregated together as well. Contigs were binned using MetaBAT (Kang et al. 2015). The resulting bins were visualized using VizBin (Laczny et al. 2015) to assess overlap between bins, only contigs over 1,000 base pairs were used to visualize the bins. Gene prediction was performed using Prodigal (Hyatt et al. 2010) and the translated amino acid sequences were used to perform a BLASTp search (Altschul et al. 1990) using DIAMOND (Buchfink et al. 2014). The BLASTp results were used to determine the taxonomy of the contigs within the bins and the abundance of the taxa identified using MEGAN (Huson et al. 2007). Additionally, we visualized the abundance of characteristic SEED metabolic pathways within each bin in MEGAN (Goesmann et al. 2005). The amino acid files generated by Prodigal were also used with GhostKOALA to annotate the contigs using KEGG (Kanehisa and Goto 2000) to identify and map specific metabolic pathways present within the contigs. BLASTp and KEGG analysis was only completed for the largest bin, bin 146 from Sample 4, generated by MetaBAT. In order to compare the relative abundance of 16 rRNA genes and metagenomic genes, the unassembled reads for each Sample and Station were compared using BLAST against a custom database made using the contigs and sequences in bin 146 from Sample 4.

CHAPTER III

RESULTS

Community composition changes using 16S rRNA gene

The dominant members of the microbial community were determined through amplification and sequencing of the 16S ribosomal RNA (rRNA) gene. Before Hurricane Harvey, the bacterial community structure of Galveston Bay was dominated by Cyanobacteria and Proteobacteria comprising 71.7% of the total community, Figure 2. Considering Proteobacteria, the group increases 10-12% from pre-Harvey levels in the first sampling and returns to normal levels by September 16 and 28, 2017. The structure within Proteobacteria was largely affected by the increase in the amount of Betaproteobacteria, specifically Burkholderiales, a highly diverse group, however, the species that make up Burkholderiales are mainly chemotrophic, capable of both heterotrophy and autotrophy. Many of the Burkholderiales bacteria are associated with both plants and the human gut. Following the storm, the community returned to mostly Alpha- and Gammaproteobacteria, specifically Pelagibacterales and Oceanospirillales. The return to mostly Pelagibacterales and Oceanospirillales is indicative of a return to a community structure dominated by marine bacteria. Pelagibacterales is known for being the most abundant group in oligotrophic waters due to its chemophototrophic nature. Oceanospirillales is best known as a hydrocarbon degrading microorganism common in marine invertebrates and in marine systems with abundant hydrocarbons, such as after the Deep Water Horizon oil spill (Cao et al. 2014).

Cyanobacteria followed the opposite trend with levels dropping to 4.0% on average in the first sampling and then returning to near pre-Harvey levels, at 26.1% abundance, in the fourth

sampling. Cyanobacteria are photosynthetic bacteria known for their ubiquity in aquatic habitats including oceans and freshwater systems.

Actinobacteria on average increased compared to pre-Harvey levels starting at approximately 8-11% abundance and effectively doubling to 18-20% in the last 2 samplings (Figure 2). When comparing the locations of sampling it can be seen that by Sampling 4 (week 3) and 5 (week 4) the Actinobacteria levels at Stations 7 and 10 had reached levels comparable to that of the controls while Stations 1 and 4 were approximately 25% of relative abundance. The composition of the Actinobacteria group changed with the first sampling after Hurricane Harvey being mostly from the order Actinobacteria unlike the controls that were predominately Acidimicrobiia. Stations 7 and 10 recovered to predominantly Acidimicrobiia while Stations 1 and 4 remained predominantly Actinobacteria. Acidimicrobiia have been found to share similar adaptations to marine life that Pelagibacterales have like chemophototrophy, small size, and small genomes (Mizuno et al. 2015) and their presence near the Gulf is expected.

The Bacteroidetes relative abundance did not see large changes over the sampling period, however, the relative amount of Unknown Actinobacteria (Incerta Sedis) and Sphingobacteriia did see major changes and neither fully recovered within the sampling time period (Figure 2). While Sphingobacteriia was more prevalent directly following Hurricane Harvey the population did go into a decline and remained below pre-Harvey levels even in the final sampling. Unknown Actinobacteria were at levels less than 10% of the levels from before Hurricane Harvey and never recovered to their original levels, the highest levels recorded in the final sampling were only 38.1% of what was found pre-Hurricane Harvey. Verrucomicrobia recovered to normal levels within the sampling periods (Figure 2). Verrucomicrobia had a similar growth to Proteobacteria following Hurricane Harvey when the average abundance for the first Sample jumped to 12.2% from 2.4%,

Figure 3, mainly due to increases in Opitutae and OPB35 Soil Group. The communities began shifting back within the first week and then in the third week a large increase in the Verrucomicrobidae and Spartobacteria caused another increase in the relative abundance of Verrucomicrobia. The final sampling saw a return to pre-Harvey levels in the phylum Verrucomicrobia. Bacteriodetes and Verrucomicrobia are most known for being human gut and soil bacteria, respectively.

