Fall 2023

Metagenomic Survey of Marine 16S Bacterial Communities off Palmer Station in Antarctica

Daniel Salter

Metagenomic Survey of Marine 16S Bacterial Communities off Palmer Station in Antarctica

A Project Presented to

Dr. Philip Heller

Dr. Shellie Bench

Dr. Wendy Lee

The Department of Computer Science

San José State University

In Partial Fulfillment of the

Requirements for the Degree

Master of Science In

Bioinformatics

By

Daniel Salter

December 2023

ABSTRACT

This project surveys the metagenomic bacterial community composition in marine surface waters off Palmer Station, Western Antarctic Peninsula and correlates findings with temperature and salinity data. Marine bacterial communities play a vital role in nutrient cycling, but data on surface waters in this region are limited. Analyzing fifteen samples of 16S sequencing data from three austral summers, consistent dominance was observed by the classes Alphaproteobacteria, Gammaproteobacteria, and Flavobacteria. Correlation analysis confirmed significant relationships between taxa and environmental conditions. The observed trends suggest varying abilities of phyla to resist and adapt to changing environmental conditions. Notably, Alphaproteobacteria demonstrated adaptability to favorable conditions, contrasting with Flavobacteria's resistance. The study contributes insights into the responses of bacterial communities in the Western Antarctic Peninsula to climate variation.

Keywords: Marine bacteria, taxonomic classification, 16S rRNA, climate variation effects, Western Antarctic Peninsula

TABLE OF CONTENTS

I. INTRODUCTION

With an ever-changing climate, it's increasingly difficult to pinpoint the effects of rising seas, increasing global temperatures, and extreme weather circumstances on the ecosystems of our planet—especially on microbial communities of life we cannot see. In spite of these challenges, researchers continue to find novel ways to study community dynamics of microbes and find correlations that connect with the sharp changes in climate among other factors that govern the niches of organisms (Pearce, 2008). Such studies help uncover grand-scale ecosystem services that large populations of microbial communities provide, such as primary production of biomass and recycling of decomposing particles that would otherwise settle in the ocean for long segments of time (Cho & Azam, 1988; J. A. Fuhrman, 2009).

The continued study of microbial communities thus presents itself as a promising avenue for helping society understand how it can preserve these important benefits from our neighboring species. The aim of this project is to further contribute to our current understanding of how climate change has or has not affected communities of marine bacteria off the coast of Palmer Station in the Western Antarctic Peninsula. This is done by collecting water samples at various time points across a three-year span and sequencing the RNA molecules present. The results will provide knowledge as to which taxonomic ranks dominate, which are rare in numbers but still present, and which ranks either resist the environmental effects or change with them. Further analysis on the results can determine statistical correlations between ranks and environmental conditions as well as provide a quantifiable understanding of the biodiversity.

II. BACKGROUND

A. Role of Global Oceans Against Climate Change

Climate change is one of the driving forces that guides changing environments on our planet. The rate at which it is occurring indicates an anthropogenic source to some unknown but significant degree (IPCC, 2022), likely caused by an unrelenting assault of greenhouse gases on the atmosphere. Before the industrial revolution, biogeochemical cycles were accustomed to natural sources of greenhouse gas emissions. However, since 1850, an increase in anthropogenic greenhouse gas emissions has caused a noticeable upward trend in global temperatures (Aristarain et al., 1986). This has far-reaching effects on not only terrestrial ecosystems, such as the melting of shelf ice and the related rise in sea level (IPCC, 2022), but also on marine systems as well.

Global oceans have long provided to the rest of the earth a buffer from spiraling levels of greenhouse gases. By acting as a carbon dioxide and heat sink, our oceans keep our ecosystems and futures in equilibrium—balancing out extreme weather circumstances of the past, present, and future (Hogg et al., 2015; Llovel & Terray, 2016). However, this has had negative long-term effects that have been documented in research. Absorbing the excess heat generated from global warming has caused surface ocean temperatures to rise in parallel (Llovel & Terray, 2016). While the increase isn't a perfect one to one ratio— with atmospheric and terrestrial temperatures rising at a faster rate that of marine systems—it is substantial enough to bring about grand-scale changes in the many oceanographic characteristics that govern the vast bodies of water (Bijma et al., 2013; Vaughan et al., 2003).

Studies report that water column stratification—in which two distinct layers of water are thermodynamically separated at specific depths below the surface—is affected considerably by an increase in heat delivery to such columns (Bijma et al., 2013). Water stratification regulates nutrient mixing—bringing biologically meaningful molecules up to surface waters from benthic depths. Microorganisms utilize these materials to produce energy for the ecosystem. Changes in this process may ultimately affect the levels of biological energy that is produced—even influencing the amount of carbon absorbed from the atmosphere. The effects, as seen in this example, can stack to global-scale consequences. The Antarctic Circumpolar Current (ACC) transports water in a clockwise direction around the continent of Antarctica, acting as a hydrodynamic buffer between the polar waters south of the current and the oceans to the north (Siegert et al., 2019). At the same time, studies report that the ACC has delivered enough heat to suggest a 6°C warming of the surface waters off the coast of Antarctica since 1950 (Ducklow et al., 2007). And in the recent decades, the once cold and dry terrestrial climate has evolved into a warmer, moister climate (Aristarain et al., 1990; Turner et al., 2005). The Western Antarctic Peninsula has been central to scientific study due to the degree in which it has been affected by global warming.

B. The Western Antarctic Peninsula Surface Oceans and Microbial Ecosystems

The Western Antarctic Peninsula is a stretch of land that extends beyond the Antarctic continent towards the polar tip of South America. It is a highly variable land mass with various islands, channels, and a mostly glacial coastline (Ducklow et al., 2007). The ACC and the continental shelf bordering the western edge of the Antarctic Peninsula interact in this region; the multitude of glacial landmasses provide ample opportunity for melting events. Consequently, the Western Antarctic Peninsula is an important source of climate change research. In the last two thousand years, the peninsula shelf has undergone more warming compared to the global mean (Vaughan et al., 2003). Additionally, seasonal variation effects are influential to the ecosystem. During winter, the ocean surface is largely covered in sea ice, preventing light from reaching to depths more commonly available in the summer (Ducklow et al., 2007). In the summer, increased solar irradiance leads to warmer surface temperatures and decreased sea ice cover. As a result, salinity decreases due to this “freshening” effect of sea ice melting (Alcamán-Arias et al., 2021). Changing temperature and salinity conditions reduce water column stability and manipulate the ecological equilibrium of the marine life. Palmer Station is located off Anvers Island in the northern region of the Western Antarctic Pacific. At this U.S. facility, research on the environmental factors and their effects on summer food webs occurs as part of the Palmer Long Term Ecological Research program (PAL-LTER)—a decades-long effort to study mechanisms of climate and ecological variation (Saba et al., 2014; Smith et al., 1995).

Despite the anthropogenic disturbances to this icy ecosystem, the Western Antarctic Peninsula is highly productive region (Ducklow et al., 2007). It harbors populations of mammals, birds, Antarctic krill, and most importantly, the marine microbial community supporting this

wide network of organisms. Marine microbial communities are important forces in the global production of biomass and recyclers of carbon, nitrogen, and other inorganic materials (J. A. Fuhrman, 2009). While Bacteria, Eukaryotes, and Archaea are all included under the umbrella term "microbes," the focus of this project is on Bacteria. Bacteria are generally known for their prolific and resilient nature—they are found in most corners of the world, in our homes, in our digestive system, and even in the most isolated, oligotrophic patches of the oceans. Bacteria are best known for their ability to decompose dead organic material and recycle it into usable resources for other living—mostly multicellular—organisms (Cho & Azam, 1988). Historically, when phytoplankton blooms generate large reserves of energy and dissolved organic carbon, an increase in phytoplankton-associated bacteria occurs in succession (Ducklow et al., 2012). The bacteria are able to metabolize the dissolved organic carbon and this builds the foundation for trophic level interactions from bacterivores like zooplankton (J. Fuhrman & Steele, 2008). Just as a concrete foundation reinforces a skyscraper, so does the bacterial community structure support a wide network of ecosystem interactions. Additionally, bacteria and other microbes maintain complex, symbiotic relationships in nature. Recent research applying time-dependent correlation analysis report both positive and negative correlations between relative abundances of bacteria, archaea, and protists (Steele et al., 2011). The negative values could be associated with some type of predation or parasitism. Ultimately, the study of bacteria in microbe community composition is important in helping to better understand nutrient cycling and ecological relationships.

The bacterial community in the Western Antarctic Pacific is a major source of bioproductivity and nutrient cycling (Nikrad et al., 2014). With the region undergoing extreme

seasonal variations and long-term climate change, it's important to understand the effects on the marine bacterial community. Studies report that increased ice retreat, reduced wind shear, and enhanced salinity gradients correlate with stable populations of phytoplankton (Ducklow et al., 2007; Saba et al., 2014). Water column stability is important for maintaining biodiversity and effects from climate change (increased "freshening" events from sea ice melt) often disrupt these gradients, indicating that bacteria community composition may be less resilient to climate change in this regard (Cram et al., 2015). With that being said, some bacterial taxa respond positively to decreased salinity (Polaribacter, Pseudoalteromonas), suggesting an adaptable nature and resilience to climate change. Overall, community composition changes will largely occur due to magnification of seasonal events like solar irradiance variance and changes in sea ice extent and retreat (with these factors affecting water column gradients). Moreover, studies have already reported that seasonal changes in the environment lead to more significant changes in the bacterial community composition than trophic interactions between bacterial taxa and other microbial groups. Correlations in abundance in bacteria-bacteria relationships are stronger when compared with bacteria-eukaryotes relationships (Gilbert et al., 2012). Temporal studies help inform our knowledge of bacterial community composition dynamics—setting the basis for understanding climate change's effects on the global ecosystem.

