MICROBIAL ECOLOGY OF THE ISLAND MASS EFFECT OFF THE WAI‘ANAE COAST OF O‘AHU

A THESIS SUBMITTED FOR PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE

IN

GLOBAL ENVIRONMENTAL SCIENCE

MAY 2025

By

DEVIN HOGATE

Thesis Advisor

DR. CRAIG NELSON

I certify that I have read this thesis and that, in my opinion, it is satisfactory in scope and quality as a thesis for the degree of Bachelor of Science in Global Environmental Science.

THESIS ADVISOR

DR. CRAIG NELSON
Department of Oceanography

For my family, Stephen, Angela, Cassie, and Lynzee, for supporting me throughout my time in college and being a base that I could rely on.

ACKNOWLEDGEMENTS

I would like to acknowledge the invaluable help and support of a couple of groups which without this work would not have been possible. My biggest thanks goes to the people of the Nelson Lab for assisting me as an undergraduate student through the struggle of scientific research and writing. My thanks also go out to my GES cohort in which I have found steadfast friends and vital support through coursework.

This work was logistically and financially supported by NSF Division of Ocean Sciences grants 2221749 (PI Glenn Carter) and 2023298 (PI Craig Nelson) as well as an award from the University of Hawaiʻi at Mānoa Undergraduate Research Opportunity Program to Devin Hogate. Thank you to the Biological Electron Microscope Facility for allowing me to use their flow cytometer, supported by Integrative Center for Precision Nutrition and Human Health COBRE grant NIH 1P20GM139753-01. Thank you to the SOEST Laboratory for Analytical Biogeochemistry for the use of their facilities to process my chlorophyll samples. This thesis reports data obtained at the University of Hawaiʻi at Mānoa Microbial Genomics and Analytical Laboratory Core, which is supported by two Institutional Development Awards (IDeA) from the National Institute of General Medical Sciences of the National Institute of Health under grant numbers P20GM125508 and P20GM139753.

ABSTRACT

While the Island Mass Effect (IME) has been widely observed and documented for decades, little research has been done regarding how it impacts heterotrophic bacterioplankton communities, and none in Hawaiʻi. My project studied how bacterioplankton communities vary with distance from shore in the context of an IME by synoptically characterizing the taxonomic composition of bacteria over depth along duplicate occupations of a transect from nearshore to offshore over a 24h period. I collected water samples from 6 depth profiles (5 to 250 meters) at three stations (roughly 2, 7, and 23 km from the Waiʻanae Coast of Oʻahu) conducted in September 2023 aboard the research vessel R/V Kilo Moana. I characterized the relative abundances of all bacterial taxa using 16S amplicon sequencing and measured absolute concentrations of the Cyanobacteria *Prochlorococcus* and *Synechococcus* as well as total heterotrophic bacterioplankton and picoeukaryotic phytoplankton using flow cytometry. Chlorophyll *a* measurements showed little variation in the mixed layer but a strengthening and shoaling of the deep chlorophyll maximum (DCM) in the nearshore. I characterized dissolved organic matter, observing vertical shifts in multivariate dissolved organic matter fluorescence signatures but no horizontal shifts along the transect. Bulk dissolved organic carbon also showed little significant horizontal variation. When comparing the community at each station, *Prochlorococcus* was more abundant offshore within the DCM and *Synechococcus* was more abundant nearshore in the upper mixed layer in both relative and absolute abundances, suggesting resource partitioning between these two major clades of picocyanobacteria. I observed the greatest horizontal differences in bacterioplankton community makeup between stations in the range of 100 to 125 meters depth; dozens of Bacteria ecotypes including the abundant taxa SAR86, *Flavobacteriaceae*, and *Cryomorphaceae* showed clear horizontal gradients. Taken together my work documents a clear IME in heterotrophic microbial communities that manifests throughout the euphotic zone, emphasizing a need for a deeper understanding of the vertical context in understanding how islands affect marine biogeochemistry and microbial ecology.

Keywords: Island Mass Effect, Microbial community dynamics, Flow cytometry, biogeochemical markers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 THE ISLAND MASS EFFECT

The Island Mass Effect (IME) describes the long-term increase in the phytoplankton biomass and productivity of waters around islands compared to the nearby oligotrophic open ocean (Doty and Oguri, 1956; Gilmartin and Revelante, 1974). This effect was first observed through a shift in chlorophyll concentrations which decreased with distance from an island. The IME is observed to be the most pronounced around coral reef island and atolls where there are increased nutrient inputs from terrestrial runoff and reef ecosystems (Gove et al., 2016; Messié et al., 2022).
The biggest drivers of the IME are thought to be an island's bathymetry, reef productivity, and human populations (De Falco et al., 2022; Gove et al, 2016). Terrestrial inputs from human activity and coral reefs transport nutrients and organic matter into nearshore waters which are thought to support the increase in productivity. Altered ocean mixing and upwelling along island bathymetric slopes brings in additional nutrients from enriched deeper waters below the euphotic zone (Hasegawa et al., 2004).

Studies of the IME across the Pacific basin have been conducted using satellite measurements of surface chlorophyll levels, capturing chlorophyll in the surface 30 or so meters, demonstrating that the IME contributed to long-term increases in phytoplankton biomass in the nearshore environment, an increase ranging from 30 to 85% over the surrounding oligotrophic waters (Gove et al., 2016). The IME also has been observed to increase surface chlorophyll levels around islands by 9%, and a greater increase of 26% in the pixels closest to islands to islands, although satellite studies cannot resolve shallow areas along to shore due to satellite image resolution on the scale of kilometers per side of pixel (Messié et al., 2022). Increased productivity from IMEs contributes to more productive ecosystems that sustain biodiversity and human populations around islands (Gove et al., 2016). To measure chlorophyll levels below the surface waters that satellites cannot reach, ship based sampling of the water column can resolve shifts in chlorophyll below the surface. Through these ship-based measurements, it was found that the IME has an influence throughout the water column beyond the upper mixed layer (Gove et al., 2016).

1.2 MICROBIAL OCEANOGRAPHY IN THE PACIFIC

Much of the ocean is open oligotrophic waters, where the ecosystem's biomass and productivity are dominated by microbial life. The influence of the IME on nearshore island ecosystems is important because of the outsized contributions of microbes to the food web of oligotrophic regions. Increased microbial productivity increases the available biomass for higher trophic level organisms and enhances the overall productivity of the region (Drazen et al., 2023). The microbial community constitutes an important role as the foundation of the marine food web as primary producers and recyclers of dissolved organic material (Pomeroy et al., 2007). Microbial osmotrophs also help to recycle dissolved organic matter (DOM) in the water too small for larger organisms to consume by absorbing them and incorporating the nutrients into their cells, which are then consumed and moved through the food web (Pomeroy et al., 2007). These compounds originate from dissolving excrement or dead cells aggregates from larger organisms or broken-down organic material. Without the microbial loop much of the nutrients would not be able to reenter the food web, restricting the growth of organisms.

Within the ocean, microbial communities mostly differ by latitude and depth, where the availability of sunlight and nutrients defines which microbes are dominant (Flombaum et al., 2013; Comstock et al., 2022). The Hawaii Ocean Time-Series at Station ALOHA is a long-term time series that studies oligotrophic ocean conditions within the Pacific Ocean subtropical gyre (Karl et al, 2001). Studies at Station ALOHA have created a picture of the vertical structure of the microbial community in the North Pacific Sub-tropical gyre just north of Hawaii (Delong et al, 2006; Mende et al., 2017, van den Engh et al., 2017). Within the upper layers of the ocean in the euphotic zone, autotrophic bacteria make use of the sunlight to photosynthesize. The most abundant genus of photosynthetic bacteria, *Prochlorococcus*, can be found in a wide range of latitudes and is most dominant in nutrient-limited oligotrophic environments (van de Engh et al., 2017). As the community shifts to deeper waters, the available light decreases, the abundance of photosynthetic taxa such as *Prochlorococcus* and *Synechococcus* decreases (Campbell and Vaulot, 1993). Moving below the euphotic zone into the mesopelagic shows a drastic shift in microbial community composition as light environmental drivers including light and nutrient availabilities change with depth (Mende et al., 2017). Distinct communities were found to occur at the depth horizons of 25, 75, 125, and 200 meters. Shifts within populations in multiple clades across the DCM including groups of high-light and low-light *Prochlorococcus* (Mende et al., 2017). The availability of organic and inorganic nitrogen was not temporally constant and was shown to drive changes in the community at 125 meters and deeper. As nutrient sources are expected to change with distance to shore and light changes with depth, so too should the microbial community reflect the change in available light and nutrients.

1.3 ISLAND NEARSHORE CHEMISTRY

Seawater conditions vary across the ocean based on geography and productivity. Near islands, the seawater generally contains elevated concentrations of nutrients (Comfort et al., 2024) and sediment (Dandonneau and Charpy, 1985) but concentrations of dissolved organic matter compared to open ocean environments are less predictable, sometimes exhibiting nearshore depletion (Nelson et al., 2011) and other times nearshore enrichment (Torréton et al., 1997). Terrestrial inputs of nutrients and particulate organic matter from the island provide a source that enriches the biological productivity of the water near the island (Hoover and Mackenzie, 2009). Along the coast of volcanic islands like O‘ahu, submarine groundwater discharge into the nearshore environment introduces nutrients and sediments (Knee et al., 2010) as well as terrigenous DOM (Nelson et al. 2015). Within the shadow of islands, the physical presence of the island mass contributes to mixing and upwelling as currents move past an island (Signorini et al., 1999). Mixing and upwelling introduces otherwise inaccessible nutrients in the deeper waters to the nutrient depleted surface waters, enriching the surface for increased productivity.

1.4 STUDIES OF THE ISLAND MASS EFFECT

Although IME’s have been studied for many decades, the ways the IME, biogeochemistry, and microbial communities influence each other is still unclear. Much of the work done on the IME has focused on phytoplankton and biochemical tracers within the water. A study of the Pacific basin with satellites found *Synechococcus* to increase within IME regions and *Prochlorococcus* to decrease (Messié et al., 2022). A previous work on the IME surrounding the Hawaiian Islands has observed similar shifts in the photosynthetic bacteria *Prochlorococcus* and *Synechococcus*. Comparing nearshore sites within 7 km of the leeward side of O‘ahu near highly populated

Waikiki area to the Station ALOHA oceanographic study area (100 km north of the northernmost point of Oʻahu) Comfort et al. (2024) found that *Prochlorococcus* concentrations increased with distance from shore, while *Synechococcus* concentrations exhibited a decrease with distance from shore. The study also investigated the biogeochemical tracers of the IME in the ocean environment, namely particulate silicate, macronutrients nitrate, nitrite, phosphate, and linked those to concentrations of chlorophyll a as a proxy for phytoplankton biomass.

Other studies of the IME in Hawaii have used a sampling scheme of 151 stations along transects that were spread through the main Hawaiian islands from Oʻahu to Hawaiʻi Island (Gilmartin and Revelante, 1974). Most of these were located between one to 20 km from shore and found that surface chlorophyll concentrations were highest nearshore on the leeward coasts of islands and decreased at offshore sites. The study also found primary productivity also increased nearshore and determined that the growth was caused by the nutrients from the island being assimilated by microorganisms. Other high-resolution studies of the IME in the Pacific have used similar methods that focus on the phytoplankton and zooplankton distribution through flow cytometry and net tows, and biogeochemical tracers in the environment (Martinez et al., 2020).

More limited are studies examining dynamics of microbial communities beyond the phytoplankton. The study on the IME around the island of Mo'orea by Comstock et al. (2022) sequenced the microbial community in the reefs around the island and the surrounding waters in a grid of 51 by 67 km and over an 11-day sampling scheme. The impact of the IME on the microbial community was found to be the greatest in the DCM (90 to 125 meters depth) and driven by mater mass densities rather than nutrient availability. The large spatial scale of this study missed any short-term local changes in the community. These previous studies miss the potential shifts within the heterotrophic microbial community that could help to explain the observed shifts in biogeochemical tracers. My study aims to close the gap in high-resolution studies of the IME by focusing on a short timeframe of the bacterial community distribution across a repeated transect survey.

