UNIVERSITY OF CALIFORNIA

Santa Barbara

Microbial community composition is a tracer for biogeochemical cycling and organic matter transformation in coastal and open ocean systems

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Marine Science

by

Jacqueline Comstock

Committee in charge:

Professor Craig Carlson, Chair

Professor Alyson Santoro

Professor David Siegel

September 2023

The dissertation of Jacqueline Comstock is approved.

_____

David Siegel

_____

Alyson Santoro

_____

Craig Carlson, Committee Chair

September 2023

Microbial community composition is a tracer for biogeochemical cycling and organic matter transformation in coastal and open ocean systems

by

Jacqueline Comstock

ACKNOWLEDGEMENTS

This dissertation represents original research I have conducted with significant help and support of my mentors, family, friends, and collaborators. I am immensely grateful for my committee members, faculty, research staff, fellow graduate students, research station staff, and R/V crew I have worked with throughout my time at UCSB for their expertise, guidance, and support. I am also very grateful to Professor Craig Carlson, who has provided excellent mentorship and support throughout my graduate career.

I would like to thank the members of the Carlson lab – Elisa Halewood, Keri Opalk, Chance English, Ranchel Sanquist, Nicholas Baetge, Emma Wear, Anna James, Shutin Liu, and Brandon Stephens - for their companionship and mentorship throughout my graduate career, which aided immensely in my development as a researcher. I would also like to thank my wonderful collaborators in Mo'orea and Bermuda – Craig Nelson, Linda Wegley Kelly, Andi Haas, Irina Koester, Zach Quinlan, Wesley Sparagon, Milou Arts, Jessica Bullington, Rachel Parsons, Steve Giovannoni, Kevin Vergin, Amy Maas, Leo Blanco-Bercial, Michelle Michelsen, Alex Worden, Fabian Wittmers, Lillian Henderson, and Hilary Close.

Lastly, I would like to thank my family. To my twin, who I am so lucky to have shared my lifetime alongside, and to my mother, who singlehandedly raised and supported me. It is a testament to her support and love and hard work that I was able to pursue my interests and complete this degree. I am eternally grateful for the support system around me and the people I have the honor to share my life with.

VITA OF JACQUELINE COMSTOCK

September 2023

EDUCATION

Bachelor of Science in Aquatic Biology, University of California, Santa Barbara, March 2016 (with high honors)

Doctor of Philosophy in Marine Science, University of California, Santa Barbara, September 2023

PROFESSIONAL EMPLOYMENT

2017 - 2023: Teaching Assistant, Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara

- EEMB 3/2LL (Introductory Biology Lab II/III)
  - Spring 2017, 2020, 2021
- EEMB 142B (Environmental Processes in Oceans and Lakes)
  - Winter 2018-2023
- EEMB 131 (Principles of Evolution)
  - Summer 2021
- EEMB 158 (Climate Marine Ecology)
  - Spring 2022, 2023

2023 - present: Instructor of Record, Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara

- EEMB 131 (Principles of Evolution
  - Summer 2023
- EEMB 7 (Introductory Biology for Psychology & Brain Sciences)
  - Summer 2023

FIRST AUTHOR PUBLICATIONS

Comstock J, Nelson CE, James A, Wear E, Baetge N, Remple K, Juknavorian A, Carlson CA. Bacterioplankton communities reveal horizontal and vertical influence of an Island Mass Effect. Environ Microbiol. 2022 Sep; 24(9): 4193-4208. doi: 10.1111/1462-2920.16092. Epub 2022 Jun 12. PMID: 35691616; PMCID: PMC9796716.

COAUTHOR PUBLICATIONS

Kelly LW, Nelson CE, Petras D, Koester I, Quinlan ZA, Arts MGI, Nothias L, Comstock J, White BM, Hopmans EC, van Duyl FC, Carlson CA, Aluwihare LI, Dorrestein PC, and Haas AF (2022) Distinguishing the molecular diversity, nutrient content, and energetic potential of exometabolomes produced by macroalgae and reef-building corals. PNAS. 119:2110283119. doi.org/10.1073/pnas.2110283119

Liu S, Baetge N, Comstock J, Opalk K, Parsons R, Halewood E, English CJ, Giovannoni S, Bolaños LM, Nelson CE, Vergin K and Carlson CA (2020) Stable isotope probing identifies bacterioplankton lineages capable of utilizing dissolved organic matter across a range of bioavailability. *Front. Microbiol.* 11:580397. doi: 10.3389/fmicb.2020.580397

Stephens BM, Opalk K, Petras D, Liu S, Comstock J, Aluwihare LI, Hansell DA and Carlson CA (2020) Organic matter composition at Ocean Station Papa affects its bioavailability, bacterioplankton growth efficiency and the responding taxa. *Front. Mar. Sci.* 7:590273. Doi: 10.3389/fmars.2020.590273

Bisson, K. M., Baetge, N., Kramer, S. J., Catlett, D., Girling, G., McNair, H., ... & Valentine, D. L. (2020). California wildfire burns boundaries between science and art. *Oceanography*, *33*(1), 16-19.

CONFERENCS

April 2023 - SoCal BOOM: Presentation

Comstock, J.; Close, H.; Worden, A.; Henderson, L; Wittmers, F.; Halewood, E.; Liu, S.; Carlson, C.: Size-fractionated characterization of organic particles and associated microbial communities over the surface 500 m at the Bermuda Atlantic Time-series Study site

February 2022 - Ocean Sciences Meeting: Presentation

Comstock, J.; Nelson, C.; Kelly, L.; Haas, A; Quinlan, Z.; Koester, I.; Sparagon, W.; Bullington, J.; Arts, M.; Carlson, C.: Bacterial community structure associated with size-fractionated organic matter over the coral reefs of Mo'orea, French Polynesia

November 2020 - AtlantECO Workshop: Presentation

Comstock, J.; Albers, J.; Santoro, A.; Carlson, C.: Comparison of bacterioplankton community structure across extraction methods and filter types

February 2020 - Ocean Sciences Meeting: Poster

Comstock, J.; Nelson, C; Haas, A.; Kelly, L.; Koester, I.; Quinlan, Z.; Carlson, C. Tracking microbial uptake of phytoplankton derived DOM and model compounds by coral reef bacterioplankton using stable isotope probing

February 2019 – ASLO Aquatic Sciences Meeting: Presentation

Comstock, J.; Carlson, C.; James, A.; Nelson, C.: Linkages between bacterial and archaeal community structure and biogeochemical variability around the islands of Mo'orea and Tahiti

ABSTRACT

Microbial community composition is a tracer for biogeochemical cycling and organic matter transformation in coastal and open ocean systems

by

Jacqueline Comstock

Marine microbes drive ocean biogeochemistry. The myriad of marine microbes are organized by which microbes respond to an environmental condition and which one are deterred or displaced by a set of environmental conditions. Environmental conditions such as presence or absence of light, concentration of limiting inorganic nutrients and the quantity and quality of bioavailable organic matter are examples of factors that can influence microbial community structure. My work used amplicon sequencing of the 16S rRNA gene a biological tracer to elucidate drivers of biogeochemical trends across coastal and open-ocean environments. In coral reefs, intense recycling of nutrients and organic matter supports highly productive coral reef systems despite being bathed in unproductive oligotrophic waters. Around the French Polynesian islands of Mo'orea and Tahiti, we combined 16S rRNA metabarcoding with a suite of biogeochemical measurements to resolve the origins and spatial influence of an Island Mass Effect, the observed enhancement in primary production surrounding coral reef atolls and islands. Additionally, within the Mo'orea coral reef system we characterized these unique reef microbial communities across

a variety of physical regimes and reef habitats, linking bacterioplankton to gradients in biogeochemistry, physical wave forcing, and organic matter composition. Transitioning to open-ocean systems, we applied these methods at the Bermuda Atlantic Time-series site (BATS), one of the most well-studied sites in the global ocean. In open-ocean systems, particulate organic matter (POM) formed in the ocean's surface plays a central role in the carbon cycle, with sinking POM acting as the dominant pathway in the biological carbon pump. We aimed to better resolve patterns of particle colonization and chemical transformation across depth and particle size, resolving drivers of organic particle attenuation and transformation as organic matter travels from the surface into the deep. Through this research, we utilize microbial community composition as a tracer to inform the origins of observed biogeochemical patterns and organic matter transformation in both coastal and open-ocean environments.

TABLE OF CONTENTS

I. Introduction ***7***

II. Bacterioplankton communities reveal horizontal and vertical influence of an Island Mass Effect ***7***

Abstract **7**

Introduction **7**

Methods **10**

Site Description *10*

Sampling Scheme *11*

DNA Collection & Extraction *12*

Amplicon Library Sequencing & Bioinformatics *13*

Statistical Analyses *14*

Results **15**

Overall Dynamics *15*

Reef Bacteria Found at Nearshore Stations with Enhanced Chlorophyll a *16*

IME resolved in the deep chlorophyll maximum *19*

Bacterioplankton variability in the mesopelagic *21*

Discussion **23**

Island Mass Effect at the surface *24*

Evidence of IME over depth *27*

Conclusion **30**

Appendix **39**

III. Wave forcing is a primary driver of coral reef bacterioplankton communities around Mo'orea, French Polynesia ***48***

Abstract **48**

Introduction **49**

Methods **51**

Site Description *51*

Sampling Scheme *52*

Physical flow rate characterization *52*

DNA collection & extraction *53*

Amplicon library sequencing & bioinformatics *54*

Statistical analyses *55*

Results **55**

Physical flow around Mo'orea *55*

Biogeochemical variation between reef systems and habitats *56*

Overall trends in ordination *57*

Effects of reef habitat on ordination by side *58*

Taxonomic trends between island sides and reef habitats *59*

Correlation between communities & biogeochemistry *61*

Discussion **62**

Overview *62*

Physical regime and bacterioplankton ordination *63*

Biogeochemicstry and bacterioplankton communities *65*

Conclusion **68**

Appendix **75**

IV. Size-fractionation of marine particles at the Bermuda Atlantic Time-series Site (BATS) reveals cryptic links between microbial communities and chemical transformations across depth 77

Abstract 77

Introduction 78

Methods 82

In situ pump and hydrographic sampling 82

Size fractionated DNA preservation and extraction 84

Amplicon library sequencing and bioinformatics 84

Particle biogeochemistry 85

Inorganic nutrients 87

Dissolved organic carbon 87

Bacterioplankton abundance 87

Statistical analyses 88

Results 88

Hydrography 88

Chemical characterization of organic particles 89

Overall trends in bacterioplankton communities 90

Community variation within the UE, DCM, and mesopelaic 91

Community variation within size fractions 92

Alpha diversity 93

Distributions of cyanobacterial lineages 94

Non-cyanobacterial taxonomic abundance by fraction 94

Heterotrophic & chemoautotrophic succession over depth 96

Discussion 96

Biogeochemistry of marine particles *96*

Drivers of community differentiation *99*

Microbial taxa distributions *102*

Cyanobacteria found in particles below their free-living range *105*

Conclusion *107*

Appendix *115*

References ***122***

I. Introduction

Marine microbes represent some of the most evolutionarily diverse organisms on the planet (Pawlowski et al., 2012) and have a disproportionate role in driving global biogeochemical cycles (eg. carbon, nitrogen, phosphorus, sulfur, and iron), thereby influencing global climate (Falkowski et al., 2008) and ecosystems. Marine microbial assemblages contain a diverse array of metabolic strategies, including photoautotrophy in the sunlit euphotic, heterotrophy, and chemoautotrophy. These microbial communities have been well documented to change in composition spatially, vertically, and temporally in the water column, correlating with shifts in physical regimes and water biogeochemistry, such as shift in light, nutrient fields, and organic matter composition (eg. Giovannoni et al., 1996, Morris et al., 2005; DeLong et al 2006; Treusch et al., 2009). Understanding the interactions and relationships between marine microbial assemblages and surrounding biogeochemistry is critical to disentangling how these assemblages mediate these global cycles.

In the ocean's surface waters, the unicellular photosynthetic activity accounts for almost as much as their multicellular terrestrial counterparts, despite maintaining a standing biomass just 1% of terrestrial photosynthesizers (Siegenthaler and Sariento, 1993). Of the 54 Pg of marine annual net primary production, approximately 21 Pg is partitioned as dissolved organic matter (DOM) through a variety of food web processes (Carlson and Hansell 2015) of which > 90% is remineralized by heterotrophic microbial processes (Pomeroy 1974; Azam et al., 1983; Azam and Malfatti, 2007). Marine DOM is the largest reservoir of reduced carbon in the ocean (Hansell et al., 2009) and comparable in size to the atmospheric CO_2 reservoir, reaching more than 660 Gt C. This DOM pool is arguably one of the most

complex exometabolomes on earth. Comprised of thousands of compounds and containing reduced carbon bound to heteroatoms such as oxygen, nitrogen, phosphorus, and sulfur (Kujawinski, 2011), DOM serves as a substrate to support diverse heterotrophic prokaryote populations (Hansell, 2013). However, our ability to directly link diverse heterotrophic microbes to the organic matter they consume remains limited. Resolving linkages between marine microbes and DOM is relevant to global element cycling since bacterioplankton are the primary trophic conduit for DOM in aquatic ecosystems (Azam et al., 1983; Nelson and Wear, 2014).

In addition to mediating the cycling of marine DOM, marine microbes play an outsized role in controlling the sequestration of marine particulate organic matter (POM) into the ocean's interior. Sinking POM plays a central role in the transport and sequestration of carbon in the ocean, acting as the dominant pathway in the BCP (Boyd et al., 2019; Nowicki et al., 2022). These particles, comprised of aggregated phytoplankton detritus, zooplankton-derived carcasses and fecal matter can be ballasted by minerals of biogenic and lithogenic origin (Alldredge and Silver, 1988; Turner, 2015). This process transports organic matter into the ocean's interior (McCave, 1975). Sinking organic particles represent approximately 80% of the 7 – 11 Pg C (Chavez and Toggweiler, 1995; Sarmiento and Gruber, 2006) exported annually from 100 -150 m depth. The export of POC from the euphotic zone into the interior is the net outcome of two competing processes. The production and source of organic particles which affects particle size, aggregation and sinking rate is offset by biological processes that intercept, solubilize and remineralize the sinking organic matter; thus attenuating POC flux (Martin et al., 1987; Ducklow et al., 2001). During transport, organic particles undergo numerous transformations processes

including direct remineralization by bacterioplankton (Ploug and Grossart, 1999) or zooplankton (Steinberg et al., 2008), fragmentation of aggregates by zooplankton (Dilling and Alldredge, 2000; Goldthwait et al., 2004; Giering et al., 2014) and solubilization of sinking particulate organic matter (POM) to dissolved organic matter (DOM) via production of hydrolytic enzymes by attached prokaryotes (Azam et al., 1993; Smith et al., 1992; Ziervogel et al., 2010; Arnosti, 2011). These processes result in the observed attenuation of sinking particles by approximately 85% within the mesopelagic zone (200-1000m) (Marin et al., 1987). Marine POM is densely colonized by bacteria and archaea distinct from the free-living community, representing "hotspots" of microbial activity in the water column in both coastal and oligotrophic environments (DeLong et al. 1993, Simon et al., 2002; Grossart, 2010). These particle-associated microbial communities play a central role in POM degradation via production of extracellular enzymes that hydrolyze POM and DOM to lower-molecular-weight compounds prior to cellular uptake (Smith et al., 1992, Arnosti 2011; Arnosti et al., 2021). Therefore, understanding diversity and transformation of particle-associated heterotrophic microbial communities is useful for disentangling degradation processes on marine POM.

The first two chapters in this thesis focuses on research conducted in and around the island of Mo'orea, French Polynesia (-17.48, -149.82). The Mo'orea Coral Reef Long Term Ecological Research (MCR-LTER) site was established in 2004 and has since regularly collected data on physical, biological, and chemical properties, making it a well-characterized and highly studied coral reef system. This system therefore provides an excellent opportunity to investigate and resolve gradients in bacterioplankton, biogeochemistry, and physical dynamics in the context of decades of characterization.

Mo'orea is a 1.5-2 million-year-old island (Neall and Trewick, 2008) located in the westward flowing South Equatorial Current in the northern portion of the South Pacific Subtropical Gyre (Rougerie and Rancher, 1994). The island is surrounded by barrier reefs cresting within 1km of the shore. Reef pass channels occur every 5-10km around the circumference of the island, corresponding to embayments of varying sizes. The forereef slope has a relatively high coral density and rapidly drops to depths exceeding 500m within 1km of the reef crest. The backreef platform consists of a shallow (<3m) lagoon comprising a mixture of corals, macroalgae, and barren sands, as well as a deeper (10-12m) fringing reef bordering the island (Nelson et al. 2011). Surface waves move water from the forereef across the reef crest and into the backreef averaging 0.2m s-1, then is drained laterally into the bay and advected offshore through the pass (Hench et al., 2008). In addition to the strong alongshore currents on each side of the island, there is a low-frequency counterclockwise flow around the island, providing a mechanism by which water transported over the backreef and advected offshore might stay in relatively close proximity to the island over a period of days to weeks (Leichter et al., 2013, James et al., 2020). Because Mo'orea is located within the South Pacific amphidromic system, its tides do not exceed 0.3 m and exhibit negligible influence in water movement across the reef (Hench et al. 2008). The four hydrologically interconnected habitats (offshore, forereef, backreef, bay) are referred to throughout this proposal, with chapters designed to investigate biogeochemical and microbial dynamics within and between the systems.

The final chapter of this thesis focuses on research conducted in the vicinity of the Bermuda Atlantic Time-Series Site (BATS) in the subtropical North Atlantic. Physical, chemical and biological data have been collected monthly at BATS since 1988 and provide

an excellent contextual backdrop to investigate compositional and metabolic differences in microbial communities. These time series data includes biological productivity rates, vertical fluxes and biogeochemical variability (Lomas et al., 2013; Steinberg et al., 2001), DOM dynamics (Carlson et al., 1994; Goldberg et al., 2009; Liu et al., 2022 Hansell et al., 2002), microbial processes (Carlson et al., 1996; Nelson and Carlson, 2012; Parsons et al., 2012), and microbial plankton diversity (Carlson et al., 2009; Giovannoni and Vergin, 2012; Giovannoni et al., 1990; Morris et al., 2005; Treusch et al., 2009; Treusch et al., 2012, Blanco-Bercial et al., 2022). These rich data sets have identified recurring seasonal patterns and point to complex mechanisms underlying DOM oxidation (Carlson et al., 1994, Goldberg et al., 2009, Liu et al., 2022) alongside well-documented vertical stratification of the free-living microbial communities over depth (Gordon et al., 1996; Giovannoni et al., 1996; Treusch et al., 2009). In this chapter, we investigated transformations in POM across depth, coupling isotopic measurements with prokaryotic community characterization to better constrain successional patterns and organic matter transformation between the free-living and particle-associated microbial communities and organic matter of varying size ranges.

In this thesis I use 16S SSU ribosomal RNA gene sequencing on next-generation Illumina sequencers to produce relative abundance data of bacterioplankton communities. Significant improvements in sequencing technology coupled with a precipitous decline in sequencing costs over the last couple decades has made DNA sequencing an accessible and relatively cheap endeavor for academic laboratory groups. 16S rRNA gene sequencing does not sequence the entire genome of the microbial communities present, and instead sequences a hypervariable region within the ubiquitous 16S rRNA gene. The primers used in all

chapters of this thesis amplify the V4 hypervariable region of the 16S rRNA gene (515F-Y and 806RB primers with custom adapters from Apprill et al., 2015; Parada et al., 2016; Wear et al., 2018) and are the primers endorsed and used by the Earth Microbiome Project. The utility of this sequencing method is not to directly look at gene diversity present within different populations, but rather to gather snapshots of microbial diversity in hundreds to thousands of individual samples relatively inexpensively. This tool provides a way to track spatial and temporal microbial taxa, which can be potential indicators of movement of water masses, or can provide insight into potential metabolic processes and biogeochemical transformations that could be occurring. We therefore utilize microbial community composition data alongside a suite of biogeochemical measurements to inform the origins of observed biogeochemical patterns and organic matter transformation in both coastal and open-ocean environments.

II. Bacterioplankton communities reveal horizontal and vertical influence of an Island Mass Effect

Abstract

Coral reefs are highly productive ecosystems with distinct biogeochemistry and biology nestled within unproductive oligotrophic gyres. Coral reef islands have often been associated with a nearshore enhancement in phytoplankton, a phenomenon known as the Island Mass Effect (IME). Despite being documented more than 60 years ago, much remains unknown about the extent and drivers of IMEs. Here we utilized 16S rRNA gene metabarcoding as a biological tracer to elucidate horizontal and vertical influence of an IME around the islands of Mo'orea and Tahiti, French Polynesia. We show that those nearshore oceanic stations with elevated chlorophyll *a* included bacterioplankton found in high abundance in the reef environment, suggesting advection of reef water is the source of altered nearshore biogeochemistry. We also observed communities in the nearshore deep chlorophyll maximum (DCM) with enhanced abundances of upper euphotic bacterioplankton that correlated with intrusions of low-density, O_2 rich water, suggesting island influence extends into the DCM.

Introduction

Phytoplankton production provides a fundamental source of energy at the base of marine food webs (Duarte & Cebrián, 1996), where the distribution of phytoplankton biomass dictates the trophic structure of marine ecosystems and the production of the world's fisheries (Chassot et al., 2010; Iverson, 1990). In oligotrophic waters, the ecological impacts of increased phytoplankton production are especially pronounced, with oceanic waters around coral reef islands and atolls displaying significantly enhanced primary production,

elevated chlorophyll *a* (Chl), elevated animal biomass, and unique biogeochemistry (Andrade et al., 2014; Dandonneau & Charpyt, 1985; Doty & Oguri, 1956; Gove et al., 2016; James et al., 2020; Jones, 1962; Martinez & Maamaatuaiahutapu, 2004; Palacios, 2002; Raapoto et al., 2019; Signorini et al., 1999; Vollbrecht et al., 2021). This observed increase in phytoplankton biomass around reef island and atoll ecosystems - the 'Island Mass Effect' (IME) - was first observed more than 60 years ago and has been attributed to an increase in bioavailable nutrients in the euphotic zone surrounding the islands and atolls (Doty & Oguri, 1956). IMEs have been observed to be a common feature throughout the tropical Pacific (Gove et al., 2016), yet much remains unknown about drivers of this phenomenon, the depth to which enhanced biological activity penetrates the water column, or its impact on microbial assemblages and metabolic potentials.

An IME has recently been documented surrounding the islands of Mo'orea and Tahiti (James et al., 2020). Located in the South Pacific Subtropical Gyre, the island of Mo'orea, French Polynesia, is surrounded by some of the least biologically productive waters in the global ocean (Longhurst et al., 1995). However, the backreef - the shallow lagoon behind barrier reefs - of this coral reef system displays unique chemical and biological signatures such as elevated nitrate and Chl, and depleted dissolved organic carbon (DOC) concentrations when compared to the open waters surrounding the island (Nelson et al., 2011; Leichter et al., 2013). Transport and retention of key biogeochemical constituents such as inorganic nutrients and organic particles within tropical reefs are thought to support the high biomass and increased productivity observed (Johannes et al., 1972; Odum & Odum, 1995). James et al. (2020) showed that the islands of Mo'orea and Tahiti have the potential to alter the biological activity and biogeochemistry of oceanic water in their

vicinity, documenting enhanced levels of net primary productivity (NPP), Chl, heterotrophic bacterioplankton productivity (BP) and particulate organic carbon (POC) in oceanic waters within 15 km of the islands relative to farther offshore.

The horizontal and vertical influence of the observed IME around Mo'orea and Tahiti has not previously been investigated. Mechanisms contributing to observed IMEs have been diverse and found to vary by location. These include nutrient and organic matter inputs from terrestrial sources, such as riverine or submarine groundwater discharge (Anderson et al., 2002; Dandonneau & Charpyt, 1985; Garrison et al., 2003; Knee et al., 2016; Street et al., 2008). Anthropogenic factors such as land-use change, urban development, and wastewater discharge can also enhance nutrient inputs into nearshore waters (Anderson et al., 2002; Dandonneau & Charpyt, 1985; Knee et al., 2016; Street et al., 2008). Beyond direct terrestrial inputs, coral reef organisms have been found to biogeochemically alter reef water (M. J. Atkinson, 1987; Falter et al., 2004; Haas et al., 2011, 2013; Wyatt et al., 2010). Additionally, physical processes can impact movement and retention of nutrients around islands. Interaction of currents with island bathymetry may form eddies and induce upwelling (Hamner & Hauri, 1981; Takahashi et al., 1981; Wolanski & Hamner, 1988), and internal waves can interact with bathymetry, creating instabilities and delivering deeper waters with distinct biogeochemical and physical properties to the surface (Carter et al., 2006; Leichter et al., 2003, 2012; Wyatt et al., 2020). Around Mo'orea, a counterclockwise flow has been observed and has been suggested as a mechanism for physical retention of biogeochemically altered reef water around the island (James et al., 2020; Leichter et al., 2013). This counterclockwise circular flow is thought to circulate advected reef water in the oceanic environment around the island, keeping this biogeochemically altered water near

Mo'orea and allowing for potential re-entrainment into the reef system (Leichter et al., 2013).

Here we provide a comprehensive synoptic survey of 16S rRNA gene metabarcoding data collected around the islands of Mo'orea and Tahiti in July 2014 in conjunction with the physical and biogeochemical parameters described in James et al. (2020). We describe the distribution of prokaryotic communities as a potential biological tracer to investigate whether the biological variability associated with this IME could be due to advection of backreef water or upwelling from depth. We also use these phylogenetic markers to assess how an IME may affect the microbial composition of the deep chlorophyll maximum (DCM) and upper mesopelagic.

Methods

Site Description

This study was carried out within a grid (51 km by 67 km) around the Society Island of Mo'orea. Mo'orea is a 1.5 - 2-million-year-old (Neall & Trewick, 2008) high volcanic island surrounded by barrier reefs within 0.5-1 km of shore. There are fringing reefs bordering the island (~10 m deep) and a shallow (<3 m) backreef lagoon with a mixture of coral, algae, and barren sands. The backreef lagoons connect to open ocean by reef pass channels that occur at approximately 5-10 km intervals around the island. These reef pass channels often correspond with embayments of varying sizes. The forereef has a high coral density and drops steeply to depths exceeding 500 m within 1 km offshore from the reef crest (Nelson et al., 2011; James et al., 2020). Waves drive water from the forereef across the reef crest into the backreef lagoon (Hench et al., 2008). Water then flows parallel to the

island into the reef pass channels and is advected offshore through the passes. Movement within the reef environment is rapid (averaging 0.2 m s^{-1} with limited tidal influence; Mo'orea is near a local M2 amphidromic point) and maintains a short residence time of hours to days within the backreef environment (Hench et al., 2008; Knee et al., 2016). Additionally, a weak counterclockwise flow has been observed around Mo'orea, and has been suggested to be a mechanism by which advected reef water can stay near the island (Leichter et al., 2013).

Sampling Scheme

Water samples were collected aboard the R/V *Kilo Moana* coincident with biogeochemical and physical oceanographic variables over the surface 500 m at 41 hydrographic stations in a grid (17.359˚S - 17.820˚S by 149.491˚W - 150.121˚W) around the Society Islands of Mo'orea and Tahiti in French Polynesia (Figure 2.1) from July 27 to August 7, 2014. Stations occupied by the R/V *Kilo Moana* are referred to as "oceanic" stations. Staying with the description in James et al. (2020), the oceanic stations were further categorized as "nearshore" or "offshore" if they were within or beyond 15 km of the islands, respectively. Samples from the R/V *Kilo Moana* were collected nominally at 10, 25, 50, 75, 100, 150, 200, 250, 300, and 500 m by a conductivity, temperature, depth (CTD) profiling rosette affixed with twenty-four 10 L Niskin bottles. We define all depths within the surface 75 m as being representative of the upper euphotic zone, depths between 90 and 125 m as the DCM and depths at 150 m and deeper as the upper mesopelagic.

During the same time period an additional forty-six stations within the reef environment on the north, west, and east sides of Mo'orea were also sampled from the surface to as deep as 25 m and are referred to as "reef" stations. Reef stations were sampled with a horizontal

Niskin bottle using a polyester rope and weighted messenger, with collections at 2 to 3 depths spanning the entire water column of the forereef, backreef, and reef pass environments (nominally 1, 10, and 25 m in forereef habitats; in backreef habitats bottom samples were collected 20 cm above the benthos, typically at 2-3 m depth). Measured biogeochemical variables included dissolved inorganic nutrients (nitrate, nitrite, ammonium, and phosphate), net primary production (NPP), chlorophyll *a* (Chl), phaeopigments, heterotrophic bacterial productivity (BP), bacterial abundance (BA), dissolved organic carbon (DOC), particulate organic carbon (POC), and particulate organic matter nitrogen isotopes (POM-δ^{15} N). A detailed description of these data and standard operating procedures are available in James at al. (2020). Mixed layer depth (MLD) was calculated as the depth where density was 0.01 kg m^{-3} greater than the local near-surface value. All CTD and biogeochemical data are available in the National Science Foundation (NSF) Environmental Data Initiative (EDI) data catalog under the package ID knb-lter-mcr.5035.1.

DNA Collection & Extraction

Two liters of whole seawater were gravity-filtered onto a 0.22 µm pore-size 47 mm polyethersulfone flat filter (Pall Supor®) housed in a polycarbonate filter cartridge that was attached directly to the Niskin bottle (nearshore and offshore samples) or after transfer to shore in a 5L HDPE carboy held in a chilled cooler (reef samples). All filters were stored dry at -40°C until further processing. DNA was extracted using the Qiagen DNEasy Blood & Tissue Kit using manufacturer instructions modified as follows: one ml of sucrose lysis buffer (40 mmol L^{-1} EDTA, 50mmol L^{-1} Tris HCl, 750 mmol L^{-1} sucrose, 400 mmol L^{-1} NaCl, pH adjusted to 8.0), 100 µl of sodium dodecyl sulfate (10% w/v) and 10 µl of 20 mg ml^{-1} proteinase K were added, samples were incubated at 55°C for 2 hours, then refrozen

until extraction. Additionally, to increase DNA yield, 750 µl of lysate mixture were extracted by running three 250 µl volumes through the same microcentrifuge column with 750 µl of Qiagen AW1 buffer. Extracted genomic DNA was stored at -40° C until amplification. Samples from 3 stations (oceanic stations 14, 15, and 25; fig 1) were compromised prior to extraction and were removed from further analysis.