C. Metagenomics Strategies for Surveying Bacterial Community Composition

There exist several strategies for surveying the metagenomic composition of an environment. In the past, studies implemented cultivation-based analysis and automated ribosomal intergenic spacer analysis (ARISA) (Kovacs et al., 2010; Suzuki et al., 1997). The former strategy ensures specificity in that only culturable bacteria from the sample will grow on the culture and be

available for analysis, but it lacks sensitivity because many species of bacteria are not able to be cultivated (Staley & Konopka, 1985). ARISA is an electrophoresis test for analyzing the length heterogeneity of the 16S-23S bacterial ribosomal RNA (rRNA) intergenic spacer. This helps ensure unculturable bacteria are accounted for in community composition analysis because it relies on a gene-based approach. However, a drawback is that some unrelated bacteria may share identical length spacer regions. A purely genetic sequence-based approach will resist the drawbacks of both these methods while preserving sensitivity. The 16S rRNA region is the RNA component of the 30S subunit of the prokaryotic ribosome. This segment is a commonly used sequencing target because of its slow rate of evolution (Lane et al., 1985) while also being highly conserved across all bacteria. Because of its large degree of conservation, it's applicable to target this region using universal primers in an ecosystem sample. Variable regions within the 16S sequence also ensure that it is highly unlikely that two species of bacteria will share the same sequence. Many databases document the 16S sequence of bacterial species and are thus used for taxonomic classification. One major drawback to this technique, however, is that sequencing technology is prone to errors, preventing complete reliance on its specificity. Despite this, 16S sequencing and analysis is a useful technique for characterizing community composition.

Sampling the 16S reads from the environment can produce millions of reads—far too many to cross-check manually. Bioinformatic techniques are applied to analyze the multitude of 16S sequences. Alignment algorithms measure the similarity between two sequences—providing a metric to assist in matching an unknown sequence of DNA with a known match documented in a database. Basic Local Alignment Search Tool (BLAST) centralizes these algorithms and implements them on a grand scale (Altschul et al., 1990). It takes as input a query sequence and

performs heuristic alignment algorithms on a chosen database of subject sequences and genomes—outputting matches as well as their degree of similarity. BLAST is also an order of magnitude faster than comparable alignment tools. However, it struggles to maintain speed when performing alignment on a massive amount of query sequences, and its output is not formatted in a user-friendly format. For this reason, other tools are available that are more specifically designed for metagenomic analysis.

Mothur is a software tool introduced in 2009 that allows analysis of metagenomic sequence data to be achieved in reasonable run-time amounts. It works by implementing sequence pre-processing strategies, screening steps, and alignment on large amounts of query reads (Schloss et al., 2009). It also contains the ability to assign reads to operational taxonomic units and calculate alpha and beta diversity metrics. Taxonomic classification is accomplished by aligning pre-processed 16S reads with a reference 16S alignment. Searching the resulting aligned 16S sequence in a formatted taxonomy database provides the user with a classified match. Reducing the complexity of the input reads through the pre-processing and screening stages decreases the run-time cost. Ultimately, the user can input millions of 16S reads and receive an output file with each read matching a taxonomic rank within a day (Schloss et al., 2009). This tool enables a faster strategy for analyzing bacterial community composition than BLAST. However, due to the pre-processing stage many reads are dropped due to low quality or long homologous regions—reducing the sensitivity of the results. Yet Mothur remains a useful tool for analyzing metagenomic data.

D. Bacteria Taxonomic Background

Proteobacteria, also known as Pseudomonadota, are a phylum of gram-negative bacteria that represent a wide variety of metabolism types. Across the surface waters, they are known to be one of the most prolific group of heterotrophic organisms (Glöckner et al., 1999; Stevens et al., 2005). Studies report their numbers at over 50% of the phyla level of all bacteria surveyed in the northern isles of the Antarctic Peninsula (Kim et al., 2021). Classes from this phylum include Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria. The same studies report Alphaproteobacteria and Gammaproteobacteria holding the highest proportion of all bacterial classes in surface waters (Gilbert et al., 2012; Kim et al., 2021). Pelagibacterales, also known as the SAR11 clade of Alphaproteobacteria, are free-living bacteria that make up almost one third of all cells in the ocean surface (Morris et al., 2002). They are adept at scavenging in oligotrophic waters and this makes them prolific and successful in isolated, nutrient-poor regions of the surface ocean.

Bacteroidetes are gram-negative, rod-shaped bacteria that are also prominent in the oceans. While they are mostly known for inhabiting gut microbiomes, more research is revealing the degree to which they dominate the marine surface waters (Kim et al., 2021). Following Proteobacteria, Bacteroidetes are also found in high proportions (Luria et al., 2014; Stevens et al., 2005). Classes from this phylum include Bacteroidia, Chitinophagia, Cytophagia, Flavobacteria, Saprospiria, and Sphingobacteria. Flavobacteria is reported in high abundance in the ocean partly due to their ability to thrive in nutrient-poor environments (González et al., 2008). Bacteroidetes are often reported to have high diversity within its subgroups. They are

often responsive to environmental variation, as reported by research in the Southern Ocean coastal waters (Ghiglione & Murray, 2012).

E. Project Description

This study analyzes surface ocean samples of 16S RNA sequences that were taken at Palmer Station in the Western Antarctic Peninsula during the austral summers of 2012-2013, 2013-2014, and 2014-2015. Seasonal variation taking place within each of these summers is accounted for by staggering the collection events across the whole summer. In addition, interannual climate variation may occur between each summer. This variation may reveal how environmental conditions affect community composition of the bacteria present in the surface waters. The samples were then filtered and sequenced to produce reads. The reads were taxonomically classified using Mothur and analyzed with visual data analysis in RStudio. Community composition relative abundance charts were generated as well as environmental condition heat maps to help find correlations between changes in community compositions and the changes in surface temperature and salinity. The results provide insight into the effectiveness of applying the command-line functions of Mothur to sequencing data, as well as the responses of bacterial community composition to both seasonal and global variation in climate.

III. METHODS

A. Sample Collection at Palmer Station

Marine surface water samples were taken at Palmer Station by Dr. Shellie Bench at 10 meters in depth. These samples were taken during three consecutive austral summers on the dates 11/27/2012, 02/08/2013, 12/27/2013, 01/23/2014, 02/03/2014, 02/10/2014, 02/28/2014, 03/04/2014, 12/01/2014, 12/11/2014, 01/12/2015, 01/19/2015, 02/09/2015, 02/23/2015, 03/09/2015. There were two Year One (2012-2013) samples, six Year Two (2013-2014) samples, and seven Year Three (2014-2015) samples. The samples were filtered to separate prokaryotes and eukaryotes (with prokaryotes being the focus of this project) and PCR amplified. The resulting amplicons were sequenced to generate paired-end FASTQ sequencing read files for each sampling date. Paired-end reads are when the same sample sequence is read by a sequencer in both the forward and the reverse direction, resulting in a "pair" of reads. The region sequenced was the V4 hypervariable region of the 16S rRNA gene, a known gene for being present in prokaryotes while being distinct enough across species to allow taxonomic classification. There were in total 15 paired-end FASTQ file pairs.

Environmental data collection of surface ocean temperature and salinity was performed by the Palmer LTER group in parallel with Dr. Bench's sampling period (Smith et al., 1995). The data was generated in Conductivity, Temperature, and Depth files (CTD). Temperature and salinity data was extracted using R (filtering for 10-meter depth and specific dates that correlate with the 16S collection period). The resulting file was used to characterize the environmental conditions in correlation with the relative abundance of marine bacterial taxa for this project.

B. Sequencing Classification with Mothur

Mothur was used to process and taxonomically classify the FASTQ reads. A command-line shell script was written to process the reads in a pipeline with little manual work needed during this step. To mitigate the computational requirements of run-time-heavy algorithms performed during the read processing stage, the San José State University High Performance Cluster (HPC) was used. The HPC consists of 72 compute nodes, 2 high memory nodes, and 4 GPU nodes for computational use. Each compute node has two 14-core Xeon E5-2680 v4 2.4GHz processors and 128Gb of RAM. These nodes were utilized to execute the shell script for each sample. Overall, 15 compute nodes were used for the 15 FASTQ pairs, running in parallel. Each sample required about a day to execute the script. The resulting file types generated by the shell script were a .taxonomy file and .tax.summary file, which were used to reveal the read counts for each classified taxon.

Pre-processing of the 16S reads was crucial in reducing the run-time cost of the alignment and classification algorithms. PCR amplification is a powerful tool but it tends to introduce errors. Sequencing is also prone to mistakes, especially towards the far end of the sequencing read (Dohm et al., 2008)—highest quality tends to be present in the beginning of the read. This can lead to disingenuous classification results. One way to mitigate this challenge is to analyze the quality score of each base pair (represented by ASCII characters in the raw FASTQ files) and remove base calls that are designated low in quality. This mainly occurs in Mothur during the merging step. Merging is when the forward and reverse reads for one sample sequence are merged by alignment algorithms. Mothur makes use of the fact that sequencing quality tends to decrease towards the end of the read to help merge the reads. If two base calls disagree, Mothur

requires that if one is a base and the other is a gap, that the base call must have a quality score of 25 or greater for the base to be considered correct. If the two base calls are both bases but they are different, the "correct" base call must have at least a quality score 6 points higher than the other. If neither of these conditions correctly pass, then the base call is set to N—meaning that there is potentially a base pair in that position, even though it could be a gap, but the identity is unknown.

Mothur provides a screening stage to reduce computational complexity. The maximum ambiguous base pairs allowed was set to 0. This means that any N calls were not allowed to continue during the succeeding stages of the pipeline. This setting was chosen because preliminary examination of the quality scores of the FASTQ using Trimmomatic reads showed high quality even towards the ends of reads (Bolger et al., 2014). To maximize the number of reads used during the classification stage and reduce false positive results, the maximum ambiguous base parameter was set to 0. The maximum length of reads was set to 275. This was chosen because the sequencer is supposed to generate reads 250 base calls in length, and anything significantly longer might cause mistakes in the merging process and further down the pipeline. The maximum homopolymer length was set to 8. Homopolymers—when one base pair repeats—are features of genes that, although are important and not dangerous in any way to the organism, are problematic for bioinformatics. They tend to disrupt sequencing accuracy and cause problems for assembly and alignment algorithms (Feng et al., 2016). For this reason, the maximum length was set to 8, and any regions exceeding this maximum were screened from the pipeline. This preprocessing stage was crucial for both reducing the run-time complexity of the alignment/classification stage and ensuring high quality reads were analyzed for the results.