1.5 CHARACTERIZING THE Oʻ‌AHU IME INFLUENCED MICROBIAL COMMUNITY

The goal of this study was to characterize the IME off of the Waiʻanae coast of Oʻahu focusing on the microbial community composition and evaluate the shifts in the community across a transect running perpendicular from shore. The current understanding of the IME is that it increases the primary productivity and chlorophyll around islands caused by island nutrient input and ocean mixing, but past studies focused on phytoplankton concentrations and biochemical tracers such as chlorophyll, dissolved organic matter, and nutrients nitrate, nitrite, phosphorous. I hypothesized that we would recreate synoptically the patterns of Comfort et al. (2024), who demonstrated that the photosynthetic microbe *Prochlorococcus* was most abundant in offshore waters at all depths and decreased in the nearshore stations, while *Synechococcus* was most abundant in nearshore waters in the upper mixed layer. Additionally, I measured the abundances of individual taxa within the heterotrophic microbial community to observe shifts within those taxa with distance from shore, to resolve potential covariation patterns with the shifts previously studied in phytoplankton. I also measured biochemical tracers, including chlorophyll and dissolved organic matter, to test for shifts in these tracers that correlated with the patterns in the microbial taxa.

2.0 METHODS

2.1 STATION SETUP

From September 23rd to the 26th of 2023, I participated in a collaborative student research cruise (KM2315) where we collected water samples aboard the *R/V Kilo Moana*. The survey region was located off the West Coast of Oʻahu and included 4 stations that were positioned at different distances perpendicular to Pili o Kahe within the Honouliuli ahupuaʻa at the boundary of ʻEwa and Waiʻanae moku (Fig. 1). Station A was situated closest to the island, only 1.75 kilometers from the shore. This station was to be as close to Oʻahu as the ship could safely approach and still represent nearshore conditions. Station C was placed furthest out from the island at 23 kilometers from Oʻahu to simulate oligotrophic conditions. Station B was located between the other stations 7 kilometers from shore to try to capture a gradient in biogeochemical and microbial properties.

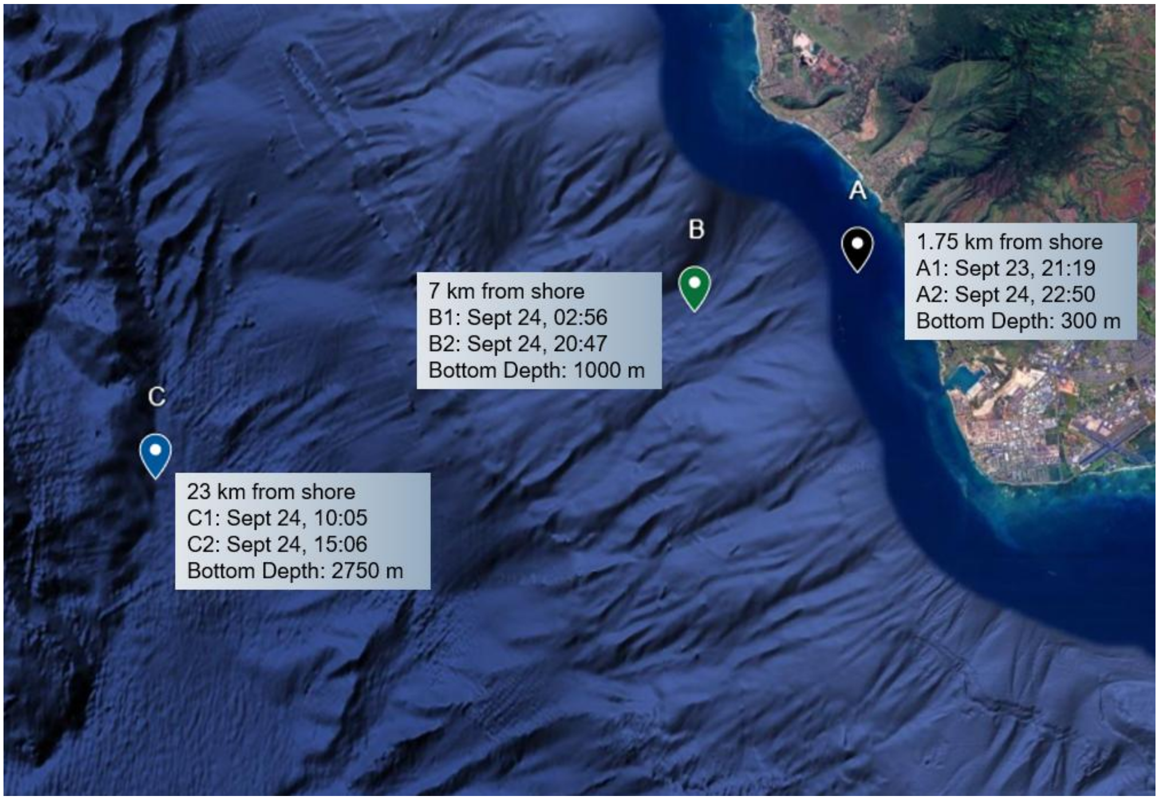


Figure 1. A map showing the locations of the stations and the Wai‘anae coast visible in the upper right corner. Each station is annotated with the distance from shore, the cast times, and the bottom depth.

Table 1. A list detailing the location, starting times of each cast in the local HST time zone, and the approximate depth of water at the location of each cast. The table is organized in chronological order of when the casts were started.

Cast	Location	Cast Starting Time and Date (HST)
A1	21°21'16"N 158°08'53"W	September 23, 21:19
B1	21°20'37"N 158°11'51"W	September 24, 02:56
C1	21°18'04"N 158°21'55"W	September 24, 10:05
C2	21°17'48"N 158°21'38"W	September 24, 15:06
B2	21°20'36"N 158°11'56"W	September 24, 20:47
A2	21°21'22"N 158°08'53"W	September 24, 22:50

Water samples were collected using a 24 position Niskin bottle rosette array with sensor package (including Conductivity-Temperature-Depth, henceforth CTD) to a depth of 250 meters. On the upcast of the CTD, a pair of niskin bottles were fired at each selected depth to capture the seawater. These depths were 5, 25, 45, 75, 100, 125, 150, 175, and 250 meters and were based on the sample depths used by the Hawaii Ocean Time-Series program (Karl et al., 2001). An additional pair of bottles were fired at the deep chlorophyll maximum (DCM) of each cast as determined by the chlorophyll peak captured from the fluorometer attached to the CTD array on the downcast.

2.2 SAMPLE COLLECTION

Water was collected from the Niskin bottles for samples for dissolved organic carbon (DOC), fluorescent DOM (fDOM), flow cytometry (FCM), Chlorophyll, 16S sequencing, particulate organic carbon (POC), and macronutrient analysis. From each Niskin Bottle we collected whole water in acid-washed and sample-rinsed 1L polycarbonate bottles. First, we collected FCM samples by storing 1 mL of whole sample seawater with 16 μM of 32% paraformaldehyde (0.5% final concentration) in 2 mL plastic tubes, then storing the tubes at -80 ºC. Next, samples for 16S amplicon sequencing were generated by passing about 900 mL of seawater over a 0.2 um sterivex filter. Any remaining water was pushed by air with a syringe to dry the filter before storing the sterivex capsule in a -80 ºC freezer. Water that had been passed through the sterivex filter was collected and separated to be used as samples for nutrient, fDOM and solid phase extraction mass spectrometry analyses. fDOM samples consisted of 20 mL of filtrate placed into a 30 mL glass vial and stored at 4 ºC. We placed 45 mL of filtrate for nutrient samples into 50ml polypropylene centrifuge tubes before being stored at -20 ºC. We acidified 500 mL of sterivex filtrate in an acid-washed polycarbonate bottle, then passed the acidified seawater over a PPL solid phase extraction resin. The PPL resin was stored at -80 degrees Celsius, and the filtrate was stored at 4 degrees Celsius. In addition, 4L of seawater was passed directly from the Niskin bottles over a combusted 47mm GF/F filter before storing the filter at -80 ºC for the POC sample. Samples for chlorophyll a were collected by passing around 150 mL of seawater over a nitrocellulose filter, then the filters were wrapped in foil and placed into storage at -80 ºC. We stored 25 mL of the POC filtrate in combusted glass vials acidified to pH2 using 50 µL HCl and stored it in a refrigerator at 4 ºC for use in DOC analysis. After the cruise, samples were transferred from the ship storage to appropriate cryogenic storage devices on land and remained in storage until the samples could be run through the appropriate analysis procedures later in 2024.

2.3 DATA COLLECTION FROM SAMPLES

Cell Concentrations - FCM samples were examined using a Beckman-Coulter CytoFLEX S flow cytometer in the Biological Electron Microscope Facility. After thawing the samples to room temperature, they were aliquotted into polystyrene 96-well U-bottom plates. An undiluted 275 uL of each sample was mixed with 7 uL of Hoescht 34580 dye (40ug/uL) resulting in a final dye concentration of 1 ug/mL and allowed to stain in the dark for about 30 minutes before being run through the flow cytometer (Selph, 2021). To analyze each sample, the cytometer drew liquid from each well at a calculated rate of 60 uL per minute for 75 seconds. Events from the first 15 seconds (15 uL) of each run were excluded from analysis to avoid any cross contamination between each well; the events analyzed were from 60uL of sample. A hierarchical gating scheme separated the events picked up by the cytometer according to size, DNA stain, and pigmentassociated autofluorescence. Events were first gated by “high” or “low” chlorophyll content, then background events were separated out from each group by comparing the distribution of events in the samples to events in both stained and unstained distilled water blanks (Figure 2b). Remaining “low chlorophyll” events were classified as heterotrophic bacteria (Figure 2c). ‘High chlorophyll” events were divided by forward scatter, with larger events classified as “picoeukaryotes,” and smaller events were classified as picocyanobacteria and further divided by phycoerythrin (PE) content: “High” PE events were called Synechococcus, and “low” PE events were called Prochlorococcus. Event counts from each gate were standardized to volume processed (60uL) to calculate cell densities as gated events per uL.

Community Composition - The 0.2um polyethersulfone filters for 16S sequencing were removed from the sterivex casing and placed into bead tubes (MPBio Lysing Matrix D) to be processed at the UH Mānoa Microbial Genomics and Analytical Laboratory (MGAL). DNA extraction was run on a Machery-Nagel NucleoMag microbiome kit on a ThermoFisher KingFisher Flex system. The V4 region of the small subunit (16S) rRNA gene was amplified from each genomic DNA template using forward-indexed primers 515F-Parada: GTGYCAGCMGCCGCGGTAA and 806R-Apprill: GACTACNVGGGTWTCTAAT (Walters et al., 2016) following the protocols of the Earth Microbiome Project (Greg Caporaso et al., 2018), with the following minor changes: we did not pool triplicate reactions and 1 µL was used of genomic template. Amplicons were cleaned and normalized to equimolar concentrations using the SequelPrep kit (Thermo Fisher Scientific) and then pooled and sequenced on the Illumina MiSeq platform (600 cycle V3 chemistry). Sequences were demultiplexed using a custom probabilistic script, and microbiome profiling was implemented in the MetaFlow|mics pipeline (Arisdakessian et al.,

2020) following the specific settings presented by Jani et al (2021). Sequences were assembled, denoised, and quality trimmed using DADA2, globally aligned with mothur to the Silva (v132) global SSU rRNA alignment database, bayesian consensus classified (70%) to the Genus level, and clustered into 99% sequence identity operational taxonomic units using vsearch refined to reduce intragenomic taxonomic splitting errors using LULU (Jani et al., 2021; Callahan et al., 2016; Schloss et al., 2009; Yilmaz et al., 2014; Wang et al., 2007; Rognes et al., 2016; Frøslev et al., 2017). Sequences that could not be classified at the Domain level or were classified as chloroplast or mitochondrial 16S were discarded, and sequencing depth was subsequently standardized to 10,000 random reads per sample. All bioinformatics were deployed on the C-MAIKI gateway at the University of Hawaiʻi at Mānoa by the Center for Microbial Oceanography: Research and Education (Cleveland et al., 2022). After the data was through the pipeline, we filtered the data to remove OTUs that were never greater than 0.1% abundance in at least one sample.

Organic matter composition - Water samples to be analyzed for fluorescent dissolved organic matter (fDOM) were removed from the refrigerator the night before processing to allow them to acclimate to room temperature. 3 ml subsamples were transferred into 1 cm quartz cuvettes, which were then loaded into the fluorometer. Prior to running samples, cuvettes were washed thoroughly with DIW and wiped clear with Kim wipes. The cuvettes were triple rinsed with DIW between each sample. After every 7 samples, a DIW blank was run. Fluorescence was measured using a Horiba Aqualog scanning fluorometer. Data processing methods and script followed the methods of Nelson et al. (2015). Excitation-Emission Matrices (EEMS) were generated with excitation every 5nm from 240 nm to 500 nm and emission values measured continuously via

CCD ranging from 250 nm to 825 nm integrated over 4 s at each excitation increment. Inner filter correction was applied to account for fluorescence quenching by absorbance (Kothawala et al., 2013). Fluorescence values were scaled to Raman Units by dividing by the integrated emission range of 381 nm to 426 nm at an excitation of 350 nm in averaged DIW blanks (Lawaetz and Stedmon 2009). An average blank EEM - calculated from DIW blanks run immediately before and after each set of samples - was subtracted from each set of 7 sample EEMs. A set of informative biogeochemical indices were calculated comprising M:C (Burdigee et al. 2004), FI (McKnight et al. 2001), BIX (Huguet et al. 2009), HIX (Zsolnay et al. 1999), M, A, C, T, and B (Coble 1996).