Amplicon Library Sequencing & Bioinformatics

Amplification of the V4 region of the 16S rRNA gene was performed using the 515F-Y (5'-GTGYCAGCMGCCGCGGTAA-3') and 806RB (5'-GGACTACNVGGGTWTCTAAT-3') primers (Apprill et al., 2015; Parada et al., 2016). PCR-grade water process blanks and mock communities (BEI Resources mock communities HM-782D and HM-783D) were included with each 96-well plate of samples as quality control checks. Amplicons were cleaned and normalized using SequalPrep plates (Invitrogen), pooled at equal volumes, concentrated using Amicon Ultra 0.5 ml centrifugal tubes (Millipore), gel extracted using the QIAquick Gel Extraction Kit to remove non-target DNA (Qiagen) and sequenced on an Illumina MiSeq using PE250 chemistry at University of California (UC), Davis DNA Technologies Core.

Samples were demultiplexed at UC Davis DNA Technologies Core. Fastq files were quality filtered and merged using the dada2 package (version 1.13) in R (Callahan et al., 2016). Chimeras were removed de novo using the removeBimeraDenovo function in the dada2 package. Amplicon sequence variants (ASVs) were given a taxonomic assignment in the dada2 package using the assignTaxonomy command and the SILVA database (version 132)(Quast et al., 2013). After sequence classification, and initial assessment of plastid abundance, plastid sequences were removed and remaining reads were subsampled to 3000

reads per sample to standardize read depth. Samples below a read depth of 3000 were removed from further analysis which amounted to 19 out of 478 environmental samples. Mock communities and negative controls were checked to confirm consistency in amplification and lack of contamination between PCR plates and then removed from further analysis. All DNA sequence data are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under project number PRJNA773129.

Statistical Analyses

For all multivariate analyses, ASV relative abundances were pre-treated using an angular transformation to normalize the dataset. Bray-Curtis dissimilarity matrices and NMDS ordinations were calculated in R (v4.0) using the vegedist and metaMDS functions in vegan (v2.5) (Oksanen et al., 2020). NMDS ordinations were calculated and visualized on 2 axes, with stress values for all ordinations falling below 0.2 and reported with each NMDS ordination. Up to 20 iterations of NMDS ordination were performed until convergence. ASVs found in only one sample were removed when clustering DCM communities to better investigate the impact of isopycnal intrusions on DCM community, but were kept for all other analyses. A "horseshoe effect" can be observed in NMDS ordinations with communities collected over several depths. The bending of dissimilar samples towards one another is because the analysis struggles with discriminating between samples that do not share common features (Morton et al., 2018). To avoid misinterpretation of overall trends due to this "horseshoe effect", dendrograms were used to emphasize separation of bacterioplankton communities by depth. Dendrograms were created using Ward dissimilarity of Bray-Curtis dissimilarity matrices and were generated using JMP Pro 15. PERMANOVA and pairwise PERMANOVA analyses were run using adonis and betadisper

functions in vegan (v2.5).T-tests were run in R using the stats package (v3.6.2). One-way ANOVAs were run in SAS using JMP Pro 15 (SAS Institute Inc., Cary, NC). Analyses run can be found at https://github.com/jacquicomstock/KM14-16-16SrDNA.git.

Results

Overall Dynamics: Bacterioplankton Communities Separate by Location and by Depth

Overall, 459 unique environmental samples were successfully sequenced containing 2980 Amplicon Sequence Variants (ASVs), with 2751 non-singleton ASVs. Bacterioplankton communities clustered strongly by depth (PERMANOVA $R^2 = 0.66$, $p < 0.001$) with distinct community clusters at every sampled depth deeper than 75 m (Figure 2.2a, b). Within the upper euphotic, samples from the upper 75 m for the oceanic stations clustered together, apart from a handful of 75 m samples that clustered separately (Figure 2.2a). Reef stations (1- 25 m) clustered separately from upper euphotic (upper 75 m) oceanic stations (Figure 2.2c; yellow filled symbols). Figure 2.3 shows the distribution of ASVs found in at least 3 samples with a relative abundance greater than 4%. All ASVs displayed variation in relative abundance by depth, with each depth below 75 m displaying enhanced abundance of a unique suite of ASVs. All depths within the upper 75 m showed similar relative abundances for the majority of ASVs resolved (Figure 2.3). Additionally, reef stations displayed strong enhancement of ASVs largely absent or in low abundance at oceanic stations.

The whole 16S rRNA gene metabarcoding community contained prokaryotes and eukaryote plastids, with chloroplast plastid sequences averaging just 1.18% of all ASVs in

the surface 75 m, 4.9% of all ASVs within the DCM (90- 135), and 0.33% of all ASVs within the upper mesopelagic (150 – 500). However, ASVs associated with cyanobacterial lineages averaged 32% of all ASVs and 96.4% of the photoautotrophic ASVs in the upper euphotic. In the DCM, cyanobacteria averaged 15.9% of all ASVs and 76.4% of photoautotroph ASVs. These data indicate that the chlorophyll variability at both depth zones were likely primarily driven by bacterioplankton rather than eukaryotes. However, it is important to note that due to a lack of a mechanical lysis step during DNA extraction, plastid abundance may be underestimated. For all further analyses, plastid sequences were removed and samples were rarefied to standardize read depth and ensure even sampling effort.

Several ASVs that were associated with the *Prochlorococcus* clade stratified vertically. The most abundant cyanobacterium in the upper 75 m of the water column was the *Prochlorococcus* ASV #1 that averaged 31.1% and 5.7% of the total prokaryotic community in the upper 75 m and DCM, respectively. The *Prochlorococcus* ASV #44, the fourth most abundant of cyanobacteria in the dataset, was found primarily within the DCM with an average relative abundance of 3.5% and reached a relative abundance of 14.4% at one DCM station. A third deep *Prochlorococcus* ASV (#22) was found primarily at the base of the euphotic zone (i.e. ~150 m), averaging 2.3% of the community at that depth (Figure 2.3).

Reef Bacteria Found at Nearshore Oceanic Stations with Enhanced Chlorophyll *a*

James et al. (2020) reported elevated Chl, NPP, POC and BP for nearshore stations relative to offshore stations. Chl distribution within the upper euphotic revealed elevated Chl

concentrations at nearshore stations in the upper 75 m relative to offshore stations. Five nearshore stations with Chl concentrations above the mean (0.12 ± 0.06 ug Chl L^{-1}) had corresponding MLDs that were shallower than the DCM, indicating that the enhanced Chl values at the surface were not a result of physical entrainment of DCM Chl at the time of sampling.

In this study, we utilized spatial distributions of various ASVs to gain a deeper understanding of the extent of influence of the observed IME on bacterioplankton communities. First, we categorized ASVs by location they were most abundant. ASVs found in significantly higher abundance (T-test $p < 0.05$, Table 2.1) at reef stations relative to oceanic upper euphotic stations, averaged greater than a 1% abundance on the reef, and displayed a greater than 2-fold difference in mean abundance between reef and oceanic stations were defined as 'reef ASVs'. Likewise, ASVs found in significantly higher abundance (T-test $p < 0.05$) with a 2-fold increase in abundance in the oceanic upper euphotic stations relative to reef or mesopelagic stations and a >1% mean abundance in oceanic upper euphotic stations were defined as 'oceanic upper euphotic ASVs'.

For our analyses, we utilized the nearshore/offshore station groupings previously defined by James et al. (2020) as a basic delineation of oceanic stations, then further delineated nearshore stations by the bacterioplankton present. At seven nearshore oceanic stations we observed reef ASVs in the upper euphotic. We defined nearshore oceanic stations with at least one reef ASV present (i.e., > 0%) as 'nearshore reef-influenced' (NRI) stations. No other nearshore stations contained reef ASVs above the level of detection. The reef ASVs that were observed at these NRI stations included *Synechococcus* (#7), Flavobacteriaceae (Cryomorphaceae #19, NS5 #29, and NS9 #47), Alteromonas (#9),

OM60(NOR5) (#18), HMIB11 (#10), and unassigned Alphaproteobacteria (#54) (Figure 2.4, Table 2.1). We note that while 21 of our 22 offshore stations did not contain any detectable reef ASVs, there was a single offshore station where *Synechococcus* (#7) was present (Appendix Figure 2.3), indicating that the reef influence could extend beyond the previously defined nearshore stations. The relative abundance of these reef ASVs increased within the reef environment as Chl concentrations and reef residence time increased moving from the forereef to the backreef to the bay (Hench et al., 2008) (Appendix Figure 2.4).

NRI stations were found to have significantly higher (one-way ANOVA, $p < 0.05$) Chl, NPP, POC, and BP than both other nearshore stations and offshore stations in the upper 75 m, with the largest differences in mean concentrations found in the upper 10 m (Table 2.2). Biogeochemical parameters and select ASV abundances in the upper 10 m are visualized in Figure 2.4. Additionally, the NRI stations also had a lower mean relative abundance of the most abundant oceanic upper euphotic ASV, *Prochlorococcus* ASV#1 (27.1 ± 1.6%) than other oceanic stations (31.7 ± 0.7%). Ordination of the microbial community structure in the surface 10 m revealed three distinct clusters; the reef community, a mixture of nearshore and offshore, and the NRI (PERMANOVA $R^2 = 0.27$, $p < 0.001$, Figure 2.2c). Separation of these three clusters was observed throughout the upper 75 m (Appendix Figure 2.2), however the degree of difference between NRI stations and other oceanic stations was greatest in the upper 10 m (Figure 2.2c) relative to the entire upper 75 m (pairwise PERMANOVA, $R^2 = 0.06$ at 10 m and 0.02 in for the upper 75 m, $p < 0.001$). This suggests the cause of the unique communities was due to influence from the surface rather than upwelling from depth, consistent with the hypothesis of James et al. (2020) that the DCM was not the source of enhanced Chl at the surface. In the upper 10 m,

NRI stations displayed a similar magnitude of difference between other oceanic stations (pairwise PERMANOVA $R^2 = 0.06$, $p = 0.002$) and reef stations (pairwise PERMANOVA $R^2 = 0.05$, $p < 0.001$). NRI stations thus represent a mixture of oceanic and reef communities, demonstrating connectivity between the reef and nearshore microbial community and biogeochemical response.

IME resolved in the deep chlorophyll maximum

The DCM ranged between 90 to 125 m during the sampling period. There was a larger phylogenetic variability within the DCM and 150 m compared to other depth ranges and reef samples (Appendix Figure 2.5). There was a weak but significant correlation between ordination of the DCM communities (NMDS ordination of Bray-Curtis distances) and water mass captured within the density surfaces between 23.6 – 24.5 kg m^{-3} (PERMANOVA $R^2 = 0.150$, $p < 0.001$). This suggests the shoaling and deepening of these isopycnals through the DCM influenced the composition of the DCM microbial community structure. A clear bifurcation of DCM samples was observed (Figure 2.5a) when DCM samples were hierarchically clustered with Ward's minimum variance method using Bray-Curtis dissimilarity values. ASVs found only in one DCM sample were removed from multivariate analyses to focus on relative abundance changes of cosmopolitan DCM bacterioplankton due to the high number of individual DCM station ASVs. The two clear groupings ordinated separately (PERMANOVA $R^2 = 0.15$, $p < 0.001$, Appendix Figure 2.6) and can be broadly characterized as influenced by "upper-euphotic-like" or more "mesopelagic- like" communities (Figure 2.5a). Upper-euphotic-like communities had a smaller magnitude of difference with upper euphotic communities (pairwise PERMANOVA $R^2 = 0.19$, adjusted $p = 0.006$) than mesopelagic-like communities had with upper euphotic

communities (pairwise PERMANOVA $R^2 = 0.45$, adjusted $p = 0.006$). Likewise, mesopelagic-like communities had a smaller magnitude of difference with 150 m communities (pairwise PERMANOVA $R^2 = 0.21$, adjusted $p = 0.006$) than upper-euphotic-like communities had with 150 m communities (pairwise PERMANOVA $R^2 = 0.24$, adjusted $p = 0.006$) (Appendix Figure 2.6). Samples within the mesopelagic-like group occupied significantly denser water masses (mean = 24.14 ± 0.11 kg m^{-3}) than samples within the euphotic-like group (mean = 23.90 ± 0.17 kg m^{-3}, T-test $p = 0.0014$) (Figure 2.5b). Thus, euphotic-like communities coincided with the vertical displacement of less-dense isopycnals to deeper depths and mesopelagic-like communities coincided with uplift of denser isopycnals to shallower depths (Figure 2.5b). Additionally, mean oxygen and DOC concentrations were significantly lower for waters surrounding the mesopelagic-like group (O_2 = 196.7 ± 2.7 µM O_2; DOC = 62.1 ± 4.4 µM C) than for those surrounding the upper euphotic-like group (O_2 = 199.7 ± 2.1 µM O_2; DOC = 66.0 ± 3.5 µM C, T-test $p = 0.0027$ and $p = 0.0173$). Pairwise PERMANOVA results and mean values for density, oxygen, and DOC are listed in Appendix Table 2.3. No significant difference in nutrient concentrations (nitrate + nitrite, phosphate) was observed between nearshore and offshore stations or between upper euphotic-like and mesopelagic-like communities, suggesting that nutrient dynamics were not primary drivers in community differentiation in the DCM.

The physical and chemical differences between the two DCM groupings suggest they are influenced by distinct water masses. Mesopelagic-like communities had an increased relative abundance of SAR86, *Prochlorococcus* (ASV #59), Flavobacteriaceae, SAR 202 (ASV #154) and SAR 406. Euphotic-like samples displayed enhanced abundance of ASVs found in higher abundance in the upper 75 m, including *Prochlorococcus* (ASV #1),

AEGEAN-169, SAR202 (ASV #60, 99), and SAR11 (#3, 4, 6, 14) (Appendix Table 2.2). These are not reef taxa found present in NRI stations -- they are taxa found in high relative abundance in the upper 75 m of all oceanic stations. While mesopelagic-like DCM communities were also found at nearshore stations, the euphotic-like DCM communities (light grey circles) were found almost exclusively at nearshore stations (red stations) that lie within 15 km of either Mo'orea or Tahiti (Figure 2.5 a, b). This indicates that proximity to the islands influenced the downward displacement of lower density isopycnals that resulted in the DCM of those stations being comprised of more surface -like taxa. Thus, IME appears to influence microbial community structure as deep as the DCM.

Bacterioplankton variability in the Mesopelagic

Permutational multivariate analyses of variance (PERMANOVAs) were used to determine if nearshore and offshore samples ordinated separately both within the entire upper mesopelagic and within each discrete depth throughout the upper mesopelagic (nominally 150, 200, 250, 300, and 500 m). While mesopelagic communities clearly separated between discrete depths (Figure 2.2a, b), these analyses did not reveal any significant difference in ordination with regard to proximity to shore for any of the mesopelagic depths. Thus, we were unable to detect an IME on microbial community structure deeper than the DCM.

Hierarchical clustering demonstrated clear vertical stratification of the prokaryotic community structure within the mesopelagic (150 - 500 m) (Figure 2.2). At the domain level, the proportion of Archaea in the overall bacterioplankton community increased with depth and reached a maximum in relative abundance at 200 and 500 m in oceanic stations

(Figure 2.6) comprising up to 55% of total community structure. Marine Group II (MGII) Euryarchaeota averaged between 3-6.5 % of the overall community at all depths from the surface to 200 m, while Marine Group III (MGIII) averaged approximately 3% through the euphotic zone (Figure 2.6). Deeper within the mesopelagic zone, the relative abundance of MGII increased to a maximum of 10.9% of the overall community structures by 250 to 300 m. MGIII reached its highest average relative abundance of 5.1% at 250 m. Thaumarchaeota were resolved in only two samples shallower than the DCM (two samples at 75 m). However, the relative contribution of Thaumarchaeota to the total prokaryotic community increased with depth starting within the DCM (90- 125 m) and reaching maximal relative contributions of 34% and 32% at 200 and 500 m, respectively (Figure 2.6). The two deep Thaumarchaeota maxima were comprised of distinct ASVs (ASVs #18, 51, and 161 at 300 m, ASVs #15, 32, 48, and 195 at 500 m, Figure 2.3).

Within the bacterial domain, the most abundant taxa in the mesopelagic included members of the SAR11, SAR202, SAR86, SAR406, and SAR324 clades (Figure 2.3, 2.6). Within the SAR11 Clade, subclades I and IV decreased in relative abundance from surface to mesopelagic, while members of SAR11 Ib increased in relative abundance from upper euphotic to DCM and mesopelagic. This is consistent with previously understood SAR11 distributions (Giovannoni, 2017). Average relative abundance of SAR11 subclade II was not significantly different between upper euphotic and mesopelagic depths; however, within the clade individual ASVs showed variable depth distributions. SAR202 was present at low relative abundances in the top 50 m, averaging less than 2% at all depths. Beginning at 75 m, the relative abundance of SAR202 began to increase with depth, reaching greatest relative abundance between 150 and 250 m. SAR86 was found in highest relative abundance

in the upper 50 m of oceanic stations before declining to its lowest relative abundance at 150 m, then increasing again between 200 m and 500 m. The SAR86 ASVs that occupied the upper euphotic and the mesopelagic were distinct (ASVs #20, 34, 36, 39 in upper euphotic, ASVs # 27, 79 in mesopelagic, Figure 2.3; Appendix Table 2.1). Similar to SAR202, SAR406 maintained a low relative abundance in the upper euphotic, with two peaks in relative abundance in the DCM and 300 m. SAR324 was in low abundance in the upper euphotic, but increased with depth, peaking at 500 m (Figure 2.6).

Discussion

Altered physical dynamics, biogeochemical patterns and enhanced biological activity observed for stations within the nearshore oceanic environment provides evidence that is consistent with the phenomenon known as an island mass effect (Andrade et al., 2014; Doty & Oguri, 1956). James et al. 2020 analyzed an extensive biogeochemical data set collected from a 51 km by 67 km grid encompassing the Society Islands of Moorea and Tahiti in the South Pacific Subtropical Gyre. They proposed that an IME influenced the biogeochemical constituents around the islands of Mo'orea and Tahiti, with nearshore stations showing enhanced levels of NPP, Chl, BP, and POC in the upper euphotic zone of the water column. They hypothesized that advection of reef influenced water may have contributed to those nearshore features. However, in the absence of a chemical tracer or biological marker they were unable to establish a direct connection between the reef water and nearshore "hot spots". In addition, the high variability associated with bulk measurements of biological rates (NPP and BP) and concentrations of biogeochemical variables limited their ability to resolve the vertical extent of the potential IME beyond the surface 75 m.

In the present study we used 16S rRNA gene metabarcoding data, collected concurrently with the previously reported biogeochemical measurements (James et al., 2020) to assess whether the metabarcoding data provided a useful biological tracer of an IME. We explored the use of bacterioplankton ASVs as biological tracers to better elucidate the sources of the observed enhancement of key biogeochemical constituents at nearshore stations. More specifically, we sought to determine whether the metabarcoding data could help resolve the potential influence of IME on the distribution, composition, and connectivity of microbial communities along horizontal (reef to offshore) gradients as well as over depth, extending from the upper euphotic zone through the DCM and into the upper mesopelagic zone.

Island Mass Effect at the Surface

Horizontal advection of reef water with elevated productivity and nutrient concentrations was hypothesized as a primary mechanism supporting the IME (James et al., 2020). An alternative hypothesis is that the elevated surface Chl and NPP at the nearshore reef-influenced (NRI) stations resulted from entrainment of photoautotrophs and nutrients from the DCM and upper mesopelagic, bringing a greater Chl signature to the surface or stimulating primary production. While previous studies have observed upwelling-induced IMEs near deep-sea islands (Furuya et al., 1986; Hasegawa et al., 2004; Heywood et al., 1990), two lines of evidence rule out upwelling as a source for enhanced NPP and surface Chl concentrations around Mo'orea and Tahiti. First, the study area was a relatively well stratified system, characteristic of Pacific subtropical gyres (Karl & Church, 2014, 2017; Wilson et al., 2015). Five nearshore stations with enhanced Chl were found to have MLD shallower than 75 m (James et al., 2020) precluding entrainment of Chl from the DCM or

nutrients from upper mesopelagic into the upper euphotic zone. Second, bacterioplankton ASVs in the surface 75 m were distinct from those in the DCM (Figure 2.3). For example, the ASVs found in highest relative abundance in the DCM (e.g. Prochlorococcus #44, #59, #131, SAR406 #31, SAR11 Ib #69, SAR324 #113) were absent from samples collected above 75 m. We observed distinct depth ranges for ASVs through the water column, with one notable example being the separation of 3 *Prochlorococcus* ASVs through the euphotic zone, all members of strain MIT9313. All Prochlorococcus ASVs found in this dataset have previously been reported in the subtropical Pacific (Berube et al., 2018). The stratification of *Prochlorococcus* ecotypes in this dataset reflects subclade adaptations to varying light and nutrient fields within the euphotic zone (Biller et al., 2014; Braakman et al., 2017; Durand et al., 2001; Rocap et al., 2003), providing unique signatures for tracking the vertical displacement of water masses.

Consistent with the hypothesis of horizontal advection of modified reef water, we observed that bacterioplankton ASVs found in significantly and substantially higher relative abundance in reef stations were located at oceanic stations with enhanced Chl, NPP, BP, and POC (Table 2.1, Figure 2.4), suggesting advection of water from the reef to the nearshore oceanic was a potential source for altered biogeochemistry and biology at these NRI stations. All individual reef taxa except *Synechococcus* (#7) average less than 1% of overall abundance in oceanic stations, suggesting they are not the major drivers of community structure ordination. However, these specific ASVs can serve as biological tracers connecting the biology of the reef environment with that of surrounding oceanic stations. The presence of reef-associated ASVs and decreased relative abundance of the oligotrophic cyanobacterium *Prochlorococcus #1* at NRI stations are associated with a separation in

ordination from other nearshore stations (Figure 2.2c). The similar magnitude of difference in ordination NRI stations displayed with other oceanic and reef stations highlight the gradient in community between reef and offshore oceanic environments that mapped onto the gradients of Chl and productivity established by James et al. (2020).

The horizontal advection of reef and inshore waters has been observed around atolls in the Northwestern Hawaiian Islands (Gove et al., 2016) and around Rangiroa Atoll in the French Polynesian Tuamotu archipelago (Vollbrecht et al., 2021), both of which have pronounced IMEs. Ocean swells can pump substantial amounts of water over the fore and backreefs of these islands and atolls that then flow through the reef system and exit through reef passes (Aucan et al., 2012), similar to the flow around Mo'orea (Hench et al., 2008). Wave forcing across atolls acts as a highly efficient flushing mechanism, advecting detritus and other nutrient sources from the coral reef into the surrounding nearshore oceanic environment (Callaghan et al., 2006). Rapid flushing rates associated with coral reef systems can exceed rates of nutrient assimilation by the coral reef benthic community (Atkinson, 2011), providing a mechanism by which the nearshore oceanic environment can experience higher nutrient input and higher phytoplankton biomass than the offshore environment. The coral reefs around Mo'orea experience a rapid flushing rate with a backreef residence time of hours to days (Hench et al., 2008; Knee et al., 2016); thus, advection of water from the coral reef to the nearshore oceanic environment is a plausible source of increase in biological productivity near the island. Additionally, the observed weak counterclockwise flow around the island of Mo'orea could act as a mechanism for retention of advected reef water in the surrounding nearshore oceanic environment (James et al., 2020; Leichter et al., 2013). However, with the data collected, it remains unknown whether the presence of reef

ASVs in the nearshore oceanic environment is solely due to advection and subsequent dilution of reef ASVs, or whether advection of chemical subsidies from the reef allowed for growth of reef-like prokaryotic populations at NRI stations. Thus, further discussion on metabolic potential and activity of these reef taxa in the oceanic environment without additional supporting data would be limited to conjecture.

Evidence of IME over depth

The metabarcoding data from the present study revealed several distinct bacterioplankton taxa typically observed within the upper euphotic zone, such as *Prochlorococcus* #1, AEGEAN-169 (#8), SAR202 (ASV #60, 99), and SAR11 (#3, 4, 6, 13, 14), occupying comparatively low density, warm water in the DCM at stations near the islands. This suggests island-influenced vertical displacement of upper euphotic water occurred around time of sampling. The cause of this downwelling cannot be determined from data in the present study; however, surface currents and eddies have been found to vertically displace local surface derived organic matter, oxygen and biological activity to mesopelagic depths in oceanic systems (Benitez-Nelson & McGillicuddy, 2008; McGillicuddy et al., 1999, 2007; Nelson et al., 2014). At time of sampling, a predominance of mesoscale anticyclonic flow was observed in the region around Mo'orea and Tahiti (James et al., 2020). On a smaller scale, a weak counterclockwise flow has been observed around Mo'orea, and was present at time of sampling (James et al., 2020; Leichter et al., 2013). This anticyclonic flow likely resulted in the observed elevated sea surface height around Mo'orea and Tahiti (James et al., 2020), causing downward displacement of the pycnocline, depressing the nutricline and upper euphotic biogeochemistry (McGillicuddy et al., 1999). We suggest that upper euphotic biology was displaced downward because of this

anticyclonic flow, with the islands acting as a physical barrier for the buildup and vertical depression of nearshore upper euphotic water masses. Nutrient concentrations did not vary significantly between upper euphotic-like and mesopelagic-like DCM communities, suggesting that changes in nutrient regimes was not a primary driver for differentiation of bacterioplankton communities within the DCM. Therefore, influence of an IME on the DCM was observed thorough downward displacement of less dense isopycnals with enhanced oxygen concentration and increased relative abundance of upper euphotic taxa.

In the mesopelagic, clear clustering of bacterioplankton taxa was observed, with ASVs displaying variation in relative abundance by depth (Figure 2.2a, b, 2.3). The vertical stratification of bacterioplankton taxa from the upper euphotic into the mesopelagic observed in the present study (Figure 2.3) is consistent with previous findings in the stratified ocean systems (Bryant et al., 2016; DeLong et al., 2006; Field et al., 1997; Giovannoni et al., 1990; Konstantinidis et al., 2009; Mende et al., 2017). While the mesopelagic microbial communities clearly separated between discrete depths, further multivariate analyses were not able to resolve any additional significant relationship with distance to shore. Thus, 16S rRNA gene metabarcoding as a tracer of IME was not useful at depths deeper than the DCM with the resolution of the present dataset.

Wide variability in core properties of microbial genomes such as GC content and genome size have been found among different archaea and bacteria (McCutcheon & Moran, 2011; Ochman & Davalos, 2006). A sharp transition zone in this genetic variability is reported to occur at the base of the euphotic zone correlating with changes in the surrounding energy and nutrient fields, supporting the hypothesis that nutrient limitation is a central driver in the evolution of genomic properties of bacterioplankton (Mende et al.,

2017). IMEs are often observed because of input of limiting nutrients into the surrounding environment, whether by riverine runoff (Anderson et al., 2002), advection by wave-forcing (Callaghan et al., 2006), groundwater discharge (Street et al., 2008), near-island upwelling (Gove et al., 2006; Hamner & Hauri, 1981), or turbulence from eddies and wakes on the downstream side of the island (Coutis & Middleton, 1999; Hernández-León, 1991; Heywood et al., 1990). Therefore, we suggest the influence of an IME is most obvious in populations above the nutricline where horizontal and vertical delivery of limiting nutrients can influence microbial activity and resulting community structure. Below the nutricline, presence of limiting nutrients in the upper euphotic such as nitrogen and phosphorus no longer become the primary limitation for microbial growth, and therefore addition of nutrients from processes near the islands may not alter microbial growth or composition.

The dark ocean (below 200 m depth) harbors approximately 65% of all bacterioplankton (Whitman et al., 1998), and dark ocean chemoautotrophic DIC fixation has been suggested to be in the same order of magnitude as heterotrophic activity, equaling 15-53% of phytoplankton export production in the North Atlantic (Reinthaler et al., 2010). Chemoautotrophic capabilities have been found in multiple common dark ocean taxa, including taxa abundant below the DCM in this dataset, such as ammonia-oxidizing Thaumarchaota (Wuchter et al., 2006), nitrite-oxidizing Nitrospinae (Lücker et al., 2013), and sulfur-oxidizing SAR324 (Swan et al., 2011). All these taxa were in low abundance or below detection in the upper 75 m and increased in abundance at or below the DCM. These taxa distribution patterns exemplify the shifting metabolic functions and nutritional strategies from photoautotrophy in the sunlit euphotic to chemoautotrophy in the dark ocean.

Conclusion

Our study utilizes 16S rRNA gene metabarcoding data to trace origins of observed physical and biogeochemical variability around the islands of Mo'orea and Tahiti. We present clear evidence of an IME, with reef bacterioplankton found in the oceanic environment at nearshore stations with enhanced Chl, NPP, POC, and BP. This suggests alteration of oceanic water by the island's reefs and the subsequent advection and retention of this microbially and biogeochemically altered water in the nearshore oceanic environment is likely a major mechanism contributing to the IME observed. Within the DCM, microbial communities with greater abundances of upper euphotic bacterioplankton were found almost exclusively in the nearshore environment. While the mechanisms for this observation are not clear, it is possible that anticyclonic flow around the island could provide a mechanism for downwelling into the DCM. Within the mesopelagic, no IME was observed. We hypothesize that island proximity no longer has a significant impact on the conditions necessary for microbial growth below the nutricline, as nutrients no longer become a limiting factor to growth. Our results demonstrate the utility of metabarcoding data as a biological tracer for physical and biogeochemical variability and provide further insight into the extent and influences of phenomena such as mesoscale perturbations (Nelson et al., 2014) and IMEs.