Once the reads are pre-processed and merged, Mothur suggests removal of duplicate reads. The reasoning behind this is two-fold. First, duplicate sequences may originate from the same organism, skewing results to favor one taxonomic rank versus the other. Instead, the goal is to search and identify unique sequences with appropriate genetic distance. Second, duplicate reads account significantly for increased run-time of bioinformatics algorithms. It is not necessary to perform the same alignment and classification on the same sequence hundreds, if not, thousands of times. Consideration of these facts contributed to the decision to allow screening of duplicate sequences. The resulting reads have been quality-assured and are unique.

The next stage is alignment. The selection of the reference database is important. The file must be formatted as aligned sequences for the alignment algorithm to perform correctly. Different database files will result in different results, and different run-times necessary. The SILVA bacterial reference alignment is consistently updated and is widely accepted and utilized in bioinformatics, being applied in software packages such as ARB. For this reason, SILVA was chosen as the alignment database. The algorithms performed on the reads included k-mer searching/suffix tree searching, Needleman-Wunsch or Gotoh pairwise alignment, and NAST reinsertion of gaps (necessary for making the candidate compatible with the original template alignment) (Schloss et al., 2009).

Another trick that Mothur implements is mapping the region of interest using start and end points and filtering out aligned reads that are not in the specified target area—then removing the overhangs. This ensures that only aligned reads appear in the 16S V4 region of the prokaryotic

genome. Not only will this increase classification accuracy, but it will also reduce computational run-time. Next, the sequences are pre-clustered in groups by distance, sorting them by abundance—allowing up to "2" differences between sequences before they are merged. Specifying the value "2" here (for the distance parameter) reduces the overall amount of sequences in the classification alignment step—but too much distance would eliminate too many unique sequences.

Once the sequences have been aligned to a reference alignment database and processed to remove extraneous base pairs and reads, they are then classified. For this step, the Ribosomal Database Project (RDP) v9 Mothur-formatted trainset database is used. The database is formatted to match aligned sequences to taxa. The command uses the Wang, k-nearest neighbor consensus, and zap algorithms (Schloss et al., 2009). Once classified, the next command removes undesirable classifications that should not be a part of the sample—this includes mitochondria, chloroplasts, archaea, eukaryotes, and unknowns. The classified read-taxonomy matches are then clustered into operational taxonomic units using distance calculation. A summary file counts the number of reads per OTU. A possible undesirable occurrence in this step is the generation of spurious OTUs. This is mitigated by consolidating the spurious OTU read counts with the genuine OTUs in a future processing step. The generated summary file marks the end of the Mothur classification pipeline.

Processing the result counts was an important step before performing analysis. The files were next parsed by a Java program developed by Dr. Philip Heller and the read-taxonomy matches within the OTUs were counted and displayed on a tree. This allowed simple visual analysis of

the read counts along different ranks of their taxonomy. The most represented phyla were noted, and their classes expanded. Excel tables of read counts per sampling date were generated.

C. Visual Analysis

Stacked bar plots of various groupings were created. The horizontal axis represents the sampling dates and the vertical axis designates the percentage of read counts per taxa grouping for each sampling date. Plots were generated for all phyla with reasonable representation, most represented classes, and phyla with low representation. The purpose of these plots is to represent the bacterial community composition at specific time points. The changes in relative abundance may reflect ecological phenomena.

Two different types of time course plots were also created. The first contains a horizontal axis that represents the sampling dates and the vertical axis that designates the percentage of read counts per taxa groups for each sampling date. The second type contains a vertical axis with logarithm (base 10) of read counts. The first plot type shows raw count data including variations in counts per sampling date. An advantage of using log read counts is that the low represented groups are not hidden, and their trends can be observed. Another benefit is that raw abundance metrics can be observed along with their trends over time. The second plot type with the percent read counts shows relative abundance trends— like stacked bar plots but with more clarity in the trend directions.

A heat map of temperature and salinity environmental conditions was generated across all the sampling dates for identifying correlations with relative abundance trends. Temperature is recorded in degrees Celsius, and salinity is recorded in parts per thousand. The heat map is created using the median of temperature and salinity readings across all sampling dates as the middle color level, with lower values generating white and higher values generating black. The heat map is attached to all time course plots.

A Pearson correlation map was generated of all the groups and environmental conditions. Red represents a negative correlation coefficient and blue represents a positive correlation coefficient. This plot allows correlations to be derived quantitatively from the relative abundance trends over the sample dates. From this plot, groups of taxa that show positive and negative correlations to temperature and salinity can be isolated and further analyzed.

D. Software

Software used for analysis can be found at https://github.com/danssalter/Metagenomic-Survey-of-Marine-16S-Bacterial-Communities-off-Palmer-Station-in-Antarctica

IV. RESULTS

A. Environmental Data

As shown in Table 1, the median temperature and salinity recorded across all austral summer data points in the phylogenetic survey are recorded by the CTD measurements as 0.62°C and 33.33 ppt. In the early summer of 2013 (November 2012 – February 2013), the temperature dropped to -0.45°C in December then rose again to 1.6°C by the beginning of February. During this time, the salinity dropped from 33.75 ppt to 33.39 ppt. Overall the temperature increased substantially while the salinity decreased in the summer of 2013. In the earlier part of the summer of 2014, a similar increase in temperature occurred from 0.19°C in late December to 1.8°C in February. A week later, however, the temperature was recorded at 0.86°C and continued to drop until it measured 0.27°C in March of 2014. During this time the salinity was initially measured at 33.92 ppt and dropped to 32.85 ppt in February. There was an increase, however, to 33.19 ppt a week later (coinciding with the drop in temperature) before dropping to 32.89 ppt by March 2014. Overall, the temperature trend in the summer of 2014 shows a peak in February with a decrease towards the end of the summer. The salinity shows a decrease with a trough coinciding with the peak in temperature in early February. The summer of 2015 begins in December with the temperature and salinity measuring -0.68°C and 33.69 ppt. The temperature decreases slightly to -0.85°C before rising to 1.01°C in mid January 2015. By March 2015, the temperature was measured at 0.56°C. The salinity remained largely the same value, varying by only a few tenths of a value, before last being measured at 33.55 ppt in March 2015. Overall, this year showed less salinity variation than the previous year but similar temperature variation. The first summer (2013) showed the greatest temperature variation at +2.05°C.

Table 1

Date	Temperature (°C)	Salinity (ppt)
2012-11-27 (Y1)	-0.4449	33.7555
2013-02-08 (Y1)	1.6186	33.3893
2013-12-27 (Y2)	0.1924	33.9241
2014-01-23 (Y2)	0.7696	32.9958
2014-02-03 (Y2)	1.8136	32.8465
2014-02-10 (Y2)	0.8569	33.1929
2014-02-28 (Y2)	0.6228	32.7221
2014-03-04 (Y2)	0.2673	32.8897
2014-12-01 (Y3)	-0.6836	33.6873
2014-12-11 (Y3)	-0.8498	33.4236
2015-01-12 (Y3)	0.449	33.3351
2015-01-19 (Y3)	1.0173	33.2011
2015-02-09 (Y3)	0.7323	33.4084
2015-02-23 (Y3)	0.753	33.2617
2015-03-09 (Y3)	0.5587	33.5496

Table 1. CTD data showing temperature in degrees Celsius (°C) and Salinity in parts per thousand (ppt) across all sampling dates. The years for each summer are designated by Year 1 = Summer 2013, Year 2 = Summer 2014, and Year 3 = Summer 2015.

B. Taxa Groupings

Classified taxa were divided into groups based on their composition features. Table 2 shows the groupings. Unclassified and Chloroplast identifications were discarded from analysis. On average 13.7% of the read counts were identified as Bacteria_unclassified. These reads were discarded from analysis. Further analysis of these read counts via BLAST cross-checking revealed that these identifications belonged to uncultured, uncategorized species, confirming their original classification by Mothur. Phyla with unsubstantial read counts were discarded. Phyla with substantial compositions were combined into the All Phyla Group. The phyla with the greatest representation were expanded into their most represented classes and combined into a class group called the Most Represented Classes Group. Finally, the taxa with low representation, but still substantial enough in their counts to show trends, were included into a

separate group called the Least Represented Group. Overall, the Summer of 2014 (Year 2) showed the greatest number of read-match counts according to Figure 1. The relative abundance of bacteria are standardized from the total read counts per sampling date.

Table 2

All Phyla Group	Actinobacteria
	Bacteroidetes
	Chloroflexi
	Deinococcus-Thermus
	Firmicutes
	Fusobacteria
	Planctomycetes
	Proteobacteria
	Spirochaetes
	Verrucomicrobia
Most Represented Classes (expanded from phyla Proteobacteria and Bacteroidetes)	Proteobacteria.Gammaproteobacteria
	Proteobacteria.Alphaproteobacteria
	Proteobacteria.Betaproteobacteria
	Bacteroidetes.Flavobacteria
	Bacteroidetes.Sphingobacteria
	Actinobacteria.Actinobacteria
Least Represented Group	Chloroflexi
	Deinococcus-Thermus
	Firmicutes
	Fusobacteria
	Planctomycetes
	Spirochaetes
	Verrucomicrobia

Table 2. Bacterial taxa groupings based on preliminary community composition analysis.