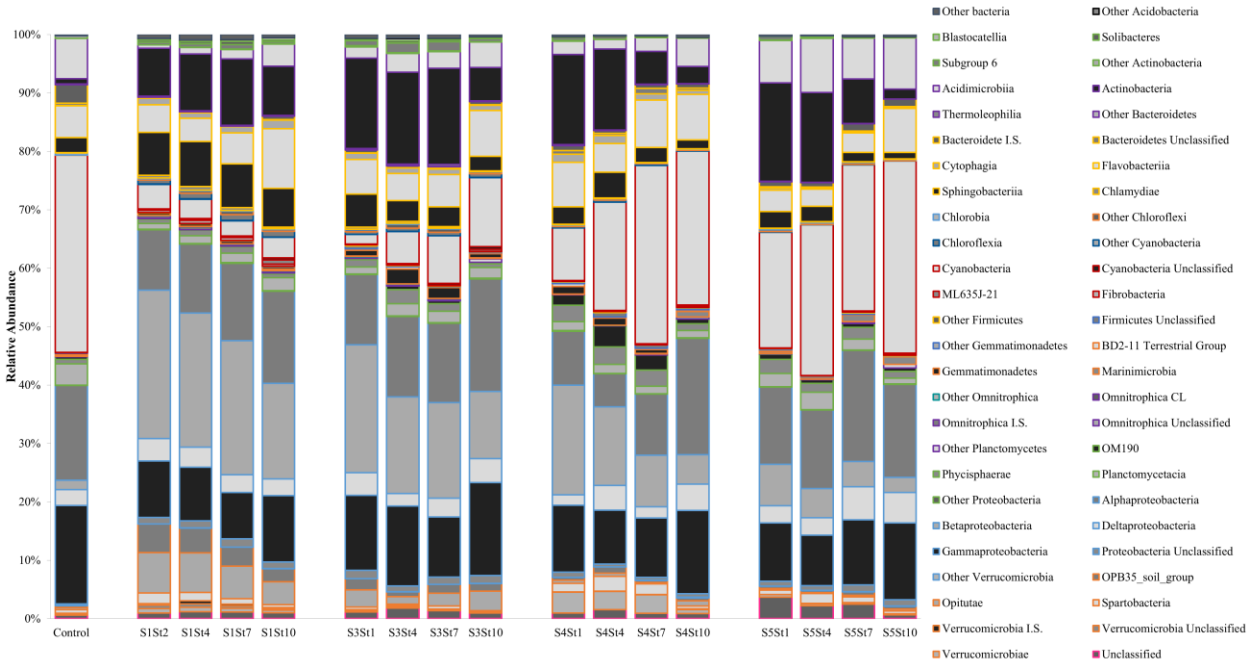

Figure 2. Bar graph of prokaryotic relative abundance of 16S rRNA genes identified to class level. Each bar represents a time and sample location. Station 2 was substituted for Station 1 in Sampling 1 due to the lack of meaningful data from Station 1.

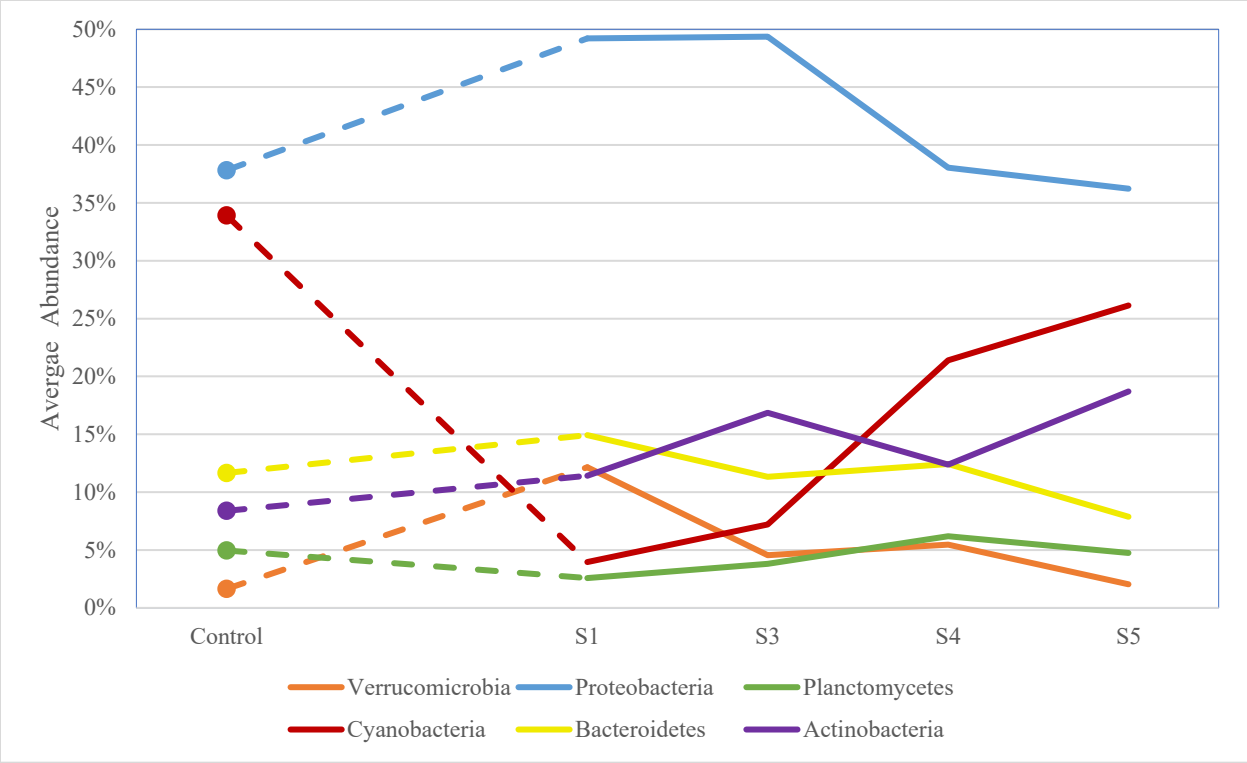

Figure 3. Average percent abundance of all Stations during each Sample date for the six most abundant Phyla.