Chlorophyll samples were run at the UH Mānoa SOEST Laboratory for Analytical Biogeochemistry (SLAB) using a Turner Designs 10AU fluorometer to measure the fluorescence of chlorophyll α of each sample. The nitrocellulose filters were removed from the freezer, and each filter was dissolved in 5 mL of 90% acetone dilution in a borosilicate vial and vortexed to dislodge cells from the filter. The vials were then frozen overnight before being analyzed the next day. The acetone mixtures were loaded into the fluorometer in small borosilicate vials. Fluorescence was measured at 680 nm. Absorbances were converted to chlorophyll a concentrations using SLAB's calibration curve. The resulting absorbance values were recorded for each sample.

2.4 STATISTICAL ANALYSIS OF DATA

Data captured by the CTD array during casts were cleaned in R to include only data from the descent portion of the cast but exclude the initial system check at the beginning. Since there were often more than 100,000 data points per cast, the data was averaged across slices of depth 0.5 m deep through the entire downcast. Depth profile and bivariate scatter plots were created with Python in Jupyter Notebook utilizing the PyPlot package. For data from discrete samples, plots featured the points and the mean of each station (n=2) and whiskers to indicate the range of the data. FCM cell count data was also plotted using similar depth profiles and compared to the depth profiles from sequencing data for the groups of *Prochlorococcus*, *Synechococcus*, picoeukaryotes, and heterotrophic bacteria. To investigate the dissimilarities between the community composition at each sample, a unifraq distance matrix was built from the 16S community data and used to create a multivariate ordination. The ordination was plotted in R and points were colored by depth and symbol-coded by station. A heatmap was created using the scaled relative abundances of microbial families. I used the dissimilarities between individual taxa across the samples to show how the taxa changed between sample distance from shore and depth. Other bottle samples for fDOM, DOC, and chlorophyll *a* were plotted as depth profiles with the same method as community data. Sequencing and FCM results comparisons were made by summing the total relative abundances of OTUs classified in the genera “Prochlorococcus” and “Synechococcus” for each plot. Relative abundances from FCM data are calculated by using the total cell count per uL of each group, then divided by the sum of the Prochlorococcus, Synechococcus, and heterotrophic bacteria counts per uL for a relative abundance from FCM cell counts.

3.0 RESULTS

3.1 ENVIRONMENTAL CONDITIONS

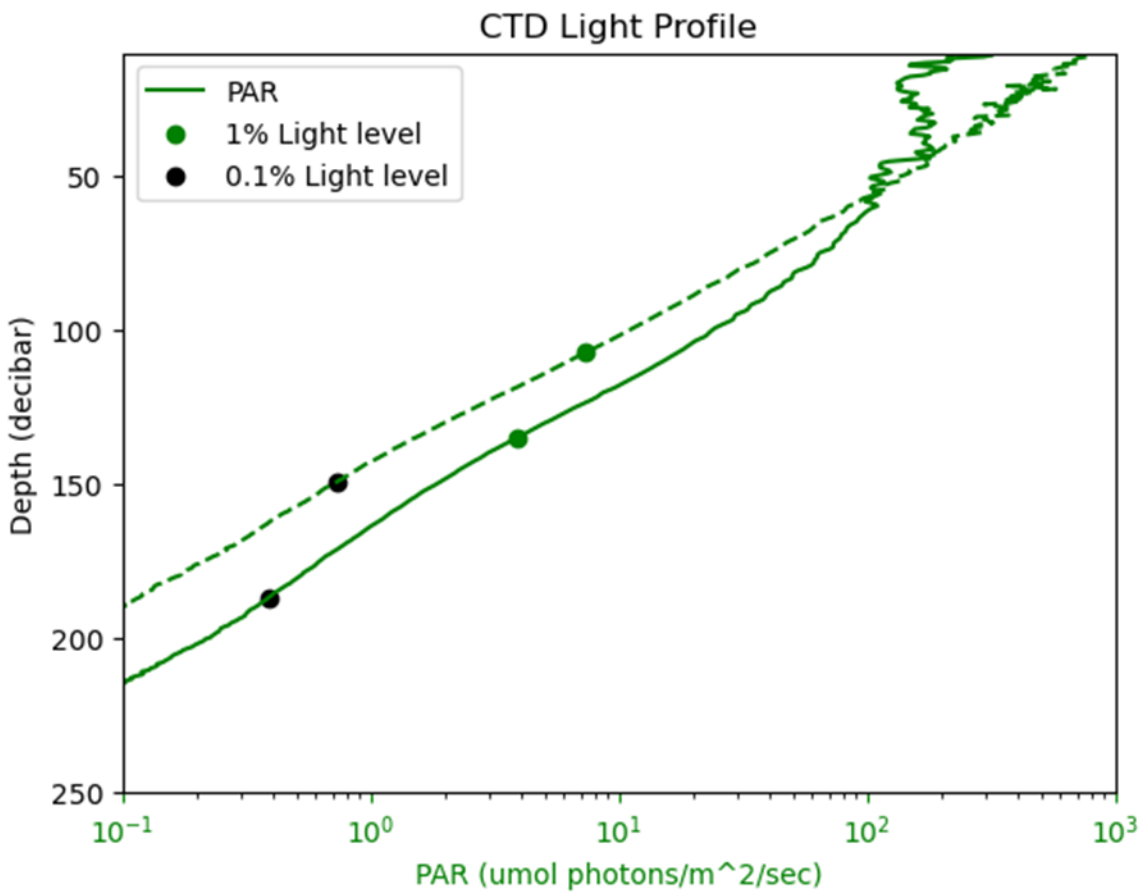


Figure 2: Daytime light levels on a log scale at Station C with points showing the depth where the light is 1% of the surface light level in the top 15 meters. Morning Cast C1 (10:00) is plotted with the solid line and Afternoon Cast C2 (15:00) is plotted with the dashed line. Stations A and B were sampled at night and PAR data are not presented.

To test our primary hypothesis that biogeochemical and microbial depth profiles vary predictably with distance from landforms, the primary independent variable used in our analyses was distance from shore using sampling stations along a transect running perpendicular the West Coast of O‘ahu (Figure 1). Light levels in the water at the stations decreased logarithmically with depth: At Station C the depth of the 1% surface irradiance ranged between 110 and 140 m, and the depth of the 0.1% light level ranged from 150 to 190 m (Figure 2); the attenuation coefficient k was calculated as 0.0461 m^-1.

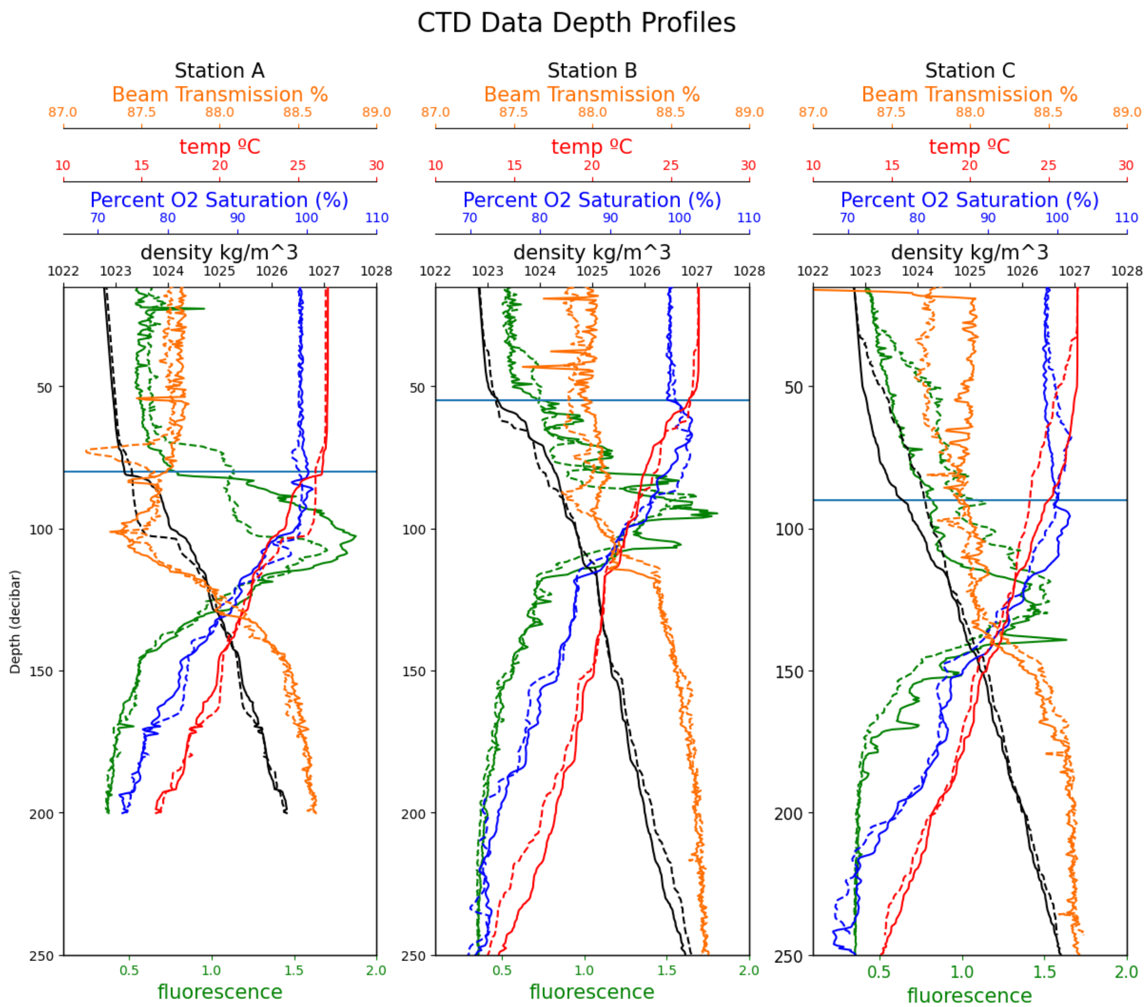


Figure 3: Depth profiles of environmental conditions captured from the CTD on each downcast. Lines for chlorophyll fluorescence (green), seawater density (black), seawater temperature (red), oxygen saturation (blue), and turbidity (orange) are plotted for each cast and separated by plots for each station. The approximate depth of the mixed layer is marked with a horizontal blue line. The first cast at each station is plotted with a solid line, the second cast is plotted with a dashed line.

The depth profiles of environmental conditions at each of the stations were closely replicated between the two casts, most notably temperature, density, and oxygen saturation (Fig. 3). The upper mixed layer depth was determined as the point of inflection of the rate of density increase with depth at each station. At station C, there is no clear inflection in the rate of density increase with depth, so the mixed layer depth was determined using the plot of oxygen saturation and marking the depth where dissolved oxygen had accumulated at the bottom of the mixed layer.

The deep chlorophyll maximum of each station is considered as the depth of the peak fluorescence. This peak occurs near 100 meters depth at station A, 90 meters depth at station B, and 125 meters depth at station C. The strength of the peak fluorescence also varies between stations, where it is strongest at station A and decreases with distance from shore. The mixed layer depths follow a similar trend to the DCM, where station B is the shallowest and station C is the deepest. Turbidity, which is represented as the inverse of beam transmission, decreases in all of the stations below the DCM: waters in the mixed layer and DCM are more turbid and have more particles disrupting the light passing through them (decreasing measured beam transmission) than in deeper waters.