Figure 1

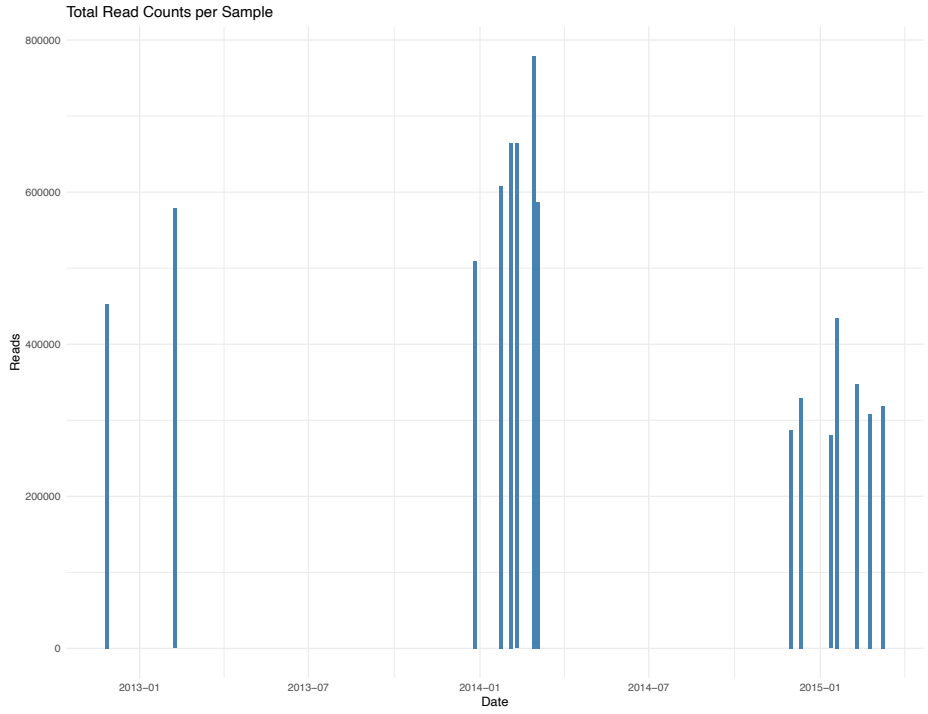


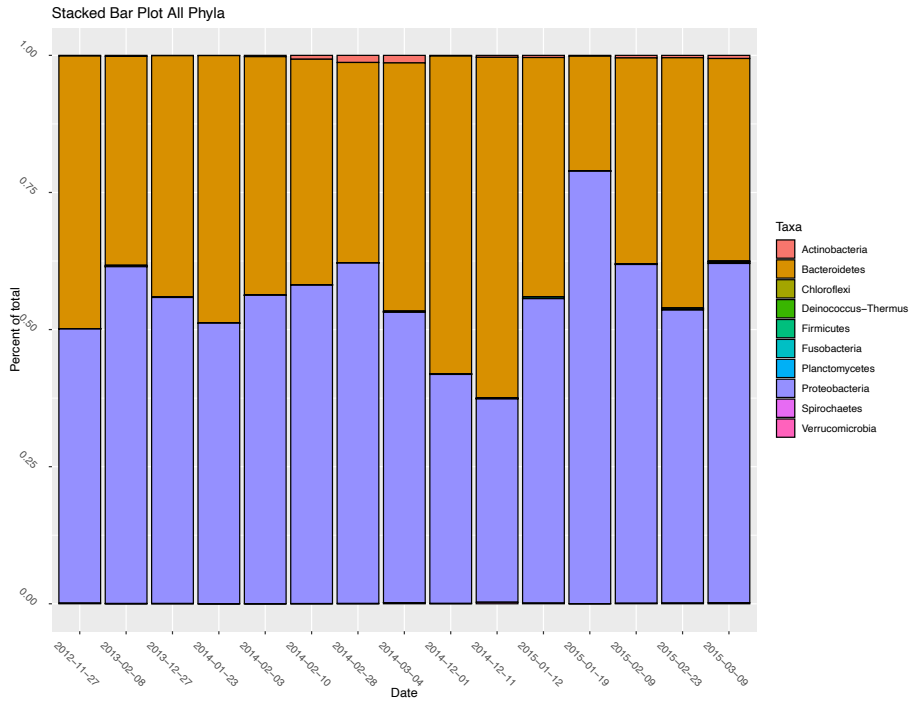
Figure 1. Raw read-match counts (counted from OTU reads in the .tax.summary file generated from the Mothur pipeline).

C. Stacked Bar Plots

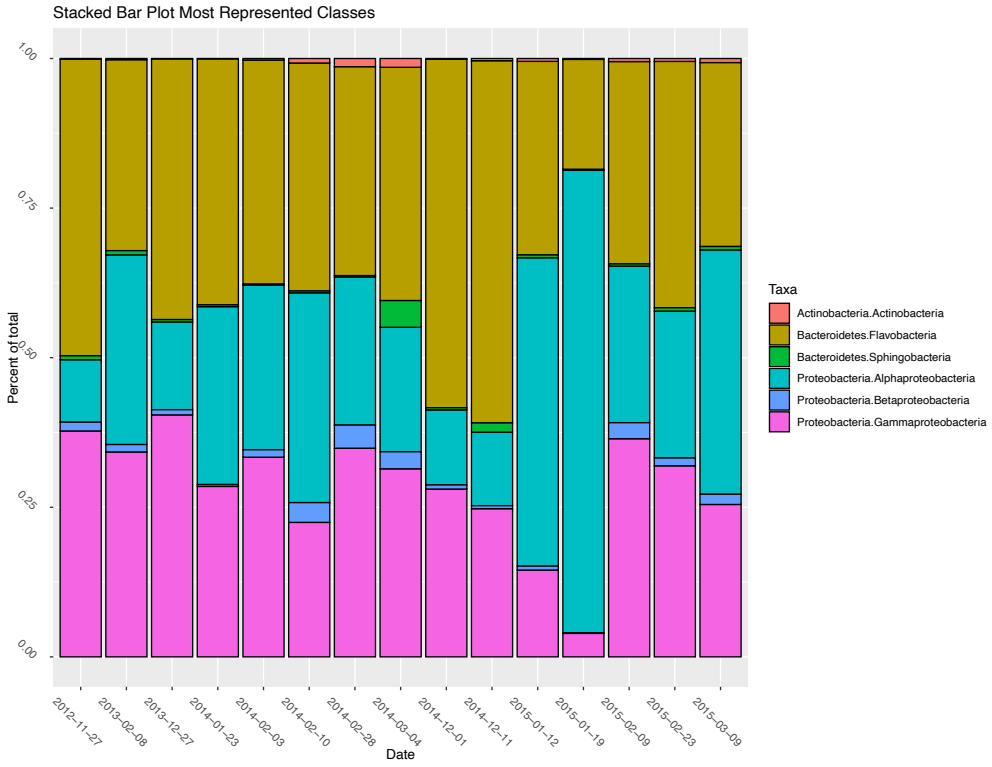
Bacterial community composition results showed that the relative abundance of phyla were dominated by two groups: Proteobacteria and Bacteroidetes—with other phyla comprising below 1% of the remaining classifications. As shown in Figure 2.A, Proteobacteria were the most dominant group with 13 out of the 15 sample dates showing them constituting a majority of the identified phyla. The second most dominant group were the Bacteroidetes. Bacteroidetes did constitute a majority during the beginning of December 2014, but by January 2015 Proteobacteria had regained the majority. Proteobacteria showed the greatest fraction of the identified taxa on 01/19/2015 with its phyla comprising 79.0% of the community composition.

Overall, the two major phyla show variation in composition across all the sample dates with the overall trends of dominance being consistent. On average Proteobacteria made up 55.9% with Bacteroidetes taking up 43.4% of the overall community composition

Figure 2.A.



B.


Stacked Bar Plot Most Represented Classes
Percent of total
1.00
0.75
0.50
0.25
0.00
2012-11-27
2013-02-08
2013-12-27
2014-01-23
2014-02-03
2014-02-10
2014-02-28
2014-03-04
2014-12-01
2014-12-11
2015-01-12
2015-01-19
2015-02-09
2015-02-23
2015-03-09
Date
Taxa
Actinobacteria.Actinobacteria
Bacteroidetes.Flavobacteria
Bacteroidetes.Sphingobacteria
Proteobacteria.Alphaproteobacteria
Proteobacteria.Betaproteobacteria
Proteobacteria.Gammaproteobacteria

C.

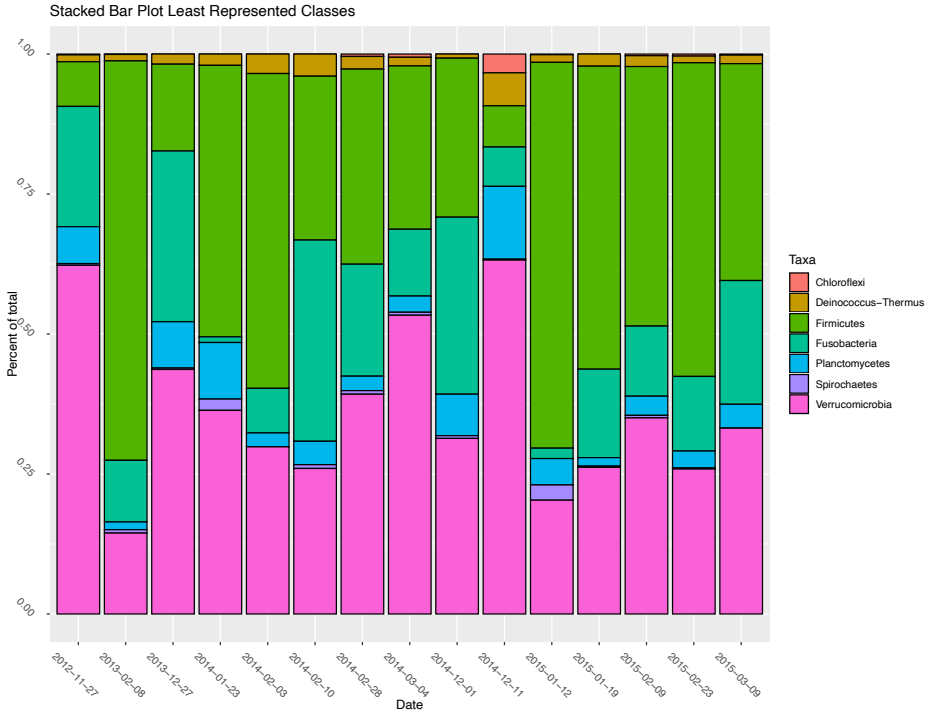

Stacked Bar Plot Least Represented Classes
Percent of total
1.00
0.75
0.50
0.25
0.00
2012-11-27
2013-02-08
2013-12-27
2014-01-23
2014-02-03
2014-02-10
2014-02-28
2014-03-04
2014-12-01
2014-12-11
2015-01-12
2015-01-19
2015-02-09
2015-02-23
2015-03-09
Date
Taxa
Chloroflexi
Deinococcus−Thermus
Firmicutes
Fusobacteria
Planctomycetes
Spirochaetes
Verrucomicrobia

Figure 2. Stacked bar plots of all groupings (A-C). A. Shows the All Phyla Group with the phyla that are most represented in the samples. B. The class groups that comprise the most dominant phyla (Proteobacteria and Bacteroidetes). C. The phyla that are least represented, shown to observe trends in their relative abundance. Each group is normalized to the total of all OTU read counts.