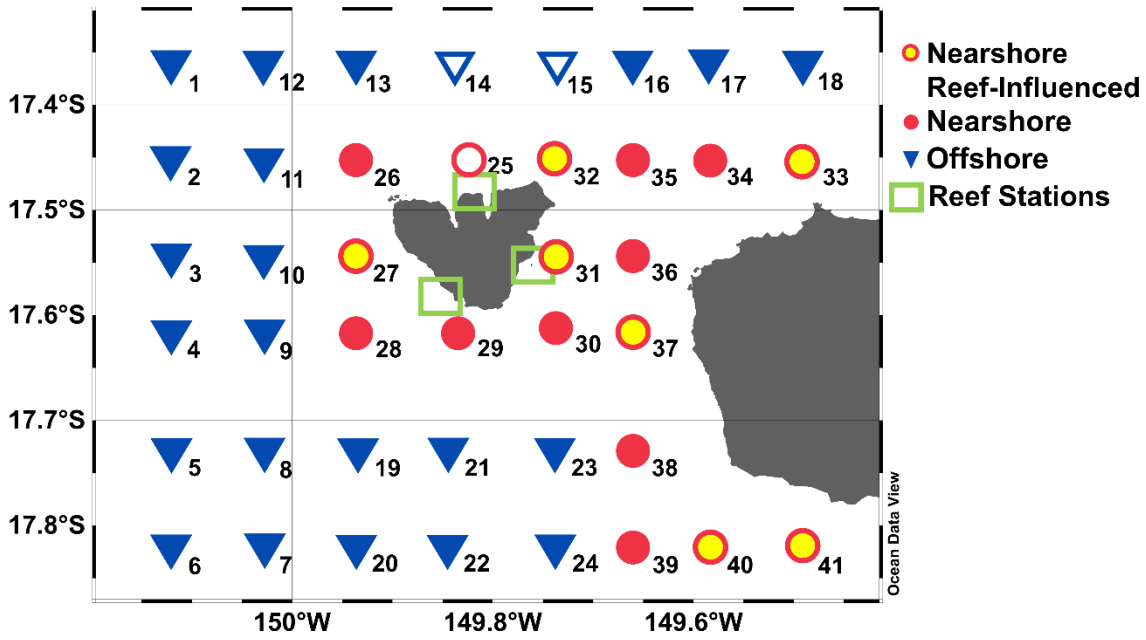


Figure 2.1. Map of study area around Mo'orea and Tahiti and shipboard sampling grid during July/August 2014. Symbol outline colors refer to spatial classifications following James et al (2020): Offshore stations (blue) and nearshore stations (red). Nearshore stations exhibiting influence of reef bacterioplankton are designated as "Nearshore Reef-Influenced" (yellow interior). An additional 46 Reef stations were sampled within the green rectangles (Map in Supplementary Figure 1). Offshore and nearshore stations with white interior represent stations with physical and biogeochemical measurements but no bacterioplankton community data.

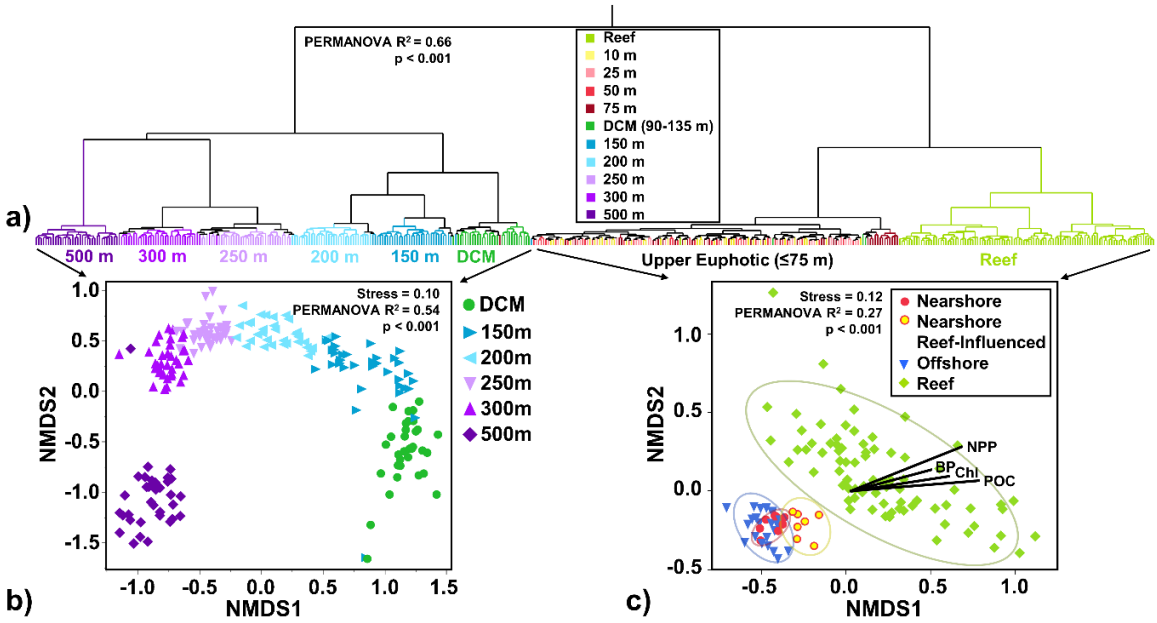


Figure 2.2: Community variation of bacterioplankton among sampling stations and depths. Hierarchical clustering of samples according to Bray-Curtis dissimilarity of community relative abundance profiles using Ward's Minimum Variance method show separation by depth and location (a). Non-metric multidimensional scaling (NMDS) ordination of bacterioplankton communities show separation by depth in samples subset from the deeper depth strata - DCM through the mesopelagic (b) - and NDMS ordination of communities from the upper 10 m and shallow reef environments with overlaid environmental vectors of net primary production (NPP). heterotrophic bacterioplankton production (BP), particulate organic carbon (POC), and chlorophyll *a* (Chl) (c).

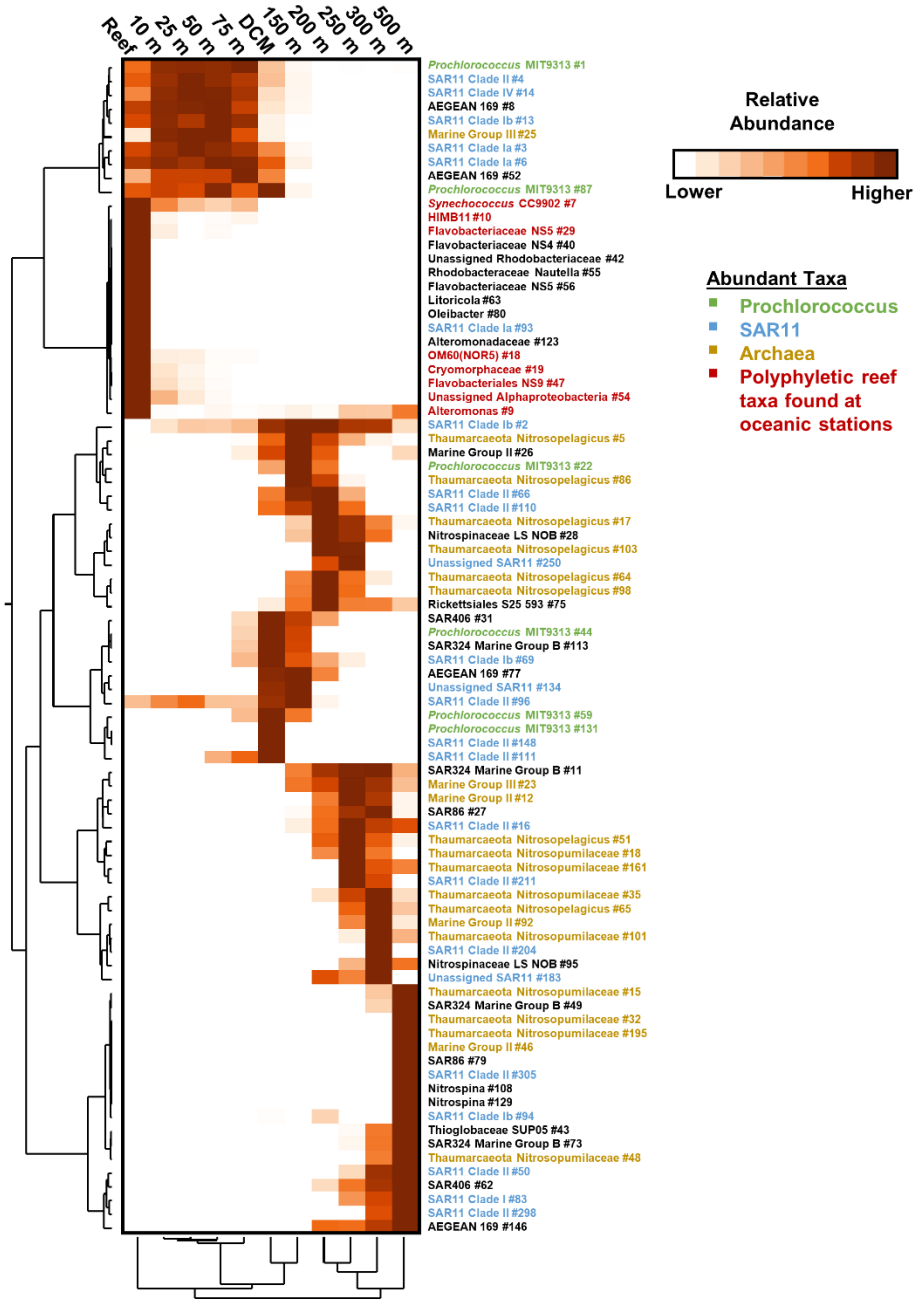


Figure 2.3: Depth enrichment stratification of abundant and widespread bacterioplankton ASVs. Heatmap of standardized mean relative abundance of bacterioplankton ASVs found in at least 3 samples with a relative abundance higher than 4%). ASVs and sampling locations are clustered using Ward's method as dendrograms at left and bottom, respectively. Ranges of relative abundances that correspond to the standardized heatmap coloration for each bacterioplankton ASV are provided in Appendix Figure 2.7.

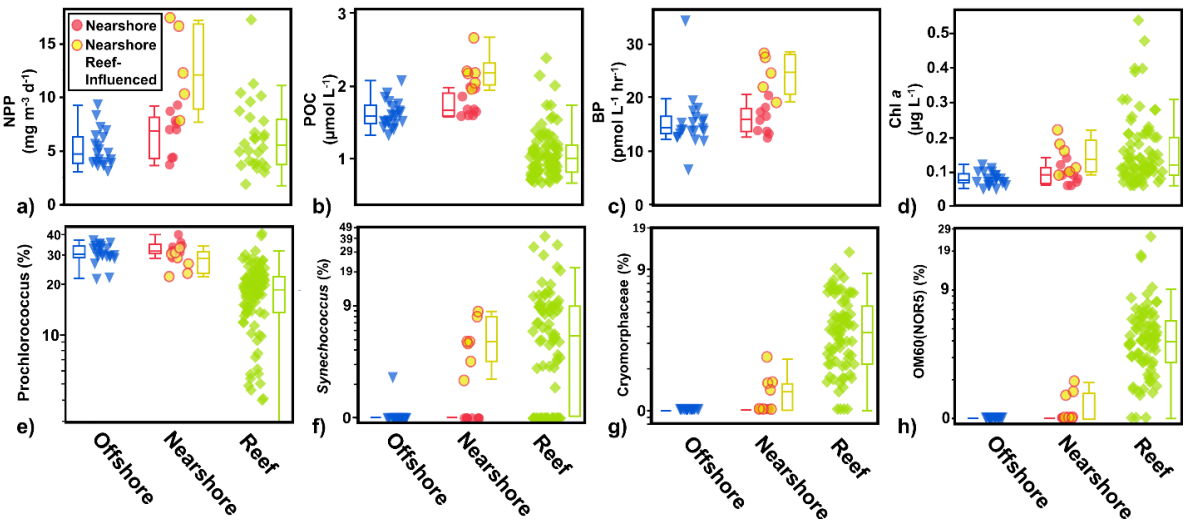


Figure 2.4: Distribution of key biogeochemical parameters and bacterioplankton taxa among Offshore, Nearshore and Reef surface water habitats. Each symbol is a sample from ≤10 m colored according to Figure 2.1. The top row comprises biogeochemical parameters: Net Primary Production (NNP; a) Particulate organic carbon (POC; b) Heterotrophic bacterial productivity (BP; c) and Chlorophyll *a* (Chl; d). The bottom row comprises selected abundant bacterioplankton taxa: *Prochlorococcus* (e) *Synechococcus* (f) Cryomorphaceae (g) and OM60(NOR5) (h).

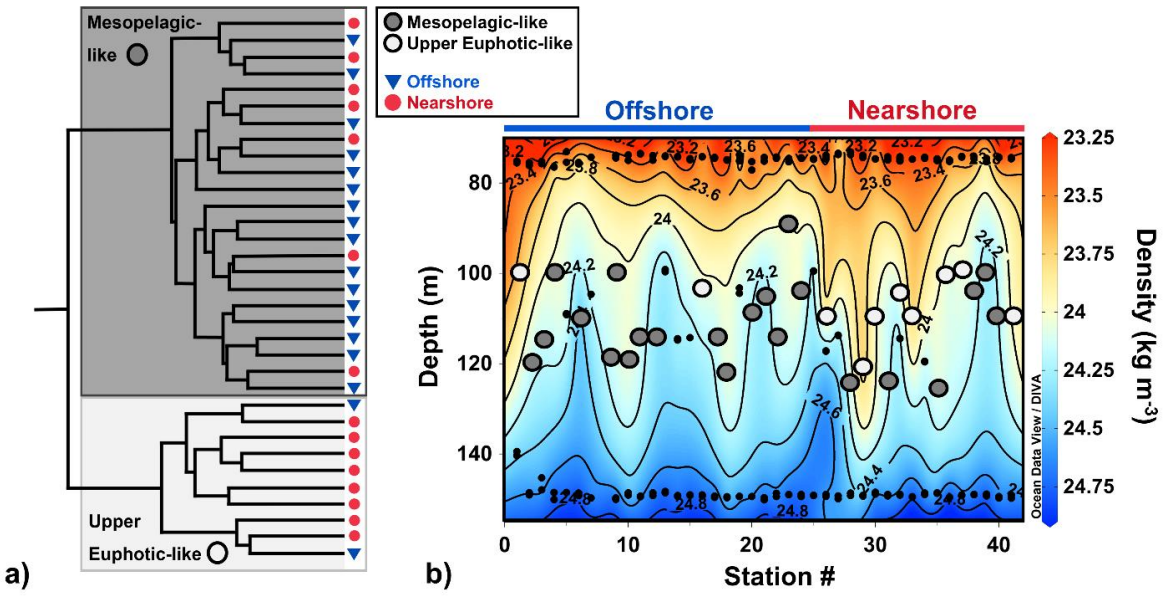


Figure 2.5: Deep Chlorophyll Maxima (DCM) samples cluster into two distinct groups according to proximity to nearshore habitats. (a) DCM samples clustered using Ward's method with Bray-Curtis dissimilarity distances; two distinct clusters are resolved and annotated according to their similarity to adjacent shallower ("upper-euphotic-like"; light grey) or deeper ("mesopelagic-like"; dark grey) depth communities. (b) Contour plot overlaying sample community types on density isobaths. Small black dots between 90 – 125 m show samples where density measurements were taken but prokaryote community data is absent. Light grey "upper euphotic-like" communities were found almost exclusively at nearshore (red) stations.

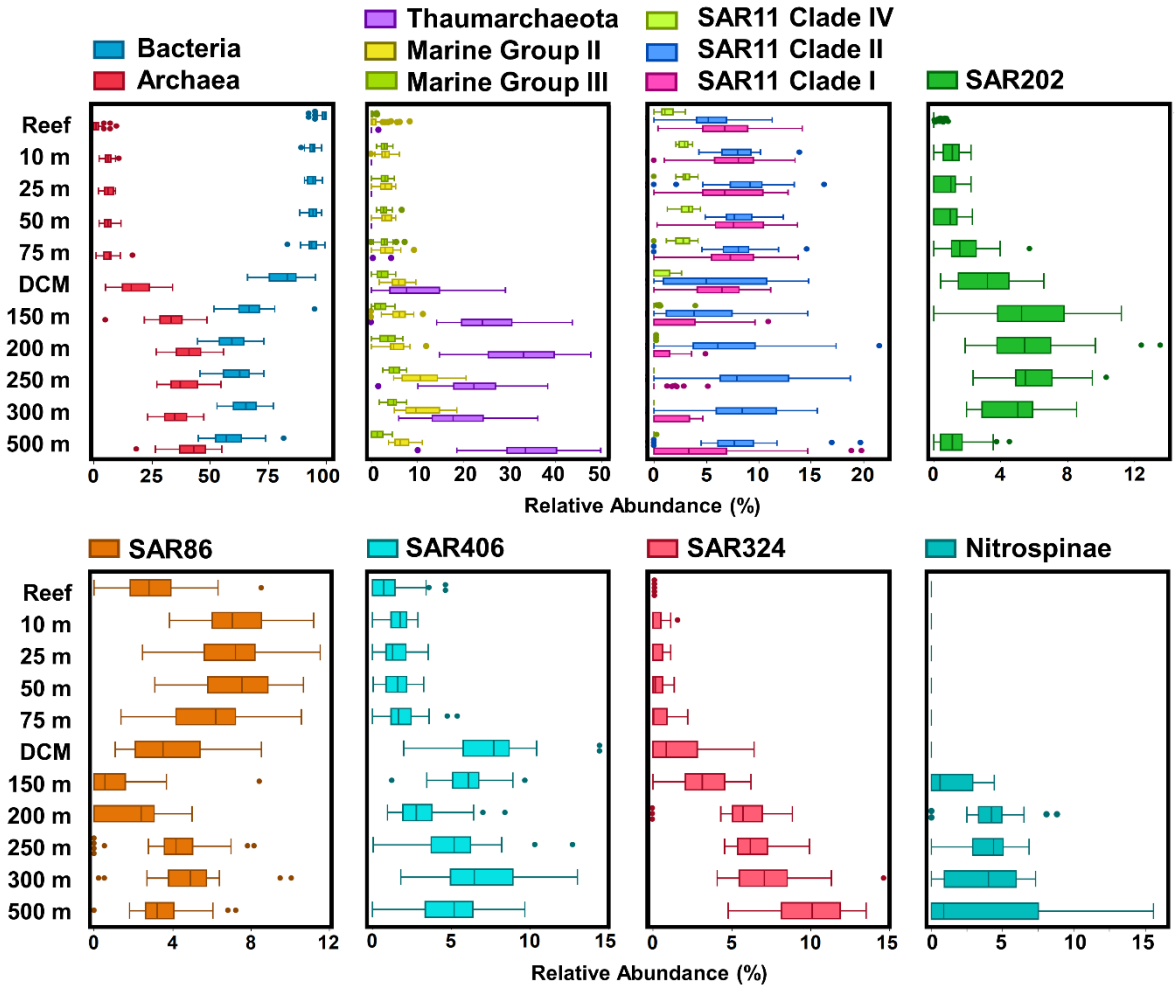


Figure 2.6: Relative abundances of major mesopelagic bacterioplankton taxa at each discrete depth horizon sampled.

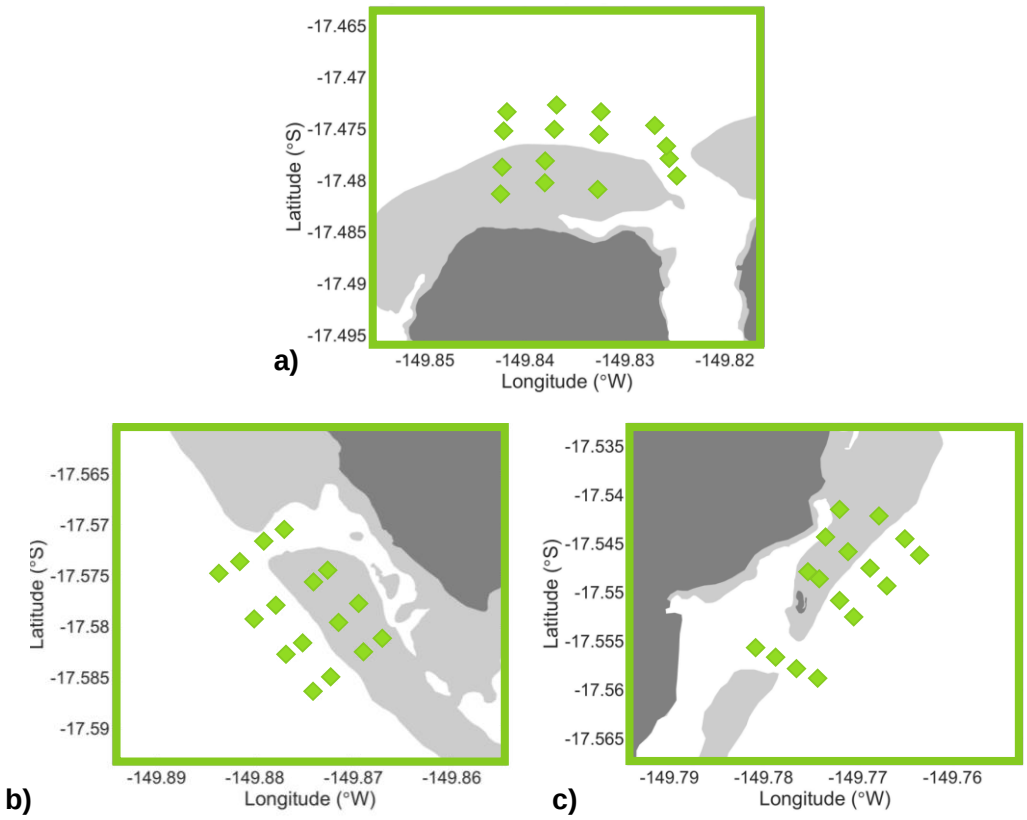
Amplicon Sequence Variant (ASV)	Reef Mean Abundance (%)	Oceanic (≤75 m) Mean Abundance (%)	T-test p-value	T-test adjusted p-value
Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_*Synechococcus*_CC9902_7	6.48 ± 6.17	0.51 ± 1.53	2.95×10^{-11}	3.06×10^{-10}
Proteobacteria_Gammaproteobacteria_AlteromoNAdales_Alteromonadaceae_Alteromonas_9	5.48 ± 6.07	0.04 ± 0.14	3.21×10^{-10}	2.80×10^{-9}
Proteobacteria_Alphaproteobacteria_Rhodobacterales_Rhodobacteraceae_HIMB11_10	5.31 ± 3.50	0.03 ± 0.17	4.26×10^{-19}	9.28×10^{-18}
Proteobacteria_Gammaproteobacteria_Cellvibrionales_Halieaceae_OM60(NOR5)_18	2.90 ± 1.92	0.03 ± 0.15	6.32×10^{-19}	1.25×10^{-17}
Bacteroidetes_Bacteroidia_Flavobacteriales_Cryomorphaceae_19	3.41 ± 2.42	0.03 ± 0.15	1.54×10^{-17}	2.80×10^{-16}
Bacteroidetes_Bacteroidia_Flavobacteriales_Flavobacteriaceae_NS5 marine group_29	1.87 ± 1.44	0.01 ± 0.10	4.83×10^{-16}	7.07×10^{-15}
Bacteroidetes_Bacteroidia_Flavobacteriales_NS9 marine group_47	1.93 ± 1.49	0.02 ± 0.11	4.86×10^{-16}	7.07×10^{-15}
Proteobacteria_Alphaproteobacteria_unclassified_54	1.39 ± 2.40	0.03 ± 0.18	1.37×10^{-5}	6.78×10^{-5}

Table 2.1: Reef ASVs with Reef/Oceanic averages and T-test p-values for Reef ASVs found in oceanic stations.

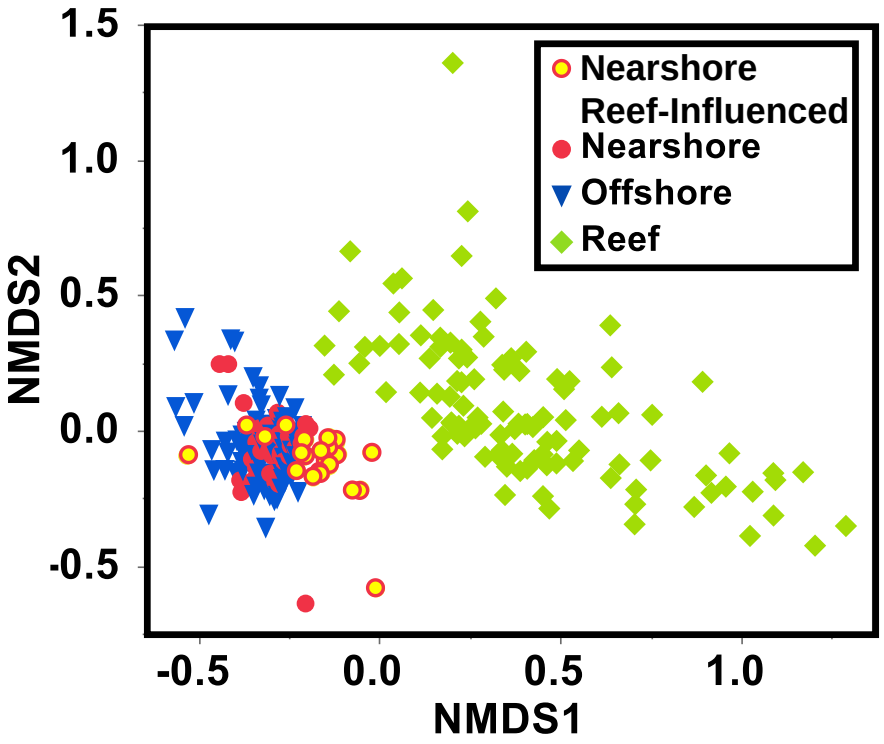
	10 m				Upper 75 m			
	Nearshore-Enhanced Mean	Nearshore Mean	Offshore Mean	ANOVA p-value	Nearshore-Enhanced Mean	Nearshore Mean	Offshore Mean	ANOVA p-value
Chl (µg L^{-1})	0.14 ± 0.05	0.09 ± 0.03	0.08 ± 0.01	< 0.0001	0.16 ± 0.04	0.10 ± 0.03	0.08 ± 0.02	< 0.0001
NPP (µM C d^{-1})	1.07 ± 0.34	0.55 ± 0.17	0.43 ± 0.14	< 0.0001	0.71 ± 0.42	0.38 ± 0.21	0.31 ± 0.17	< 0.0001
POC (µM C d^{-1})	2.19 ± 0.25	1.73 ± 0.16	1.60 ± 0.18	< 0.0001	2.10 ± 0.23	1.73 ± 0.13	1.61 ±0.17	< 0.0001
BP (µM C d^{-1})	0.07 ± 0.01	0.05 ± 0.01	0.05 ± 0.01	0.0012	0.07 ± 0.01	0.05 ± 0.01	0.05 ± 0.01	0.0005

Table 2.2: Mean concentrations and ANOVA p-values for chlorophyll (Chl), particulate organic carbon (POC), net primary production (NPP), and heterotrophic bacterioplankton production (BP) at 10 m and in the upper 75 m of offshore, nearshore, and nearshore-modified stations.

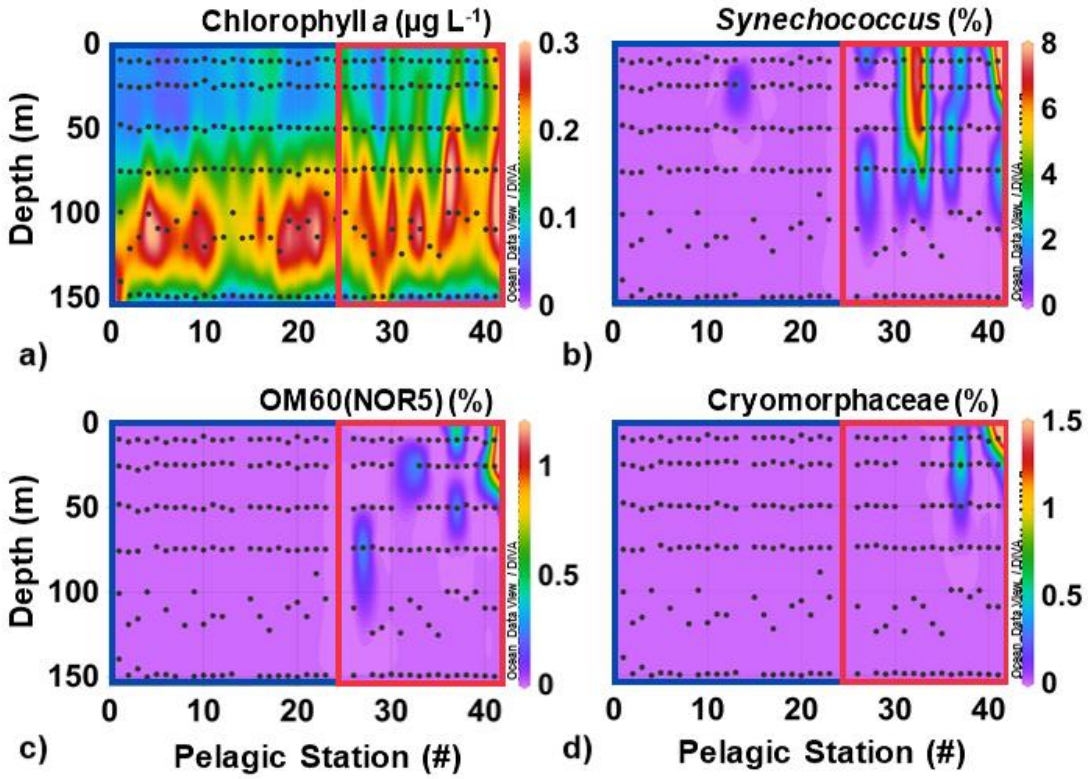
Appendix



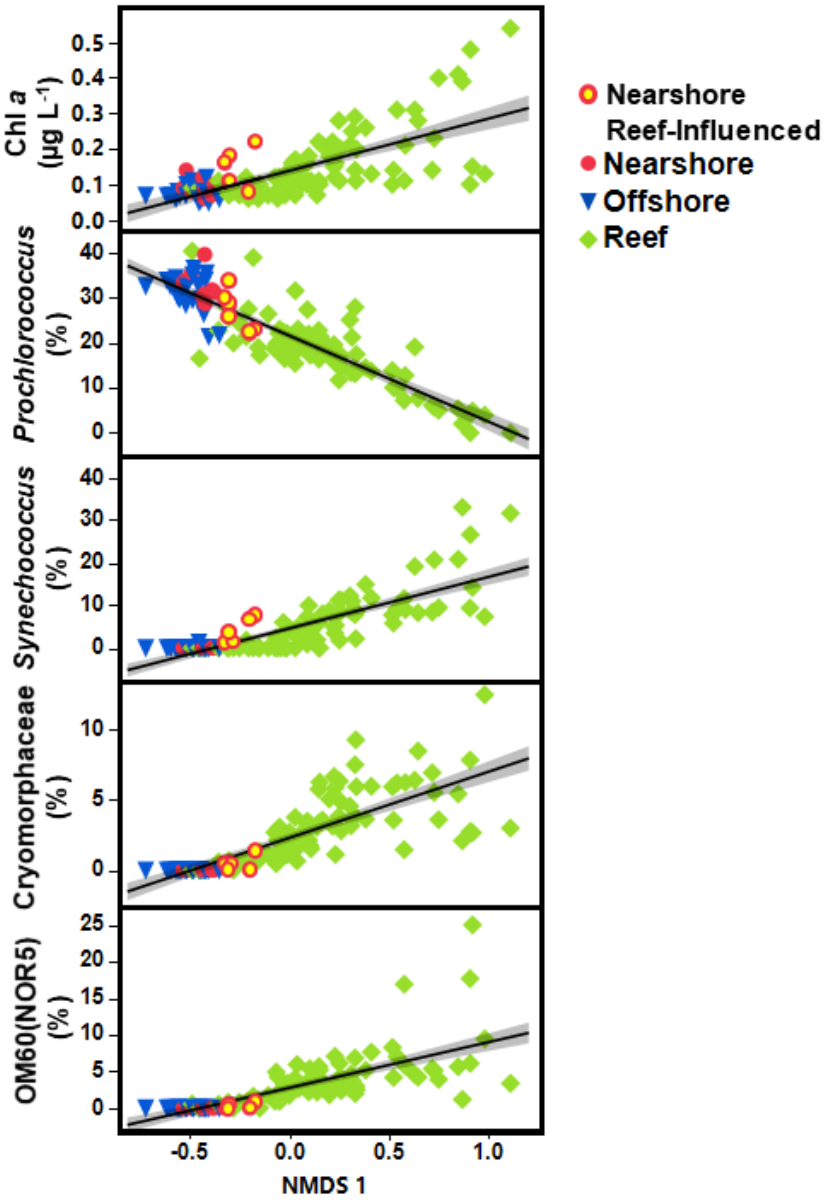
Appendix Figure 2.1: Reef sampling locations on the north (a), west (b), and east (c) sides of Mo'orea. Light grey shading indicates the reef platform, nominally 3 m deep. Dark grey shading indicates land.