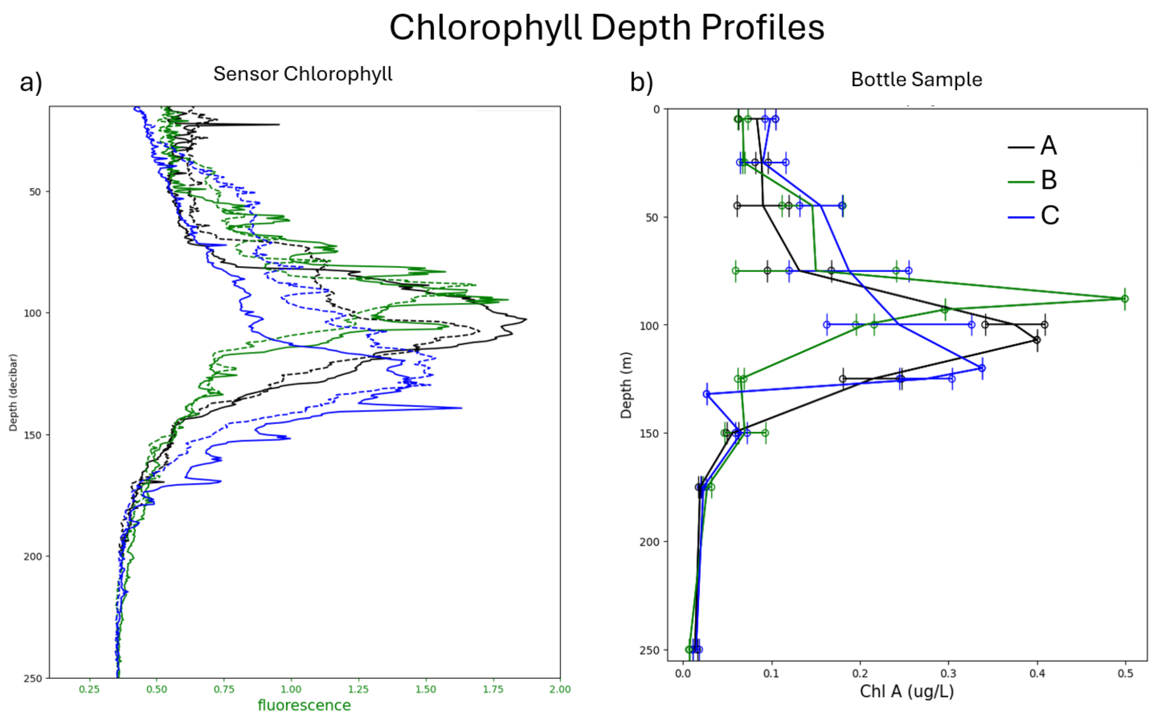


Figure 4: Depth profiles of chlorophyll from the (a) CTD sensor array and(b) bottle samples analyzed in a lab after the cruise. The profiles from each stations are separated by color. Lines for the individual casts from the CTD sensors are plotted with a solid line for the first cast and a dashed line for the second cast. Bottle data is plotted with whiskers to show the range of the casts from each station. Note that the x-axis for each plot are different units.

The peak of the deep chlorophyll maximum weakens and deepens at the offshore station C (120-140 m) compared to the two stations closer to shore (80-110 m). This shift in the DCM is reflected in both methods used to measure the fluorescence. The sensor data shows the strongest peak at station A at 100 m depth, but the bottle data shows that the strongest peak was at station B around 85 m depth, although only present at a single cast, and it is not much stronger than the peak at station A. The trend in peak fluorescence strength shifts from its strongest at station A and weakening further out from shore at stations B and C. There was an increase of 14% between the fluorescence maxima measured by the CTD sensor array (4a) at station A and station C, and a 21% increase in the measured chlorophyll *a* concentration from bottle samples (4b) at the same stations.

3.2 COMMUNITY DYNAMICS

3.2.1 FLOW CYTOMETRY ANALYSIS

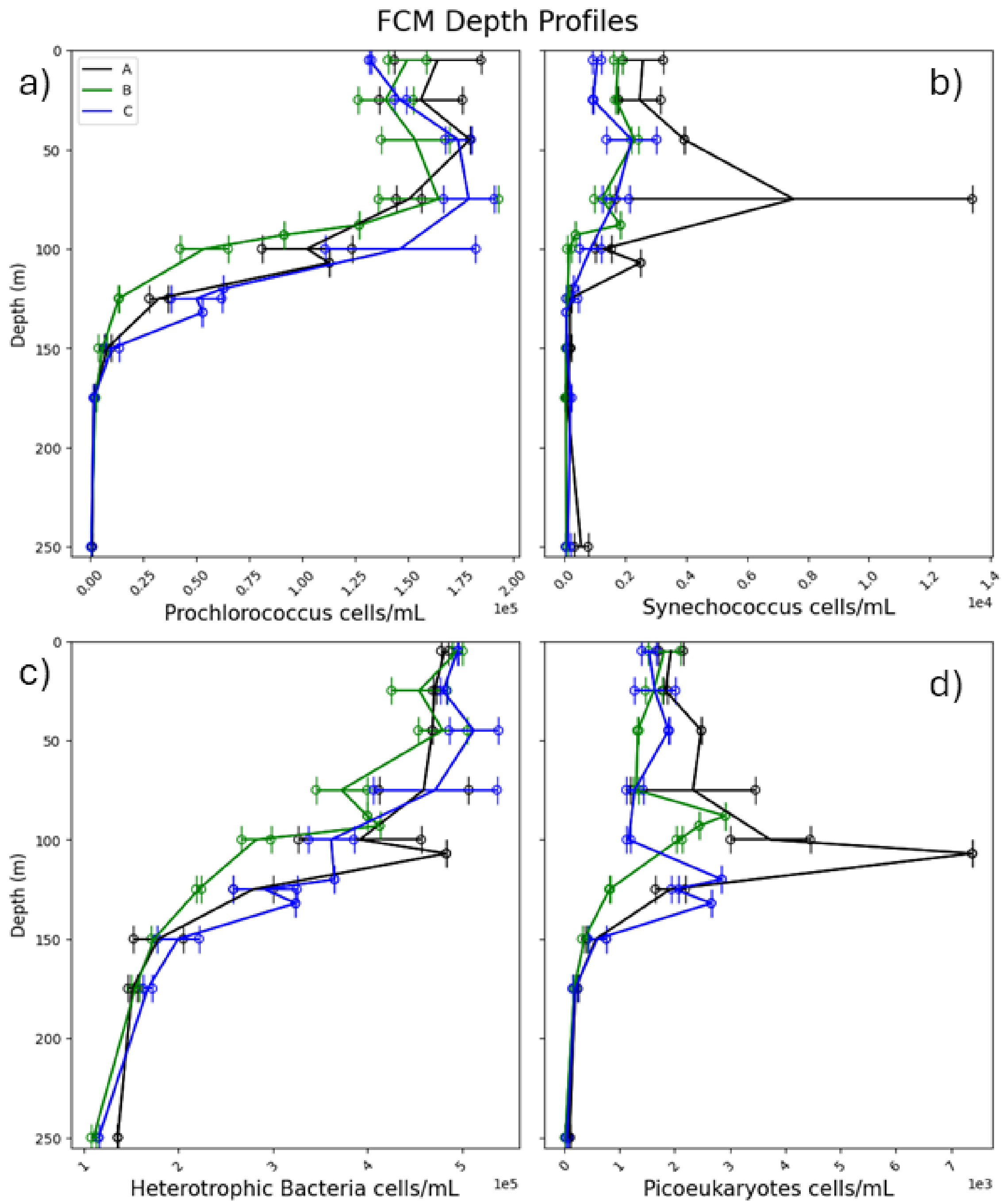


Figure 5: Microbial taxa cell counts from FCM analysis for *Prochlorococcus, Synechococcus*, Heterotrophic Bacteria, and Picoeukaryotes. Cell counts are in scientific notation and are different for each group of microbes. The line shows the mean of each station, with whiskers and open circles showing the range in cell counts of the two individual casts.

FCM analysis of cell counts illustrated differences between the stations (Fig. 4). Comparing the groups, the smaller photosynthetic microbes *Prochlorococcus* and *Synechococcus* are most abundant in the mixed layer and DCM and decrease to almost zero in the deeper waters. The concentration of *Prochlorococcus* is one order of magnitude higher than *Synechococcus*. Picoeukaryotes are most abundant in the DCM, although they are still present in 2-3 fold lower concentrations within the mixed layer. Heterotrophic bacteria are generally present at a similar concentration throughout the mixed layer but begin to decrease in the DCM and deeper, but do not completely reach zero like the photosynthetic groups do at 250 meters.

Comparing across stations within each group, picoeukaryotes show a large difference between their concentrations at station A and the other stations (Fig. 4). They increase greatly in the DCM and show a peak in concentration at station A over twice as much as the peaks at stations B and C. They also appear more abundant in the lower mixed layer at station A than at the other stations. *Synechococcus* shows a 59% decrease in the average cell concentration at 5 m from station A to station C, and an 88% decrease at the DCM between the two stations. There is an outlier at station A at 75 meters in depth (Fig. 4; also confirmed with genomic analyses in Figure 5 below) but when ignoring that point, there is little difference in concentration between stations at the DCM. The lines showing *Prochlorococcus* have a lot of overlap, but station C decreases in the surface waters while stations A and B appear to maintain the same concentration throughout the entire mixed layer. Moving deeper from the mixed layer, station B has a much greater decrease at the DCM than the other stations. The concentrations of

Prochlorococcus in the DCM at station A and C are comparable, but the average at station C is slightly greater.

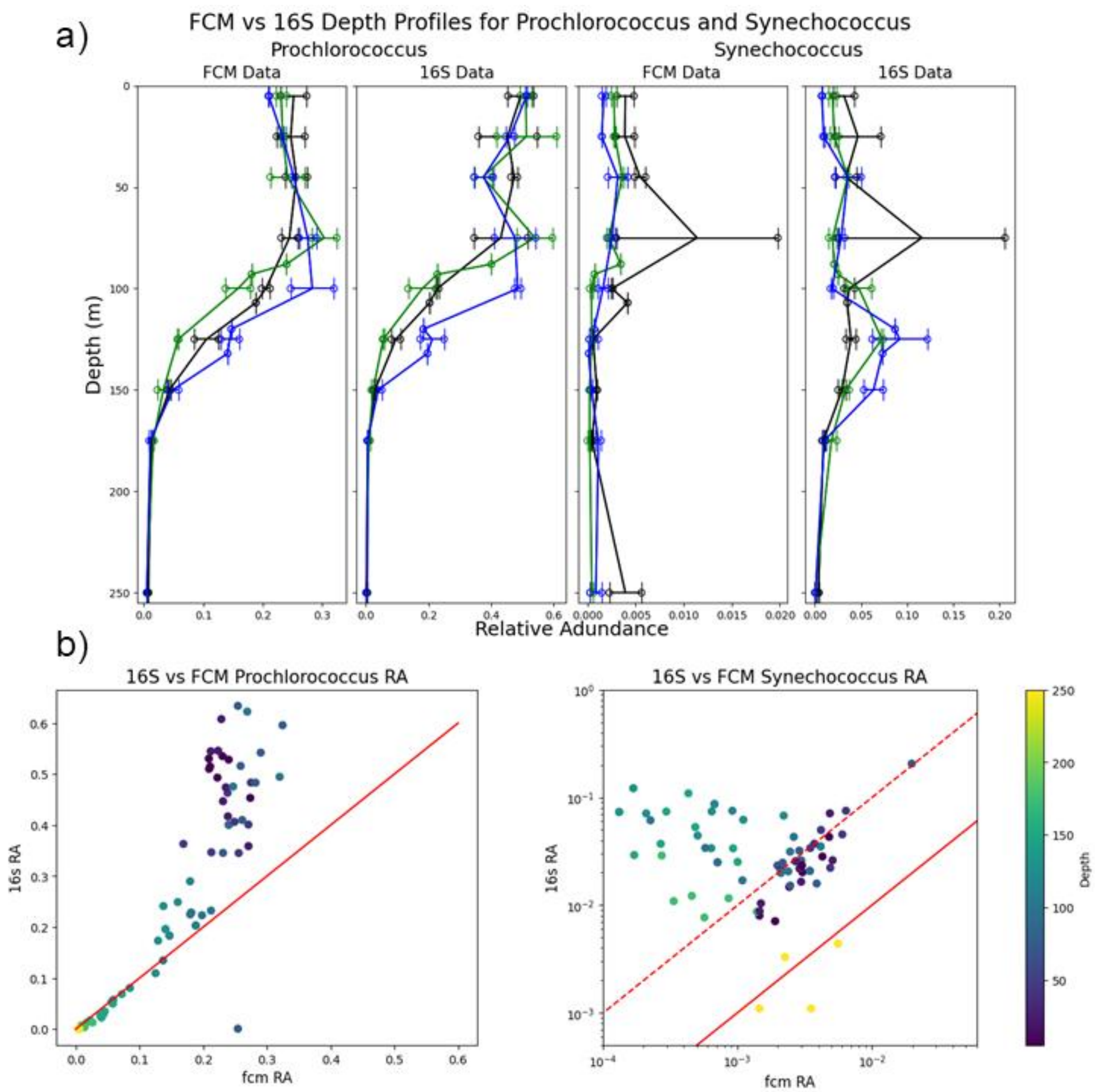


Figure 6: Comparing relative abundances for *Prochlorococcus* and *Synechococcus* using 16S sequencing data and FCM data. (a) compares depth profiles for both methods separately side-by-side to show similar patterns in abundances. (b) compares relative abundances from FCM on the x-axis to sequencing on the y-axis across all depths and stations for *Prochlorococcus* and *Synechococcus*. *Synechococcus* is log scale on both axes. A solid red 1:1 relative abundance line is plotted on both graphs, and a dashed 1:10 line is plotted for *Synechococcus*.