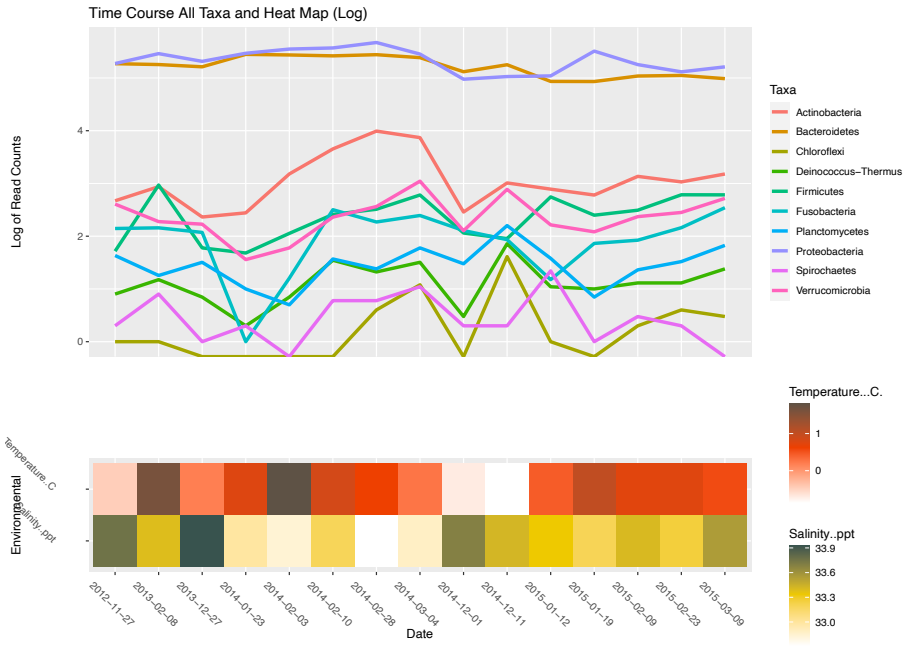
The classes Alphaproteobacteria, Gammaproteobacteria, and Betaproteobacteria make up most of the phylum Proteobacteria. Phylogenetic analysis shows Alphaproteobacteria and Gammaproteobacteria with the highest fraction (Figure 2.B), although the variation in their fractions makes it difficult to discern which is the dominating group. Betaproteobacteria make up a smaller fraction, constituting 0.05% to 4% of the most represented classes. Within Bacteroidetes are the two dominating classes Flavobacteria and Sphingobacteria (the latter to a much reduced extent). Figure 2.B shows Flavobacteria being the most prevalent class among the Most Represented Classes Group, making up on average 39.4% of this group. The next most prevalent class is Alphaproteobacteria with on average 29.3% of the community composition. Between the six classes, Alphaproteobacteria shows the most variation.

Figure 2.C shows the Least Represented Group with mostly consistent fractions of taxa—there is some variation but mostly all the phyla make up a consistent fraction of the group composition. Of this group, the groups with the highest fractions are Firmicutes (39.6%) and Verrucomicrobia (36.2%). In February 2013, Firmicutes comprised of 71.3% of the community composition, coinciding with an increase in temperature. In December 2014, Verrucomicrobia was observed making up most of the community composition at 63.2%. This spike coincides with the simultaneous dominance of Bacteroidetes over Proteobacteria at that point in time. Overall the Least Represented Group shows phyla with relatively consistent community

composition fractions across the three summers, with variation occurring sporadically and temporarily.

D. Time Course Plots and Environmental Heat Maps

Figure 3. A.



B.

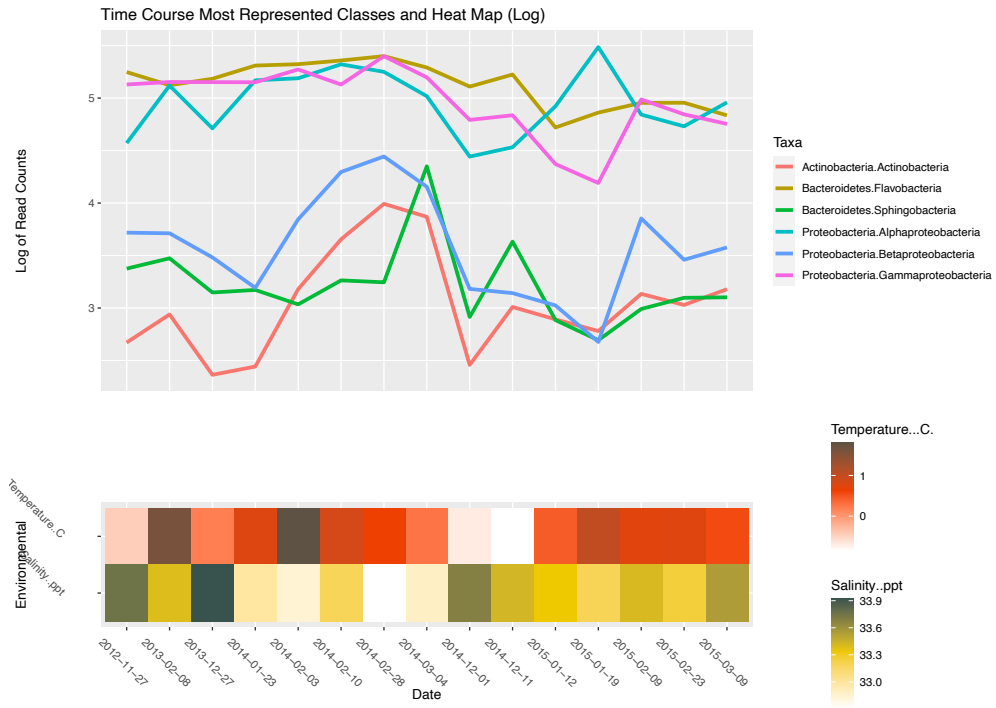
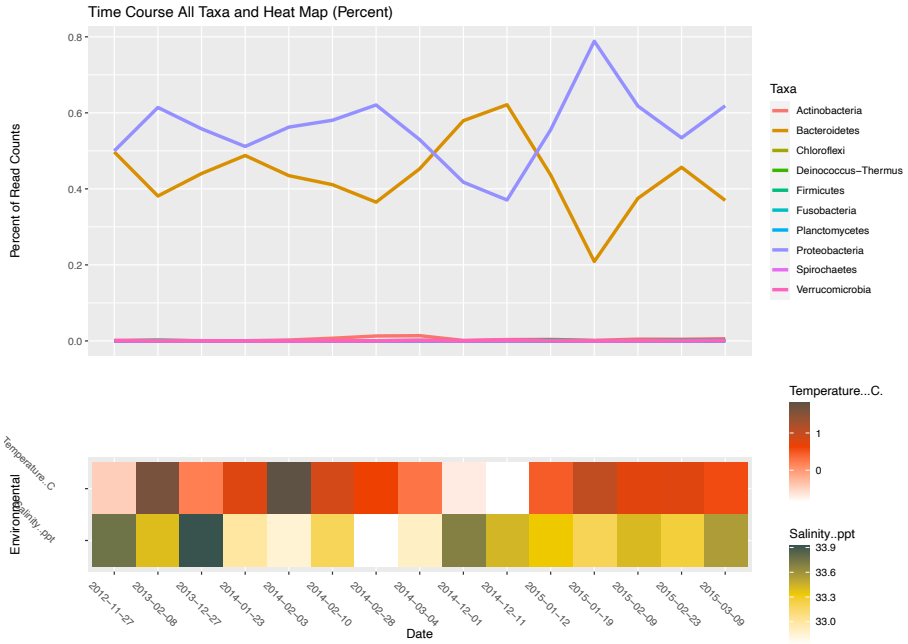


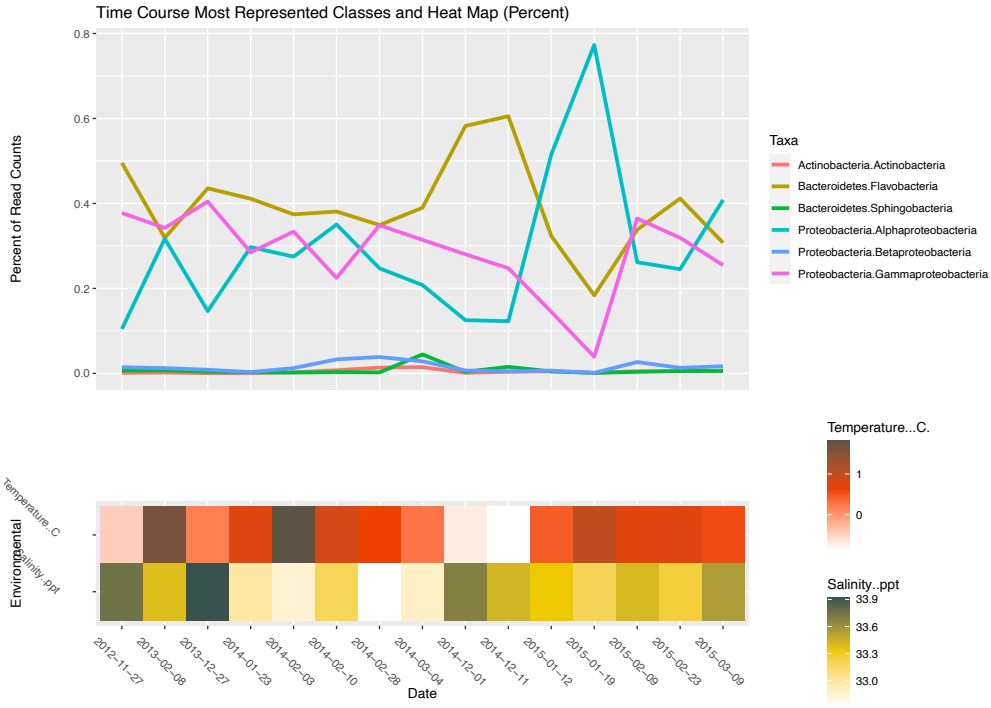
Figure 3. Logarithm of Read Count. Time course line plots with temperature and salinity heat maps below for the All Phyla Group (A) and the Most Represented Class Group (B). In the heat map, color (red for temperature, yellow for salinity) represents the median, white represents low values, and black represents high values. For the vertical axis, logarithm of read count is shown. Horizontal axis represents the sample dates.