Metagenomic assembly and binning

Comparing the methods used for assembly in this study versus that of assembling each sample and station individually, the assemblies for the individual stations did generally return higher maximum contiguous DNA segment (contigs) lengths as well as higher amounts of base pairs assembled. The total amounts of contigs for separate assemblies was due to the overlap of bins that would have been one bin in the aggregate assemblies. Average contig size and the minimum contig length that covers 50% of the contigs (N50) were always higher in the assembly of all Stations within a Sample together, as shown in Table 1.

Table 1. Comparison of Megahit assembly for all Stations in each Sample (Aggregates) and the sum or average statistics of the assembling each Station in each Sample separately (Totals).

	Total Contigs	Total bp	Max Length	Average	N50
Total S1 Megahit	4,139,696	2,621,803,209	322,552	631	646
Aggregate S1 Megahit	3,851,584	2,776,161,248	236,527	721	794
Total S3 Megahit	4,382,833	3,078,061,078	543,787	706	776
Aggregate S3 Megahit	3,851,584	2,776,161,248	236,527	721	794
Total S4 Megahit	4,174,109	2,897,867,719	238,672	697	764
Aggregate S4 Megahit	3,100,250	2,402,876,539	141,772	775	918
Total S5 Megahit	3,981,537	2,870,010,303	173,418	721	807
Aggregate S5 Megahit	2,818,699	2,195,504,751	186,444	779	930
Total VC Megahit	2,110,634	1,510,161,294	138,312	716	789
Aggregate VC Megahit	2,048,575	1,534,336,996	142,297	749	858

Binning of the samples resulted in the generation of 1,136 total species bins, 202 in Sample 1, 259 in Sample 3, 279 in Sample 4, 259 in sample 5, and 142 in the control samples. Vizbin employs the Barnes-Hut Stochastic Neighborhood Embedding of center log-ratio in order to normalize and reduce compositional data into a two-dimensional space. Visualization of the bins using this method returned similar results across all samples and the controls with apparent overlap between numerous contigs that aggregate in the center of the image while other bins are clearly defined on the outer edges of the image seen in Figure 4.