The depth profiles of relative abundances of *Prochlorococcus* and *Synechococcus* from both direct concentrations (FCM counts) and genomic analyses (16S amplicons) show similar vertical and horizontal distributional patterns (Figure 5a-b). However, the two taxa differ in calculations of relative abundance between the two methods, with Prochlorococcus methods largely in agreement and genomic analyses appearing to overrepresent Synechococcus by a factor of nearly 10 fold (Figure 5c-d). Relative abundances of *Prochlorococcus* from 16S and FCM data are in close agreement with each other until it reaches about 25 percent of the community, at which point the sequencing data begins to report a higher abundance than FCM (Figure 5c). Using 16S amplicons *Synechococcus* relative abundance in the surface waters is roughly 10 times higher than direct counts by flow cytometry. In the DCM this enrichment increases to nearly 100-fold, while in the deeper waters the methods become more congruent and closer to unity (Figure 5d). As noted above, the one outlier point of high abundance of *Synechococcus* at 75 meters depth is present in both the sequencing and FCM depth plots, confirming a localized enrichment of this taxon at 75 m on the second cast at Station A.

3.2.2 16S SEQUENCING DATA ANALYSIS

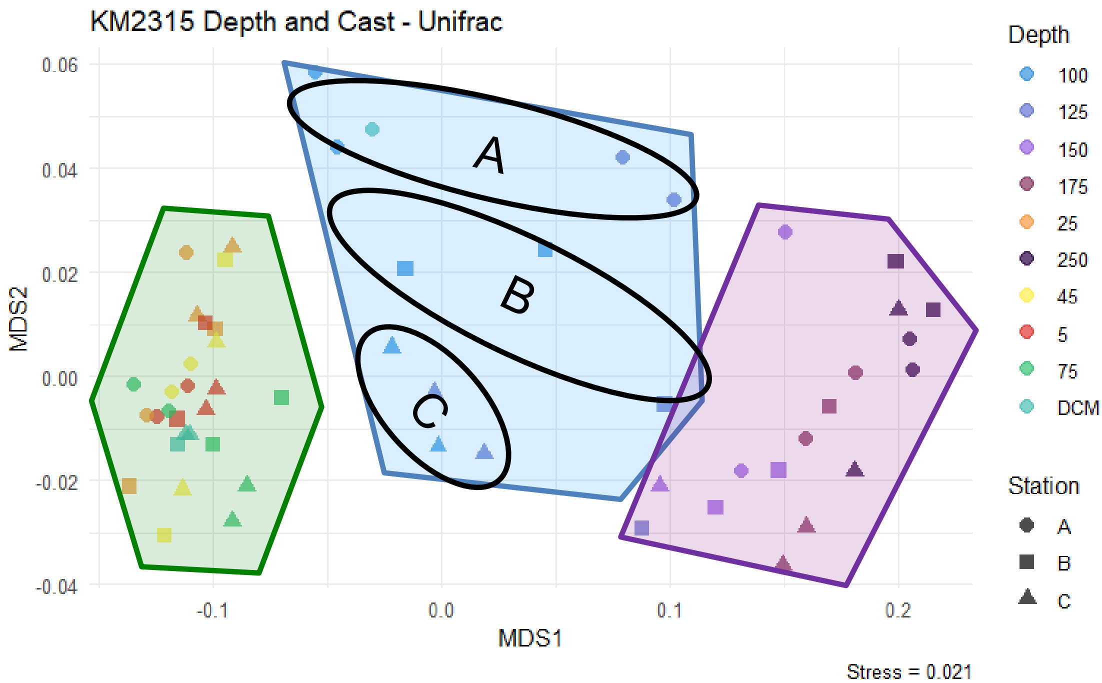


Figure 7: An ordination plot of all the samples marked with color by depth and with shape by station. Three depth groups are roughly shown with colored polygons. The upper mixed layer (5-75 m samples) is in green, the DCM group (100m, 125m, and DCM samples) is blue with the stations highlighted with black circles and labeled with the station, and the mesopelagic (150-250 m samples) is purple.

The ordination has a visible grouping of samples by depth along the x axis. Samples from the upper mixed layer from 5 meters to 75 meters are clustered together on the left side of the plot, As the points shift to the right, they also increase in depth, showing a strong correlation of microbial diversity with depth along the x axis.

Samples also shift with station along the y axis, with samples from station A appearing towards the top of the plot and samples from station C appearing towards the bottom of the plot. When exploring the taxonomic families in these samples with a heatmap (Figure 7), I observed shifts across stations in the same depth in multiple families including *Actinomarinaceae*, DEV007, and the OCS116 clade. The biggest distance between samples at similar depths is within the DCM depth group (blue), which includes the DCM samples as well as the 100- and 125-meter depth samples, in which the samples separate out by station along the y-axis. The samples in the upper mixed layer (green) cluster together very closely and do not show a shift in the depth or station parameter along either axis. In the mesopelagic group (purple) the samples are more widespread than in the UML and appear to shift in depth, with shallower samples to the left side and deeper samples more towards the right of the group.

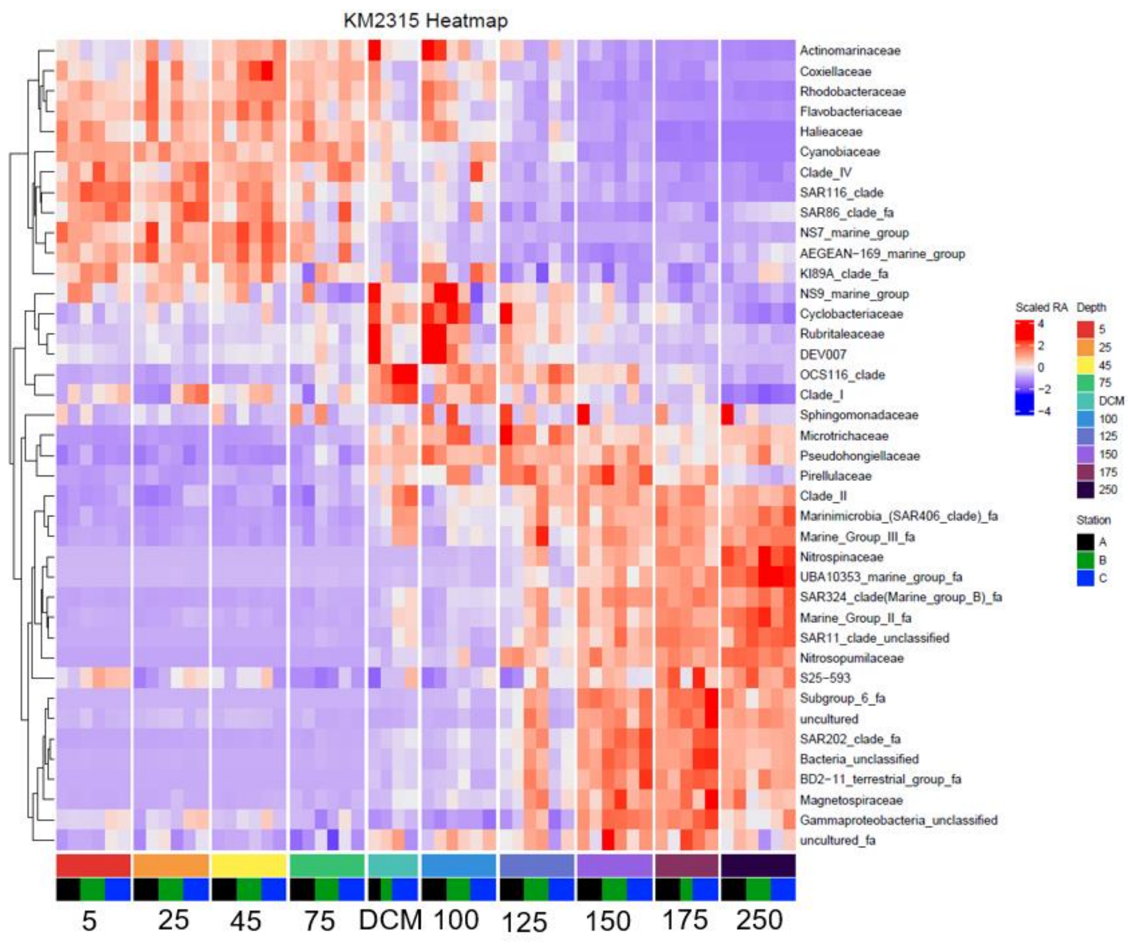


Figure 8: A hierarchical heatmap based on the microbial community composition of each sample. The heatmap shows the scaled mean relative abundance of the 40 most abundant families across all the samples. The samples are clustered into column groupings based on the depth of each sample, ranging from 5 to 250 meters, with the

DCM samples in their own group. Within each of the column groupings the stations are arranged A, B, and C from left to right.

The heatmap is grouped by the depth of each sample to emphasize the shift in the abundance of a family between stations at each depth horizon. Each depth horizon falls into a grouping based on the grouping of samples in the ordination (Figure 6), where different families are abundant within these groups. As highlighted in the ordination, these groups are the upper mixed layer (5 to 75 meters depths), the DCM region (100 and 125 meter depths), and the mesopelagic (150 to 250 meter depths), with samples from the DCM belonging to the DCM region grouping. Focusing on the DCM region which showed the most dissimilarity between stations, I observed some families that exhibited a notable shift in relative abundance along nearshore to offshore gradients in the 100 and 125 meter depth horizons, including the families *Actinomarinaceae*, DEV007, and the OCS116 clade. The families *Actinomarinaceae* and DEV007 are more abundant in the nearshore at station A than at stations B or C. The OCS116 clade exhibits an opposite pattern and was present in higher abundance at station C in the offshore environment than at the stations closer to shore.

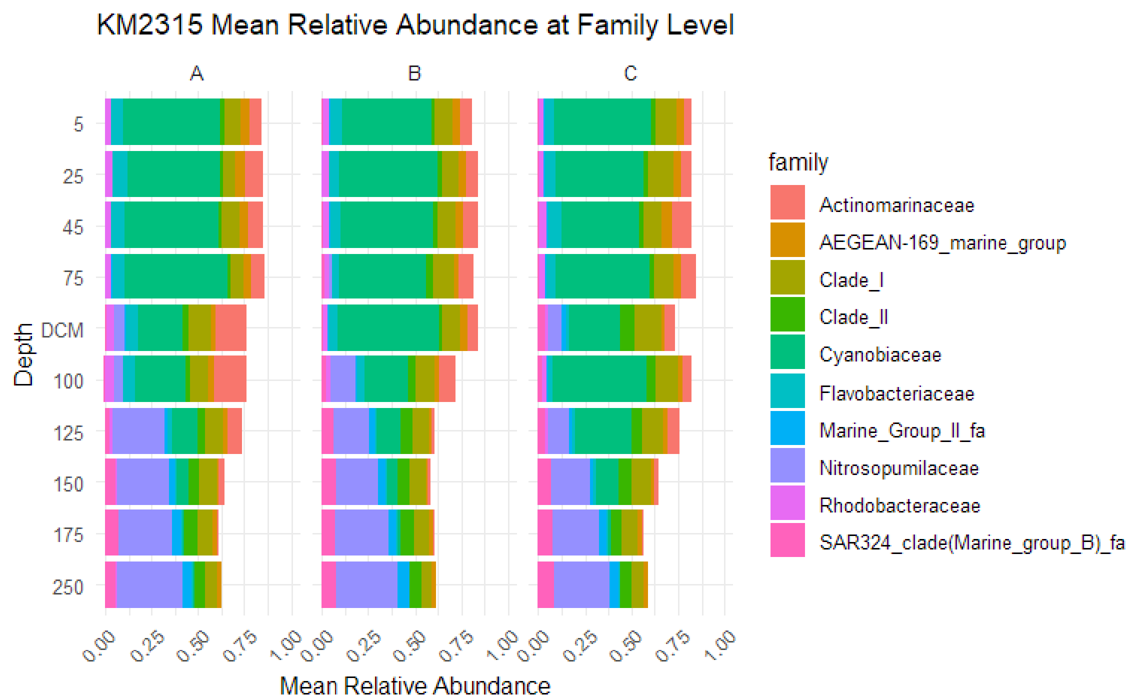


Figure 9: A plot showing the mean relative abundance of different bacterial families across each depth at each station.