Raw read counts for each of the taxa respond strongly to a relative temperature decrease in December 2014 as shown in Figure 3. The decrease, for example in Actinobacteria, is almost two orders of magnitude (value 10). Some phyla, like Chloroflexi and Verrucomicrobia follow the sudden decrease with an instant spike (12/11/2014). Additionally, the dominant phyla Proteobacteria and Bacteroidetes follow a decreasing trend with the cooling period.

Figure 4. A.


Time Course All Taxa and Heat Map (Percent)
Percent of Read Counts
0.8
0.6
0.4
0.2
0.0
Taxa
Actinobacteria
Bacteroidetes
Chloroflexi
Deinococcus−Thermus
Firmicutes
Fusobacteria
Planctomycetes
Proteobacteria
Spirochaetes
Verrucomicrobia
Temperature...C.
1
0
Salinity..ppt
33.9
33.6
33.3
33.0
Environmental
Temperature..C
Salinity..ppt
2012−11−27
2013−02−08
2013−12−27
2014−01−23
2014−02−03
2014−02−10
2014−02−28
2014−03−04
2014−12−01
2014−12−11
2015−01−12
2015−01−19
2015−02−09
2015−02−23
2015−03−09
Date

B.


Time Course Most Represented Classes and Heat Map (Percent)
Percent of Read Counts
0.8
0.6
0.4
0.2
0.0
Taxa
Actinobacteria.Actinobacteria
Bacteroidetes.Flavobacteria
Bacteroidetes.Sphingobacteria
Proteobacteria.Alphaproteobacteria
Proteobacteria.Betaproteobacteria
Proteobacteria.Gammaproteobacteria
Temperature...C.
1
0
Salinity..ppt
33.9
33.6
33.3
33.0
Environmental
Temperature..C
Salinity..ppt
2012−11−27
2013−02−08
2013−12−27
2014−01−23
2014−02−03
2014−02−10
2014−02−28
2014−03−04
2014−12−01
2014−12−11
2015−01−12
2015−01−19
2015−02−09
2015−02−23
2015−03−09
Date

C.

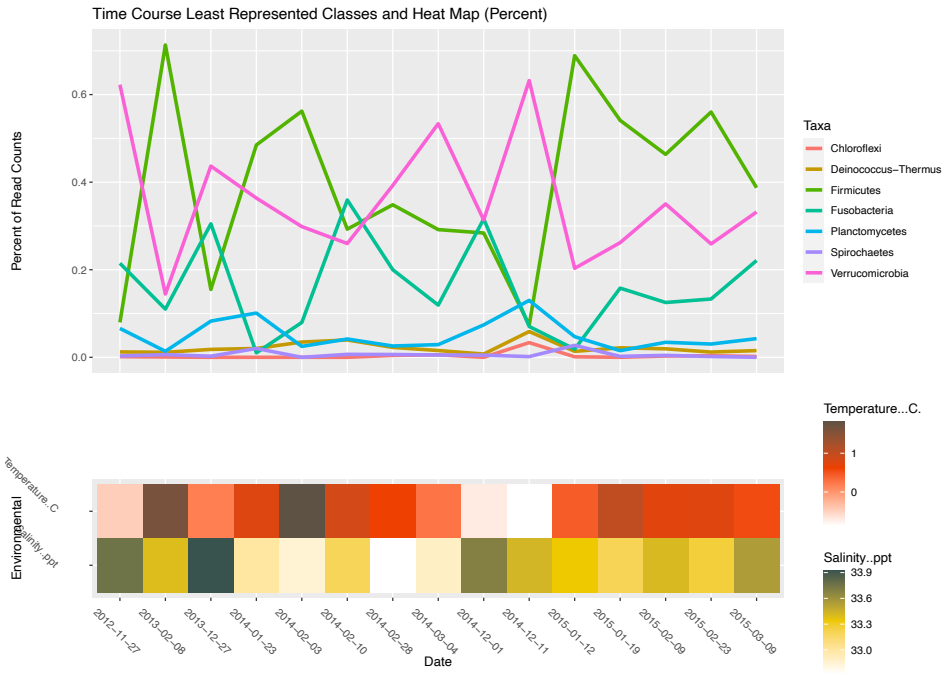


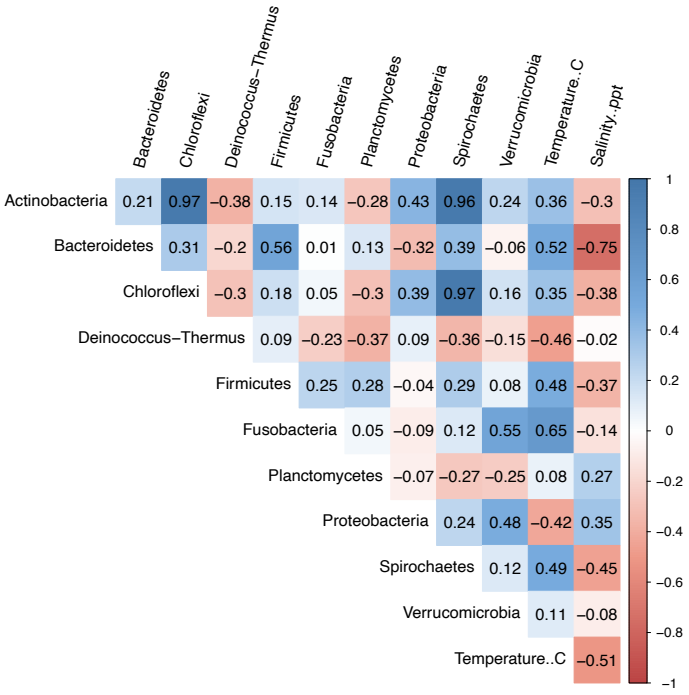
Figure 4. Percent of Read Counts. Time course line plots with temperature and salinity heat maps below. For the vertical axis, percentage of read count is shown. Each plot (A-C) corresponds with a taxa grouping. A. Shows the All Phyla Group. B. Most Represented Classes. C. Least Represented Taxa

Proteobacteria is initially recorded around 60% of the read counts, with Bacteroidetes observed at around 40%. Percent of read count results shows Bacteroidetes being more dominant than Proteobacteria during December 2014, but reverts to Proteobacteria being more dominant by Jan 2015 (Figure 4.A). This also occurs during the cooling period where the temperature drops almost a degree compared to the median. Looking at the Most Represented Classes Group its noted that the sudden relative abundance of Bacteroidetes can be attributed to a rise in Flavobacteria read counts (Figure 4.B). Figure 4.C shows great variation in the relative

abundance of each of the Least Represented Phyla, without any phylum showing consistent composition across the time points.

E. Pearson Correlation Coefficient Analysis

Figure 5 A.



B.

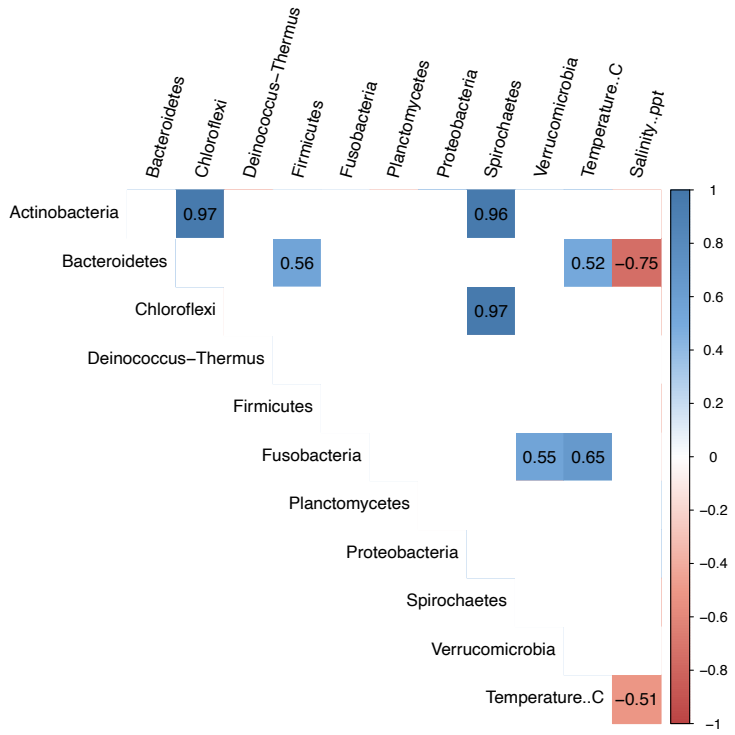
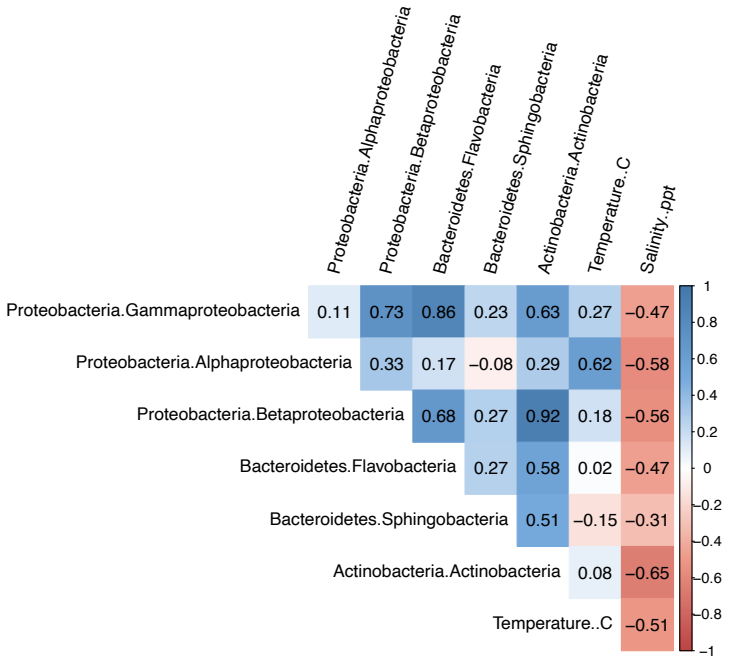


Figure 5. Pearson Correlation Heat Map for the All Phyla Group. A. Significance level = 1.0. B. Significance level = 0.05. Blue represents a positive correlation. Red represents a negative correlation.