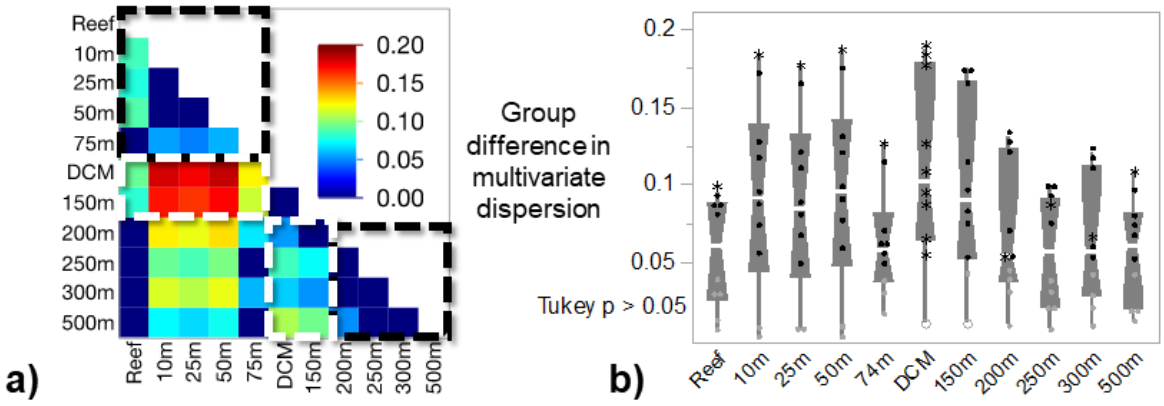
Appendix Figure 2.2: Non-metric multidimensional scaling (NMDS) ordination of bacterioplankton communities in the top 75 m in the reef and oceanic environments



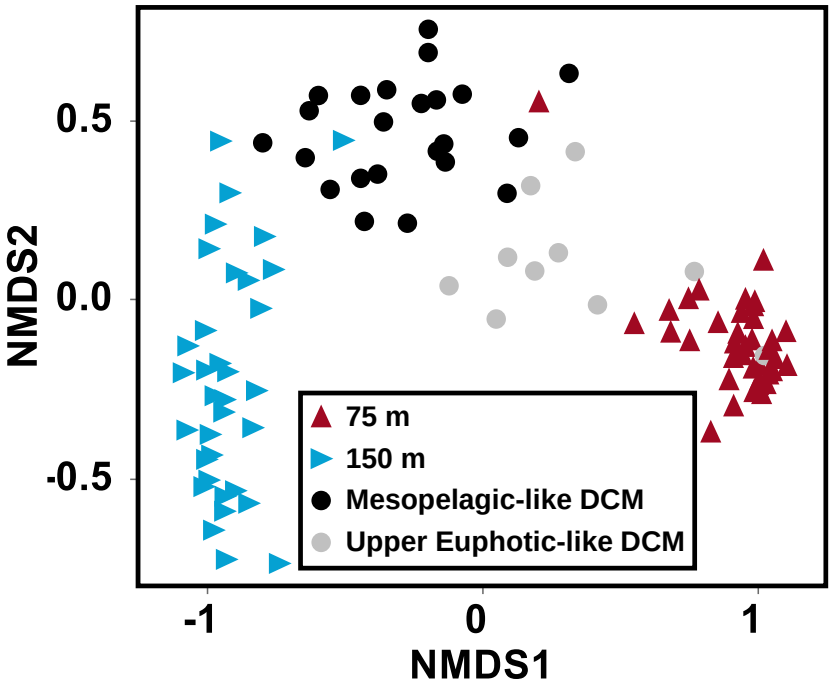
Appendix Figure 2.3: Contour plots of chlorophyll *a* (a) *Synechococcus* relative abundance (b) OM60(NOR5) relative abundance (c) and Cryomorphaceae (d) relative abundance in the upper 150 m of the water column in the pelagic environment. Black dots indicate sites of sample collection. Color boxes indicate offshore (blue) and nearshore (red) stations as in Figure 2.1.



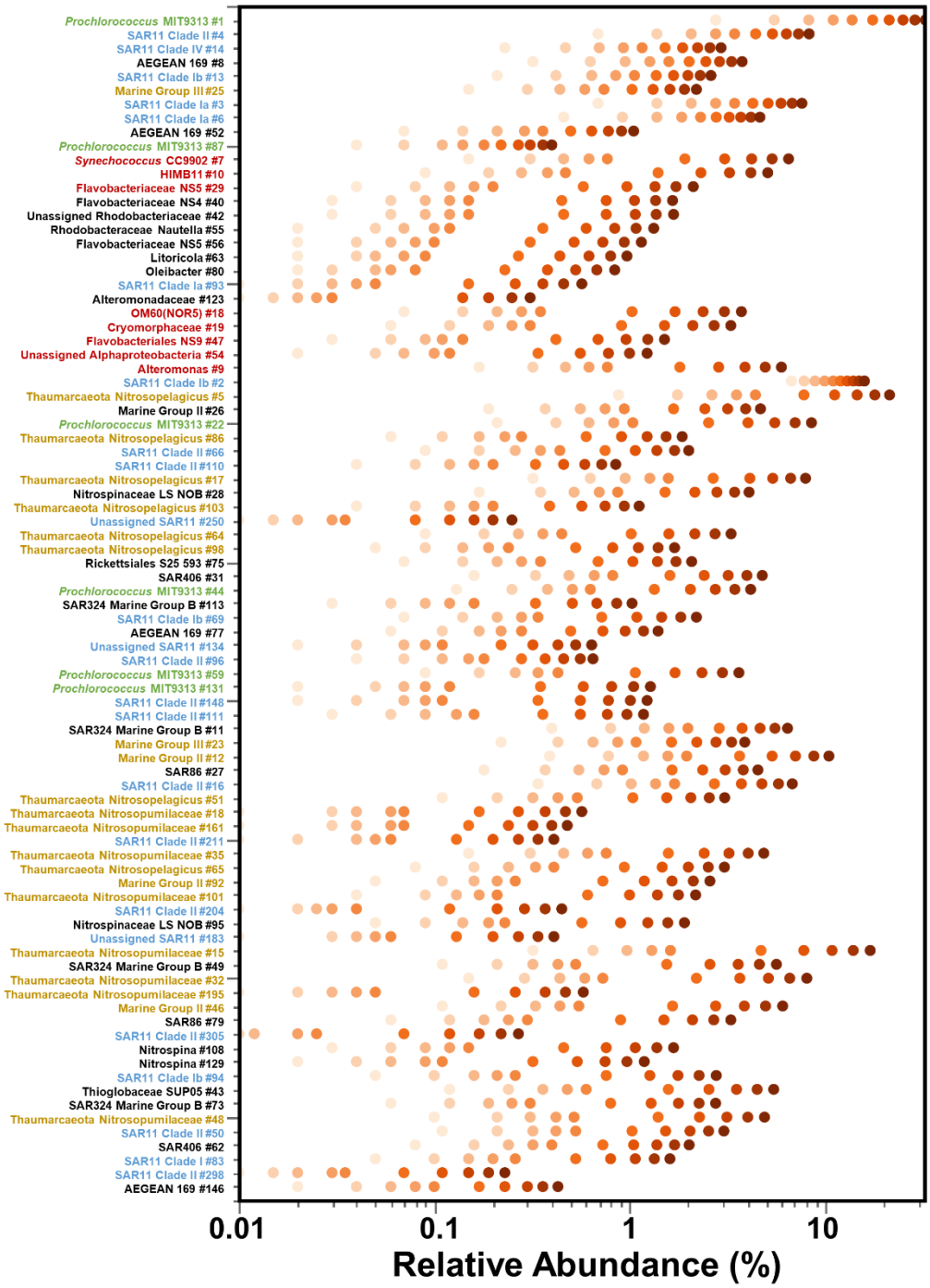
Appendix Figure 2.4: Change in chlorophyll *a* (Chl *a*), *Prochlorococcus*, *Synechococcus*, Cryomorphaceae, and OM60 (NOR5) in the surface 10 m along NMDS axis 1.



Appendix Figure 2.5: The DCM exhibits greater microbial community variation than other depth intervals. Panel (a): Heatmap of pairwise microbial community dispersion. While there were minimal differences in multivariate dispersion of microbial communities among upper euphotic samples (above 90 m) or among mesopelagic samples (below 200 m) [black dashed lines], dispersion at the DCM (90-135 m) and 150 m was significantly greater than all other samples [white dashed lines]. Panel (b): The difference in dispersion between DCM/150 m samples and other depth intervals was greater than differences among other depth intervals; for each depth pairwise differences in dispersion with other depths are represented as symbols (black if significant, grey if not significant by Tukey post hoc $p < 0.05$ on dispersion ANOVA) – asterisks represent pairwise differences from the DCM and open circles are the pairwise difference between DCM and 150 m (note that asterisks are generally the greatest pairwise difference in dispersion).



Appendix Figure 2.6: Non-metric multidimensional scaling (NMDS) ordination of bacterioplankton communities at 75 m, 150 m, and in the mesopelagic-like and upper euphotic-like groups in the deep chlorophyll maximum (DCM) communities. Upper euphotic-like DCM communities cluster more closely to communities at 75 m than 150 m, and mesopelagic-like DCM communities cluster more closely to communities at 150



Appendix Figure 2.7: Ranges of relative abundances that correspond to the standardized heatmap coloration for each bacterioplankton ASV shown in Figure 2.3.

Amplicon Sequence Variant (ASV)	Reef	10 m	25 m	50 m	75 m	DCM	150 m	200 m	250 m	300 m	500 m
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_1	17.79 ± 7.74	30.95 ± 4.12	31.17 ± 4.63	30.97 ± 4.98	32.87 ± 6.41	6.96 ± 10.38	0.85 ± 5.01	0 ± 0	0.07 ± 0.34	0 ± 0	0.18 ± 0.62
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ib_2	5.68 ± 2.31	8.99 ± 1.73	8.18 ± 2.59	8.06 ± 2.21	8.83 ± 3.52	15.19 ± 6.22	15.94 ± 4.53	15.94 ± 4.42	14.72 ± 4.58	14.65 ± 3.96	7.31 ± 5.6
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ia_3	5.65 ± 2.04	7.05 ± 1.59	7.63 ± 1.28	7.28 ± 1.5	6.7 ± 1.7	3.38 ± 2.64	0.22 ± 1.29	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_4	5.13 ± 2.36	7.81 ± 2.02	8.3 ± 3.07	7.82 ± 1.97	6.89 ± 2.53	1.87 ± 2.71	0.3 ± 0.99	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_5	0.02 ± 0.16	0 ± 0	0 ± 0	0 ± 0	0.13 ± 0.72	9.36 ± 7.28	21.43 ± 6.82	14.09 ± 5.5	2.95 ± 1.9	0.46 ± 0.66	0.05 ± 0.32
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ia_6	4.04 ± 1.79	4.48 ± 1.47	4.19 ± 1.73	4.65 ± 1.39	4.8 ± 1.75	3.13 ± 2.64	0.29 ± 1.2	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Synechococcus_CC9902_7	6.52 ± 7.32	0.8 ± 1.9	0.43 ± 1.53	0.3 ± 1.32	0.46 ± 1.25	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodospirillales_AEGEAN-169 marine group_8	2.97 ± 1.6	3.84 ± 0.89	3.71 ± 0.84	3.77 ± 0.89	2.9 ± 1.18	0.36 ± 0.62	0.1 ± 0.58	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Gammaproteobacteria_AlteromoNAdales_Alteromonadaceae_Alteromonas_9	5.99 ± 6.2	0.04 ± 0.16	0.02 ± 0.09	0.08 ± 0.19	0.03 ± 0.11	0.06 ± 0.16	0.11 ± 0.44	0.05 ± 0.18	0.38 ± 0.71	0.37 ± 0.93	1.32 ± 3.22
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodobacterales_Rhodobacteraceae_HIMB11_10	5.1 ± 3.43	0.04 ± 0.19	0 ± 0	0.02 ± 0.09	0.05 ± 0.27	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Deltaproteobacteria_SAR324 (Marine group B)_11	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	2.24 ± 2.22	5.51 ± 2.03	6.38 ± 1.18	6.44 ± 1.94	1.33 ± 1.15
Archaea_Euryarchaeota_Thermoplasmata_Marine Group II_12	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	2.69 ± 3.22	10.49 ± 4.15	8.19 ± 3.7	0.18 ± 0.76
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ib_13	1.83 ± 2	2.49 ± 2.08	2.2 ± 2.25	2.81 ± 2.39	2.4 ± 2.48	0.13 ± 0.52	0.01 ± 0.03	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade IV_14	1.17 ± 0.77	2.79 ± 0.65	2.94 ± 0.7	2.92 ± 0.7	2.55 ± 1.2	0.34 ± 0.71	0.09 ± 0.51	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_15	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.77 ± 2.96	17.06 ± 7.06
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_16	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.24 ± 0.81	2.76 ± 3.34	6.81 ± 4.34	4.98 ± 3.28	3.8 ± 3.3
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_17	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.67 ± 1.34	7.06 ± 2.88	6.92 ± 2.47	1.77 ± 2.01	0.1 ± 0.62
Bacteria_Proteobacteria_Gammaproteobacteria_Cellvibrionales_Halieaceae_OM60(NOR5)_18	3.74 ± 3.55	0.05 ± 0.19	0.05 ± 0.2	0.01 ± 0.06	0.01 ± 0.06	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_Cryomorphaceae_19	3.3 ± 2.52	0.07 ± 0.25	0.03 ± 0.14	0.01 ± 0.04	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_22	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.79 ± 2.68	8.54 ± 7.91	2.11 ± 2.91	0 ± 0	0 ± 0	0 ± 0
Archaea_Euryarchaeota_Thermoplasmata_Marine Group III_23	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.46 ± 1.64	2.59 ± 1.85	3.89 ± 1.3	3.38 ± 1.76	0.6 ± 1.18
Archaea_Euryarchaeota_Thermoplasmata_Marine Group III_25	0.15 ± 0.3	2.14 ± 1.05	2.2 ± 1.22	2.2 ± 0.91	1.37 ± 1.35	0.09 ± 0.52	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Euryarchaeota_Thermoplasmata_Marine Group II_26	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.14 ± 0.43	2.87 ± 1.49	4.68 ± 1.74	2.28 ± 2.55	0 ± 0	0 ± 0	0.32 ± 0.59
Bacteria_Proteobacteria_Gammaproteobacteria_SAR86_27	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.04 ± 0.26	1.62 ± 1.75	3.88 ± 1.94	4.54 ± 1.84	0.09 ± 0.51
Bacteria_Nitrospinae_Nitrospinia_NitrospiNAles_Nitrospinaceae_LS-NOB_28	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.44 ± 1.13	4.09 ± 1.89	3.56 ± 1.8	1.48 ± 2.18	0 ± 0
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_Flavobacteriaceae_NS5 marine group_29	2.12 ± 1.93	0.03 ± 0.18	0 ± 0	0.01 ± 0.05	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Marinimicrobia_SAR406 clade_31	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.27 ± 0.66	4.78 ± 2.48	3.35 ± 1.06	0.85 ± 0.87	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_32	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	8.08 ± 2.01
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_35	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.2 ± 0.45	3.11 ± 1.24	4.87 ± 2.31	0.22 ± 1.32
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_Flavobacteriaceae_NS4 marine group_40	1.66 ± 1.51	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodobacterales_Rhodobacteraceae_42	1.66 ± 2.93	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Gammaproteobacteria_Thiomicrospirales_Thioglobaceae_SUP05_cluster_43	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.03 ± 0.17	1.13 ± 0.99	5.47 ± 1.71
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_44	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.24 ± 1.47	4.17 ± 4.69	2.57 ± 3.69	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Euryarchaeota_Thermoplasmata_Marine Group II_46	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	6.08 ± 1.84
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_NS9 marine group_47	1.51 ± 1.45	0.04 ± 0.19	0.02 ± 0.11	0 ± 0.03	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_48	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.84 ± 1.37	4.91 ± 2.03
Bacteria_Proteobacteria_Deltaproteobacteria_SAR324 (Marine group B)_49	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.2 ± 0.76	5.65 ± 1.6
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_50	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.17 ± 1.03	2.6 ± 3.8	3.04 ± 2.88
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_51	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.37 ± 1.75	3.07 ± 1.17	1.41 ± 1.21	0.08 ± 0.37	
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodospirillales_AEGEAN-169 marine group_52	0.24 ± 0.59	0.76 ± 1.4	0.76 ± 1.39	0.76 ± 1.3	1.05 ± 1.57	0.37 ± 1.05	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_unclassified_54	1.23 ± 2.08	0.08 ± 0.33	0.02 ± 0.14	0.01 ± 0.03	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodobacterales_Rhodobacteraceae_Nautella_55	1.33 ± 4.09	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_Flavobacteriaceae_NS5 marine group_56	1.15 ± 1.5	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_59	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.24 ± 1.49	3.83 ± 4.44	0.84 ± 2.4	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Marinimicrobia_SAR406_62	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.13 ± 0.44	0.63 ± 1.23	1.66 ± 1.59	2.02 ± 1.34
Bacteria_Proteobacteria_Gammaproteobacteria_Oceanospirillales_Litoricolaceae_Litoricola_63	0.97 ± 2.05	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_64	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.55 ± 0.9	3.31 ± 1.13	0.97 ± 1.41	0.08 ± 0.35	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_65	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.22 ± 2.41	3.06 ± 3.43	0.2 ± 0.83
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_66	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.56 ± 2.26	1.92 ± 3.34	2.01 ± 2.54	0.3 ± 0.85	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ib_69	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.2 ± 0.88	2.2 ± 2.7	1.08 ± 2.08	0.25 ± 0.75	0.05 ± 0.29	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Deltaproteobacteria_SAR324 (Marine group B)_73	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.03 ± 0.15	0.62 ± 0.82	2.76 ± 1.24
Bacteria_Proteobacteria_Alphaproteobacteria_Rickettsiales_SCV-693_75	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.08 ± 0.2	0.61 ± 1.03	2.08 ± 1.87	0.48 ± 1.04	0.44 ± 1.01	0.18 ± 0.5
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodospirillales_AEGEAN-169 marine group_77	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.34 ± 1.94	1.4 ± 2	0.31 ± 1.03	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Gammaproteobacteria_SAR86_79	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	3.3 ± 1.5
Bacteria_Proteobacteria_Gammaproteobacteria_Oceanospirillales_Saccharospirillaceae_Oleibacter_80	0.84 ± 1.75	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade I_83	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.23 ± 0.99	0.97 ± 1.89	1.61 ± 1.74
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_86	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.86 ± 2.2	1.23 ± 1.73	0.03 ± 0.15	0 ± 0	0 ± 0
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_87	0.26 ± 0.91	0.31 ± 0.23	0.29 ± 0.23	0.4 ± 0.22	0.26 ± 0.25	0.4 ± 1.23	0.02 ± 0.09	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Euryarchaeota_Thermoplasmata_Marine Group II_92	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.28 ± 0.61	2.58 ± 1.12	0.05 ± 0.3
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ia_93	0.57 ± 1.62	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade Ib_94	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.01 ± 0.03	0 ± 0	0.1 ± 0.46	0 ± 0	0 ± 0	2.78 ± 4.84
Bacteria_Nitrospinae_Nitrospinia_Nitrospinales_Nitrospinaceae_LS-NOB_95	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.15 ± 0.87	1.91 ± 2.31	0.49 ± 0.81
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade I_96	0.1 ± 0.52	0.18 ± 0.76	0.29 ± 1.01	0.1 ± 0.44	0.09 ± 0.67	0.58 ± 1.42	0.65 ± 1.15	0.02 ± 0.09	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_98	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.34 ± 0.91	1.7 ± 1.74	0.56 ± 1.06	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_101	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.03 ± 0.2	2.18 ± 2.95	0.14 ± 0.68
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_Candidatus Nitrosopelagicus_103	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.12 ± 2.09	1.11 ± 2.04	0 ± 0	0 ± 0
Bacteria_Nitrospinae_Nitrospinia_Nitrospinales_Nitrospinaceae_Nitrospina_108	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.68 ± 3.88
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_110	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.33 ± 1.89	0.64 ± 2.67	0.85 ± 2.37	0.33 ± 1.97	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_111	0 ± 0	0 ± 0	0 ± 0	0.11 ± 0.83	0.44 ± 1.86	1.18 ± 3.25	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Deltaproteobacteria_SAR324 (Marine group B)_113	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.07 ± 0.37	1.03 ± 1.33	0.61 ± 1.31	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Gammaproteobacteria_Alteromonadales_Alteromonadaceae_123	0.31 ± 1.13	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Nitrospinae_Nitrospinia_NitrospiNAles_Nitrospinaceae_Nitrospina_129	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.19 ± 3.43
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_131	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.28 ± 3.1	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_unclassified_134	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.58 ± 1.34	0.64 ± 1.81	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodospirillales_AEGEAN-169 marine group_146	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.18 ± 1.04	0.17 ± 1	0.33 ± 1.37	0.43 ± 1.48
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_148	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	1.23 ± 3.43	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_161	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.48 ± 1.23	0.23 ± 0.67	0.09 ± 0.52
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_181	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.07 ± 0.39	0.57 ± 1.46	0.14 ± 0.81	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_unclassified_183	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.21 ± 1.23	0.07 ± 0.44	0.41 ± 1.75	0 ± 0
Archaea_Thaumarchaeota_Nitrososphaeria_Nitrosopumilales_Nitrosopumilaceae_195	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.58 ± 1.46
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_204	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.46 ± 1.92	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_211	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.41 ± 1.71	0.24 ± 1.1	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_unclassified_250	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.14 ± 0.79	0.25 ± 1.46	0 ± 0	0 ± 0
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_298	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.12 ± 0.72	0.23 ± 1.37
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_305	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0.27 ± 1.61

Appendix Table 2.1: Mean Relative abundances and standard deviations for ASVs with an abundance greater than 4% in 3 samples, or 8% in one sample.

Amplicon Sequence Variant (ASV)	Mesopelagic-like DCM abundance (%)	Upper Euphotic-like DCM abundance (%)	T-test p-value	T-test adjusted p-value
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_1	2.06 ± 4.67	18.23 ± 11.30	0.0013	0.0205
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade II_4	0.58 ± 1.17	4.83 ± 2.98	0.0013	0.0205
Bacteria_Proteobacteria_Alphaproteobacteria_Rhodospirillales_AEGEAN-169 marine group_8	0.07 ± 0.15	1.02 ± 0.79	0.0043	0.0459
Bacteria_Proteobacteria_Alphaproteobacteria_SAR11_Clade IV_14	0.00 ± 0.00	1.13 ± 0.90	0.0034	0.0459
Bacteria_Marinimicrobia_SAR406 clade_31	5.65 ± 2.18	2.78 ± 1.88	0.0010	0.0205
Bacteria_Cyanobacteria_Oxyphotobacteria_Synechococcales_Cyanobiaceae_Prochlorococcus_MIT9313_59	4.83 ± 4.54	0.87 ± 2.75	0.0046	0.0459
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_Flavobacteriaceae_NS5 marine group_139	1.32 ± 1.15	0.14 ± 0.43	0.0002	0.0168
Bacteria_Bacteroidetes_Bacteroidia_Flavobacteriales_Flavobacteriaceae_NS2b marine group_151	1.27 ± 1.51	0.00 ± 0.00	0.0006	0.0168
Bacteria_Chloroflexi_Dehalococcoidia_SAR202_154	0.58 ± 0.58	0.06 ± 0.18	0.0005	0.0168
Bacteria_Proteobacteria_Gammaproteobacteria_SAR86_198	0.61 ± 0.74	0.00 ± 0.00	0.0006	0.0168

Appendix Table 2.2: Mean relative abundances for DCM ASVs enhanced in Mesopelagic-like and Upper Euphotic-like stations with T-test p-values.

Group	Average density (kg m^{-3})	Oxygen T-test p-value	Average DOC (µM C)	pairwise PERMANOVA with upper 75 m R^2	pairwise PERMANOVA with upper 75 m adjusted p-value	pairwise PERMANOVA with 150 m R^2	pairwise PERMANOVA with 150 m adjusted p-value
Mesopelagic-Like	24.14 ± 0.11	196.7 ± 2.7	62.1 ± 4.4	0.45	0.0006	0.21	0.0006
Euphotic-like	23.90 ± 0.17	199.7 ± 2.1	66.0 ± 3.5	0.19	0.0006	0.24	0.0006

Appendix Table 2.3: Mean density, oxygen, and DOC concentrations alongside pairwise PERMANOVA R^2 and p-values for the mesopelagic-like and upper euphotic-like DCM groups.

III. Wave forcing is a primary driver of coral reef bacterioplankton communities around Mo'orea, French Polynesia

Abstract

The presence of highly productive coral reefs within the unproductive oligotrophic waters of subtropical gyres implies intensive recycling and retention of both macronutrients and organic matter by a unique microbial community. Identifying and characterizing the unique bacterioplankton communities in reef waters is important for understanding reef metabolism and productivity. In this study, we conducted a high-resolution synoptic oceanographic profile survey of the offshore and nearshore reef waters surrounding Mo'orea, French Polynesia, coupling 16S rRNA gene metabarcoding of microbial communities with a suite of biogeochemical measurements and physical oceanographic parameters. We document consistent differences in bacterioplankton communities among forereef, backreef, reef passes and offshore surface waters, characterizing the key taxa that define the reef planktonic microbiome. Notably, within our dataset we find variation in wave energy regime, captured across three sides of Mo'orea, to be the largest driver of bacterioplankton community structure, with more than three times more variation captured between the island the north, west and east side than between reef habitats i.e. forereef, backreef, reef pass. Our data provide insights into the physical and biogeochemical drivers of reef bacterioplankton communities in the context of a well-studied long-term ecological research site.

Introduction

In tropical and subtropical waters worldwide, coral reefs are one of the most abundant and productive coastal ecosystems despite being nestled within some of the most oligotrophic and unproductive waters in the ocean (Longhurst et al., 1995). Coral reef systems display unique chemistry and biology when compared to the surrounding open waters, including elevated chlorophyll *a*, elevated primary production, elevated animal biomass, and unique biogeochemistry (Andrade et a. 2014; Dandonneau & Charpyt, 1985; Doty & Oguri, 1956; Gove et al., 2016; Jones, 1962; Martinez & Maamaatuaiahutapu, 2004; Palacios, 2002; Raapoto et al., 2019; Signorini et al., 1999; Vollbrecht et al., 2021). Transport and retention of key biogeochemical constituents such as inorganic nutrients and organic particles within the reefs are thought to support the observed elevated biomass and productivity (Johannes et al., 1972; Odum and Odum, 1995). To support its high productivity, there is a long-standing hypothesis that coral reefs display rapid biogeochemical cycling to support nutrient recycling and retention (Duarte & Cebrian, 1996; Odum & Odum, 1955). In these reef ecosystems, microbially dominated food webs are largely supported by rapid recycling of dissolved compounds back into inorganic nutrients by bacterioplankton (Nelson, Wegley Kelly & Haas, 2023 and references therein). Understanding the sources and controls on organic matter and nutrients within this ecosystem is central to predicting and managing how these ecosystems will respond to global change (Sorokin, 1990).

In both the open ocean and reef environment, microbial communities play a central role in biogeochemical cycling, with bacterioplankton recycling more than half of the ocean's net primary production (Williams, 1984; Cho and Azam, 1990; Ducklow, 1990). Tropical reef ecosystems support diverse and active microbial communities both directly

associated with corals and in the surrounding water column (Ducklow, 1990). Research over the past decade has aimed to characterize reef bacterioplankton communities both spatially and temporally (e.g Nelson et al 2011, 2013; Kelly et al., 2019), however relationships between bacterioplankton communities and reef biogeochemical cycling and physical dynamics remain understudied. Developing a clear understanding of the role reef bacterioplankton communities play in biogeochemical cycling over the reef requires that we better resolve gradients of both bacterioplankton communities and biogeochemical variables across these systems.