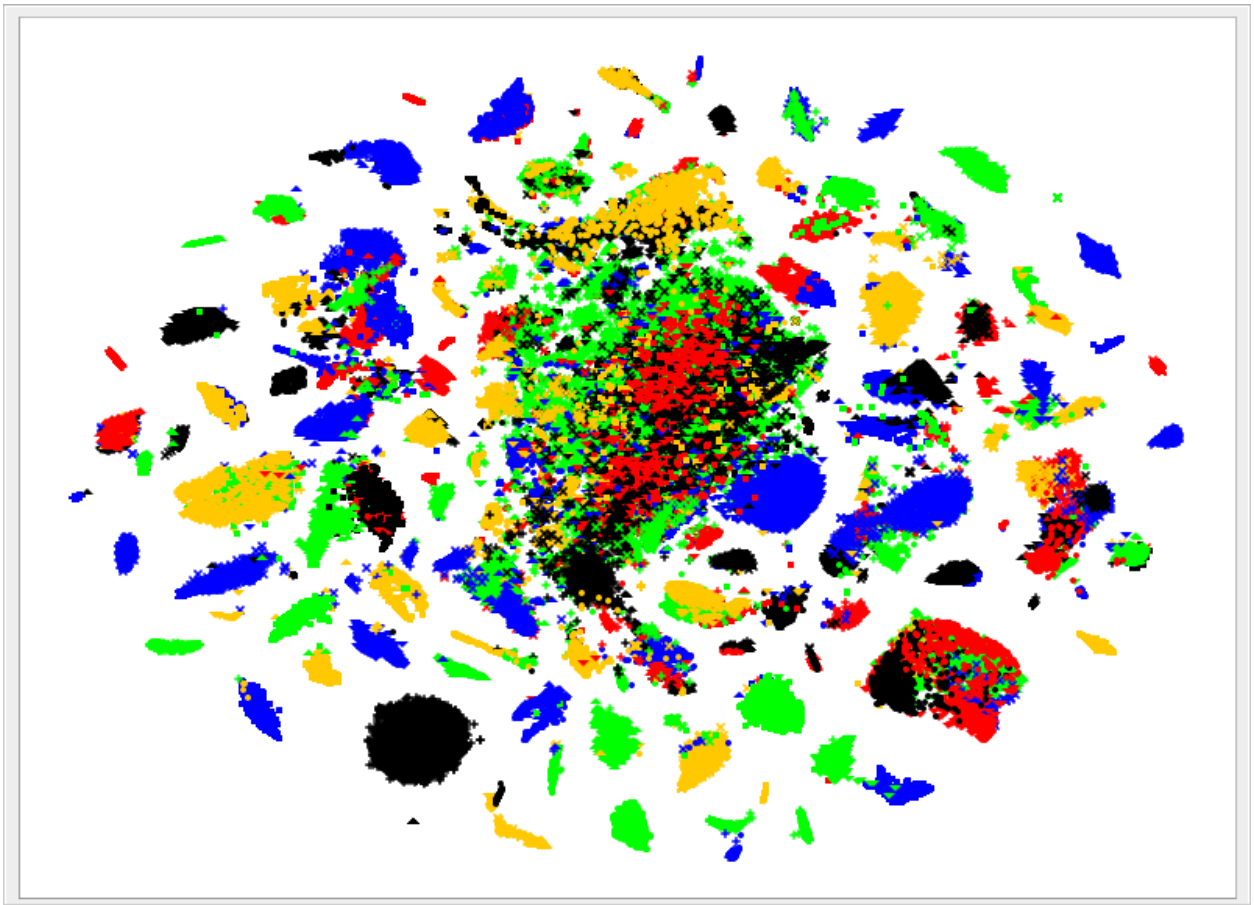

Figure 4. Visualization of the bins generated for Sample 4. Each color is not representative of a bin, the large number of bins required redundant use of colors and symbols. Clusters on outer edges made of a single color are single bins.

Taxonomic and metabolic analysis of bins

We chose to focus on bin 146 from Sample 4 because it is the largest bin, with 43,389 contigs. BLASTp results for bin 146 from Sample 4 identified the genes as predominantly from the order Pelagibacterales, formerly SAR11 clade, in class Alphaproteobacteria from phyla Proteobacteria, (Figure 5). MetaBAT uses a tetranucleotide frequencies and abundance distance probabilities to assess the similarity between bins. While this method is capable of handling large complex datasets, sequences from different species but with similar tetranucleotide frequencies

could possibly end up being combined. Quality control of each bin can be completed by the removal of contigs associated with other species post-binning.

Using the SEED analysis in MEGAN allowed for the visualization of the major types of genes found within bin 146 (Figure 6). A wide variety of genes were identified. However, a majority of the genes identified were related to typical cell function and respiration and a very small number of genes were identified as being autotrophic or photosynthetic. Pelagibacterales is a strictly heterotrophic organism (Giovannoni et al. 2005) and the presence of autotrophy would indicate the need for further quality control of the bins. KEGG analysis of the amino acid files indicated similar results as seen in the metabolic pathway map (Figure 7) for the same bin. The results of BLAST comparisons of the unassembled reads to the database made from contigs and sequences in bin 146 from Sample 4 resulted in similar trends throughout the data set, seen in Figure 8. Changes are not always proportional. However, generally if there is an increase the comparable group will increase or stay the same and if there is a decrease there will be a decrease in the equivalent group. One exception to this is Sample 3 Station 7 in which metagenomic genes decreased by 1% while the 16S rRNA genes increased by 2%. Despite the minimal exception the

similarity between the two datasets reflects an accurate assessment of the community by both the 16S rRNA and metagenomic data.