Samples separate into groups based on depth, with three main groups appearing from the multivariate analysis. These groups are the mixed layer above the DCM (5 to 75 meters depths), samples near the DCM (the DCM, 100-, and 125-meters depths), and samples below the DCM (150-250 meters). Across stations, the community shows the most differences when comparing abundances within the DCM region (100 and 125 meter depths). The mixed layer and the mesopelagic groups are similar across the three stations. At stations B and C, the DCM is dominated by *Cyanobiaceae*, resembling the mixed layer group community, but at station A the DCM is more similar to the community at 100 meters depth. The groups *Actinomarinaceae* and Flavobacteriaceae as discussed in the heatmap (Figure 7) shows a decrease in abundance in the 100 meter depth horizon from station A to station C.

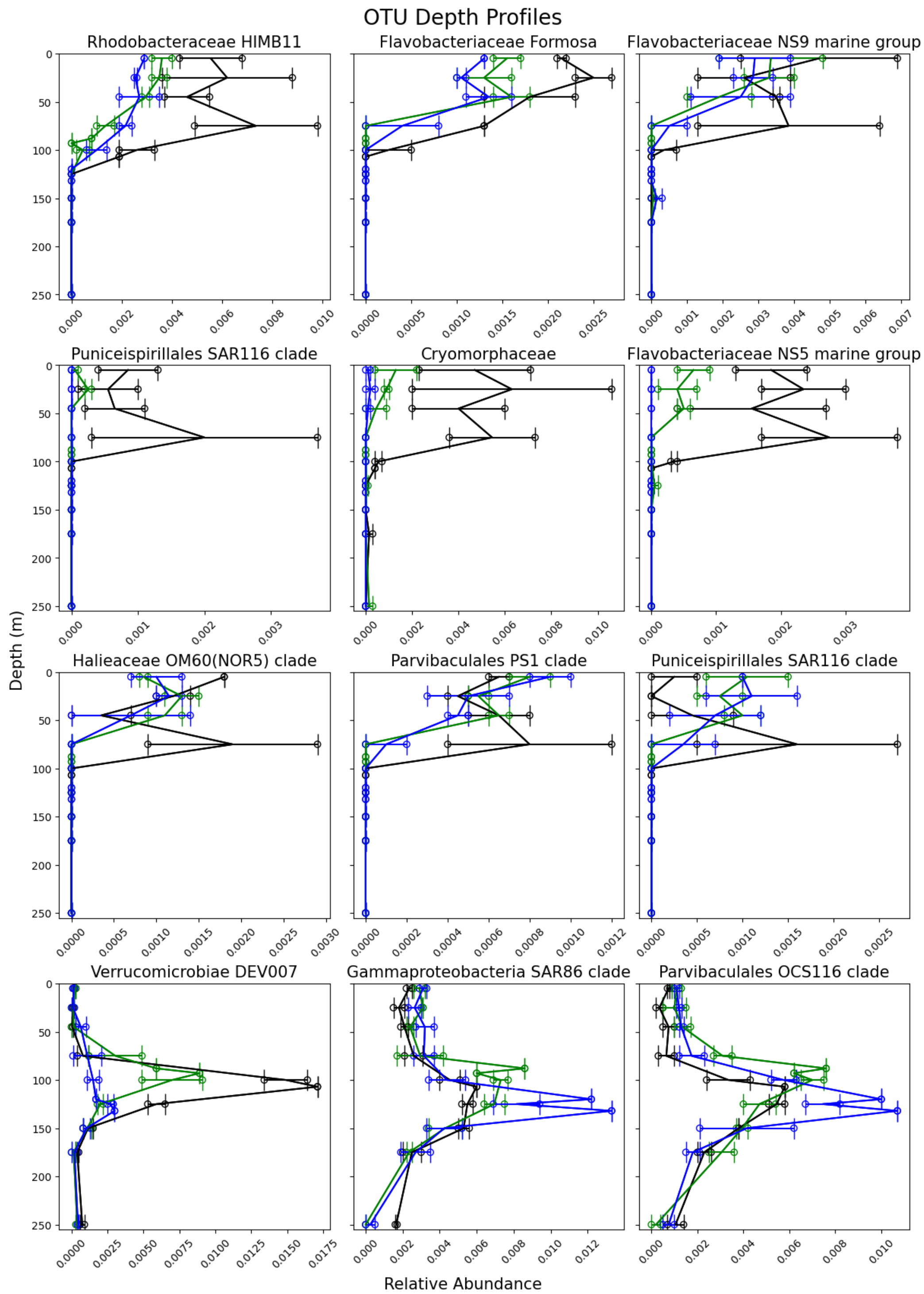


Figure 10: Relative abundances of 12 selected OTUs with notable differences or patterns across the three study stations. The OTUs are generally grouped by row by similar patterns. The depth profiles show the mean for each station with error bars and open circles showing the individual points from each cast.

These plots show a subset of abundant heterotrophic Bacterial taxa that were statistically differentiated among stations at selected depths. Row one includes OTUs that are present in the upper mixed layer at all the stations but are slightly more abundant in the nearshore samples. This includes *Roseobacter* HIMB11, Formosa of the family *Flavobacteriaceae*, and the NS9 marine group of Flavobacteriales. Row two includes OTUs that are mostly present only at station A and are not found or found at an insignificant abundance at station C. These OTUs are SAR116, the family *Cryomorphaceae*, and the NS5 marine group. Row three consist of OTUs that have a peak in abundance within the DCM at station A. These OTUs are the OM60(NOR5) clade within class *Gammaproteobacteria*, and OTUs belonging to the PS1 clade and SAR116 clade within the class *Alphaproteobacteria*. Row four includes OTUs that have a peak in their abundance at the DCM. For the group DEV007 this peak is greatest at station A and decreases across stations to become more subdued at station C. Groups SAR86 and *Parvibaculales* OCS116 show this pattern but in the reverse direction from shore, that their greatest abundance occurs at the DCM at station C, and decreases as the stations get closer to shore.

3.2 BIOCHEMICAL TRACERS

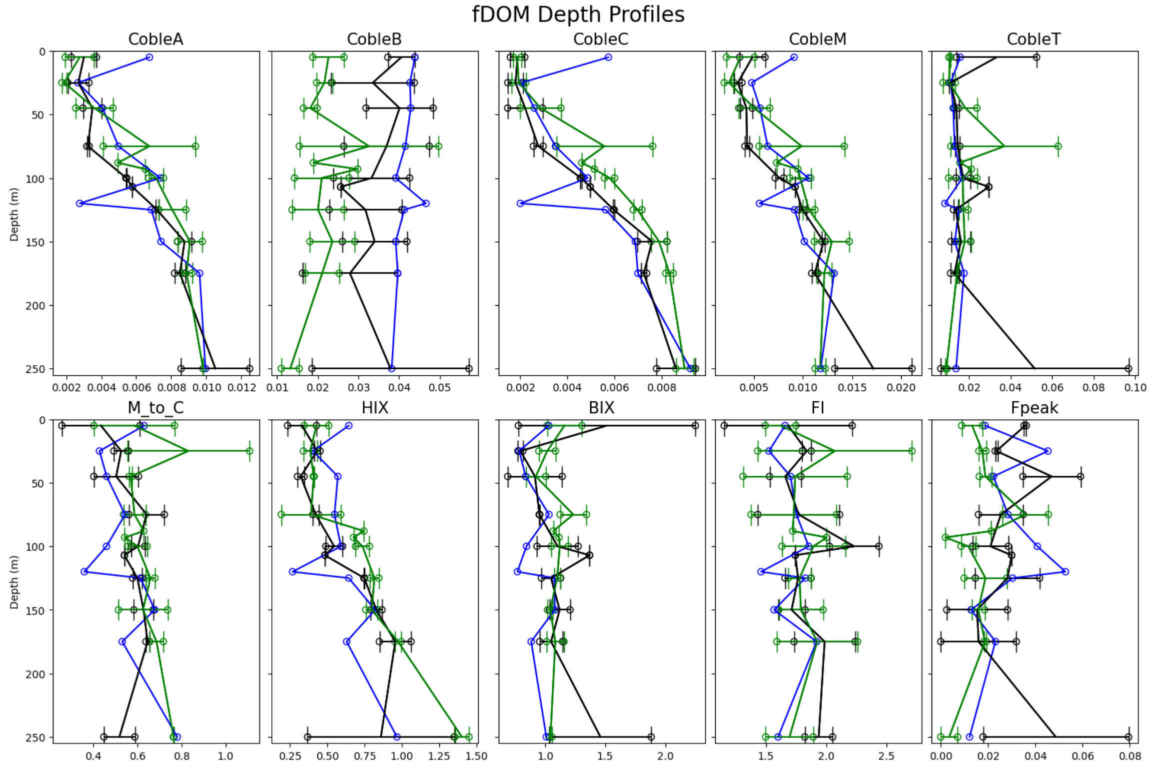


Figure 11. fDOM indices. The top row of plots contains raw raman units of fluorescence of individual peaks extracted from excitation/emission indices (Coble 1996; Coble 2007; Yamashita 2003), and the bottom row contains indices that represent various ratios of different humic types (Coble 1996; Huguet 2009; McKnight 2001; Zsolnay et al. 1999).

The fluorescence composition of the DOM pool did not vary appreciably across stations, but there are visible patterns with depth that occur. There is an increase in the levels of humic-like indices with depth (Coble A, C, M) but the protein indices (Coble B, T) show a consistent level across all depths. The ratio index HIX represents humification, which increases with depth similar to the humic-like indices. The other indices are consistent across depths.

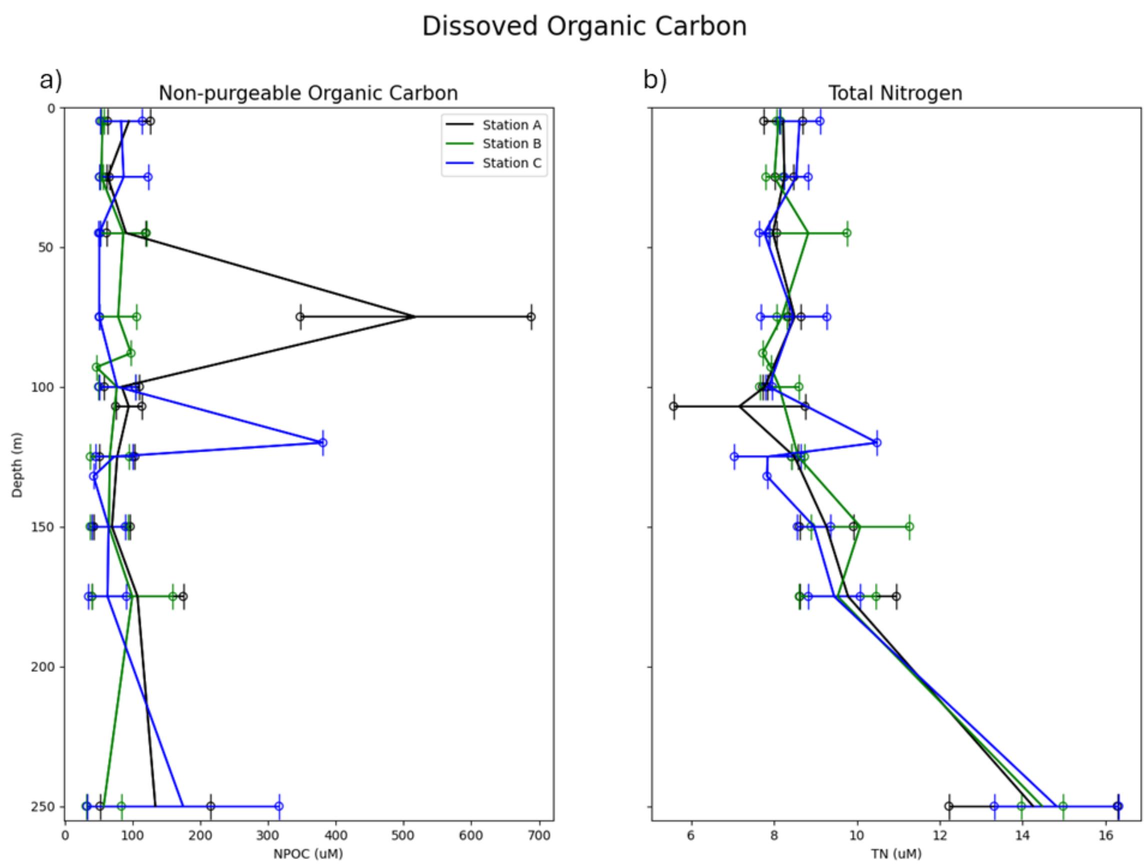


Figure 12: Depth profiles of dissolved organic carbon (a) and nitrogen (b) from bottle samples.