The Mo'orea Coral Reef Long Term Ecological Research (MCR-LTER) site was established in 2004 and has since regularly collected data on physical, biological, and chemical properties, making it a well-characterized and highly studied coral reef system. This system therefore provides an excellent opportunity to investigate and resolve gradients in bacterioplankton, biogeochemistry, and physical dynamics in the context of decades of characterization. Previous studies have demonstrated differences in both biogeochemical constituents and bacterioplankton community structure between the various reef environments of Mo'orea, French Polynesia (Nelson et al., 2011, Haas et al., 2011, McCliment et al., 2012; Leichter et al., 2013). We hypothesized that patterns of bacterioplankton community differentiation between reef habitats relative to the open ocean are consistent among a diversity of physical regimes and sought to test this hypothesis by conducting high spatial resolution sampling across three well-characterized reefs around Mo'orea subject to different physical oceanographic regimes. We synoptically sampled water from multiple depths at 47 locations spanning forereef, backreef, and reef pass environments on the reefs of three sides around the island of Mo'orea. Water was

additionally collected from 10 depths in the upper 500 m of 41 oceanic stations surrounding the island over the same 9-day period. A suite of physical and biogeochemical measurements were taken in conjunction with bacterioplankton DNA. Physical measurements allowed us to derive differential magnitudes of wave forcing across the three island sides investigated. We describe the distribution of bacterioplankton communities between reef environments (sides) and across reef habitats (forereef, backreef and reef pass) that experience varying physical forcing regimes. The resulting analyses provide insight into drivers of community structure and function within the reef environment.

Methods

Site Description

This study was conducted within a 51 km by 67 km grid around the Society Island of Mo'orea. Mo'orea is a 1.5 - 2 million year old high volcanic island surrounded by barrier reefs within 0.5-1 km of shore (Neall and Trewick, 2008). There are fringing reefs bordering the island (~10 m deep) and a shallow (<3 m) backreef lagoon with a mixture of coral, algae, and barren sands. The backreef lagoons connect to open ocean via reef pass channels that occur at approximately 5-10 km intervals around the island. These reef pass channels often correspond with embayments of varying sizes. The forereef has a high coral density and drops steeply to depths exceeding 500 m within 1 km offshore from the reef crest. Waves drive water from the forereef across the reef crest into the backreef lagoon. Water then flows parallel to the island into the reef pass channels and is advected offshore through the pass. Movement within the reef environment is rapid (averaging 0.2 m s^{-1} with negligible tidal

influence) and maintains a short residence time of hours to days within the backreef environment (Hench et al., 2008; Knee et al. 2016).

Sampling Scheme

Water samples were collected at forty-six stations within the reef environment on the north, west, and east sides of Mo'orea from the surface to as deep as 25 m. Reef stations were sampled with a horizontal Niskin bottle using a rope and weighted messenger, with collections at two to three depths spanning the entire water column of the forereef, backreef, and reef pass environments (nominally 1, 10, and 25 m in forereef habitats; in backreef habitats bottom samples were collected 20 cm above the benthos, typically at 2-3 m depth). Biogeochemical variables measured included dissolved inorganic nutrients (nitrate, nitrite, ammonium, and phosphate), net primary production (NPP), chlorophyll *a* (Chl), phaeopigments, heterotrophic bacterial productivity (BP), bacterial abundance (BA), dissolved organic carbon (DOC), particulate organic carbon (POC), and particulate organic matter nitrogen isotopes (POM-δ^{15} N). A detailed description of these data and standard operating procedures are available in Nelson et al. (in prep) and James et al., (2020).

Physical flow rate characterization

Wave data have been collected continuously since 2005 as part of the Mo'orea Coral Reef Long-Term Ecological Research (MCR-LTER) project. These data are collected at three sites on the exposed forereef of Mo'orea, with one site on each of the island's three sides (Washburn, 2019). At each site, data from bottom-mounted Wave & Tide Recorders

(SBE 26 plus: Sea-Bird Electronics, Bellevue, Washington, USA) and bottom-mounted Acoustic Doppler Current Profilers (Sentinel ADCPs: Teledyne RD Instruments, Poway, California, USA) were used to estimate significant wave height (H_s) and dominant wave period (T_p) at 2-h sampling intervals. Daily averages of H_s and T_p were then calculated and total wave power (P) per linear meter of reef was estimated as in Adam et al. (2020) using deep-water approximations as $P=\rho g^2 H_s^2 T_p/32\pi$ where ρ is seawater density (1,025 kg m^{-3}) and g is the acceleration of gravity (9.81 m/s^2). Data is available at mcrlter.msi.ucsb.edu/data/db/physoDataCoverage.php.

DNA Collection & Extraction

Two liters of whole seawater was collected into in a 5L HDPE carboy held in a chilled cooler from each station and concentrated by gravity onto a 0.22 µm pore-size 47 mm polyethyersulfone flat filter (Pall Supor) housed in a polycarbonate filter cartridge after transfer to shore. All filters were stored dry at -40°C until further processing. DNA was extracted using the Qiagen DNEasy Blood & Tissue Kit using manufacturer instructions modified as in Comstock et al. (2022). Briefly, one ml of sucrose lysis buffer (40 mmol L^{-1} EDTA, 50 mmol L^{-1} Tris HCl, 750 mmol L^{-1} sucrose, 400 mmol L^{-1} NaCl, pH adjusted to 8.0), 100 µl of sodium dodecyl sulfate (10% w/v) and 10 µl of 20 mg ml^{-1} proteinase K were added, samples were incubated at 55°C for 2 hours, then refrozen until extraction. Additionally, to increase DNA yield, 750 µl of lysate mixture were extracted by running repeated volumes through the same column with 750 µl of Qiagen AW1 buffer. Extracted genomic DNA was stored at -40° C until amplification.

Amplicon Library Sequencing & Bioinformatics

Amplification of the V4 region of the 16S rRNA gene was performed using the 515F-Y (5'-GTGYCAGCMGCCGCGGTAA-3') and 806RB (5'-GGACTACNVGGGTWTCTAAT-3') primers (Parada et al., 2016; Apprill et al., 2015). PCR-grade water process blanks and mock communities (BEI Resources mock communities HM-782D and HM-783D) were included with each 96-well plate of samples as quality control checks as described in Wear et al. (2018). Amplicons were cleaned and normalized using SequalPrep plates (Invitrogen), pooled at equal volumes, concentrated using Amicon Ultra 0.5 ml centrifugal tubes (Millipore), gel extracted to remove non-target DNA (Qiagen) and sequenced on an Illumina MiSeq using PE250 chemistry at University of California (UC), Davis DNA Technologies Core.

Sequence data were demultiplexed at UC Davis DNA Technologies Core. Fastq files were quality filtered and merged using the dada2 package (version 1.13) in R (Callahan et al., 2016). Chimeras were removed de novo using the removeBimeraDenovo function in the dada2 package. Amplicon Sequence Variants (ASVs) were given a taxonomic assignment in the dada2 package using the assignTaxonomy command and the SILVA database (version 132) (Quast et al., 2013). After sequence classification, and initial assessment of plastid abundance, chloroplast sequences were removed and remaining reads were subsampled to 3000 reads per sample to standardize read depth. Samples below a read depth of 3000 were removed from further analysis. Mock communities and negative controls were checked to confirm consistency in amplification and lack of contamination between PCR plates and then removed from further analysis. All DNA data are available in the National Center for

Biotechnology Information (NCBI) Sequence Read Archive (SRA) under project number PRJNA773129

Statistical Analyses

For all multivariate analyses, ASV relative abundances were pre-treated using an angular transformation to normalize the dataset. Bray-Curtis dissimilarity matrices were calculated in R (v4.0) using the vegan package v2.5 (Oksanen et al., 2019) and used to generate NMDS ordinations. PERMANOVAs and pairwise PERMANOVAs were run using the adonis and pairwiseAdonis functions in the vegan package v2.5. T-tests were run in R using the stats package (v3.6.2). Simple linear models were run using the stats package v3.6 in R. One-way ANOVAs were run using the lapply function in the BioGenerics package in R.

Results

Physical Flow around Mo'orea

The island of Mo'orea has a well-documented unidirectional flow of water across the reef crest from forereef to backreef, then advection out the reef passes (Hench et al., 2008). These reef passes typically correspond with embayments of varying sizes, of which Paopao bay on the island's north side is one of the island's largest. Water flow across the reef is primarily driven by wave forcing, with tidal forces being comparatively weak due to the island's proximity to amphidromic points (Hench et al., 2008). Therefore, the volume of

water being pushed across the reef platform is strongly influenced by the magnitude of wave forcing. During this study, we measured significant differences in wave power between the three island sides. At the time of sampling from 29 July to 3 August 2014, wave power across the reef platform was greatest on the west side (56.78 ± 25.99 kW m^{-1}), intermediate on the east side (31.97 ± 7.96 kW m^{-1}), and least on the north side (2.36 ± 0.71 kW m^{-1}; Figure 3.1, Appendix Figure 3.1). The relative magnitude of these values align with wave power values recorded on the three island sides during the austral winter in years prior to and following 2014, suggesting wave power during this study was representative of typical conditions around Mo'orea during the austral winter (Appendix Figure 3.2) (Adam et al., 2020). We use forereef, backreef and reef pass to describe the reef habitat.

Biogeochemical variation between reef systems and habitats

Significant trends (1-way ANOVA $p < 0.05$) in chlorophyll *a*, POC, DOC, bacterial abundance, phosphate, and Nitrate + Nitrite were observed between reef systems and habitats and the surrounding offshore environment (Figure 3.2). Chlorophyll *a*, phosphate, and nitrate + nitrite displayed increased concentrations within the reef system relative to oceanic waters, with highest concentrations observed in the backreef and reef pass environments on the north side. POC and bacterial abundance were elevated in the offshore oceanic environment and depleted in the reef system, with lowest values being observed in the backreef and on the west island side. DOC was also elevated at oceanic stations, with lowest values being observed in the backreef and reef pass habitats. Averages and significant differences for the listed parameters are reported in Figure 3.2.

Overall trends in Ordination

Strong trends in ordination of bacterioplankton communities were observed between island side and reef habitat, as visualized in Figure 3.3a with a non-metric multidimensional scaling (NMDS) ordination of Bray Curtis dissimilarities. Separation in community structure by island side is largely captured by NMDS axis 1, while separation between habitats is largely captured along NMDS axis 2. To support the visually observed trends in community structure, PERMANOVA analyses were run to identify groups that ordinated significantly ($p < 0.05$) from each other and quantify the magnitude of difference (R^2) between bacterioplankton communities, with greater R^2 values denoting a greater magnitude of difference in ordination and community structure (Figure 3.3b). Pairwise PERMANOVA analyses were run to elucidate significance and magnitude of difference in ordination between pairs of sample types. PERMANOVAs of Bray-Curtis dissimilarity showed separation between communities by both reef habitat and island side ($p < 0.001$) which are visualized in the NMDS ordination in Figure 3.3a. The magnitude of difference observed in bacterioplankton community structure between sides of the island was triple ($R^2 = 0.31$, $p < 0.001$) the magnitude of difference observed between habitats ($R^2 = 0.10$, $p < 0.001$) (Figure 3.3). This suggests that differences (whether chemical, physical, or biological) between island sides are greater drivers of bacterioplankton community structure than differences between reef habitats. Pairwise PERMANOVA analyses between each reef habitat and island side revealed that overall, reef communities strongly significantly ($p < 0.001$) ordinated by reef habitat and side of the island, with the one exception being the north pass & backreef ($p = 0.003$). This Pairwise PERMANOVA emphasizes that each habitat within

each island side displayed a bacterioplankton community structure that was significantly different from every other habitat on every island side.

Effects of reef habitat on ordination by side

The greatest magnitude of difference in ordination of bacterioplankton communities between the three reef habitats is found on the north side of the island (PERMANOVA R^2 = 0.31) (Figure 3.22a). This suggests that the bacterioplankton communities of the three habitats were more dissimilar from each other with the north side than within the east or west island sides. Within the east and west sides of the island, the magnitude of difference between the reef habitats were roughly equivalent (pairwise PERMANOVA east R^2 = 0.23, west R^2 = 0.23) (Figure 3.2a). This suggests that a larger amount of variation in community structure on the island's north side can be explained by variation in reef habitat than the other sides of the island.

On the north side of the island the greatest magnitude of difference in ordination was found between the forereef and the pass communities (pairwise PERMANOVA R^2 = 0.30) and the weakest magnitude of difference was found between the pass and backreef (pairwise PERMANOVA R^2 = 0.17) (Figure 3.2). Because the R^2 value between the pass and the backreef is the smallest, the magnitude of difference in bacterioplankton community composition between these two habitats is the smallest, so they are the most 'similar' to each other. In contrast, on the east and west sides of the island the greatest magnitude of difference was found between the pass and backreef (pairwise PERMANOVA east R^2 = 0.25, west R^2 = 0.21) and the smallest magnitude of difference was found between the

forereef and backreef (pairwise PERMANOVA east $R^2 = 0.13$, west $R^2 = 0.12$) (Fig 2b). These results imply a difference in dynamics between habitats within the north side relative to the east and west, with the north having a backreef that is more 'pass-like', and the east and west having a backreef that is more 'forereef-like'. This may be explained in part by the increased wave forcing doesn't allow enough time for reef-associated taxa to increase in their relative abundances. In contrast, wave forcing on the north side of Mo'orea was an order of magnitude smaller than either the east or west sides during the sampled time period (Figure 3.1, Appendix Figures 3.1 and 3.2). On the north side, backreef bacterioplankton communities are more similar to pass communities than forereef communities. This may potentially be caused by reduced wave forcing transporting a reduced volume of water across the reef crest allowing for greater alteration in reef bacterioplankton communities from the surrounding oligotrophic gyre. These data suggest that as wave forcing increases, as seen on the east and west sides versus the north, the backreef bacterioplankton community becomes more homogenized and thus shifts from being 'pass-like' to more 'forereef-like' (Figure 3.3). Likewise, reduced wave forcing over the reef platform can allow for a more unique community within the backreef relative to the forereef or open ocean.

Taxonomic trends between island sides and reef habitats

Simple linear models were run to identify taxa with significant (adjusted p-value <0.05) side and habitat effects. In general, ASVs that showed effects were represented in families that showed effect. Therefore, for further analyses ASVs were grouped at the family level to simplify results, with one notable exception being that Cyanobacteria remained grouped at the genus level because the highly abundant genuses *Prochlorococcus* and

Synechococcus displayed inverse trends in relative abundance between island sides and reef habitats. Overall, 28 cosmopolitan (found in ≥10 samples) families were identified as having a strong island side effect, habitat effect, or side & habitat effect (Figure 3.3). These families include members of Actinobacteria (Actinomarinaceae, Microtriachaceae), Bacteroidetes (Cyclobacteriaceaea, Cryomorphaceae, Flavobacteriaceae, NS7, NS9), Planctomycetes (Pirellulaceae, Rubinisphaeraceae), and Proteobacteria (SAR116, Rhodobacteraceae, AEGEAN-169, SAR11, Alteromondaceae, Halieaceae, Litoricolaceae, Oleiphilaceae, Saccharospirillaceae, SAR86). Both *Prochlorococcus* and *Synechococcus* also displayed strongly significant (p adjust <0.001) side and habitat effects. A full list of taxa that showed a strong effect between sides, habitats, or both, is shown in Figure 3.3.

Of the 28 cosmopolitan families identified, 11 had a significantly enhanced (Tukey HSD $p < 0.05$) relative abundance on the north side relative to the west side. These included the families Flavobacteriaceae, Cryomorphaceae, AEGEAN-169, Halieaceae, SAR406, Litoricolaceae, Marine Group II, Pirellulaceae, Microtrichaceae, SAR116, Rubinisphaeraceae, and the genus *Synechococcus*. Eight families displayed significantly enhanced relative abundance on the west side relative to the north side, and included SAR11 Clade I and II, NS9, Alteromonadaceae, Saccarospirillaceae, Marine Group III, Oleiphilaceae, an Unassigned Alphaproteobacteria family, and the genus *Prochlorococcus* (Figure 3.6).

While significant differences in relative abundance were found between the north and west island sides for several of the above taxa, differences between the north and east sides or the west and east sides were not always significant. In general, the east side

displayed mean relative abundances that were intermediate or similar to relative abundances found on the north and west sides (Fig 3). This aligns with an intermediate degree of wave forcing observed on the east side (Figure 3.1). The exceptions to this include SAR86, Actinomarinaceae, and SAR 11 Clade IV which had weakly significantly enhanced relative abundance on the east side relative to either the north or west sides.

When evaluating the differences in relative abundance between reef habitats comparisons were made between habitats within each island side rather than between habitats across all island sides. This was done because the greatest magnitude of ordination was explained by island side, so trends between reef habitats had to be investigated with each side separately. Families that were found consistently enriched in the pass relative to the forereef included SAR406, Microtrichaceae, Rubinisphaeraceae, Pirellulaceae, and the genus *Synechococcus*. Families found enriched in the backreef on the north side included the families Halieaceae, Litoricolaceae, Alteromonadaceae, SAR116. Families that were found consistently enhanced in the forereef relative to the backreef or pass included SAR11 Clade I, II, and IV, AEGEAN-169, SAR86, NS9, Actinomarinaceae, and the genus *Prochlorococcus*. A majority of taxa that were found enriched in the forereef were also found with significantly higher relative abundance in surrounding oceanic stations relative to average reef abundance (Comstock et al., 2022), suggesting that the taxa enriched in the forereef relative to the pass have a more oligotrophic oceanic lifestyle rather than being reef associated.

Correlation between bacterioplankton communities and biogeochemistry

Several biogeochemical parameters were found significantly ($p < 0.05$) correlated with overall ordination of bacterioplankton communities along NMDS axis 1, including total chlorophyll (Chl), primary production (PP), nitrate + nitrite, particulate organic carbon (POC), dissolved organic carbon (DOC), bacterial carbon per cell (BC), and bacterial abundance (BA) (Figure 3.4). This suggests that significant differences in biogeochemical parameters observed between island side and habitat in Figure 3.2 also correlate with separation in ordination of associated bacterioplankton communities.

Of all biogeochemical parameters, total chlorophyll explained the greatest magnitude of difference in ordination (PERMANOVA $R^2 = 0.23$, $p < 0.001$) of bacterioplankton communities within the reef environment. When plotting the relative abundance of taxa found to have significant side and habitat effects, taxa with increased relative abundance on the north side and within the pass and backreef environments were found to have a positive correlation with chlorophyll concentration (Figure 3.6). Likewise, taxa with increased relative abundance in the forereef, west side, and surrounding oceanic waters displayed a negative correlation with total chlorophyll, including Prochlorococcus, SAR11 (Clades I, II & IV), SAR86, and Actinomarinaceae (Appendix Figure 3.3). These results imply that there is coupling between reef bacterioplankton community structure and water column biogeochemistry, with chlorophyll acting as a rough proxy for relative abundance of taxa enhanced on the reef.

Discussion

Overview

Through high resolution sampling across reef habitats on three sides of Mo'orea, we were able to investigate the effect of physical regime on bacterioplankton structure with coincident biogeochemical variability. Our synoptic survey comprised >100 bacterioplankton 16S rRNA gene amplicon samples, and clearly demonstrated separation of community structure by island side and reef habitat, with greater differentiation between island sides. Separation in bacterioplankton community structure between reef habitats had previously been reported (Nelson et al., 2011; McCliment et al., 2012); however, the present study is the first to directly compare the relative influence of physical regime and reef habitat together. We show that magnitude of wave forcing over the reef was the primary driver of community structure, with stations displaying a higher relative abundance of several taxa previously observed to be reef associated at stations with comparatively less wave forcing. These taxa include Flavobacteriaceae, Cryomorphaceae, AEGEAN-169, Halieaceae, SAR406, Litoricolaceae, Marine Group II, Pirellulaceae, Microtrichaceae, SAR116, Rubinisphaeraceae, and the genus *Synechococcus.* This increase in relative abundance of reef-associated taxa correlated with a suite of changes in biogeochemical parameters (i.e. Chl, nitrate + nitrite, phosphorus, DOC, POC) highlighting connectivity between physical regime, biogeochemistry, and biology in the water column of coral reef systems.

Physical Regime and Bacterioplankton ordination

During time of sampling, there was a 24-fold difference in the magnitude of wave forcing between the north side and the west side of Mo'orea. Wave forcing is the primary driver of circulation patterns across the Mo'orea coral reef platform (Hench et al., 2008).

Thus, substantial differences in wave forcing implies differing physical regimes between island sides. We demonstrate that bacterioplankton communities differed by more than three times as much between island sides than between reef habitats, with separation between the north, east, and west island sides aligning with magnitude of wave forcing (Figure 3.1). This suggests that wave-driven circulation over the reef platform is a primary driver of reef bacterioplankton community structure around Mo'orea. However, it is important to note that despite such a large difference in wave forcing regimes between island sides, all sampled reef locations ordinate separately from the surrounding oceanic waters (Comstock et al., 2022), suggesting that despite the relatively rapid flow on the reef platform, especially on the west side of the island, the reef maintains a distinct bacterioplankton community from the offshore oligotrophic waters. A counterclockwise flow has previously been observed around Mo'orea and was present at time of sampling (James et al., 2020). This counterclockwise flow has been suggested as a mechanism for physical retention of biogeochemically altered reef water around the island (James et al., 2020; Leichter et al., 2013) whereby water advected though the reef passes travels counterclockwise within the oceanic environment around the island rather than advecting out and away from the island. This keeps biogeochemically and biologically altered reef water near Mo'orea, allowing for potential re-entrainment into the reef system (Leichter et al., 2013). Additionally, this phenomenon can explain why forereef communities that experience a high magnitude of wave forcing still maintain distinct community structure compared to the surrounding offshore waters.

The Mo'orea fringing reef system has a well-documented unidirectional flow of water across the reef crest from the forereef to the backreef, then subsequent travel to and

advection out of the reef platform via reef passes (Hench et al., 2008). Therefore, water parcels in the reef pass have spent a greater time within the reef system than water parcels within the backreef, which in turn have had a greater amount of time in the reef system than water parcels on the forereef. On all island sides, there were significant differences in ordination between forereef, backreef, and reef pass communities, consistent with previous findings around Mo'orea (Nelson et al., 2011). However, with our dataset, we can further investigate how wave forcing regimes impacts separation in community structure between reef habitats on each island side.

Biogeochemistry and Bacterioplankton Communities

The ordination of microbial community structure between island side and reef habitat are coupled with gradients in a variety of biogeochemical measurements and wave forcing regime (figure 3.2a). Consistent with our own measurements, time-series data collected through the LTER on the north side of Mo'orea show consistent and significant gradients in Chl a, bacterioplankton abundance, DOC, nitrate, and POC between forereef, backreef, and pass reef habitats (Leichter et al., 2013; Nelson et al., 2011). Within these historical datasets and at the time of sampling for the present study, Chl, nitrate, and POC were consistently higher concentration in the reef pass relative to the forereef, while DOC and bacterial abundance were depleted in the backreef and pass. It is hypothesized that coral reefs exhibit rapid biogeochemical cycling to support retention and rapid recycling of nutrients (Duarte & Cebrián, 1996; Odum & Odum, 1955). Coral reefs exhibit some of the highest gross primary production rates on the planet (Sorokin, 1990); however, the lack of detrital buildup suggests extremely rapid removal and recycling of organic matter on the reef (Nelson, Kelly & Haas,

2023). Autochthonous organic exudates released by reef primary producers such as coral and algae are highly labile and can be high in limiting nutrients such as nitrogen and phosphorus (D'Elia 1977; Wegley Kelly et al., 2022). This organic matter is rapidly consumed and remineralized by surrounding bacterioplankton assemblages (Gast et al., 1998; Moriarty et al., 1985; Torréton & Dufour, 1996). However, while this labile DOM can fuel rapid microbial growth, microbial biomass over rapidly flushing reefs has repeatedly been observed to be lower than that of the open ocean (Gast et al., 1998; Moriarty 1979, Nelson et al 2011; Tanaka & Nakajimi, 2018), likely due to rapid grazing by benthic suspension feeders or planktonic micrograzers (McNally et al., 2017; Legendre et al., 1988).

The sources of organic and inorganic nutrients on the reef platform create distinct biogeochemical gradients on the reef platform compared to the oligotrophic oceanic water washing onto the platform (Nelson et al., 2011; Leichter et al., 2013). This unique biogeochemistry has been demonstrated both in Mo'orea and in reefs around the globe to foster unique bacterioplankton communities, with a suite of distinct taxa consistently found enriched on reefs (Nelson et al., 2011, 2013; McCliment et al 2012; Tout et al., 2014; Wegley Kelly et al., 2014; Weber et al., 2019, 2020b; Frade et al., 2020; Weber & Apprill, 2020, Apprill et al., 2021). The prokaryotic phylogenetic trends found in this enrichment of taxa such as *Synechococcus*, Flavobacteriaceae, NS9, Cryomorphaceae, Alteromonadaceae, and Rhodobacteraceae relative to offshore waters observed in the present study is consistent with previous reports of from other coral reef systems (Nelson, Wegley Kelly, and Haas, 2023). These reef-associated taxa largely showed enhanced relative abundance on the north side and in the backreef and pass environments relative to the forereef on all island sides. Given the unidirectional flow of water across the reef and the reduced wave forcing on the

north side, it is possible that these stations with enhanced reef-associated taxa experienced a greater residence time within the reef system than other stations sampled. Notable exceptions to this observed trend were the copiotrophic Alteromonadaceae and the Flavobacteria NS9, which historically have been found to be reef-associated organic matter but were found in highest relative abundance on the west side and in forereef and backreef (Figure 3.4). The relative enhancement of these reef-associated taxa within the forereef and backreef underscores the ability for coral reefs to retain unique chemistry and biology throughout the reef environment, and suggests that certain reef-associated taxa may be specialized for variable biogeochemical and physical regimes or benthos cover.

Taxa found enhanced in surrounding oligotrophic oceanic waters, such as *Prochlorococcus*, SAR86, and SAR11 (Comstock et al., 2022) exhibited enhanced relative abundance on the west side of the island and in the forereef environments on all island sides relative to the reef pass environment or north island side. These findings suggest that there is a gradient observed in bacterioplankton community, with strong wave forcing reef habitats located earlier along the unidirectional flow path associated with more oceanic-associated communities, while weak wave forcing and stations found further along the reef flow path were associated with an increased abundance of several reef-associated bacterioplankton. This dataset captures the “oceanic” to “reef” continuum within the bacterioplankton communities. This continuum aligns with gradients in biogeochemical parameters and physical flow across the reef platform. Additionally, some taxa found in high abundance in the reef system (*Synechococcus*, Cryomorphaceae, Rhodobacteraceae, Alteromonadaceae) were found in Chl hotspots in the surrounding nearshore oceanic environment during the sampling time period (Comstock et al., 2022). This suggests

physical connectivity between the reef and nearshore oceanic waters through advection of reef water and associated biology and biogeochemistry out of the passes and into the surrounding oceanic waters.

Conclusion

Our results clearly demonstrated community separation by island side and reef habitat, with greater differentiation between island sides, suggesting that physical regime of reefs plays a central role in community structure of water column bacterioplankton. The captured gradients in bacterioplankton communities align with gradients in observed biogeochemistry, suggesting connectivity between physical, biogeochemical, and biological structure within the reef. Simple linear models revealed several bacterial taxa that were associated with both island sides and reef habitats. A majority of taxa that have previously been associated with reef environments were found in highest relative abundance at stations further along the reef flow path and with reduced wave forcing, while more oligotrophic open-ocean taxa were found to decrease under the same conditions. Our results identify and characterize the unique bacterioplankton communities in reef waters, providing insight into the physical and biogeochemical drivers of bacterioplankton communities within the reef environment.

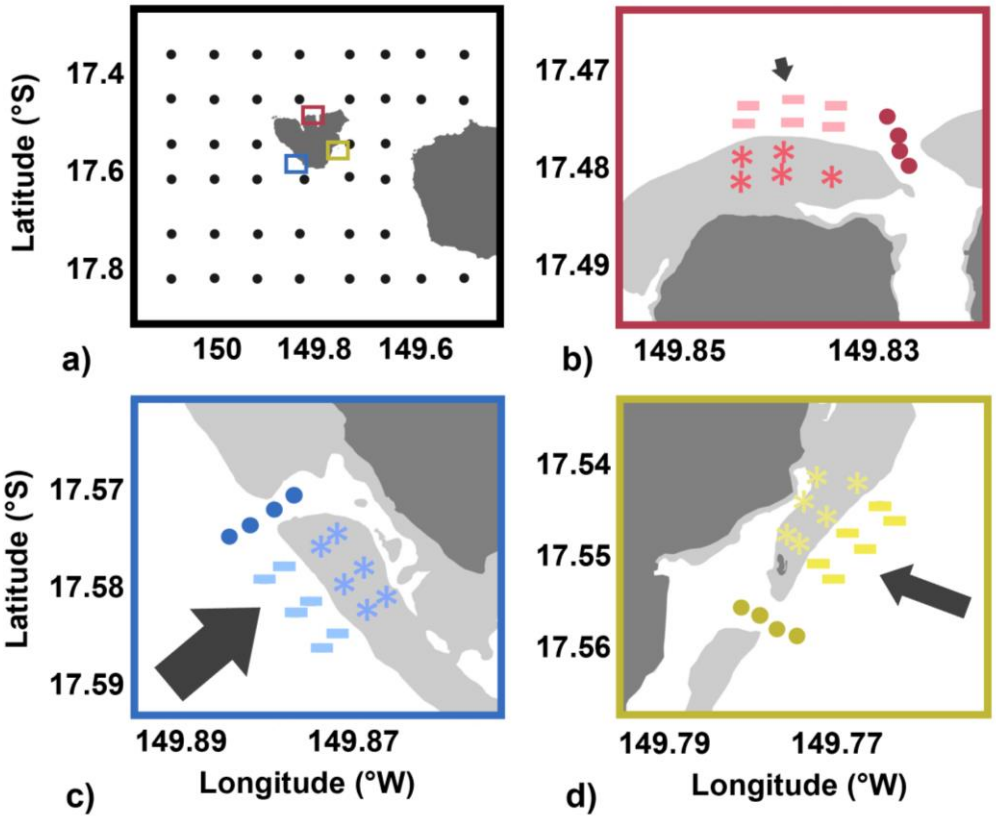


Figure 3.1: Sampling scheme a) around Mo'orea and Tahiti, b) on the north side, c) the west side, and c) the east side of Mo'orea. Arrows are proportionally to the mean wave power on each of the three sides during the 5-day sampling period.