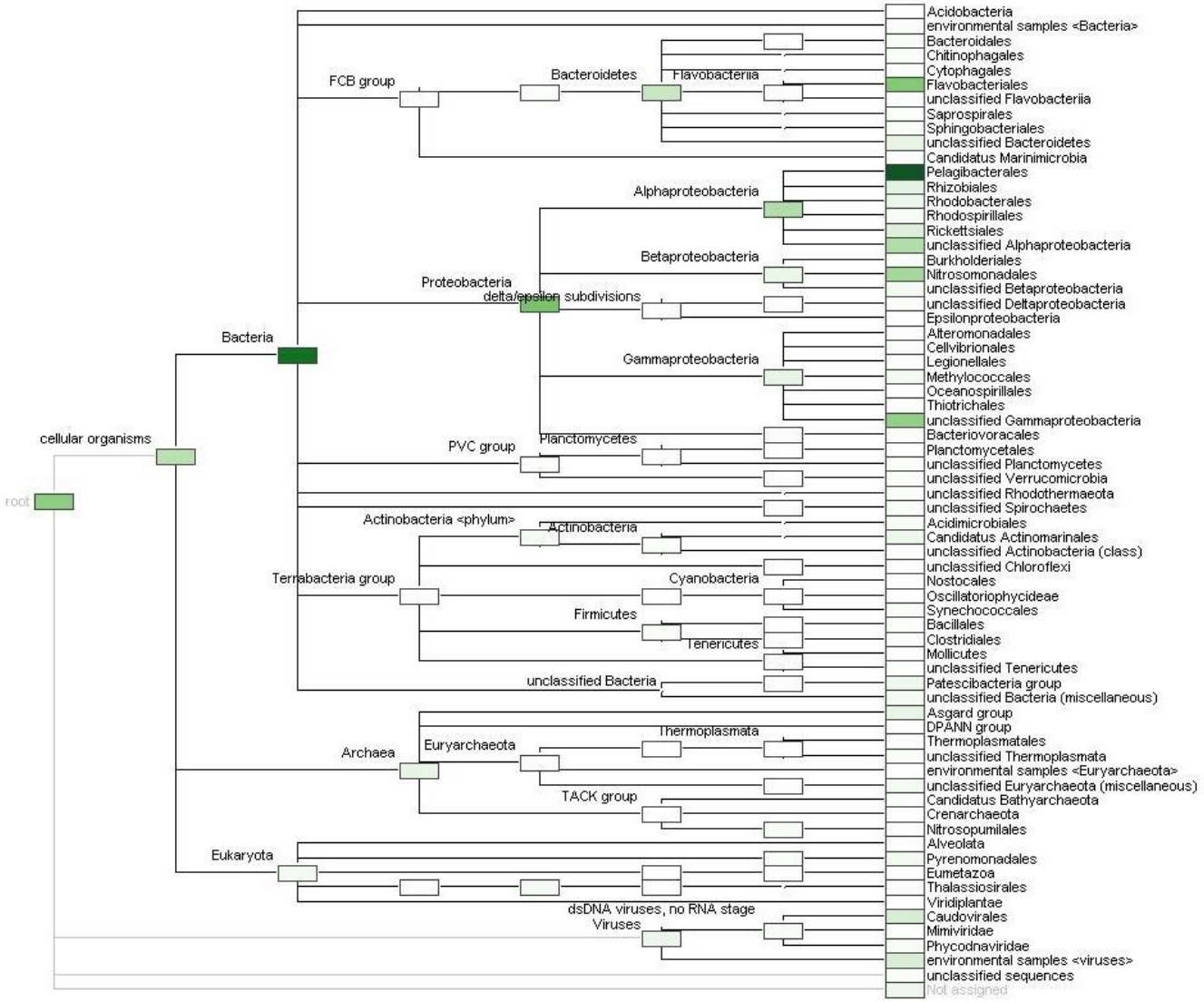

Figure 5. Heatmap of the phylogenetic BLASTp results from Sample 4, bin 146.

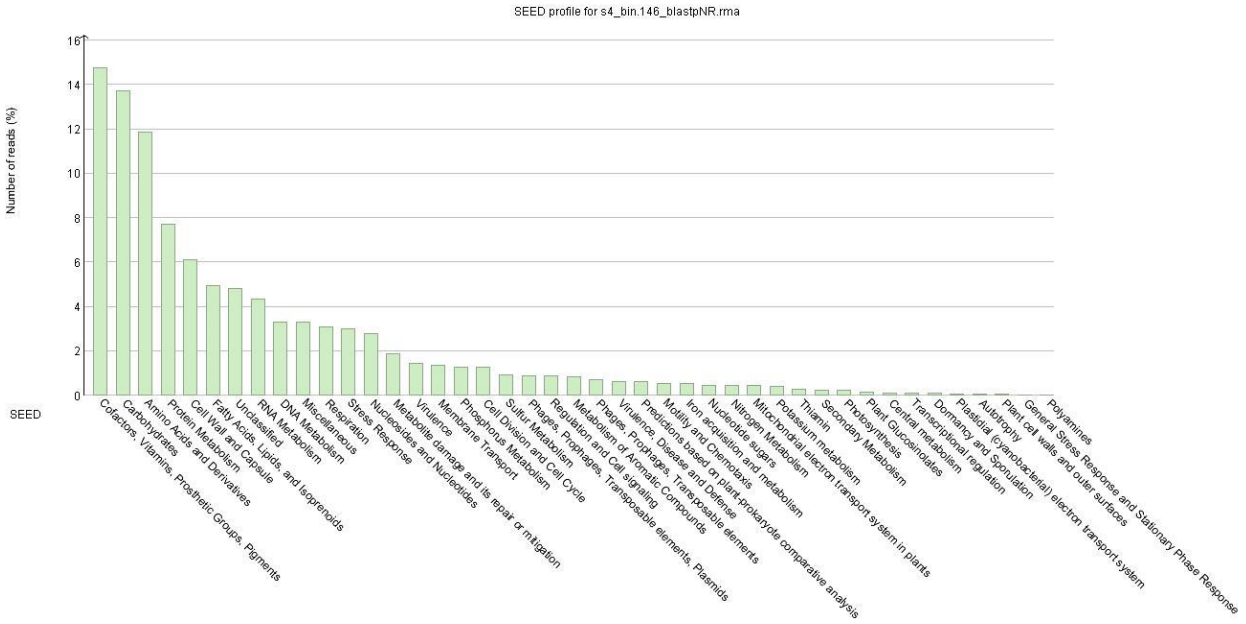

Figure 6. Relative abundance of major SEED functions in Sample 4, bin 146.