The phylum Bacteroidetes is observed following a positive correlation (Pearson coefficient=0.52) with temperature and a negative correlation with salinity (Pearson coefficient=-0.75). Additionally, Fusobacteria is observed following a positive correlation (Pearson coefficient=0.65) with temperature, albeit to a lower degree. Importantly, most of the phyla do not exhibit a significant correlation with temperature or salinity. There are correlations between phyla, with the strongest positive interaction occurring between Actinobacteria, Spirochaetes and Chloroflexi. These taxa are observed at lower assemblage levels than the two dominant phyla previously discussed.

Figure 6. A.



B.

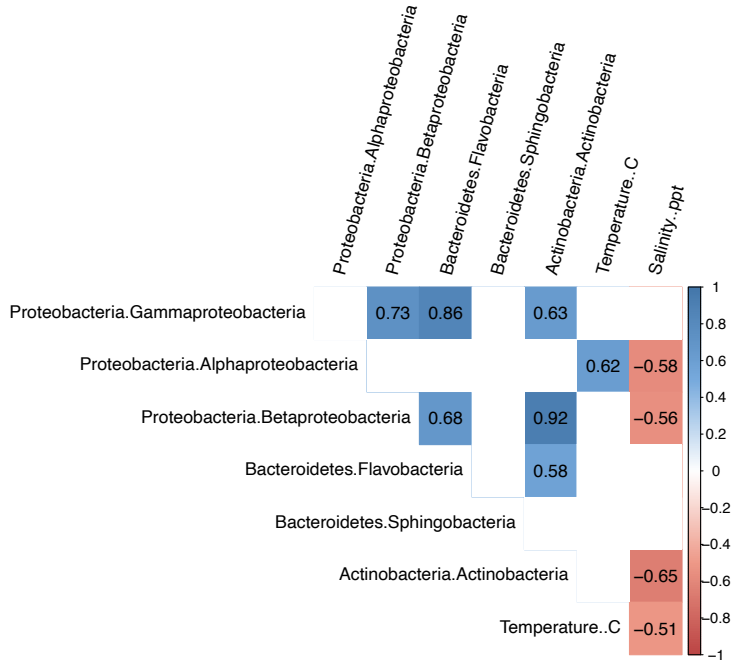


Figure 6. A. Pearson Correlation Heat Map for the Most Represented Classes Group. A. Significance level = 1.0. B. Significance level = 0.05.

Within Bacteroidetes, it's observed a lack of correlation with temperature or salinity. On the other hand, Alphaproteobacteria is shown following positive and negative correlations with temperature (Pearson coefficient=0.62) and salinity (Pearson coefficient=-0.58). Betaproteobacteria is observed with a Pearson coefficient of -0.56 with salinity as well. Most of the positive interactions occur between taxa, with the negative interactions occurring between taxa and salinity

V. DISCUSSION

A. Bacterial Community Composition off Palmer Station Surface Ocean

Bacterial community composition in the Western Antarctic Pacific is predominantly composed of the two main phyla Proteobacteria (55.9%) and Bacteroidetes (43.4%). This is in line with previous research observing the two phyla dominating in both Arctic and non-Arctic waters (Gilbert et al., 2012; Kim et al., 2021; Maas et al., 2012; Nikrad et al., 2014; Stevens et al., 2005). Within Proteobacteria are the classes Alpha-,Gamma-, and Betaproteobacteria. Alpha- and Gammaproteobacteria exhibit stronger numbers but intra- and interseasonal variation suggests that these two classes respond differently to changing conditions. The commonly observed SAR11 group (Pelagibacterales) was not recorded in this sample set. The lack of the prominent SAR11 clade suggests that perhaps the region is more oceanographically isolated than previously thought. This could be attributed to the existence of the powerful Antarctic Circumpolar Current, which buffers the Polar waters from the Southern Ocean (Siegert et al., 2019). Another hypothesis for the lack of the SAR11 clade could be due to the sequencing specificity. Resequencing with more SAR11 specific primers would ascertain the existence of this clade. One strongly represented clade from the phylum Bacteroidetes is the class Flavobacteria. As previous research as suggested, the results of this project supports the hypothesis that Flavobacteria thrives in changing environmental conditions and nutrient-poor depths (González et al., 2008). Other groups with lower representation like Chloroflexi, Actinobacteria, and Verrucomicrobia showed great variation across time. This suggests that these taxa possess the ability to adapt meaningfully to changing niche conditions. Overall these groups persistence over time lends support to the hypothesis that bacterial community composition in the Western Antarctic Peninsula can resist collapse from changing environmental conditions.

B. Correlation with Temperature and Salinity

The correlation analysis of this project reported meaningful trends by different taxa between groups and different environmental conditions. During Year One (Summer of 2013), there were significant changes in temperature and salinity. However, little meaningful variation in the bacterial community composition occurred during this time. That is not to say that there wasn't any variation at all, but rather that the macroscopic composition of the water column stayed persistent during this time. This suggests that water column instability due to environmental changes does little to affect relative abundance of bacteria. This is in line with previous research that reported strong resistance to seasonal variation in marine bacterial group composition in the Western Antarctic Peninsula (Saba et al., 2014).

Another meaningful trend to report occurs during Year Three (Summer of 2015), in which a sudden decrease in temperature caused substantial changes in the relative abundance of taxa. During December 2014 there was a drop in temperature and increase in salinity, resulting in conditions that normally destabilize water columns. Figure 3 shows that many taxa drop in read counts during this period. However, Bacteroidetes read counts resisted the drop more than other groups. Within Bacteroidetes shows Flavobacteria maintaining read counts while Proteobacteria classes Alphaproteobacteria and Gammaproteobacteria numbers drop. This change in read count numbers led to a period where Bacteroidetes dominated over all other phyla including Proteobacteria. This supports that Flavobacteria does in fact strongly resist unfavorable changes in environmental conditions (cooling periods). Later in January, an increase in temperature and decrease in salinity shows an increase in Alphaproteobacteria and decrease in Flavobacteria. This

trend suggests that while Alphaproteobacteria is less resistant to changing conditions in Arctic waters, it can adapt more quickly to favorable conditions than Flavobacteria (which is more resistant to unfavorable conditions). Pearson coefficient analysis substantiates this trend as it reports Alphaproteobacteria having significant correlations with temperature and salinity while Flavobacteria is observed without significant correlations.

Strong correlations between taxa also suggest that while these groups may successfully resist seasonal variation, their strongest challenge exists in the form between interspecies interactions. Proteobacteria and Alphaproteobacteria do exhibit a negative correlation (Pearson coefficient= -0.32), however not at a meaningful significance level. Other taxa report positive and negative correlations between each other. Some stronger positive correlations also occur with meaningful significance levels, suggesting that perhaps these taxa share similar life cycles and respond similarly to changing conditions.

Long-term climate change correlations are not discernable in this project due to the length of the project (three years); however, meaningful intraseasonal variation results are observed. The ability of specific groups like Flavobacteria to resist change and maintain persistent community composition combined with other groups' (like Alphaproteobacteria) ability to respond strongly to favorable changes in the environment suggest an overall resilience of the bacterial community. However, the existence of two dominant phyla could spell trouble for potential significant climate changes as ecosystems with more diversity tend to resist climate extremes than less diverse ones (Isbell et al., 2015). However, trends observed in this project lend support to the

idea that many taxa possess the ability to both resist and adapt to changing environmental conditions.

C. Future Research Direction

One avenue to continue this research is to delve deeper into taxonomy of the reads and look at the Order and Family level. This could provide more insight into the types of Bacteria that are present in the surface ocean of the Western Antarctic Peninsula. Future research could also explore the relationship between bacterial and eukaryotic community composition, and how phytoplankton blooms affect bacterial populations. In order to further emphasize the implications of this project, more samples from subsequent years would help benefit the project's ability to observe ecological trends and how it relates to more long-term climate change effects. Additionally, seasonal variation analysis would benefit from samples taken during both summer and winter time points, as this could help illuminate how environmental conditions and water column stability changes affect bacterial community composition. However, it is understood that it is logistically very difficult to obtain more time point samples in the Western Antarctic Peninsula, therefore more emphasis is put on gathering samples across more years rather than seasons. Also, the databases selected for the sequence processing stage of this project could be exchanged for different ones to analyze how the results change and what that would mean for the implications of this project.