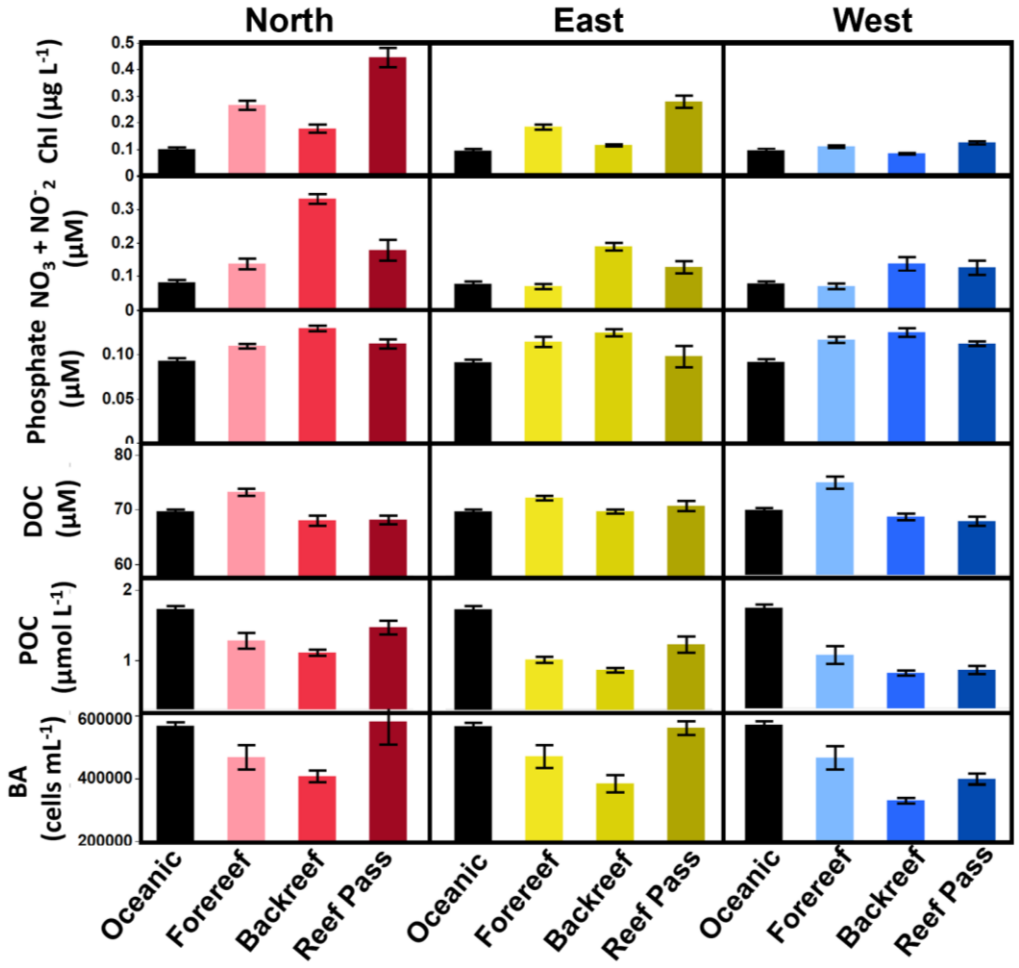


Figure 3.2: Differences in biogeochemical parameters between habitats among three reef systems. ANOVA p-values indicate significant differences among means of each habitat (annotated at bottom) within each reef system (annotated at top) and means sharing the same letter within a reef system are not significantly different by Tukey post hoc tests. Whiskers denote standard error.

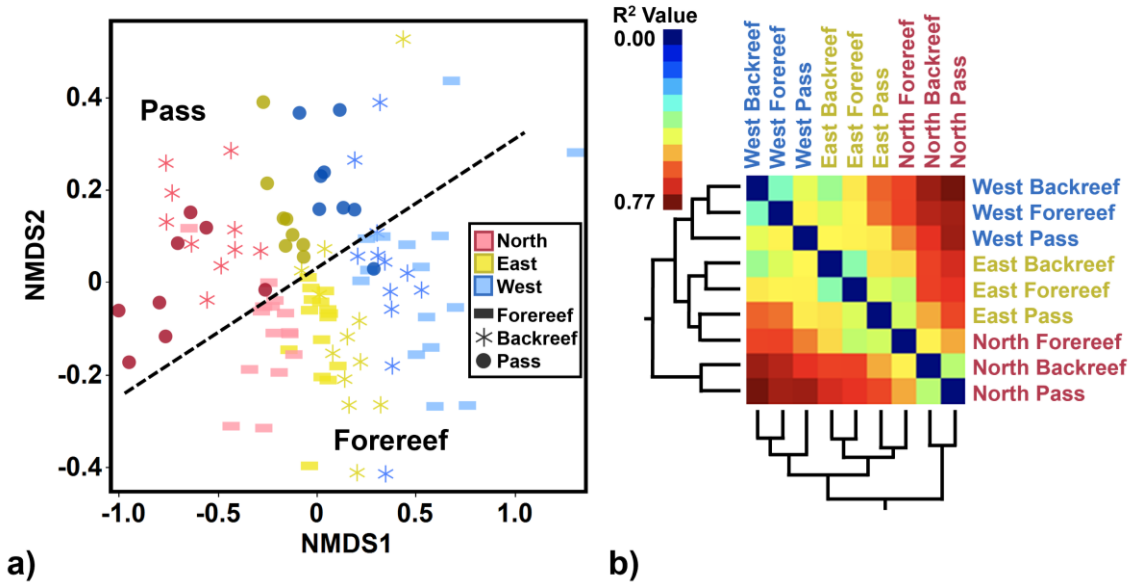


Figure 3.3: a) NMDS ordination of bacterioplankton communities within the reef environment. Red shades denote communities on the north side of the island, yellow shades denote the east side, and blue shades denote the west side of the island. Dashed line denotes visual separation of the reef pass and forereef habitats b) Heatmap of R^2 values derived from a pairwise PERMANOVA comparing prokaryotic community structure captured from each island side and reef habitat.

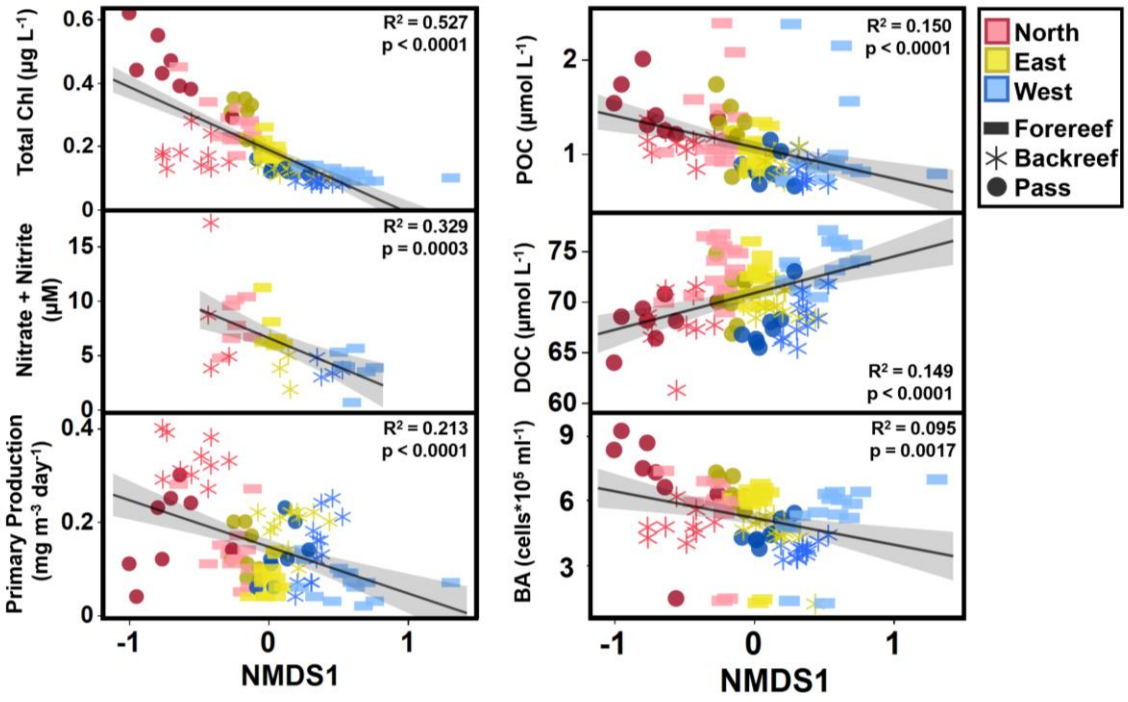


Figure 3.4: Biogeochemical parameters found significantly (PERMANOVA $p < 0.05$) correlated with bacterioplankton community ordination plotted against NMDS 1. NMDS 1 represents the strongest separation in bacterioplankton community structure and best captures separation between island side. Symbols and colors are the same as defined in Figure 3.2a.

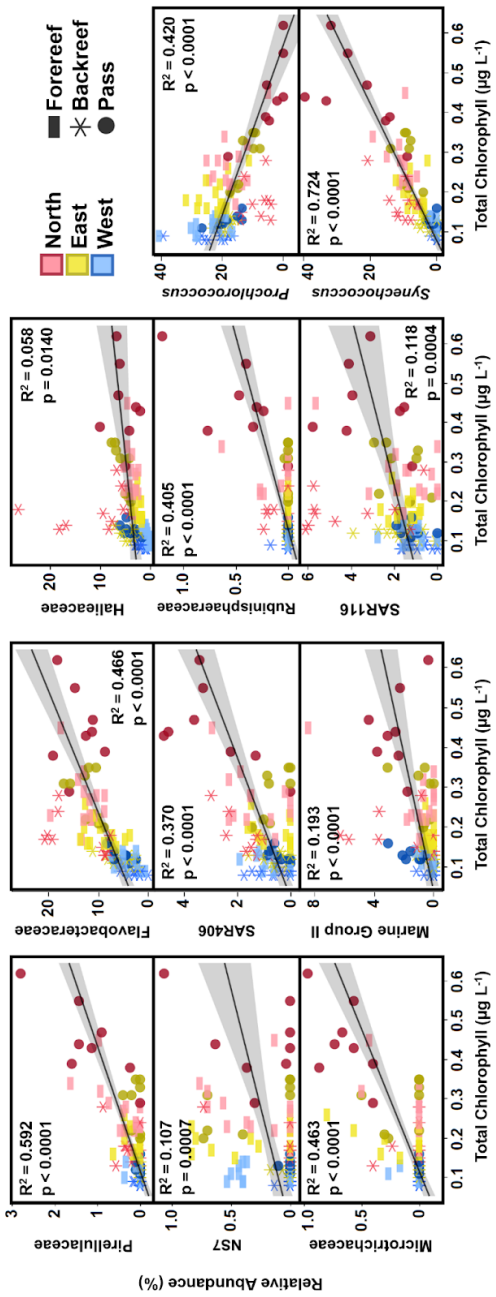


Figure 3.5: Relative abundances of reef-associated families and two most abundant cyanobacteria (*Prochlorococcus* and *Synechococcus*) plotted against total chlorophyll. Chlorophyll was the biogeochemical parameter correlated most strongly with bacterioplankton community ordination. Colors and shapes are the same as defined in fig 3.2a.

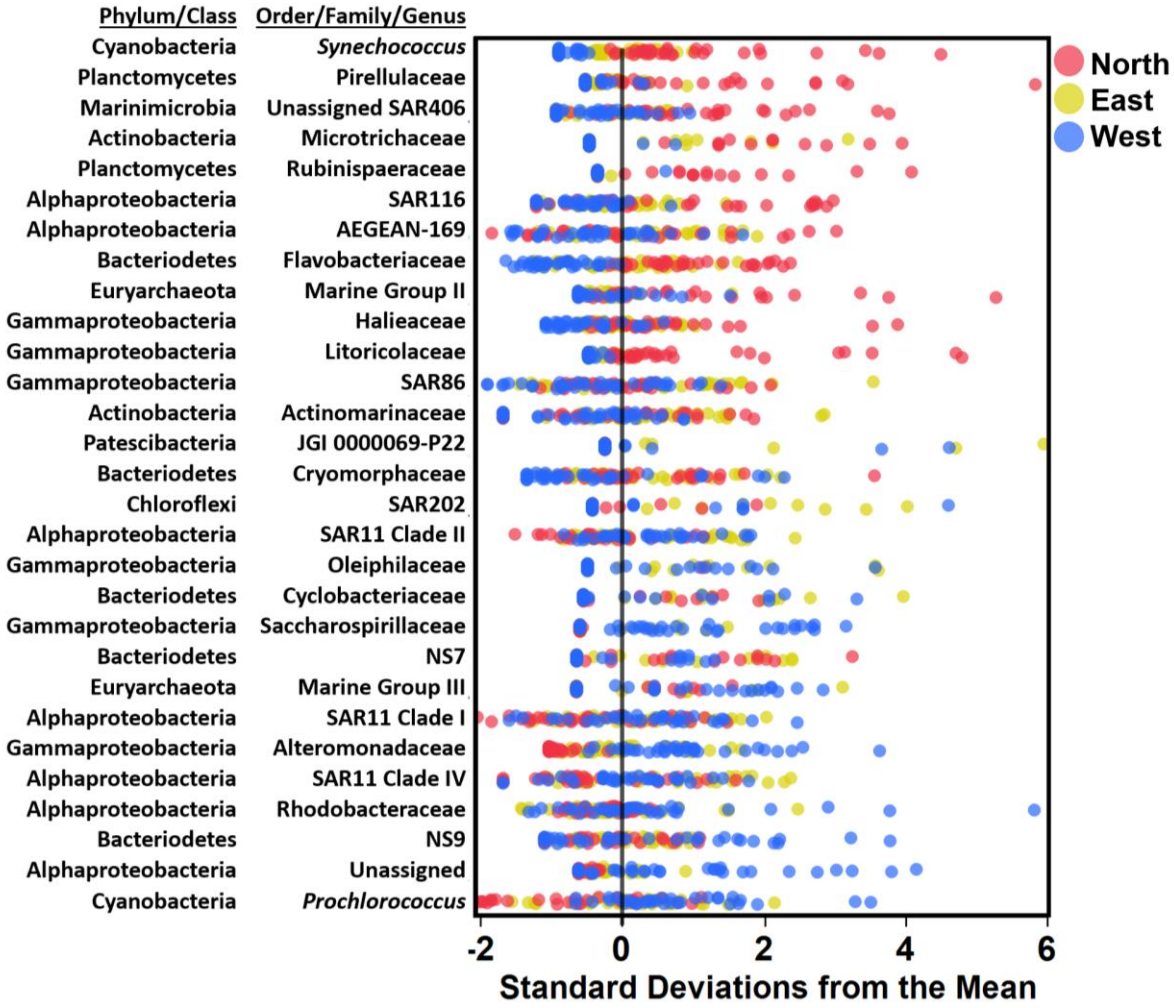
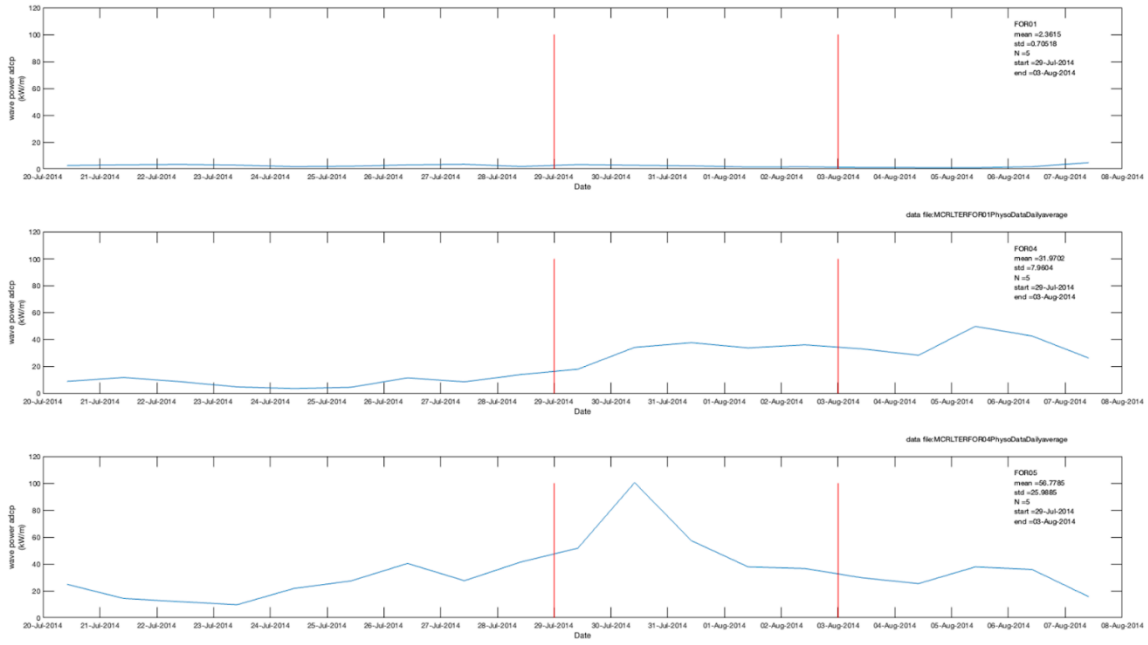
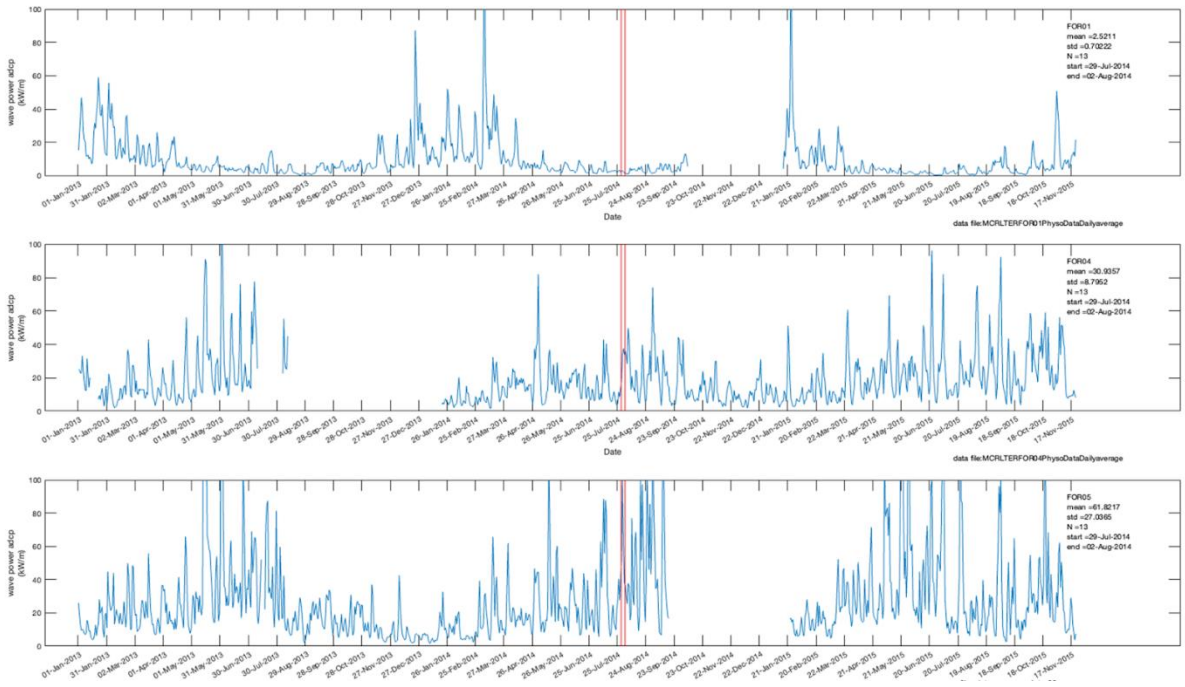


Figure 3.6: Standard deviations from the mean (z score) for the cosmopolitan (found present in >10 samples) taxa found to have significant reef habitat and island side effects. Z score was used to compare relative changes in taxa with varying relative abundance ranges. Black line denotes a z score of 0. Colors and symbol shapes are defined in fig 3.2a.

Appendix



Appendix Figure 3.1: daily average wave power (kW/m) on the a) north, b) east and c) west sides of the island. Red bars denote the beginning and end of the sampling periods.



Appendix Figure 3.2: daily average wave power on the a) north, b) east and c) west sides of the island. Red bars denote the beginning and end of the sampling period.

IV. Size-fractionation of marine particles at the Bermuda Atlantic Time-series Site (BATS) reveals cryptic links between microbial communities and chemical transformations across depth

Abstract

Particulate organic matter (POM) in the ocean's surface plays a central role in the carbon cycle, with sinking POM acting as the dominant pathway of the biological carbon pump. POM acts as "hotspots" of microbial activity, colonized by distinct communities of particle-associated bacteria and archaea. In this study, we aimed to better resolve patterns of particle colonization and chemical transformation across depth and between varying particle sizes. We evaluated microbial community structure via 16S rRNA gene amplicon sequencing alongside isotope characterization of POM collected at the Bermuda Atlantic Time-series Study (BATS) site. Four particle size fractions ranging from 0.2 – 1.2 µm, 1.2 – 5 µm, 5 - 20 µm, and >20 µm were collected via in situ pumps across the upper 500 m of the water column. Vertical stratification of the free-living bacterioplankton communities is well established at this site. Here we evaluate how particle-associated microbial community structure varies over depth and between size fractions. We demonstrate that microbial community structure differs as much between free-living and particle-associated communities as across the upper 500 m of the water column. We observed significant changes in $\delta^{15}N$ and trophic position of particles in all fractions. Smaller particles displayed greater chemical transformation via microbial heterotrophy, while larger particles revealed isotopic signatures that indicate they were disproportionately transformed by zooplankton grazing. These findings indicate that the residence time of bulk particles in all size fractions

is long enough to allow us to visualize successional patterns in microbial community structure and particle composition. Thus, while the fractions may include rapidly sinking particles, that signal is largely obfuscated by slowly sinking or suspended material. Our results provide insight into the niche partitioning of particle-associated bacterioplankton and potential causes of chemical transformations in particle composition through the upper 500 m of the water column.

Introduction

Sinking organic particles play a central role in the transport and sequestration of carbon in the ocean, acting as the dominant pathway in the Biological Carbon Pump (BCP) (Boyd et al., 2019; Nowicki et al., 2022). These particles, comprised of aggregated phytoplankton detritus, zooplankton-derived carcasses and fecal matter can be ballasted by minerals of biogenic and lithogenic origin (Alldredge and Silver, 1988; Turner, 2015). This process transports organic matter from the sunlit surface to the deep ocean (McCave, 1975). Sinking organic particles represent approximately 80% of the 7 – 11 Pg C (Chavez and Toggweiler, 1995; Sarmiento and Gruber, 2006; Hansell et al., 2009) exported annually from 100 -150 m depth. However, particulate organic carbon (POC) flux rapidly attenuates with depth, with only ~15% of particles reaching a depth of 1,000 m to be sequestered on timescales of centuries to millennia (Martin et al., 1987). More recent studies have shown that the transfer efficiency of POC from the surface to depth has significant spatial variability (Buesseler et al., 2007, 2020; Henson et al., 2012; Guidi et al., 2015; Marsay et al., 2015; Weber et al., 2016; DeVries and Weber, 2017), though the factors regulating this spatial heterogeneity remain subjects of debate.

The export of POC from the ocean's photic zone into the interior is the net outcome of two competing processes. The production and source of organic particles which affects particle size, aggregation and sinking rate is offset by biological processes that intercept, solubilize and remineralize the sinking organic matter; thus attenuating POC flux (Martin et al., 1987; Ducklow et al., 2001). During transport, organic particles undergo numerous transformations by a number of processes including direct remineralization by bacterioplankton (Ploug and Grossart, 1999) or zooplankton (Steinberg et al., 2008), fragmentation of aggregates by zooplankton (Dilling and Alldredge, 2000; Goldthwait et al., 2004; Giering et al., 2014) and solubilization of sinking particulate organic matter (POM) to dissolved organic matter (DOM) via production of hydrolytic enzymes by attached prokaryotes (Azam et al., 1993; Smith et al., 1992; Ziervogel et al., 2010; Arnosti, 2011). These processes result in the observed attenuation of sinking particles by approximately 85% within the mesopelagic zone (200-1000m) (Marin et al., 1987).

Marine POM produced by aggregation of various organic material is densely colonized by bacteria and archaea distinct from the free-living community, representing "hotspots" of microbial activity in the water column in both coastal and oligotrophic environments (e.g. DeLong et al. 1993, Simon et al., 2002; Grossart, 2010). These particle-associated (PA) microbial communities can play a central role in POM degradation due to their potential for producing extracellular enzymes that hydrolyze POM and DOM to lower-molecular-weight compounds prior to cellular uptake (Smith et al., 1992, Arnosti 2011; Arnosti et al., 2021). POM and DOM are compositionally distinct pools of organic matter, with recently-produced organic particles being a potentially fresher and less diagenetically altered source of organic matter compared to accumulated surface DOM (Skoog and

Benner., 1997, Goldberg et al., 2009, Kaiser and Benner 2012). PA microbial communities show higher cell densities, growth rates, extracellular enzymatic activities, biogeochemical transformation rates, metabolic diversity, and overall microbial activity per unit volume than surrounding free-living (FL) bacterioplankton communities (Caron et al., 1986; Karner and Herndl, 1992; Smith et al., 1992; Grossart et al., 2007; Ziervogel et al., 2010; Ganesh et al., 2014; Satinsky et al., 2014; Dang and Lovell, 2016). Therefore, marine POM may provide unique microenvironments for microbial communities, leading to microbial compositions that are distinct from the surrounding free-living (FL) communities (e.g. DeLong et al., 1993, Allen et al., 2013). The quality and quantity of sinking organic particles is highly variable, due to nutrient availability, seasonality, geography, and distributions of organisms in the water column that produce, consume, and transform POM (Alldredge and Silver, 1988; Francois et al., 2002; Hensen et al., 2012; Riley et al., 2012). Thus PA microbial communities are highly diverse and may be substantially influenced by several factors including particle types and quality (Ortega-Retuerta et al., 2013; Ganesh et al., 2014; Duret et al., 2019; Boeuf et al., 2019; Preston et al., 2020; Baumas et al., 2021), depths, and seasonality (Poff et al., 2021; Karl et al., 2012; Grabowski et al., 2019; Rieck et al., 2015). Size-fractionated collection of organic particles via large volume transfer systems (McLane in situ pumps) is an effective approach to evaluate partitioning of microbial communities across organic matter size gradients.

The Bermuda Atlantic Time series Study (BATS) has collected physical, chemical, and biological data on this seasonally oligotrophic ecosystem in the subtropical North Atlantic gyre nearly every month since 1988 and is among the most established oceanic study sites. BATS provides an excellent opportunity to investigate compositional and

metabolic differences in PA microbial communities in the context of the time series data that includes biological productivity, vertical fluxes and biogeochemical variability (Lomas et al., 2013; Steinberg et al., 2001), DOM dynamics (Carlson et al., 1994; Goldberg et al., 2009; Goldberg et al., 2011; Hansell et al., 2002), microbial processes (Carlson et al., 1996; Carlson et al., 2004; Carlson et al., 2002; Nelson and Carlson, 2012; Parsons et al., 2012), and microbial plankton diversity (Carlson et al., 2009; Giovannoni and Vergin, 2012; Giovannoni et al., 1990; Morris et al., 2005; Sowell et al., 2009; Treusch et al., 2009; Treusch et al., 2012, Blanco-Bercial et al., 2022). These rich data sets have identified recurring seasonal patterns and point to complex mechanisms underlying DOM oxidation (Carlson et al., 1994, Goldberg et al., 2009, Liu et al., 2022) alongside well-documented vertical stratification of the free-living microbial communities over depth (Gordon et al., 1996; Giovannoni et al., 1996; Field et al., 1997, Treusch et al., 2009). Additionally, the microbial communities associated with sinking particles at BATS have been found to significantly differ from the surrounding free-living community (Cruz et al., 2021). However, gradients in community structure within bulk particles (both sinking and suspended) across particle size and over depth within the upper euphotic and mesopelagic remains unresolved at this site. This study aims to further characterize these PA microbial communities in the context of changing particle biogeochemistry at the BATS site; thus, increasing our understanding of particle transformation processes.

We combined 16S rRNA metabarcoding of four distinct size fractions (0.2 - 1.2 µm, 1.2 - 5 µm, 5 - 20 µm, and >20 µm) alongside compound-specific isotope analysis of amino acids (CSI-AA) and bulk $\delta^{15}N$ characterization of size-fractionated organic particles across multiple depths from the surface into the mesopelagic at BATS. Our results demonstrate that

differing particle fractions undergo differing transformation processes, with smaller fractions having elevated $\delta^{15}N$ values below the upper euphotic, and larger fractions displaying consistently higher trophic positions. Based on prior observations of particle community connectivity across depth (Mestre et al., 2018), we hypothesized that we would observe greater community similarity in PA communities across depth compared to the FL fractions. Our prokaryotic community data demonstrates that particle size fraction was the largest driver of community structure in the upper 500 m of the water column. Notably, there was an equivalent degree of transformation of community structure across depth in all size fractions, suggesting that PA communities undergo as much transformation across depth as their FL counterparts. Separation of community structure was most pronounced between organic particle size fractions in the mesopelagic compared to the same particle size fractions in the upper euphotic. We explore the hypothesis that PA communities within the mesopelagic become more compositionally distinct from FL communities as the organic particles are diagenetically transformed. The present study helps to better constrain drivers of particle transformation and connect observed biogeochemical transformations to community structure and metabolic potential of associated microbial taxa, emphasizing the various roles microbes play in marine biogeochemical cycling.