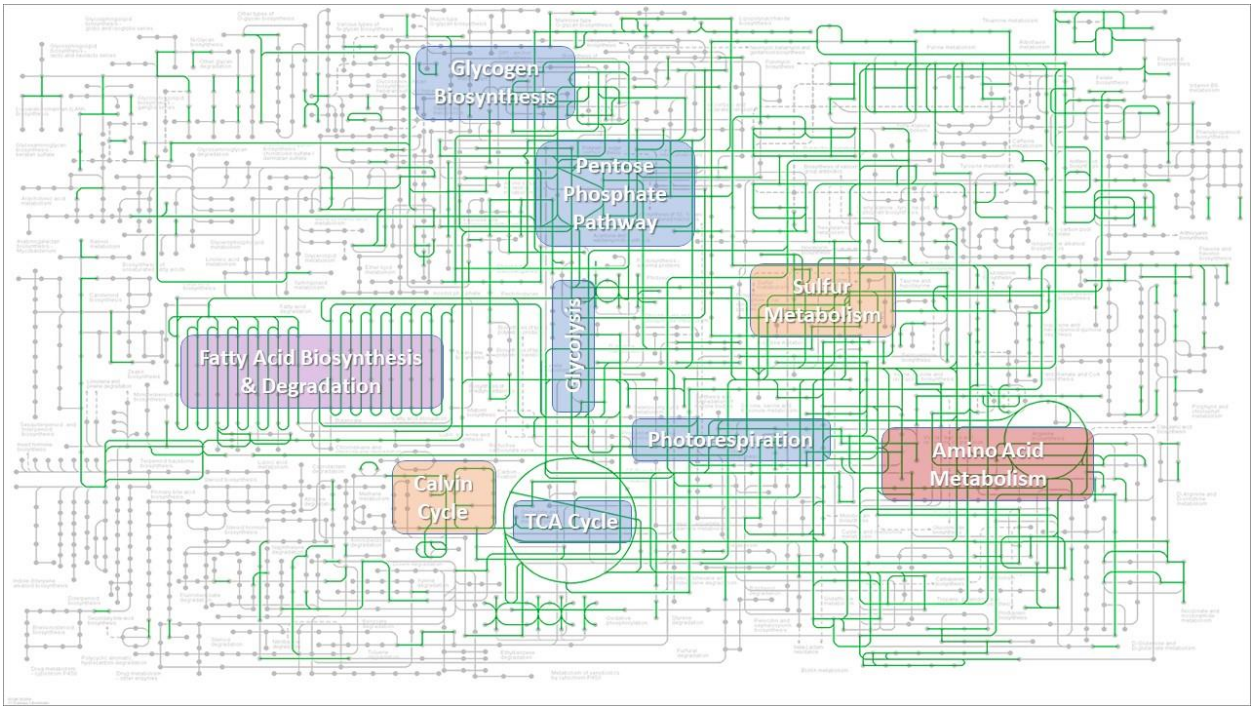

Figure 7. KEGG reference pathway with pathways identified in Sample 4, bin 146 highlighted in green. Major pathways that were nearly complete were labeled, blue labels indicating carbohydrate metabolisms and orange labels indicating energy metabolism.