VI. REFERENCES

Alcamán-Arias, M. E., Fuentes-Alburquenque, S., Vergara-Barros, P., Cifuentes-Anticevic, J., Verdugo, J., Polz, M., Farías, L., Pedrós-Alió, C., & Díez, B. (2021). Coastal Bacterial Community Response to Glacier Melting in the Western Antarctic Peninsula. *Microorganisms*, *9*(1), 88. https://doi.org/10.3390/microorganisms9010088

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Aristarain, A. J., Jouzel, J., & Lorius, C. (1990). A 400 years isotope record of the Antarctic Peninsula climate. *Geophysical Research Letters*, *17*(13), 2369–2372. https://doi.org/10.1029/GL017i013p02369

Aristarain, A. J., Jouzel, J., & Pourchet, M. (1986). Past Antarctic Peninsula climate (1850?1980) deduced from an ice core isotope record. *Climatic Change*, *8*(1), 69–89. https://doi.org/10.1007/BF00158970

Bijma, J., Pörtner, H.-O., Yesson, C., & Rogers, A. D. (2013). Climate change and the oceans – What does the future hold? *Marine Pollution Bulletin*, *74*(2), 495–505. https://doi.org/10.1016/j.marpolbul.2013.07.022

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Cho, B. C., & Azam, F. (1988). Major role of bacteria in biogeochemical fluxes in the ocean's interior. *Nature*, *332*, 441–443.

Cram, J. A., Chow, C.-E. T., Sachdeva, R., Needham, D. M., Parada, A. E., Steele, J. A., & Fuhrman, J. A. (2015). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *The ISME Journal*, *9*(3), 563–580. https://doi.org/10.1038/ismej.2014.153

Dohm, J. C., Lottaz, C., Borodina, T., & Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research*, *36*(16), e105. https://doi.org/10.1093/nar/gkn425

Ducklow, H. W., Baker, K., Martinson, D. G., Quetin, L. B., Ross, R. M., Smith, R. C., Stammerjohn, S. E., Vernet, M., & Fraser, W. (2007). Marine pelagic ecosystems: The West Antarctic Peninsula. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1477), 67–94. https://doi.org/10.1098/rstb.2006.1955

Ducklow, H. W., Schofield, O., Vernet, M., Stammerjohn, S., & Erickson, M. (2012). Multiscale control of bacterial production by phytoplankton dynamics and sea ice along the western Antarctic Peninsula: A regional and decadal investigation. *Journal of Marine Systems*, *98–99*, 26–39. https://doi.org/10.1016/j.jmarsys.2012.03.003

Feng, W., Zhao, S., Xue, D., Song, F., Li, Z., Chen, D., He, B., Hao, Y., Wang, Y., & Liu, Y. (2016). Improving alignment accuracy on homopolymer regions for semiconductor-based sequencing technologies. *BMC Genomics*, *17*(S7), 521. https://doi.org/10.1186/s12864-016-2894-9

Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature*, *459*(7244), 193–199. https://doi.org/10.1038/nature08058

Fuhrman, J., & Steele, J. (2008). Community structure of marine bacterioplankton: Patterns, networks, and relationships to function. *Aquatic Microbial Ecology*, *53*, 69–81. https://doi.org/10.3354/ame01222

Ghiglione, J. F., & Murray, A. E. (2012). Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environmental Microbiology*, *14*(3), 617–629. https://doi.org/10.1111/j.1462-2920.2011.02601.x

Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., Huse, S., McHardy, A. C., Knight, R., Joint, I., Somerfield, P., Fuhrman, J. A., & Field, D. (2012). Defining seasonal marine microbial community dynamics. *The ISME Journal*, *6*(2), 298–308. https://doi.org/10.1038/ismej.2011.107

Glöckner, F. O., Fuchs, B. M., & Amann, R. (1999). Bacterioplankton Compositions of Lakes and Oceans: A First Comparison Based on Fluorescence In Situ Hybridization. *Applied and Environmental Microbiology*, *65*(8), 3721–3726. https://doi.org/10.1128/AEM.65.8.3721-3726.1999

González, J. M., Fernández-Gómez, B., Fernàndez-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., Del Campo, J., Escudero, L., Rodríguez-Martínez, R., Alonso-Sáez, L., Latasa, M., Paulsen, I., Nedashkovskaya, O., Lekunberri, I., Pinhassi, J., & Pedrós-Alió, C. (2008). Genome analysis of the proteorhodopsin-containing marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proceedings of the National Academy of Sciences*, *105*(25), 8724–8729. https://doi.org/10.1073/pnas.0712027105

Hogg, A. McC., Meredith, M. P., Chambers, D. P., Abrahamsen, E. P., Hughes, C. W., & Morrison, A. K. (2015). Recent trends in the S outhern O cean eddy field. *Journal of Geophysical Research: Oceans*, *120*(1), 257–267. https://doi.org/10.1002/2014JC010470

IPCC. (2022). *Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009157940

Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T. M., Bonin, C., Bruelheide, H., De Luca, E., Ebeling, A., Griffin, J. N., Guo, Q., Hautier, Y., Hector, A., Jentsch, A., Kreyling, J., Lanta, V., Manning, P., … Eisenhauer, N. (2015). Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature*, *526*(7574), 574–577. https://doi.org/10.1038/nature15374

Kim, H.-R., Lim, J.-H., Kim, J.-H., & Kim, I.-N. (2021). Collection of Environmental Variables and Bacterial Community Compositions in Marian Cove, Antarctica, during Summer 2018. *Data*, *6*(3), 27. https://doi.org/10.3390/data6030027

Kovacs, A., Yacoby, K., & Gophna, U. (2010). A systematic assessment of automated ribosomal intergenic spacer analysis (ARISA) as a tool for estimating bacterial richness. *Research in Microbiology*, *161*(3), 192–197. https://doi.org/10.1016/j.resmic.2010.01.006

Lane, D. J., Pace, B., Olsen, G. J., STAHLt, D. A., SOGINt, M. L., & Pace, N. R. (1985). Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. USA*.

Llovel, W., & Terray, L. (2016). Observed southern upper-ocean warming over 2005–2014 and associated mechanisms. *Environmental Research Letters*, *11*(12), 124023. https://doi.org/10.1088/1748-9326/11/12/124023

Luria, C., Ducklow, H., & Amaral-Zettler, L. (2014). Marine bacterial, archaeal and eukaryotic diversity and community structure on the continental shelf of the western Antarctic Peninsula. *Aquatic Microbial Ecology*, *73*(2), 107–121. https://doi.org/10.3354/ame01703

Maas, E., Simpson, A., Martin, A., Thompson, S., Koh, E., Davy, S., Ryan, K., & O'Toole, R. (2012). Phylogenetic analyses of bacteria in sea ice at Cape Hallett, Antarctica. *New Zealand Journal of Marine and Freshwater Research*, *46*(1), 3–12. https://doi.org/10.1080/00288330.2011.579981

Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A., & Giovannoni, S. J. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*, *420*(6917), 806–810. https://doi.org/10.1038/nature01240

Nikrad, M. P., Cottrell, M. T., & Kirchman, D. L. (2014). Growth activity of gammaproteobacterial subgroups in waters off the west A ntarctic P eninsula in summer and fall. *Environmental Microbiology*, *16*(6), 1513–1523. https://doi.org/10.1111/1462-2920.12258

Pearce, D. A. (2008). Climate Change and the Microbiology of the Antarctic Peninsula Region. *Science Progress*, *91*(2), 203–217. https://doi.org/10.3184/003685008X332534

Saba, G. K., Fraser, W. R., Saba, V. S., Iannuzzi, R. A., Coleman, K. E., Doney, S. C., Ducklow, H. W., Martinson, D. G., Miles, T. N., Patterson-Fraser, D. L., Stammerjohn, S. E., Steinberg, D. K., & Schofield, O. M. (2014). Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula. *Nature Communications*, *5*(1), 4318. https://doi.org/10.1038/ncomms5318

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G.

G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Siegert, M., Atkinson, A., Banwell, A., Brandon, M., Convey, P., Davies, B., Downie, R., Edwards, T., Hubbard, B., Marshall, G., Rogelj, J., Rumble, J., Stroeve, J., & Vaughan, D. (2019). The Antarctic Peninsula Under a 1.5°C Global Warming Scenario. *Frontiers in Environmental Science*, *7*, 102. https://doi.org/10.3389/fenvs.2019.00102

Smith, R., Baker, K., Fraser, W., Hofmann, E., Karl, D., Klink, J., Quentin, L., Prezelin, B., Ross, R., Trivelpiece, W., & Vernet, M. (1995). The Palmer LTER: A Long-Term Ecological Research Program at Palmer Station, Antarctica. *Oceanography*, *8*(3), 77–86. https://doi.org/10.5670/oceanog.1995.01

Staley, J. T., & Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology*, *39*(1), 321–346. https://doi.org/10.1146/annurev.mi.39.100185.001541

Steele, J. A., Countway, P. D., Xia, L., Vigil, P. D., Beman, J. M., Kim, D. Y., Chow, C.-E. T., Sachdeva, R., Jones, A. C., Schwalbach, M. S., Rose, J. M., Hewson, I., Patel, A., Sun, F., Caron, D. A., & Fuhrman, J. A. (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *The ISME Journal*, *5*(9), 1414–1425. https://doi.org/10.1038/ismej.2011.24

Stevens, H., StÃ¼bner, M., Simon, M., & Brinkhoff, T. (2005). Phylogeny of Proteobacteria and Bacteroidetes from oxic habitats of a tidal flat ecosystem. *FEMS Microbiology Ecology*, *54*(3), 351–365. https://doi.org/10.1016/j.femsec.2005.04.008

Suzuki, M. T., Rappé, M. S., Haimberger, Z. W., Winfield, H., Adair, N., Ströbel, J., & Giovannoni, S. J. (1997). Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. *Applied and Environmental Microbiology*, *63*(3), 983–989. https://doi.org/10.1128/aem.63.3.983-989.1997

Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., Lagun, V., Reid, P. A., & Iagovkina, S. (2005). Antarctic climate change during the last 50 years. *International Journal of Climatology*, *25*(3), 279–294. https://doi.org/10.1002/joc.1130

Vaughan, D. G., Marshall, G. J., Connolley, W. M., Parkinson, C., Mulvaney, R., Hodgson, D. A., King, J. C., Pudsey, C. J., & Turner, J. (2003). *Recent Rapid Regional Climate Warming on the Antarctic Peninsula*.