Methods

In Situ Pump and Hydrographic Sampling

Size-fractionated samples were collected by McLane WTS-LV in situ pumps (McLane Research Laboratories, Inc) outfitted with dual filter heads each containing four-

tiered 142 mm filter holders. One filter head was dedicated to the serial collection of microbial DNA fractions and the other was used to collect size fractionated particles for chemical characterization (Henderson et al., submitted). Fifty-two discrete samples were collected between the surface and 500 m over four cruises conducted in the vicinity of BATS in July 2018, July 2019, Aug 2021, November 2021. The greatest depth resolution was sampled within the euphotic zone and upper mesopelagic zone. Three to four in situ pumps were deployed on a hydrowire simultaneously, and between 47 and 367 L of seawater was filtered through each pump over the course of two to three hours per cast. DNA was collected from particulate fractions separated as follows: > 20 µm fraction was collected on 20 µm Nytex screen, particles that passed through the Nytex screen and retained on a 5 µm polycarbonate Isopore (TMTP) filter represented the 5 – 20 µm fraction, particles that passed the 5 µm filter but retained on 1.2 µm polyethersulfone (PES, Pall Corporation) represent the 1.2 - 5 µm fraction and the particles that passed the 1.2 µm filter but retained on the 0.2 µm PES filters represent the free living 0.2 - 1.2 µm fraction. Water column hydrography from the same location of the in situ pump deployment was determined within hours before or after the in situ pump deployments. The hydrographic data includes CTD sensor profiles of density, temperature, salinity, oxygen, and chlorophyll fluorescence along with discrete measurements of dissolved organic matter (DOM), particulate organic matter (POM), inorganic nutrients ($NO_2^- + NO_3^-$; PO_4^{3-}), and bacterial abundance from Niskin bottles. Parameters captured with CTD and bottle sampling were averaged across the duration of each 4-day cruise to provide hydrographic context for each sampling period. The deep chlorophyll maximum (DCM) was defined as the depth (±10 m) with the highest chlorophyll fluorescence as captured by the CTD and varied within and between cruises.

Size Fractionated DNA Preservation and Extraction

Immediately after in situ pump recovery, the pump heads were drained of standing volume by gentle vacuum and filters were removed from each filtration tier and placed in polyethylene bag, heat sealed and stored flat at -80°C until further processing. DNA was extracted by first thawing filters and adding 6 ml of sucrose lysis buffer (40 mmol L^{-1} EDTA, 50 mmol L^{-1} Tris HCl, 750 mmol L^{-1} sucrose, 400 mmol L^{-1} NaCl, pH adjusted to 8.0), 600 µl of sodium dodecyl sulfate (10% w/v) and 10 µl of 20 mg ml^{-1} proteinase K to each polyethylene bag. Bags were resealed and incubated at 37°C for 30 min and then increased to 55°C for another 30 min for cell lysis. Approximately 1.1 ml of lysate was removed by sterile serological pipet to 2 mL microcentrifuge tubes and the remaining lysate was stored at -80°C. DNA was extracted from the lysate following the phenol and chloroform protocol of Giovannoni et al. (1996) at the University of California Santa Barbara.

Amplicon Library Sequencing & Bioinformatics

Amplification of the V4 region of the 16S rRNA gene was performed using the 515F-Y (5'-GTGYCAGCMGCCGCGGTAA-3') and 806RB (5'-GGACTACNVGGGTWTCTAAT-3') primers with custom adapters (Apprill et al., 2015; Parada et al., 2016; Wear et al., 2018). PCR-grade water process blanks and mock communities (BEI Resources mock communities HM-782D and HM-783D and a custom

community from the Santa Barbara Channel; Wear et al., 2018) were included with each 96-well plate of samples as quality control checks. Amplicons were cleaned and normalized using SequalPrep plates (Invitrogen), pooled at equal volumes, concentrated using Amicon Ultra 0.5 ml centrifugal tubes (Millipore), gel extracted using the QIAquick Gel Extraction Kit to remove non-target DNA (Qiagen) and sequenced on an Illumina MiSeq using PE250 chemistry at University of California (UC), Davis DNA Technologies Core.

Samples were demultiplexed at UC Davis DNA Technologies Core. Fastq files were quality filtered and merged using the dada2 package (version 1.16) in R (Callahan et al., 2016). Chimeras were removed de novo using the removeBimeraDenovo function in the dada2 package. Amplicon sequence variants (ASVs) were given a taxonomic assignment in the dada2 package using the assignTaxonomy command and the SILVA database (version 138.1 with species) (Quast et al., 2013). Samples were rarefied to 8000 reads, and samples with fewer than 8000 reads were removed from further analysis which amounted to 25 out of 468 environmental samples. Mock communities and negative controls were compared to confirm consistency in amplification and lack of contamination between PCR plates and then removed from further analysis.

Particle Biogeochemistry

Organic particles were chemically characterized for the four size fractions according to the following procedures. Particles in the > 20 µm fraction were first washed off Nitex filters with 0.2 µm filtered seawater and collected onto a 0.7 µm GF/F glass fiber filter. The GF/F filters and other size fraction filters were then freeze-dried and quantitatively split,

acidified via direct application of saturated sulfurous acid and then immediately dried at 60°C. Bulk POM carbon and nitrogen concentrations and N stable isotope ratios were measured via elemental analyzer coupled to an isotope ratio mass spectrometer (EA-IRMS, Thermo Scientific). Standards of known carbon and nitrogen isotope composition (Arndt Schimmelmann, Indiana University) were used to calibrate instrument response; analytical uncertainty in bulk carbon and nitrogen isotope composition was 0.2‰.

In addition, splits from 0.3 μm GF-75 and 1.2 μm GF/C filters were freeze-dried, hydrolyzed, purified, derivatized, and analyzed via gas chromatography coupled with isotope ratio mass spectrometry (GC-IRMS) as previously described (Hannides et al., 2013, Doherty et al., 2021). Norleucine and aminoadipic acid co-injection standards and a standard mixture of 14 amino acids with known $\delta^{13}C$ and $\delta^{15}N$ values were prepared and run alongside samples to correct for instrument drift, potential peak size relationships, and isotope effects associated with derivatization.

We also report trophic position (TP) of POM from the various size fraction based on measured $\delta^{15}N$ values of glutamic acid (Glu) and phenylalanine (Phe) as in Chikaraishi et al. 2009:

$$TP = (\delta^{15}N_{Glu} - \delta^{15}N_{Phe} - 3.4)/7.6 + 1$$

Trophic position uncertainty was propagated as in Jarman et al., (2017).

Inorganic nutrients

Unfiltered seawater samples were collected from 12 L Niskin bottles into 20 mL HDPE vials, frozen at -20 ˚C, and analyzed using flow injection analysis on a QuickChem 8000 (Lachat Instruments, Zellweger Analytics, Inc.) by the University of California, Santa Barbara Marine Science Institute Analytical Laboratory (detection limits: $NO_2^- + NO_3^-$, 0.2 µmol L^{-1}; PO_4^{3-}, 0.1 µmol L^{-1}).

Dissolved Organic Carbon

Seawater was filtered through precombusted (450°C) 47 mm GF/F filters with a nominal pores size of 0.7 µm set in polycarbonate filter holders attached to Niskin-type bottles. This filtered seawater was collected into precombusted 40 mL borosilicate glass vials with polytetrafluoroethylene (PTFE) coasted silica septa. All samples were acidified with 4N HCl to a pH of ~3. DOC concentrations were analyzed, referenced, and standardized according to Halewood et al., 2022 using high temperature combustion (HTC) on a modified Shimatzu TOC-V or TOC-L.

Bacterioplankton Abundance

Ten to forty mL of seawater was collected in sterile conical tubes (Falcon) and fixed with 0.2 µm filtered formalin to a final concentration of 1% and flash frozen at -80°C until processing. To create slides, seawater was filtered via gentle vacuum pressure (~100 mm Hg) onto 0.2 µm polycarbonate filters stained with Irgalan Black. Cells were stained with 5 µg mL^{-1} 4’, 6’-diamidino-2-phenylindole dihydrochloride (DAPI, Sigma-Aldrich). Cells were counted via manual counting with an epifluorescence microscope (Olympus AX70 or

Olympus BX51) at 1000 x magnification (Porter and Feig, 1980). At least 10 fields per slide were counted.

Statistical Analyses

ASV relative abundances were pretreated using an angular transformation to normalize the dataset, and singleton ASVs were removed for all multivariate analyses. Bray-Curtis dissimilarity matrices and nonmetric multidimensional scaling (NMDS) ordinations were calculated in R (v4.1) using vegedist and metaMDS in vegan (v2.5) (Oksanen et al., 2020). Permutational multivariate analyses of variance (PERMANOVA) and multivariate homogeneity of group dispersion analyses were run using 'adonis' and 'betadisper' functions in the 'vegan' package in R (v2.5). Pairwise PERMANOVAs were run using the pairwiseAdonis package in R. T-test, ANOVA, and Tukey HSD analyses were run using the R stats package (v3.6.2). Analyses run can be found at https://github.com/jacquicomstock

Results

Hydrography

The depth of the surface mixed layer varied between and within cruises, from 9 m during the stratified summer period in July 2019, to 81 m during deeper mixing in November 2021. The depth of the DCM ranged from 69 m to 134 m across the four sampling periods. The highest average surface temperatures were observed during the August 2021 sampling period (28.9°C) and lowest during November 2021 (24.6°C). During

July 2018 the DCM ranged from 69 to 88 m, 81 to 100 m in July 2019, 101 to 134 m in August 2021, and 77 to 93 m during November 2021. Microbial communities classified as being associated with the DCM fell within these depth ranges ±10 m for each sampled timeframe. Profiles for temperature, chlorophyll, bacterial abundance, nitrate + nitrite, particulate organic carbon (POC), and dissolved organic carbon (DOC) during all sampled time periods can be found in Appendix Figure 4.1 and broadly match values and trends previously reported at BATS (Steinberg et al., 2001; Lomas et al., 2013)

Chemical characterization of organic particles

Values for bulk $\delta^{15}N$ and trophic position calculated from compound-specific isotope analysis of amino acids (CSIA-AA) of POM were measured alongside bacterioplankton DNA from all size fractions during July 2018. There were consistent and significant differences in $\delta^{15}N$ values and trophic position for particles in different size fractions. Bulk $\delta^{15}N$ values were used to interpret degradation of organic matter from a more labile to a more refractory form, with higher values indicating greater utilization and reworking of organic matter as consumers preferentially removed the lighter ^{14}N isotope (Saino and Hattori, 1980). Bulk $\delta^{15}N$ values ranged from -2‰ to 4‰, varying between particle size fractions and depth (Figure 4.1a). Values between fractions in the upper euphotic were more similar than those observed between fractions at mesopelagic depths. Bulk $\delta^{15}N$ values in the 0.3 – 1.2 and 1.2 – 6 µm were lowest near in the upper 100 m, then increased sharply from 120 to 190 m. Within the >20 µm fraction no significant increase in $\delta^{15}N$ was observed within the mesopelagic. Instead, a $\delta^{15}N$ value ≤ 0‰ was observed at all mesopelagic stations except at 250 m, which had a $\delta^{15}N$ value of 2.3‰.

Trophic position of particles, as calculated through CSIA-AA, was measured for the 0.3 - 1.2 µm, 1.2 - 6 µm and >20 µm fractions in July 2018 (Figure 4.1b). Trophic position was used to identify the degree to which particles were being consumed and transformed via metazoan processes, with a higher calculated trophic position indicating increased contributions from zooplankton. The lowest trophic positions were observed closest to the surface and increased with depth. Trophic position was only measured & calculated from three samples within the 0.3 -1.2 µm fraction, however at each depth (31 m, 69 m, and 87 m) this fraction displayed a lower trophic position than the two larger particle fractions. The highest values were observed at 190 m (1.6 ± 0.2) in the 1.2 - 6 µm fraction and at 112 m (2.07 ± 0.2) in the >20 µm fraction. At all depths sampled, the >20 µm fraction displayed a higher calculated trophic position than the two smaller fractions, suggesting greater metazoan influence on the processing of the larger particle sizes.

Overall Trends in Bacterioplankton Communities

Overall, 432 environmental samples were successfully sequenced containing 10,907 non-singleton ASVs after rarefaction. Cyanobacterial contributions were most pronounced in the 0.2 -1.2 µm and 1.2 - 5 µm size fraction in the upper euphotic and DCM but were present in all size fractions at all depths to 500 m (see details below). Nonmetric multidimensional scaling (NMDS) ordination of ASVs demonstrated a similar pattern with and without the inclusion of cyanobacteria (Figure 4.2 and Appendix Figure 4.2). Unless otherwise specified, for subsequent analyses cyanobacterial taxa were removed to directly compare FL and PA heterotrophic and chemoautotrophic communities. Permutational multivariate analyses of variance (PERMANOVA) were used to identify potential

environmental drivers of prokaryotic community variability. NMDS ordinations revealed distinct clustering by size fraction and sample depth. PERMANOVAs indicate that size fraction ($R^2 = 0.15$, $p < 0.001$) and depth ($R^2 = 0.12$, $p < 0.001$) were strong drivers of community composition in this dataset. While sampling date was found to be a significant driver of community variability, it explained the lowest proportion of variance ($R^2 = 0.06$, $p < 0.001$). When these analyses were conducted with cyanobacterial taxa included, environmental predictors of community variability did not differ ($R^2 = 0.15$, $p < 0.001$; $R^2 = 0.15$, $p < 0.001$; $R^2 = 0.06$, $p < 0.001$), suggesting that size fraction, depth, and date all significantly influence photoautotrophic, heterotrophic, and chemoautotrophic communities. Pairwise PERMANOVA analyses revealed significant differences in community composition between all pairs of size fractions in the 0.2 - 1.2 µm, 1.2 - 5 µm, and > 5 µm size fractions at all depths; however, differences between the 5 - 20 µm and >20 µm fraction communities were not significantly different over depth.

Community Variation within the Upper Euphotic, DCM and Mesopelagic Depths

Sample depths were grouped into three main bins to investigate broad patterns across gradients in light, organic, and inorganic nutrient fields over the upper 500 m of the water column. Sample depths at least 10 m shallower than the observed DCM were binned as "Upper Euphotic". Samples collected within the observed DCM ± 10 m were binned as "DCM" samples, and samples collected deeper than 10 m below the DCM + 10 m were binned as "Mesopelagic". Within all sample depth bins, PERMANOVA analysis indicated that size fraction of the sample was a significant driver of community variation

(PERMANOVA R^2 = 0.32 - 0.33, $p < 0.001$); suggesting that PA and FL prokaryotic community structure differed in all depth bins.

A multivariate analog of Levine's test for homogeneity of variances was used to assess the similarity of communities within each depth, followed with Tukey post-hoc analyses to assess significance. Communities at each depth within the upper euphotic zone bin had smaller average distances from the mean (average distance from the mean = 0.42) than communities at depths in the mesopelagic (average distance from the mean = 0.54) (Appendix Figure 4.6), implying that size-fractionated communities in the upper euphotic were more similar to each other than size-fractionated communities at greater depths. Within the upper euphotic, 13.1% of ASVs were present in all size fractions (Figure 4.3). However, within the DCM and mesopelagic, only 6.6 - 7.3% of ASVs were shared between all fractions, supporting the finding that there was less overlap in FL and PA taxa below the upper euphotic.

Community Variation within Size Fractions

Clear vertical stratification of prokaryote community structure was observed over the surface 500 m within every size fraction sampled (Figure 4.2a). PERMANOVA analyses confirmed that depth was a strong driver of community structure variability within each size fraction (R^2 = 0.38 - 0.43, $p < 0.001$). The degree of variance between bacterial communities in the 0.2 - 1.2 µm, 1.2 - 5 µm, and >20 µm fractions (PERMANOVA R^2 = 0.41 - 0.43, $p < 0.001$) was similar to that observed over depth and slightly reduced in the 5 - 20 µm fraction (R^2 = 0.37, $p < 0.001$). For all four size fractions, there was only minor

overlap of ASVs (i.e, 3.6 - 5.5%) in the upper euphotic, DCM and mesopelagic depth bins (Figure 4.4). This implies that there was significant specialization across depth in all size fractions.

Alpha Diversity

Shannon diversity indices within all size fractions from FL to PA increased significantly over depth (ANOVA $p < 0.001$) as well as seasonally from early summer to late autumn (ANOVA $p < 0.001$) (Figure 4.5a). Within the upper euphotic, there was a weak but significant increase in Shannon diversity with increasing fraction size (ANOVA $p = 0.013$). However, no significant differences in Shannon diversity were observed between size fractions within the DCM and mesopelagic (ANOVA $p = 0.238$).

Trends in Chao1 were similar to those observed with Shannon diversity (Figure 4.5b). A significant difference in Chao1 diversity was observed between sampling timepoints, with July and November of 2021 showing elevated Chao1 relative to July of 2018 and 2019. Communities within the mesopelagic also displayed significantly (ANOVA $p < 0.001$) increased Chao1 diversity from upper euphotic and DCM communities. There was a weak but significant (ANOVA $p = 0.012$) increase in Chao1 values with increasing size fraction. Trends in alpha diversity were consistent with cyanobacteria included (Appendix Figure 4.3).

Distributions of cyanobacterial lineages

Cyanobacteria comprised up to 54.5% of the prokaryotic community in the 0.2 - 1.2 µm and 1.2 - 5 µm fractions. In all size fractions, cyanobacteria were found in higher abundance in the upper euphotic and DCM compared to the mesopelagic, though cyanobacterial taxa were found in samples as deep as 507 m in all size fractions (Appendix Figure 4.4). The greatest relative abundances of *Prochlorococcus* and *Synechococcus* were observed in the 0.2 - 1.2 µm and 1.2 - 5 µm fractions, with corresponding ASVs having a slight but significantly greater relative abundance in the 1.2 - 5 µm fraction (Tukey HSD $p < 0.05$). *Prochlorococcus* MIT9313 (ASV#1) and *Synechococcus* CC9902 (ASV #5) were at highest relative abundances in the upper 100 m of the water column. A greater number of less abundant cyanobacterial ASVs had higher relative abundances within and just below the DCM from 100 to 150 m, including *Prochlorococcus* MIT9313 ASVs #13, 21, 24, 77, 125, and 126 (Figure 4.6). While all cyanobacterial ASVs were found in highest relative abundance in the 0.2 - 1.2 and 1.2 - 5 µm size fractions, many cyanobacterial ASVs increased in their relative abundances in the 5 - 20 and >20 µm PA fractions size fractions at the base of the euphotic zone into the upper mesopelagic at depths several meters below their respective FL ranges.

Non-cyanobacterial taxonomic abundance by fraction

Between size fractions, significant differences were observed in the relative abundance of all major taxonomic groups. The free-living 0.2 - 1.2 µm fraction was significantly enhanced with Rhodospirillales (AEGEAN-169, Magnetospiraceae), SAR11 (Clade I, III, IV), and SAR324 compared to every other size fraction (Figure 4.7, Appendix Figure 4.4, 4.5). Rhodospiralles averaged 14.5% of the 0.2 - 1.2 µm fraction community in

the upper euphotic, while SAR11 averaged 32.1%. SAR324 was most abundant in the mesopelagic FL community, averaging 10.4%. Additionally, the relative contribution of Flavobacteriales was significantly (Tukey HSD $p < 0.05$) lower in the 0.2 - 1.2 µm fraction relative to all other fractions. The 0.2 - 1.2 µm and 1.2 - 5 µm fractions, were enriched with SAR202, SAR406, SAR324, SAR11, Puniceispirillales, Rhodobacterales, Rhodospirillales, Nitrospinales, MG II, MG III, and Nitrosopumilales compared to the 5 - 20 and >20 µm fractions, suggesting these taxa are largely free-living (Figure 4.7). The 1.2 - 5 µm fraction was enriched with MG II and MG III, and were maximal at an average 16.5% and 7.9% respectively in the mesopelagic 1.2 - 5 µm fraction.

The 5 - 20 and >20 µm fractions were each enhanced in Proteobacteria (Enterobactales), Verrucomicrobiota (Kiritimatiellales, Opitutales, Verrucomicrobiales), Planctomycetota (OM190, Phycisphaerales, Pla3, Pirellulaes, Planctomycetales), and Bacteroidota (Chitinophagles, Flavobacteriales). Enterobactales was most abundant in the 5 - 20 µm fraction in the mesopelagic, averaging 14.7% of the non-cyanobacterial community. A member of the genus *Alteromonas*, part of the class Enterobactales, represented the single most abundant ASV found in the particle-associated fractions, reaching up to 29.2% of the non-cyanobacterial PA communities. Verrucomicrobiota was highest in the upper euphotic and DCM, averaging 13.8% and 28.5% of the communities in the upper euphotic 5 - 20 and >20 µm fractions, and 26.0% and 31.5% in the DCM communities. Planctomycetota averaged 26.0% of the 5 - 20 µm community and 36.4% of the >20 µm community within the mesopelagic. Bacteroidota displayed enhanced abundance in all >1.2 µm fractions relative to the FL fraction, averaging between 22.9 - 25.4% of the >1.2 µm fractions within the upper euphotic where it was most abundant.

Heterotrophic and chemoautotrophic taxonomic succession over depth

There was strong vertical stratification of heterotrophic and chemoautotrophic bacteria and archaea over the surface 500 m. While microbial community composition varied across size fractions, each size fraction demonstrated similar vertical stratification in taxonomic relative abundance over the surface 500 m. Within the upper euphotic of all size fractions, members of the orders Opitutales, Flavobacteriales, Chitinophagles, and various Proteobacteria (Pseudomanadles, Puniceispirillales, Rhodospirllales, and SAR11) were observed in significantly higher relative abundance than at depths in the mesopelagic (Figure 4.7, Appendix Figure 4.4, 4.5). Members of SAR202, SAR406, SAR324, *Alteromonas*, Verrucomicrobiales, Nitrospinales, NB1-j, Bdellovibrionales, Planctomycetota (OM190, Pla3, Pirellulales, and Planctomycetales), and Archaea including Nitrosopumilales, MG II, MG III clades had a greater relative abundance in the DCM and upper mesopelagic in all size fractions.

Discussion

Biogeochemistry of marine particles

Similar changes in suspended particle $\delta^{15}N$ values to those reported here (increase in $\delta^{15}N$ of ~4‰ from the upper euphotic into the mesopelagic) have been observed around the global ocean (Saino and Hattori, 1980; Altabet at al., 1991; Casciotti et al., 2008; Emeis et al., 2010; Hannides et al., 2013). This isotopic shift has been reported to accompany changes in particle composition, with surface POM being dominated by phytoplankton while

mesopelagic particles include enhanced abundances of detrital material (Bishop et al., 1977; Gowing and Wishner, 1992). The enrichment of bulk $\delta^{15}N$ is attributed to the removal of labile ^{14}N organic matter that originated from the surface, leaving behind degraded organic matter enriched in $\delta^{15}N$ within the mesopelagic (Saino and Hattori, 1980). Hannides et al. (2013) reported the primary driver of ^{15}N-enrichment in suspended particles was microbial heterotrophic degradation rather than a change in trophic status by metazoan grazing or a shift in N source with depth. This suggests that the smaller particle size fractions in this dataset were largely transformed by PA microbial communities.

In contrast to the smaller (< 20 μm) fractions, we did not observe significant changes in $\delta^{15}N$ values in large (>20 μm) particles from the upper euphotic into the mesopelagic. Large sinking particles have previously been observed to increase only slightly (Casciotti et al., 2008) or decrease (Altabet et al., 1991) in $\delta^{15}N$ values with depth despite rapid attenuation of POC flux in the mesopelagic. Thus, the processes that transform larger particles differ from microbial processes affecting smaller particles (Altabet et al., 1991). These findings are consistent with previous reports of small, suspended particles being chemically distinct from large, sinking particles (Wakeham and Canuel, 1988; Abramson et al., 2010). Therefore, bulk $\delta^{15}N$ data alone could not differentiate whether large particles were sinking rapidly through the water column relatively untransformed, or whether the bulk $\delta^{15}N$ signal did not fully capture the biogeochemical transformations in large particles.

To better distinguish drivers of ^{15}N composition and particle degradation within the varying particle size fractions, we used compound-specific isotope analysis of amino acids (CSIA-AA). The interpretive power of this method is derived from the differing responses

of amino acids (AA) to trophic transfer (McClelland and Montoya, 2002). One group of "trophic" AAs (i.e. alanine, aspartic acid, glutamic acid, isoleucine, leucine, proline, and valine) becomes significantly enriched in ^{15}N with every trophic transfer. Another group of "source" AAs (i.e. glycine, lysine, methionine, phenylalanine, serine, and threonine) experience little change in ^{15}N during trophic transfers and instead retain their source N isotopic value from primary producers at the base of the food chain (Popp et al., 2007; Chikaraishi et al., 2009). Trophic position of particles can therefore be calculated through differences in $\delta^{15}N$ between the source AA glutamic acid/glutamine (Glx), and the trophic AA phenylalanine (Phe) (Chikaraishi et al., 2009). We observed higher trophic positions at all depths in the largest (>20 μm) fraction, with highest trophic positions found below 100 m in all size fractions (Figure 4.1b). Independent of depth, the higher trophic position of the larger particles suggests disproportionate transformation by metazoan metabolic processes compared to smaller particles. The importance of zooplankton grazing, metabolism and migration for organic particle packaging, transformation and flux has been widely recognized (e.g. Longhurst and Harrison, 1988; Longhurst et al., 1990; Michaels and Silver, 1988; Dam et al., 1995; Steinberg et al., 2000). Through consumption of phytoplankton and other detrital particles, zooplankton can produce dense fecal pellets, thereby repackaging small, suspended particles into larger, denser, and potentially faster sinking particles (Lan and Marchal, 2015). Disaggregation of these fecal pellets can influence the trophic position of smaller size fractions (Wakeham and Canuel, 1988); however, the observed higher trophic position of the >20 μm particle size fraction in this study suggests a greater magnitude of metazoan transformation of larger particles relative to smaller particles.

Drivers of community differentiation

Differences in community structure and metabolic potential between PA and FL bacterioplankton communities are well-established (DeLong et al., 1993) and found in diverse aquatic systems worldwide. Differences in organic matter source, composition, and micro-environmental conditions on particles compared to the surrounding water column have been reported to correlate with community structure variability (e.g. Crump et al., 1999, 2004; Simon et al., 2002; Ghiglione et al., 2007, 2009; Doxaran et al., 2012; Ortega-Retuerta et al., 2013; Bižić-Ionescu et al., 2014; Simon et al., 2014; Fontanez et al., 2015; Rieck et al., 2015; Cruz et al 2021; Valencia et al., 2021). Variability in community structure is coupled with differences in physiology and metabolic potential of PA taxa that are typically comprised of larger cells and genome sizes (Smith et al., 2013), higher respiration rates (Grossart et al., 2007), and expanded gene repertoires for polysaccharide degradation (Smith et al., 2013; Rieck et al., 2015; Kappelmann et al., 2019; Schultz et al., 2020; Reintjes et al., 2023).

Within the FL community, successional patterns in bacterioplankton taxa from the upper euphotic into the mesopelagic are well-established in subtropical gyres (Giovannoni et al., 1996, Morris et al., 2005; DeLong et al 2006; Treusch et al., 2009). Our data indicate that the PA communities underwent as extensive of a community shift as their FL counterparts between the upper euphotic and the mesopelagic. This contrasts with Mestre et al., (2018), which reported that communities on larger particles displayed a smaller change in community structure over depth compared to that of FL communities. Rapid successional patterns of particle colonization by bacteria and archaea (Datta et al., 2016) have been found to primarily correspond with changes in organic and inorganic nutrient composition and associated bioavailability (Ortega-Retuerta et al., 2013; Ganesh et al., 2014; Duret et al.,

2019; Boeuf et al., 2019; Preston et al., 2020; Baumas et al., 2021). If most particles were rapidly sinking and passed through the water column relatively untransformed, then we might expect more PA communities to be more similar over depth than FL communities. However, all particle sizes collected from the BATS site appeared to undergo significant compositional transformation through the water column. This transformation is likely driven by processes known to impact marine POM, including microbial reworking and zooplankton consumption, fragmentation, and repackaging (Smith et al., 1992; Azam et al., 1993; Ploug and Grossart, 1999; Dilling and Alldredge, 2000; Goldthwait et al., 2004; Steinberg et al. 2008; Ziervogel et al., 2010).

The increase in alpha diversity in the FL and PA communities within the mesopelagic is consistent with trends previously reported by Baumas et al. (2021). Mesopelagic organic matter consists of a diverse array of compounds (Wakeham et al., 2009) that are chemically distinct from upper euphotic organic matter and less labile (Goldberg et al., 2009; Kaiser & Benner 2008). This may support increased functional and phylogenetic diversity of prokaryotic communities with increasing depth (Wilson et al., 2017). Therefore, increasing alpha diversity suggests a shift in organic matter quality from labile, low molecular weight compounds to more recalcitrant high molecular weight compounds with depth (Goldberg et al., 2009; Kaiser & Benner 2008) in which numerous taxa are required for breakdown & utilization of this material.

Seasonal variability in FL and PA communities is widely observed both at BATS and in the global ocean (e.g. Morris et al., 2005; Fuhrman et al., 2006; Treusch et al., 2009, Cruz et al 2021). However, the magnitude of variability observed between the four sampling

timepoints was less than half of that observed between particle fraction or depth, suggesting that interannual variability, at least during the times sampled, may play a more tertiary role in community composition of FL and PA communities in the upper 500 m of the water column (Figure 4.2b).

We hypothesize that the greater degree of dissimilarity in microbial communities between PA and FL within the mesopelagic is due to extensive reworking of organic matter due to differing degradation processes. Mesopelagic DOM and POM at BATS are compositionally distinct and more diagenetically altered in the mesopelagic compared to that in the upper euphotic (Kaiser and Benner, 2009, 2012; Goldberg et al., 2009, Liu et al., 2022). Both POM and DOM in the surface ocean originate largely from phytoplankton, however labile POM and DOM are rapidly consumed, with only a fraction resisting degradation long enough to be exported into the mesopelagic (Pomeroy 1974; Azam et al., 1983; Martin et al., 1987; Azam and Malfatti, 2007). This reduction in bioavailability may cause a shift towards microbial communities capable of exploiting more recalcitrant organic matter. Additionally, zooplankton consumption and repackaging disproportionately influence larger particles, with higher trophic positions observed in the largest size fraction. In contrast, consumption and transformation of DOM is primarily performed by bacterioplankton, with microbial hydrolysis and remineralization affecting the chemical transformation of smaller size fractions (Hannides et al., 2013). While both POM and DOM are consumed and transformed by microbes, the communities associated with FL and PA lifestyles are significantly different and contain differing metabolic potentials, including more polysaccharide degradation and amino acid transport genes within the PA communities (Smith et al., 2013; Poff et al., 2021; Leu et al., 2022). We hypothesize that communities

within the mesopelagic become more specialized between size fractions as organic matter becomes more diagenetically altered by differing processes. This resulting recalcitrant material may require specialized microbial communities to break down the remaining material, potentially causing greater dissimilarity between the FL and PA communities in the mesopelagic (Figure 4.2a).