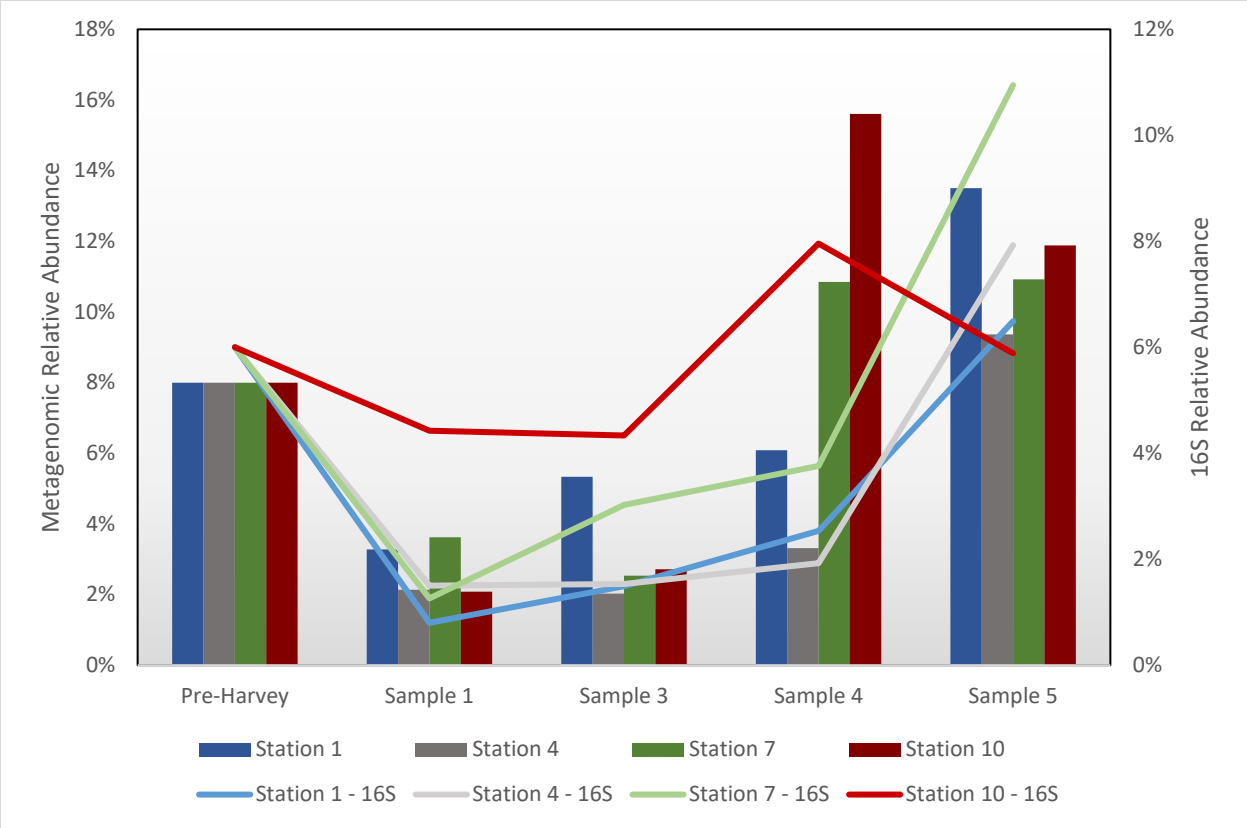

Figure 8. Comparison of 16S rRNA gene relative abundance to metagenomic gene relative abundance for Pelagibacterales (SAR11 clade).

CHAPTER IV

CONCLUSION

Drivers of change in the microbial community

Changes in Cyanobacteria levels can be explained largely due to the changes in Galveston Bay's temperature and salinity. Hurricane Harvey decreased the temperature of the Bay by approximately 8°C and the salinity fell as low as 0.0 PSU directly following the storm (Quigg 2019). Previous studies have shown that decreases in temperature, seasonal decreases of sunlight, and shifts in salinity can greatly hinder the ability of Cyanobacteria to thrive (Butterwick et al. 2005; Xia et al. 2015; Kim et al. 2018). As rainwater flooded into Galveston Bay the overall temperature and salinity of the Bay was driven down creating conditions unfavorable to marine Cyanobacteria. The introduction of stormwater that has accumulated soil and anthropogenic microorganisms explains the high levels of Betaproteobacteria, Actinobacteria, Bacteriodetes, and Verrucomicrobia following the storm. The sediment mixing in the Bay's water would have increased the amount of soil bacteria present and as the sediment descended to the bottom of the Bay the levels of these phyla should have fell. Additionally, the nutrients introduced by the stormwater runoff into the Bay would have been able to support these phyla mainly consisting of chemotrophic organisms.

Future research

Initial results from metabolic analysis have yielded a fully functional pipeline capable of isolating taxa to the order level in bins. Additionally, the similarity between the Pelagibacterales metagenomic and 16S rRNA relative abundances supports the ability of the pipeline to generate an accurate representation of the groups binned. Continued use of the pipeline to identify and

characterize the bins created in each Sample will enable for the comparison of specific taxa over the course of the sampling period. There will likely be overlap found as the bins are identified which would allow us to group together bins from different samples and build even larger collections of genomic data for each species. Utilizing the largest possible collection of genomic data for each species continued analysis of the metabolic potential of these groups using KEGG and SEED will be done. Assuming an almost complete genome can be assembled, for some of the species it should be possible to begin working on culturing the unculturable. Having a nearly complete genome will allow for the identification of the organic and inorganic needs of the species. Media can then be "tailor-made" to include everything the species will need to survive and replicate.

Completed SEED analysis of all of the bins will allow for comparisons of specific metabolic pathways across the entire sampling. Correlation between the specific conditions of the Bay that promoted specific metabolic potentials can be analyzed at the species level. Further analyses into these correlations will help to predict future impacts on microbial communities and how changes in salinity and temperature may promote some groups and exclude others from an environment.

Microbial communities are ubiquitous in marine ecosystems providing globally important functions including producing most of the carbon in the oceans and fixing nitrogen to support primary production (Jiao et al. 2010; Beman et al. 2011). Microbial communities constantly respond to changing ocean conditions and threats from global climate change and anthropogenic waste in coastal environments present new challenges to these communities (Hutchins et al. 2009; Rohwer et al. 2010). In order to understand how global and regional communities will respond to these challenges more work is needed to establish baseline community structures, how those

communities interact in a healthy environment, and how they respond to stress. Despite the usefulness of metagenomics, the ability to create pure cultures of these organisms is the only way to develop a complete picture of these organisms. The use of metagenomics should be able to inform the creation of media specific to unculturable organisms and contribute to the overall understanding of each taxa's role in the environment. This study has shown that the communities are capable of rebounding quickly from dramatic shifts in their environment and continued analysis of the data will provide greater insight into how they rebounded.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403-410 doi https://doi.org/10.1016/S0022-2836(05)80360-2

Armbrust EV, Palumbi SR (2015) Marine biology. Uncovering hidden worlds of ocean biodiversity. Science 348: 865-867 doi 10.1126/science.aaa7378

Azam F, Fenchel T, Field JG, Gray JS, Meyerreil LA, Thingstad F (1983) The Ecological Role of Water-Column Microbes in the Sea. Marine Ecology Progress Series 10: 257-263 doi DOI 10.3354/meps010257

Beman JM, Chow C-E, King AL, Feng Y, Fuhrman JA, Andersson A, Bates NR, Popp BN, Hutchins DA (2011) Global declines in oceanic nitrification rates as a consequence of ocean acidification. Proceedings of the National Academy of Sciences 108: 208-213 doi 10.1073/pnas.1011053108

Buchfink B, Xie C, Huson DH (2014) Fast and sensitive protein alignment using DIAMOND. Nature Methods 12: 59 doi 10.1038/nmeth.3176 https://www.nature.com/articles/nmeth.3176#supplementary-information