Dissolved organic carbon is mostly constant throughout the water column except for two notable peaks, at station A at 75 meters depth and at station C in the DCM sample at 120 meters depth. Total nitrogen exhibits some small variation between stations in the DCM region (100 to 125 meters depths) but is mostly consistent across stations at all depths.

4.0 DISCUSSION

In this study, I characterized shifts in the microbial community relative to distance from the shore to study the influence of the IME on the community. Community composition was also related to biogeochemical tracers in the environment to understand the potential drivers of shifts. I found that the most notable island mass effects, expressed as horizontal shifts in community structure, occurred in the DCM, where different taxa were enriched or depleted nearshore relative to the offshore waters. This study used methods that other studies used to characterize the IME, but also included sequencing the microbial community at each station to observe shifts in diverse individual taxa that could not be observed with flow cytometry.

Classifying the community highlighted major shifts in multiple taxa in three main depth zones: the upper mixed layer (5 to 75 meters), the DCM region (100 and 125 meters), and the mesopelagic (150 to 250 meters). The distribution of well-studied picophytoplankton *Prochlorococcus* and *Synechococcus* showed changes in the DCM region and upper mixed layer respectively (Fig 5a), which aligned with previous data on the depths of these shifts around Oʻahu (Comfort et al., 2024). *Prochlorococcus* exhibited a decrease in abundance in the nearshore DCM stations A and B compared to station C. Comfort et al found a difference in *Prochlorococcus* abundances between their nearshore station (2km offshore) to Station ALOHA (100 km offshore) at 75 to 125 meters depths, within their range for the DCM. *Synechococcus* showed a shift in the upper mixed layer opposite that of *Prochlorococcus*, increasing in abundance closer to shore in the upper 75 meters. The distribution of picoeukaryotes followed the pattern of *Synechococcus*, greatly increasing in abundance in the UML closer to shore but mostly at 75 meters depth (Fig 4).

Compared to the study by Comfort et al., the magnitude of the shift in abundance between my nearshore and offshore stations was smaller, although the locations and distances for offshore stations were different. My station C was located 23 km off of the Waiʻanae Coast, while Comfort compared their data to Station ALOHA, 100 km off of the North side of Oʻahu. Picoeukaryotes which showed a similar shift as *Synechococcus*, an enrichment at the nearshore station in the mixed layer, which was also observed off of Waikiki (Comfort et al., 2024).

The overall microbial distributions followed the expected trends from previous studies of the IME in the Pacific that quantified phytoplankton abundances (Comfort et al., 2024) and heterotrophic microbial communities (Comstock et al., 2022). These studies attributed the shifts to stratification between the layers and different nutrient and chemical properties at each depth. Comparing the density profiles of each station (Fig 3) there is a strong stratification between the mixed layer and the DCM, especially at stations A and B. Differences in the overall community composition at each depth across stations (Fig 7) showed the most significant differences occurred in the DCM, while the mixed layer was mostly similar down to 75 meters. A study of the microbial community in the IME found homogeneity between samples in the mixed layer (surface to 75 m) and communities stratified by depth in the deeper waters (Comstock et al., 2022). The community in the top 75 meters was distinct from the community at depths below and the high abundance ASVʻs different between each sampled depth (Comstock et al., 2022). The communities in my samples were less distinct between individual sample depths but still separated into three distinct depth groups (mixed layer 5 to 75 meters, DCM region

100 and 125 meters, mesopelagic 150-250 meters). Studying the abundances of individual taxa (Fig 9) results in similar patterns exhibited by multiple taxa, showing shifts within the upper mixed layer and the DCM. These shifts coincide with the distributions of *Prochlorococcus* and *Synechococcus*, suggesting similar environmental conditions driving the shifts in the heterotrophic and autotrophic community.

Measurements of chlorophyll levels, both by filtrate and CTD in-situ methods, captured a decrease in concentration offshore in the DCM, but no shift at the surface (Fig 4). A lack of a visible shift in surface chlorophyll concentrations contrasts with the established method of measuring the IME via satellite chla fluorescence which only captures the upper 30 meters of the water column (Gove et al., 2016; Messié et al., 2022). Messié et al. found that surface chla levels nearest to islands increased by 26% over the oligotrophic waters outside of the influence of the IME, an increase greater than the increase in DCM peaks at station A and station C I observed along my study's transect (14% and 21% differences). Although the other study used a much larger scale (4 km pixel grids) I expected to see a difference between the nearshore and offshore stations at the surface, but I did not observe any notable difference. This suggests that the scale influence of the IME differs at difference depth horizons.

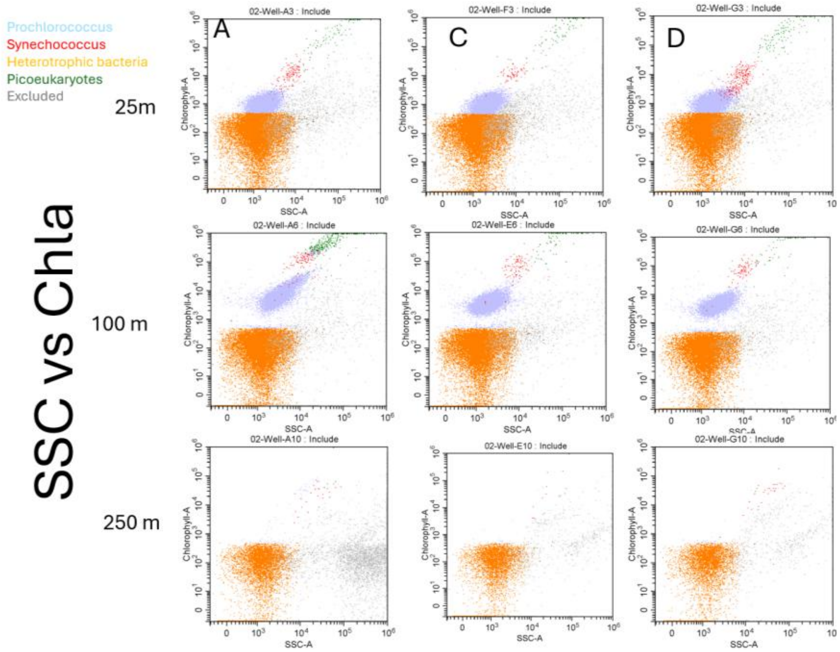
A study in Kane‘ohe Bay on O‘ahu took samples at 2 meters depth and found a rapid decrease of chla concentrations offshore from within the bay to outside of the bay's channels (Tucker et al., 2025). The study also compared the chla levels of the waters around Kane‘ohe Bay to Station ALOHA, and found a 7-fold increase just offshore of the bay over the open ocean. I did not measure a shift anywhere as large, although the stations were not near an area like Kane‘ohe bay which has a large population center and high amounts of runoff. Elevated chla levels can be attributed to nutrients in runoff from human agriculture and development and in the outflow from the bay’s waters (Gove et al., 2016). Kaneohe Bay has a much greater population and more urban development than the Wai‘anae Coastal area, resulting in different impacts on the nutrients in runoff. (Hoover and Mackenzie, 2009). The bay is also located on the rainy windward side of O‘ahu, while Wai‘anae is on the drier leeward side, so there is less rainfall to drive runoff into the nearshore ecosystems. Another study found elevated surface chlorophyll off the southern side of O‘ahu offshore from Waikiki, a densly populated tourist area (Comfort et al., 2024). The closest station to shore (2 km) had significantly higher chlorophyll concentrations than the offshore station (Station ALOHA 100 km) from 5 to 75 meters depths. Constrating with these two studies which occured near more highly developed and wetter areas of O‘ahu, I found no notable shift in surface chlorophyll with distance to shore off the leeward coast. Instead, I found a shift with distance from shore within the DCM from 90 to 125 meters depths, where the DCM shifted with the MLD, which was also found to be the case off of Waikiki (Comfort et al., 2024). At 150 meters depth, I found that chlorophyll concentrations had returned to being similar across all stations

Individual taxa showed shifts in abundance with distance from shore at specific depths, either in the mixed layer (5 to 75 m) or the DCM region (100 to 125 meters). Off the Wai’anae Coast the group *Halieaceae* OM60(NOR5) were enriched in the surface mixed layer, decreased in the lower mixed layer offshore (Fig 9). The *Flavorbacteriaceae* groups *Cryomorphaceae* and NS5 were found almost exclusively at the nearshore station A in the upper mixed layer, and barely detected at station B, while *Flavobacteriaceae* NS9 was abundant in the upper 50 meters at all stations but decreased at B and C at 75 meters while it remained abundant at station A. These groups were categorized as "reef ASV's by Comstock in their study of Moorea, where these ASV's were in high abundance in samples from reefs but not detected at distance from the shore (Comstock et al., 2024). This makes sense since all of these taxa were most abundant at the nearshore station A.

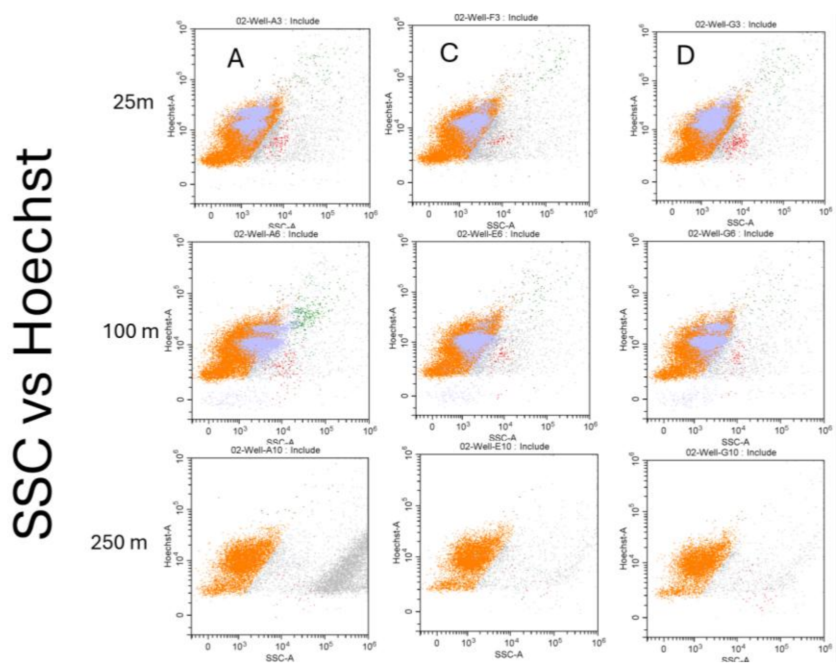
5.0 CONCLUSION

This study analyzed the microbial community composition through FCM and 16S sequencing methods and measured biogeochemical tracers including chlorophyll a and fDOM to understand the influence of the IME on the community. Shifts in phytoplankton communities followed expected trends within the IME, including nearshore enrichment of *Synechococcus* and picoeukaryotic phytoplankton in the euphotic zone (Figures 4-5), that have been documented previously both using satellite imagery (Messié et al, 2022) and in the waters around Hawaiʻi (Comfort et al., 2024). This study reinforced these broad spatial and temporal studies in surface waters with a more detailed synoptic snapshot of the IME across depths, expanding our understanding of the IME by highlighting deeper horizons at which communities shift most drastically nearshore: the mixed layer and the DCM. Using 16S sequencing to quantify the relative abundance of heterotrophic bacteria taxa and investigate the shifts in taxa, I found groups of taxa that exhibited similar patterns of enrichment from nearshore to offshore in the mixed layer and the DCM (Figures 6-9). A key next step is documenting the biogeochemical shifts that underpin these patterns of microbial enrichment nearshore: our explorations of the fluorescence composition of DOM did not resolve any evidence of IME, but other biogeochemical properties are known to shift along these gradients (such as inorganic N and P; Comfort et al. 2024 or dissolved oxygen; Comstock et al. 2022). Future work should seek to better understand the ecology of the specific groups enriched nearshore in the DCM (*Actinomarinaceae*, DEV007, and *Rubritaleaceae*) and which biogeochemical properties of the water column dictate these enrichment patterns.