Butterwick C, Heaney SI, Talling JF (2005) Diversity in the influence of temperature on the growth rates of freshwater algae, and its ecological relevance. Freshwater Biology 50: 291-300 doi 10.1111/j.1365-2427.2004.01317.x

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME journal 6: 1621

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappé MS, Short JM, Carrington JC, Mathur EJ (2005) Genome Streamlining in a Cosmopolitan Oceanic Bacterium. Science 309: 1242-1245 doi 10.1126/science.1114057

Goesmann A, McHardy AC, Osterman A, Hanson A, Linke B, Rückert C, Iwata-Reuyl D, Rodionov DA, Frank ED, Glass EM, Meyer F, Olsen G, Pusch GD, Chuang H-Y, Neuweger H, Thiele I, Steiner J, Choudhuri JV, Krause L, Cohoon M, Fonstein M, Kubal

M, Diaz N, Jamshidi N, Larsen N, Vassieva O, Zagnitko O, Butler RM, Stevens R, Edwards R, Olson R, Overbeek R, Jensen R, Gerdes S, Begley T, Disz T, de Crécy-Lagard V, Portnoy V, Vonstein V, Ye Y (2005) The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes. Nucleic Acids Research 33: 5691-5702 doi 10.1093/nar/gki866 %J Nucleic Acids Research

Handelsman J (2004) Metagenomics: Application of Genomics to Uncultured Microorganisms68: 669-685 doi 10.1128/MMBR.68.4.669-685.2004 %J Microbiology and Molecular Biology Reviews

Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. Genome research 17: 377-386 doi 10.1101/gr.5969107

Hutchins DA, Mulholland MR, Fu F (2009) Nutrient Cycles and Marine Microbes in a CO_2-Enriched Ocean. Oceanography 22: 128-145

Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJJBb (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification11: 119

Jiao N, Herndl GJ, Hansell DA, Benner R, Kattner G, Wilhelm SW, Kirchman DL, Weinbauer MG, Luo T, Chen F, Azam F (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. Nature Reviews Microbiology 8: 593 doi 10.1038/nrmicro2386

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28: 27-30

Kang DD, Froula J, Egan R, Wang Z (2015) MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3: e1165 doi 10.7717/peerj.1165

Khan MA, Bhatia P, Sadiq MJIJRTE (2012) BBTool: a tool to generate the test cases1: 192-197

Kim Y, Jeon J, Kwak MS, Kim GH, Koh I, Rho M (2018) Photosynthetic functions of Synechococcus in the ocean microbiomes of diverse salinity and seasons. PLOS ONE 13: e0190266 doi 10.1371/journal.pone.0190266

Laczny CC, Sternal T, Plugaru V, Gawron P, Atashpendar A, Margossian HH, Coronado S, Van der Maaten L, Vlassis N, Wilmes PJM (2015) VizBin-an application for reference-independent visualization and human-augmented binning of metagenomic data3: 1

Li D, Liu C-M, Luo R, Sadakane K, Lam T-W (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31: 1674-1676 doi 10.1093/bioinformatics/btv033

Liberto TD (2017) Reviewing Hurricane Harvey's catastrophic rain and flooding. NOAA, https://www.climate.gov

Mallin MA, Ensign SH, McIver MR, Shank GC, Fowler PK (2001) Demographic, landscape, and meteorological factors controlling the microbial pollution of coastal waters. Hydrobiologia 460: 185-193 doi Doi 10.1023/A:1013169401211

Pace N, Stahl D, Lane D, Olsen G (1986) The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences. Advances in Microbial Ecology. Springer, Boston, MA

Parada AE, Needham DM, Fuhrman JA (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environmental microbiology 18: 1403-1414

Quigg A (2019) Hurricane Harvey's Effects on Phytoplankton in Galveston Bay. Unpublished Raw Data. Texas A&M University at Galveston

Rappé MS, Giovannoni SJ (2003) The Uncultured Microbial Majority. Annual Review of Microbiology 57: 369-394 doi 10.1146/annurev.micro.57.030502.090759

Rohwer F, Youle M, Vosten D (2010) Coral reefs in the microbial seas. Plaid Press Granada Hills

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537-7541

Schmidt JL, Deming JW, Jumars PA, Keil RGJL, Oceanography (1998) Constancy of bacterial abundance in surficial marine sediments43: 976-982

Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C (2013) Computational meta'omics for microbial community studies. Molecular systems biology 9: 666-666 doi 10.1038/msb.2013.22

Suttle CA (2007) Marine viruses--major players in the global ecosystem. Nat Rev Microbiol 5: 801-812 doi 10.1038/nrmicro1750

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428: 37 doi 10.1038/nature02340
https://www.nature.com/articles/nature02340#supplementary-information

Williamson KE, Harris JV, Green JC, Rahman F, Chambers RM (2014) Stormwater runoff drives viral community composition changes in inland freshwaters. Front Microbiol 5: 105 doi 10.3389/fmicb.2014.00105

Xia X, Vidyarathna NK, Palenik B, Lee P, Liu H (2015) Comparison of the Seasonal Variations of Synechococcus Assemblage Structures in Estuarine Waters and Coastal Waters of Hong Kong. Applied and Environmental Microbiology 81: 7644-7655 doi 10.1128/aem.01895-15