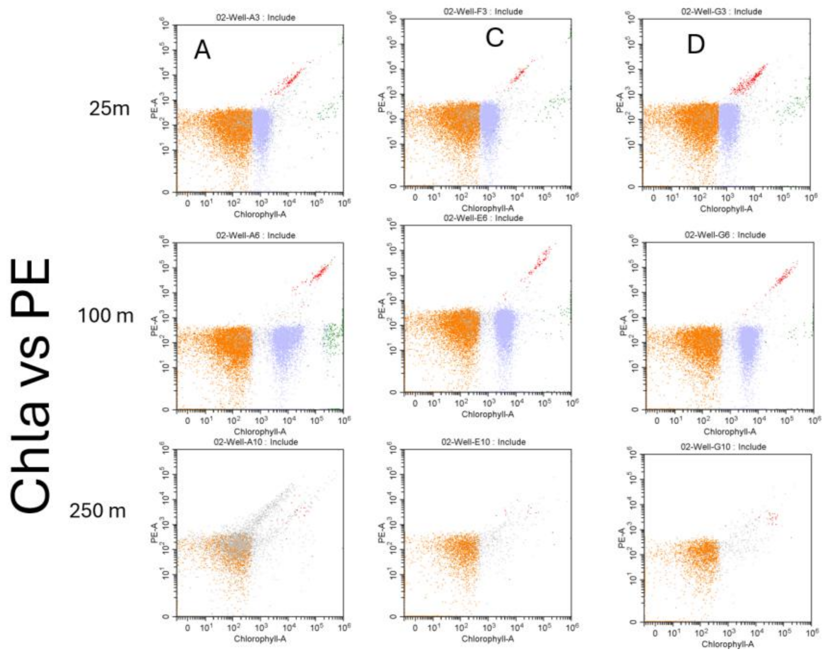
APPENDIX



Appendix 1: Flow cytometry plots showing the distribution of events across sample depths and stations. The axes are comparing Side Scatter on the x, a proxy for "roughness" of a cell, and Chlorophyll content. The columns are stations A, C, and D, and the rows are depths 25, 100, and 250 meters. Events are colored according to their groups: orange are heterotrophic bacteria, red are Synechococcus, light purple are Prochlorococcus, green are Picoeukaryotes, and grey are events gated out as junk.



Appendix 2: Flow cytometry plots showing events on the axes comparing Side Scatter to the DNA stain Hoechst.



Appendix 3: Flow cytometry plots comparing Chlorophyll content to Phycoerythrin of a cell. Events in the Synechococcus gate are visible as the red blob in the upper right corner of each plot.

LITERATURE CITED

Arisdakessian C, Cleveland SB, Belcaid M. 2020. MetaFlow|mics: scalable and reproducible nextflow pipelines for the analysis of microbiome marker data. Practice and Experience in Advanced Research Computing 2020: Catch the Wave (PEARC '20). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3311790.3396664

Campbell, L., & Vaulot, D. (1993). Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii (station ALOHA). *Deep Sea Research Part I: Oceanographic Research Papers*, *40*(10), 2043–2060. https://doi.org/10.1016/0967-0637(93)90044-4

Cleveland S, Arisdakessian C, Nelson C, Belcaid M, Frank K, Jacobs G. 2022. The C-MĀIKI gateway: a modern science platform for analyzing microbiome data. https://doi.org/10.1145/3491418.3530291

Coble, P. G. (1996). Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. *Marine Chemistry*, *51*(4), 325–346. https://doi.org/10.1016/0304-4203(95)00062-3

Comfort, C. M., Ostrander, C., Nelson, C. E., Karl, D. M., & McManus, M. A. (2024). A 7-yr spatial time series resolves the island mass effect and associated shifts in picocyanobacteria abundances near O’ahu, Hawai’i. *Limnology and Oceanography*, *69*(12), 2830–2845. https://doi.org/10.1002/lno.12711

Comstock, J., Nelson, C. E., James, A., Wear, E., Baetge, N., Remple, K., Juknavorian, A., & Carlson, C. A. (2022). Bacterioplankton communities reveal horizontal and vertical influence of an Island Mass Effect. *Environmental Microbiology*, *24*(9), 4193–4208. https://doi.org/10.1111/1462-2920.16092

Dandonneau, Y., & Charpy, L. (1985). An empirical approach to the island mass effect in the south tropical Pacific based on sea surface chlorophyll concentrations. Deep Sea Research Part A. Oceanographic Research Papers, 32(6), 707–721. https://doi.org/10.1016/0198-0149(85)90074-3

De Falco, C., Desbiolles, F., Bracco, A., & Pasquero, C. (2022). Island Mass Effect: A Review of Oceanic Physical Processes. *Frontiers in Marine Science*, *9*. https://doi.org/10.3389/fmars.2022.894860

DeLong, E., Preston, C., Mincer, T., Rich, V., Hallam, S., Frigaard, N.-U., Martinez, A., Sullivan, M., Edwards, R., Brito, B., Chisholm, S., & Karl, D. (2006). Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior. Science (New York, N.Y.), 311, 496–503. https://doi.org/10.1126/science.1120250

Doty, M. S., & Oguri, M. (1956). The Island Mass Effect. *ICES Journal of Marine Science*, *22*(1), 33–37. https://doi.org/10.1093/icesjms/22.1.33

Drazen, J. C., Clark, B. H., Gove, J. M., Phipps, J. E., Copeland, A. M., Lecky, J., Green, J. A. M., Kobayashi, D. R., Turner, J. R., Whitney, J. L., & Williams, G. J. (2023). Near-island enhancement in mesopelagic micronekton assemblages off Hawaiʻi. *Deep Sea Research Part I: Oceanographic Research Papers*, *199*, 104107. https://doi.org/10.1016/j.dsr.2023.104107

Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., Karl, D. M., Li, W. K. W., Lomas, M. W., Veneziano, D., Vera, C. S., Vrugt, J. A., &

Martiny, A. C. (2013). Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *Proceedings of the National Academy of Sciences*, *110*(24), 9824–9829. https://doi.org/10.1073/pnas.1307701110

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications*, *8*(1), 1188. https://doi.org/10.1038/s41467-017-01312-x

Gilmartin, M., & Revelante, N. (1974). The 'island mass' effect on the phytoplankton and primary production of the Hawaiian Islands. Journal of Experimental Marine Biology and Ecology, 16(2), 181–204. https://doi.org/10.1016/0022-0981(74)90019-7

Gove, J. M., McManus, M. A., Neuheimer, A. B., Polovina, J. J., Drazen, J. C., Smith, C. R., Merrifield, M. A., Friedlander, A. M., Ehses, J. S., Young, C. W., Dillon, A. K., & Williams, G. J. (2016). Near-island biological hotspots in barren ocean basins. *Nature Communications*, *7*(1), 10581. https://doi.org/10.1038/ncomms10581

Greg Caporaso J, Ackermann G, Apprill A, Bauer M, Berg-Lyons D, Betley J, Fierer N, Fraser L, Fuhrman JA, Gilbert JA, et al. 2018. EMP 16S Illumina amplicon protocol v1. PLOS One. https://doi.org/10.17504/protocols.io.nuudeww

Hasegawa, D., Yamazaki, H., Lueck, R. G., & Seuront, L. (2004). How islands stir and fertilize the upper ocean. *Geophysical Research Letters*, *31*(16). https://doi.org/10.1029/2004GL020143

Hoover, D. J., & Mackenzie, F. T. (2009). Fluvial Fluxes of Water, Suspended Particulate Matter, and Nutrients and Potential Impacts on Tropical Coastal Water Biogeochemistry: Oahu, Hawai‘i. Aquatic Geochemistry, 15(4), 547–570. https://doi.org/10.1007/s10498-009-9067-2

Jani, A. J., Bushell, J., Arisdakessian, C. G., Belcaid, M., Boiano, D. M., Brown, C., & Knapp, R. A. (2021). The amphibian microbiome exhibits poor resilience following pathogen-induced disturbance. *The ISME Journal*, *15*(6), 1628–1640. https://doi.org/10.1038/s41396-020-00875-w

Karl, D., Dore, J., Lukas, R., Michaels, A., Bates, N., & Knap, A. (2001). Building the Long-Term Picture: The U.S. JGOFS Time-Series Programs. *Oceanography*, *14*(4), 6–17. https://doi.org/10.5670/oceanog.2001.02

Knee, K. L., Street, J. H., Grossman, E. E., Boehm, A. B., & Paytan, A. (2010). Nutrient inputs to the coastal ocean from submarine groundwater discharge in a groundwater-dominated system: Relation to land use (Kona coast, Hawaii, U.S.A.). Limnology and Oceanography, 55(3), 1105–1122. https://doi.org/10.4319/lo.2010.55.3.1105

Kothawala Dolly N. , Murphy Kathleen R. , Stedmon Colin A. , Weyhenmeyer Gesa A. , Tranvik Lars J. , (2013), Inner filter correction of dissolved organic matter fluorescence, Limnol. Oceanogr. Methods, 11, doi:10.4319/lom.2013.11.616

Lawaetz, A. J., and C. A. Stedmon. 2009. Fluorescence Intensity Calibration Using the Raman Scatter Peak of Water. Applied Spectroscopy 63: 936–940. https://doi:10.1366/000370209788964548

Martinez, E., Rodier, M., Pagano, M., & Sauzède, R. (2020). Plankton spatial variability within the Marquesas archipelago, South Pacific. Journal of Marine Systems, 212, 103432. https://doi.org/10.1016/j.jmarsys.2020.103432

McKnight, D. M., Boyer, E. W., Westerhoff, P. K., Doran, P. T., Kulbe, T., & Andersen, D. T. (2001). Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. Limnology and Oceanography, 46(1), 38–48. https://doi.org/10.4319/lo.2001.46.1.0038

Mende, D. R., Bryant, J. A., Aylward, F. O., Eppley, J. M., Nielsen, T., Karl, D. M., & DeLong, E. F. (2017). Environmental drivers of a microbial genomic transition zone in the ocean's interior. Nature Microbiology, 2(10), 1367–1373. https://doi.org/10.1038/s41564-017-0008-3

Messié, M., Petrenko, A., Doglioli, A. M., Martinez, E., & Alvain, S. (2022). Basin-scale biogeochemical and ecological impacts of islands in the tropical Pacific Ocean. *Nature Geoscience*, *15*(6), 469–474. https://doi.org/10.1038/s41561-022-00957-8

Nelson, C. E., Alldredge, A. L., McCliment, E. A., Amaral-Zettler, L. A., & Carlson, C. A. (2011). Depleted dissolved organic carbon and distinct bacterial communities in the water column of a rapid-flushing coral reef ecosystem. *The ISME Journal*, *5*(8), 1374–1387. https://doi.org/10.1038/ismej.2011.12

Nelson, C. E., Donahue, M. J., Dulaiova, H., Goldberg, S. J., La Valle, F. F., Lubarsky, K., Miyano, J., Richardson, C., Silbiger, N. J., & Thomas, F. I. M. (2015). Fluorescent dissolved organic matter as a multivariate biogeochemical tracer of submarine groundwater discharge in coral reef ecosystems. *Marine Chemistry*, *177*, 232–243. https://doi.org/10.1016/j.marchem.2015.06.026.

Pomeroy, L., leB. Williams, P., Azam, F., & Hobbie, J. (2007). The Microbial Loop. *Oceanography*, *20*(2), 28–33. https://doi.org/10.5670/oceanog.2007.45

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4:e2584. https://doi. org/10.7717/peerj.2584

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

Selph, K. E. (2021). Enumeration of marine microbial organisms by flow cytometry using near- UV excitation of Hoechst 34580-stained DNA. *Limnology and Oceanography: Methods*, *19*(10), 692–701. https://doi.org/10.1002/lom3.10454

Signorini, S. R., McClain, C. R., & Dandonneau, Y. (1999). Mixing and phytoplankton bloom in the wake of the Marquesas Islands. Geophysical Research Letters, 26(20), 3121–3124. https://doi.org/10.1029/1999GL010470

Torréton J, Page`s J, Dufour P, Cauwet G. (1997). Bacterioplankton carbon growth yield and DOC turnover in some coral reef lagoons. In: Proceedings of the 8th International Coral Reef Symposium. Smithsonian Tropical Research Institute: Panama, pp 947–952 van den Engh, G. J., Doggett, J. K., Thompson, A. W., Doblin, M. A., Gimpel, C. N. G., & Karl, D. M. (2017). Dynamics of Prochlorococcus and Synechococcus at Station ALOHA Revealed through Flow Cytometry and High-Resolution Vertical Sampling. *Frontiers in Marine Science*, *4*. https://www.frontiersin.org/articles/10.3389/fmars.2017.00359

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, *73*(16), 5261–5267. https://doi.org/10.1128/AEM.00062-07

Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JAJansson JK, Caporaso JG, Fuhrman JA, Apprill A, Knight R2016.Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems1:10.1128/msystems.00009-15. https://doi.org/10.1128/msystems.00009-15

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F. O. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, *42*(D1), D643–D648. https://doi.org/10.1093/nar/gkt1209