Microbial taxa distributions

We observed gradients in relative abundances of taxa between and within all fractions by depth. The 1.2 - 5 µm fraction includes both FL and PA taxa due to differences in cell sizes of microbes, which complicates the interpretation of microbial "lifestyles". Taxa such as Verrucomicrobiota and Planctomycetota were in greater relative abundance in the > 5 µm fractions but were also observed at elevated relative abundance in the 1.2 - 5 µm fraction. These taxa likely represent PA microbes that were either colonizing small (<5 µm) particles or colonized larger particles that were fragmented during the filtration process. Conversely, several abundant taxa in the 0.2 - 1.2 µm fraction, including members of the SAR324, SAR11, Rhodospirallales, and Nitrosopumilales clades, were also found in elevated relative abundance in the 1.2 - 5 µm fraction relative to the larger fractions, suggesting that these taxa have a FL lifestyle but were at least partially caught by the 1.2 µm filter.

Marine Group II, Marine Group III, SAR202, Flavobacteriales and Cyanobacteria such as *Prochlorococcus* and *Synechococcus* were abundant in the intermediate 1.2 - 5 µm fraction. Recent studies have found *Synechococcus* to be significant sources of transparent

exopolymer particles (TEP) (Ortega-Retuerta et al., 2019; Zamanillo et al., 2019), which can aid in particle aggregation of surrounding cells, increasing their effective size and density (Burd and Jackson, 2009).

The particle-associated 5 - 20 µm and >20 µm fractions displayed an enhanced abundance of members of Bacteriodota, Proteobacteria, Verrucomicrobiota, Bdellovibrionota, and Planctomycetota, which have been found to typically be the most dominant taxa found in PA communities (Salazar et. al, 2015). Deep sea ecotypes of *Alteromonas* can consume urea, and are hypothesized to colonize large, rapidly sinking particles (Ivars-Martinez et al., 2008). The most abundant heterotrophic ASV in the 5-20 and >20 µm size fractions in this dataset was the deep-sea ecotype *Alteromonas mediterranea*. This r-strategist has been associated with fast-sinking, large particulate material (López-López et al., 2005; López-Pérez et al., 2013, 2014). *Alteromonas* is a typical copiotroph that can grow very fast and is known to take advantage of sporadic inputs of organic matter (Ivanova et al., 2015; López-Pérez and Rodruigez-Valeria, 2016).

Previous studies suggest that various taxa will exploit niches within the PA community as the compounds found within marine particles become altered. While members of Alphaproteobacteria can be more efficient in the incorporation of monomers and amino acids, members of Bacteriodota (specifically *Flavobacteriaceae*) developed an enzymatic repertoire that appears to degrade higher molecular weight compounds (Cottrell and Kirchman, 2000; Reintjes et al., 2019). Members of Gammaproteobacteria are important degraders of algal polysaccharides, an abundant component of phytoplankton-derived POM (Francis et al., 2021). Plactomycetota have been predominantly found in larger particle

fractions (DeLong et al., 1993; Fuchsman et al, 2012) with genomic analyses revealing several members within that clade to be capable of degrading complex organic matter (Wegner et al., 2013).

Rhodobacteraceae, a member of Alphaproteobacteria typically found in high abundance on marine particles (Salazar et al., 2015), was found in higher abundance in smaller 0.2 - 1.2 and 1.2 - 5 um fractions, though it was present in larger fractions as well. Several other taxa, including members of Bacteriodia (*Flavobacteriaceae*), Gammaproteobacteria, Bdellovibrionia, and Planctomycetota, were found in detectable abundances in all fractions. This could be due to a "stick or swim" lifestyle switch, a mechanism previously shown in various marine taxa including *Rodeobacter*, *Gammaproteobacteria*, and *Flavobacteriia* (Fernandez-Gomez et al., 2013; Marin, 2014; Michael et al., 2016). It could also be caused through accidental attachment due to particle stickiness (Passow, 2002), niche partitioning between members of the same taxa (Moormann et al., 1997; Thompson et al., 2005; Dadon-Pilosof et al., 2017) or due to accidental retention of FL taxa on larger filters and the concomitant fragmenting of particle communities into smaller size fractions during filtration.

A previous study of sinking POM collected via sediment traps at the BATS site revealed that Cytophagales, Chitinophagales, and Cellvibrionales were indicator taxa of sinking particles (Cruz et al., 2021). The *in situ* pump collection method used here demonstrated relatively low abundance of these taxa within PA communities. This sampling approach did not differentiate between sinking and suspended particles within the water column and instead collected bulk organic particles from depths in the upper 500 m at BATS

in four size fractions. Differences in the microbial community structure of suspended and sinking particles in the Scotia Sea have been reported (Duret et al., 2018) and support the hypothesis that bulk particle communities are likely composed primarily of suspended particles and could display communities distinct from those of actively sinking particles. The observed differences in particle community composition between our study and Cruz et al. (2021) likely emphasize that a majority of organic particles at BATS are suspended or slowly sinking rather than rapidly sinking particles like those collected in sediment traps. However, consistent with Cruz et al. (2021), we observed *Alteromonas* and Vibrionales to be in the greatest relative abundance in the > 5 µm fractions. These size fractions would include some fast-sinking particles that could also be caught in drifting sediment traps.

Cyanobacteria found in particles below their free-living range

Large particles can be fragmented via physical shear, destruction via zooplankton feeding (eg. Dilling and Alldredge, 2000; Giering et al., 2014), and enzymatic solubilization by attached microbial communities (Cho and Azam, 1988; Smith et al 1992). While many ASVs found in highest abundance in the PA fractions were also present in low abundance in the <5 um fractions, we were not able to resolve any PA taxa transition to a primarily FL fraction for any depths sampled. The FL communities observed in this study are largely similar to whole communities previously sampled at BATS (Morris et al., 2005; Treusch et al 2009), suggesting that the abundance of FL microbes is great enough to largely obscure PA communities during amplicon sequencing. Due to the scarcity of marine particles in the water column, PA bacteria make up only around 1% of the total prokaryotic community (Alldredge et al., 1986; Heins et al., 2021). Therefore, while fragmentation of particles may

be occurring, these taxa could be vastly outnumbered by the ambient FL communities and therefore any transition to a FL lifestyle cannot be captured effectively by our approach.

Synechococcus and *Prochlorococcus* were observed shifting highest relative abundance from small to larger size fractions with depth. For both, ASVs, which were found at highest relative abundances in the 0.2 - 1.2 and 1.2 - 5 µm fractions, were found to have a significantly higher relative abundance in PA (5 - 20 µm and >20 µm) fractions at depths directly below their FL (0.2 - 1.2 µm and 1.2 - 5 µm) range. This suggests that members of the FL community can aggregate or be repackaged into larger particles. Cyanobacteria have been found to contribute to carbon export in a manner proportional to their net primary production despite their small size (Richardson and Jackson, 2007), and more recent metabarcoding of particle microbial communities has revealed sequences of *Synechococcus* to be associated with particulate carbon export (Amacher et al., 2013; Guidi et al., 2016; Valencia et al., 2021). Lomas et al., (2010) observed increased export of picophytoplankton at BATS, possibly due to increases in effective size and sinking rate. When nutrients are scarce, unbalanced carbon fixation in the absence of nutrient uptake has been found to increase the production of exopolymeric substances by phytoplankton, including picophytoplankton (Klut & Stockner et al., 1991). Exopolymeric substances such as TEP can increase particle aggregation rates (Alldredge et al., 1993; Burd & Jackson 2009; Guidi et al., 2009, Passow, 2002), thus these exopolymeric substances released by phytoplankton may enhance their aggregation (Chow et al., 2015) and therefore subsequent particle size and sinking rate. These aggregates may also serve as a food source for zooplankton grazers, which can repackage this material as fecal pellets, a process that contributes substantially to export production (Wilson and Steinberg 2010; Stukel et al., 2013; Ebersback et al., 2014).

Conclusion

Our study combines 16S rRNA gene metabarcoding with biogeochemical characterization of size-fractionated marine particles to better constrain oligotrophic PA microbial communities and transformation processes over depth at the BATS site. We present clear and consistent differences in biogeochemical composition, potential reworking processes, and associated microbial communities between size fractions. Smaller organic particles were found to be more heavily influenced by microbial processes, while larger particles experienced higher amounts of trophic transfer through zooplankton feeding and repackaging. Additionally, there was clear separation in microbial community structure through depth in all fractions, with various gradients in taxa relative abundance through depth and across size fractions. We demonstrated that PA microbial communities undergo as much community differentiation across depth as the FL fraction, implying extensive transformation of particle composition and particle communities through depth. Finally, we observed increased specialization between FL and PA communities within the mesopelagic, hypothesized to be caused by the reworked, more recalcitrant mesopelagic organic matter requiring more specialized microbial communities to be able utilize the material. Our data suggests that organic matter of varying size ranges undergo compositional transformations across depth, and differing processes driving these observed transformations are represented by differing microbial communities and particle biogeochemistry.

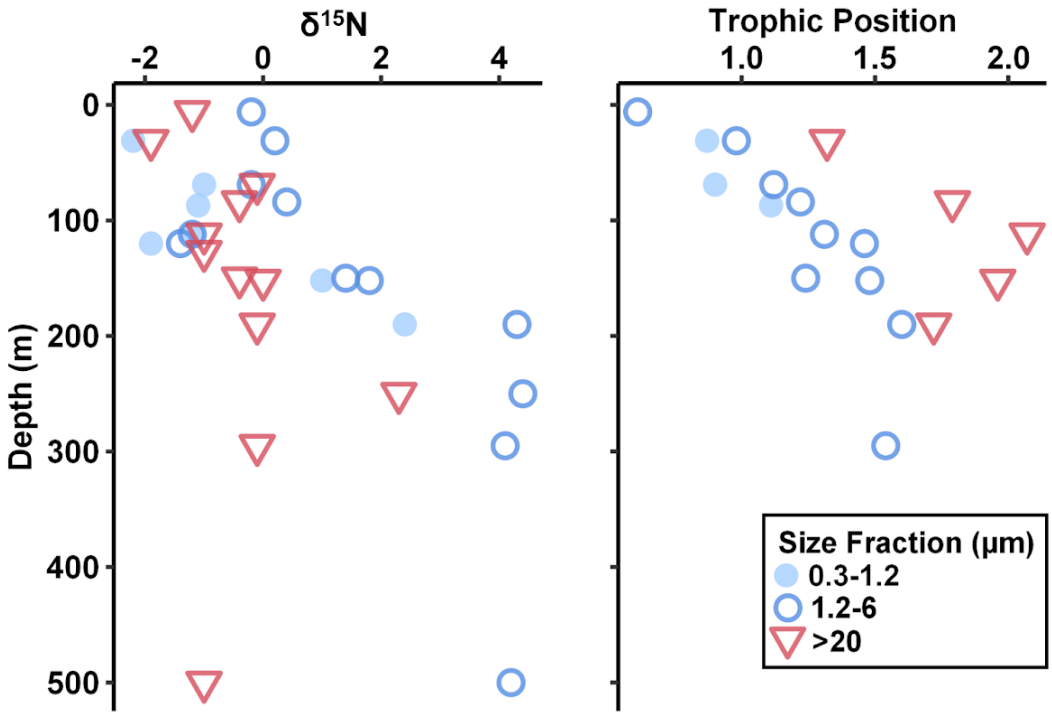


Figure 4.1: $\delta^{15}N$ and trophic position of particles in each size fraction over depth in the July 2018.

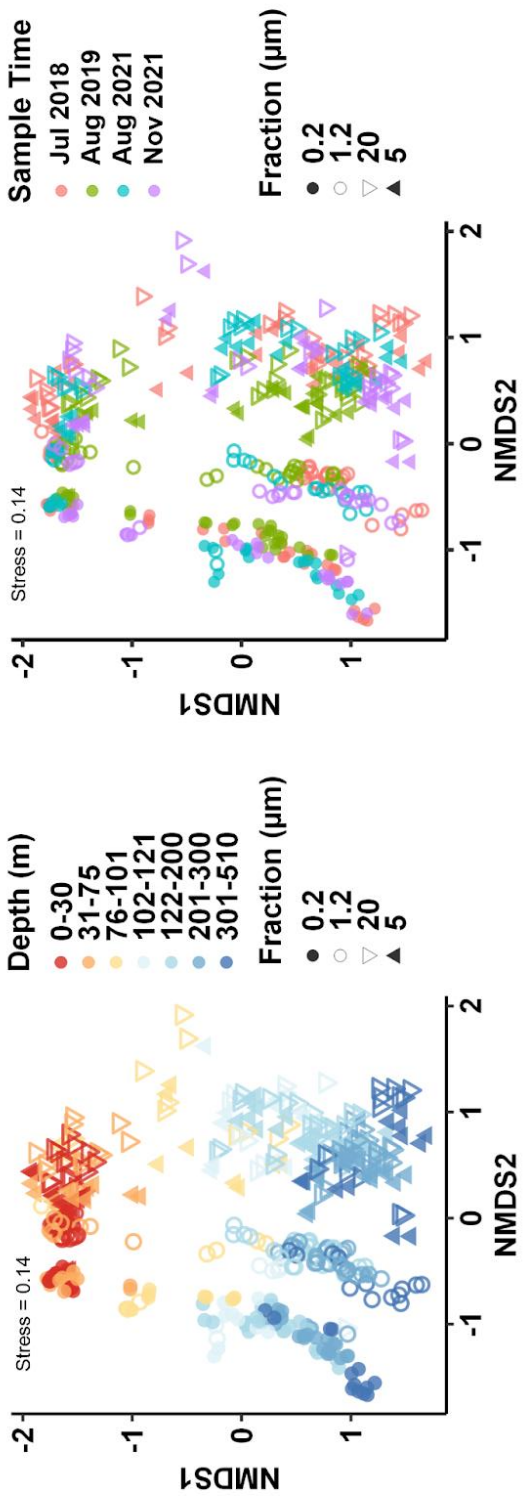


Figure 4.2: NMDS ordination of 16S community excluding cyanobacteria. a) Colors denote depth within the water column, shapes denote the size fraction. b) Colors denote time of collection.

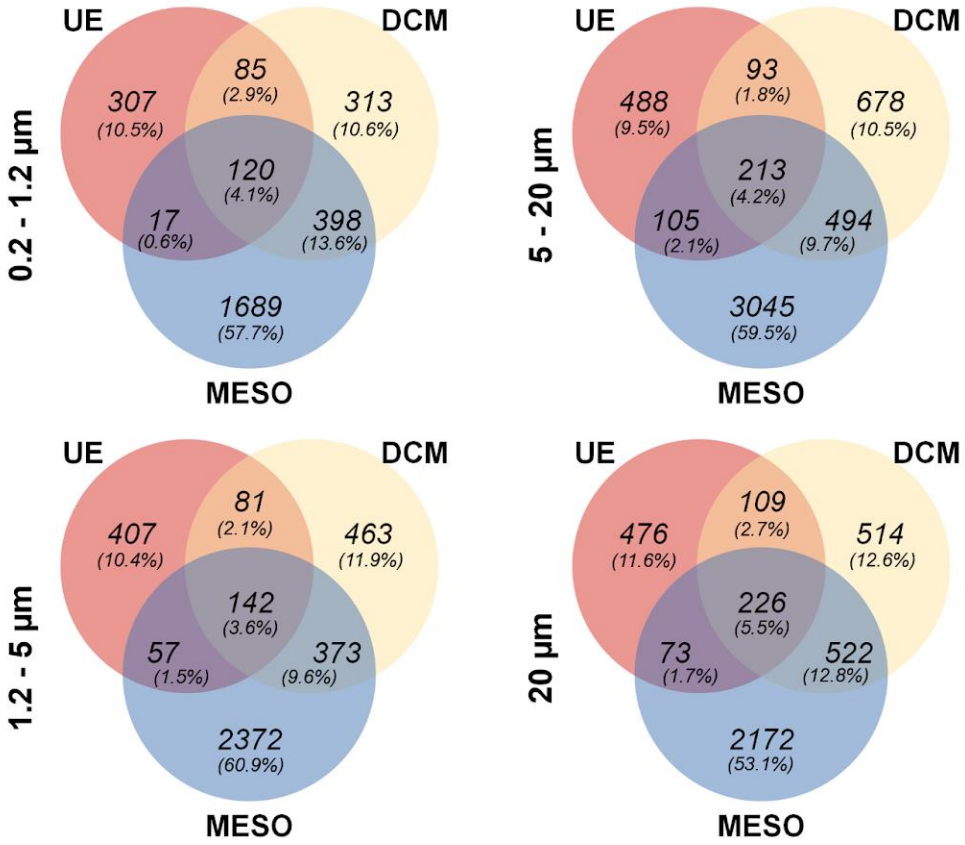


Figure 4.3: Venn diagrams showing unique and overlapping amplicon sequence variants (ASVs) between depth bins associated with the upper euphotic (EU), the deep chlorophyll maxima (DCM) and the mesopelagic (MESO) in a) the 0.2 - 1.2 µm fraction, b) the 1.2 - 5 µm fraction, c) the 5 - 20 µm fraction, and d) the >20 µm fraction.

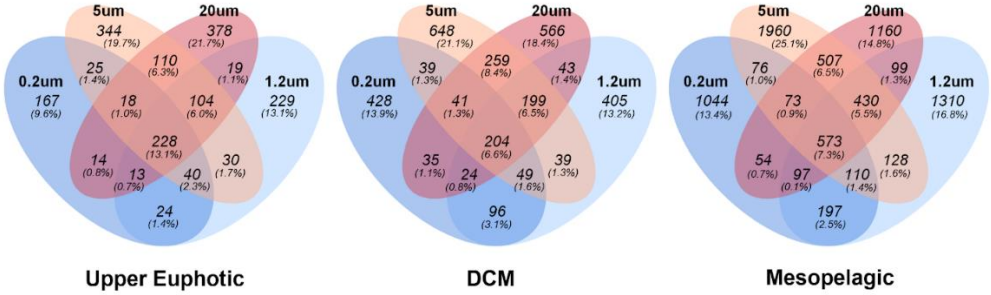


Figure 4.4: Venn diagrams showing unique and overlapping amplicon sequence variants (ASVs) between the 0.2- 1.2µm, 1.2 - 5µm, 5 - 20µm, and >20 µm size fractions in a) the upper euphotic, b) the deep chlorophyll maximum (DCM) and, c) the mesopelagic.

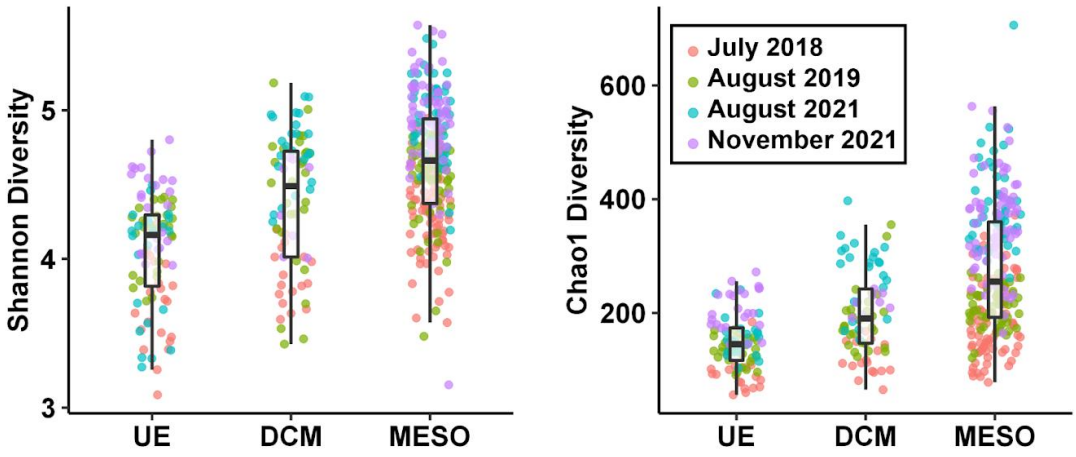


Figure 4.5: Alpha diversity of the 16S community in all size fractions excluding cyanobacteria.

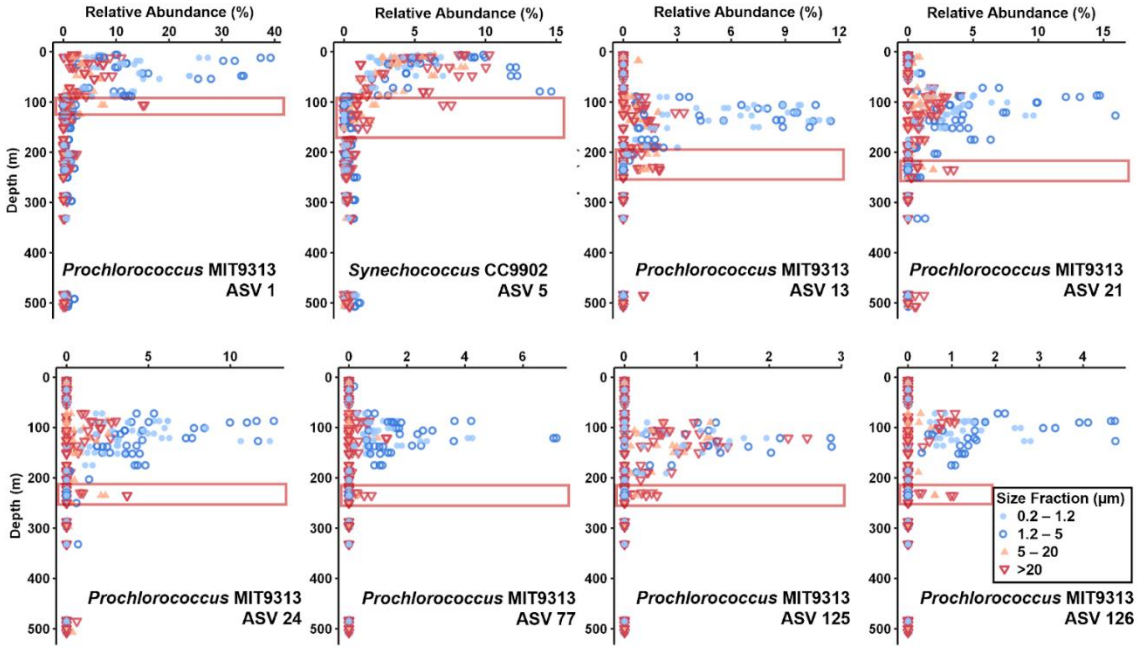


Figure 4.6: Cyanobacteria ASV relative abundance over depth. Red boxes denote depths where cyanobacteria are found with greater relative abundance in the particle-associated size fractions compared to free-living fractions.

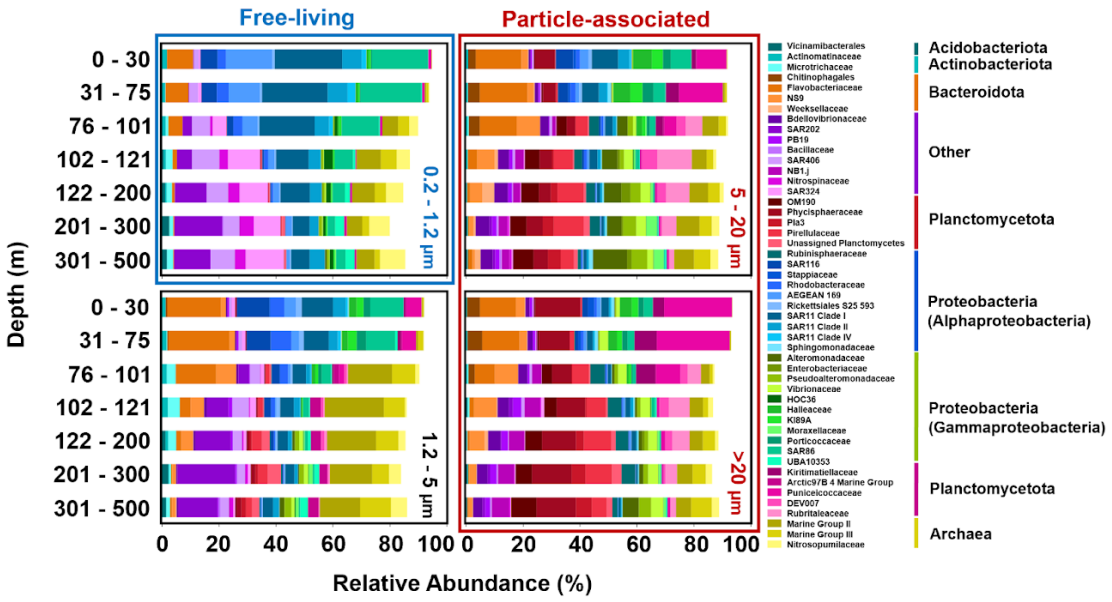
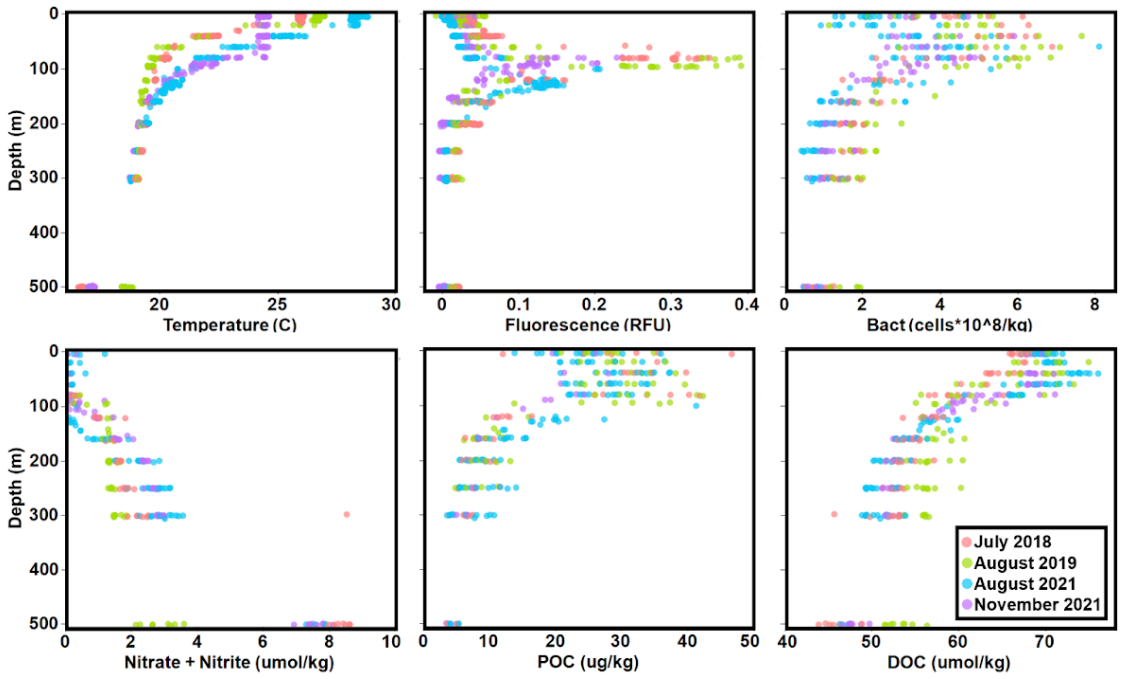
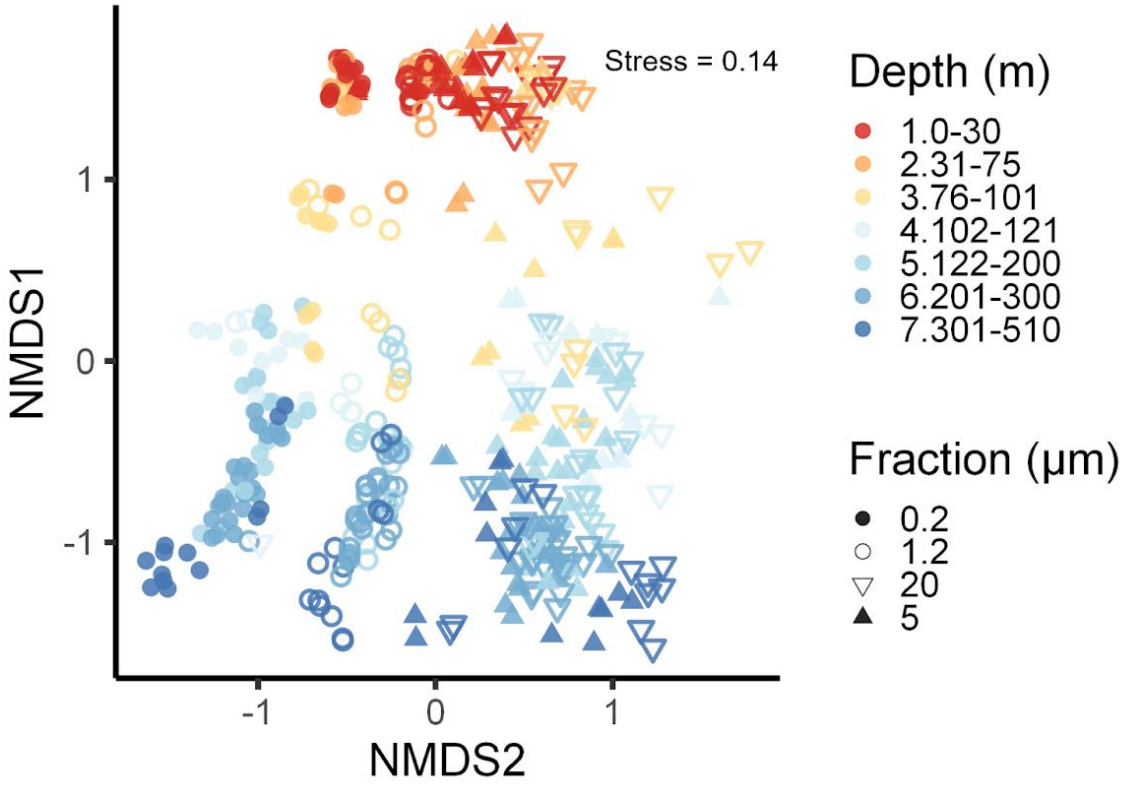


Figure 4.7: Microbial taxa (excluding cyanobacteria) grouped by family showing change in community composition across fraction and depth bins.

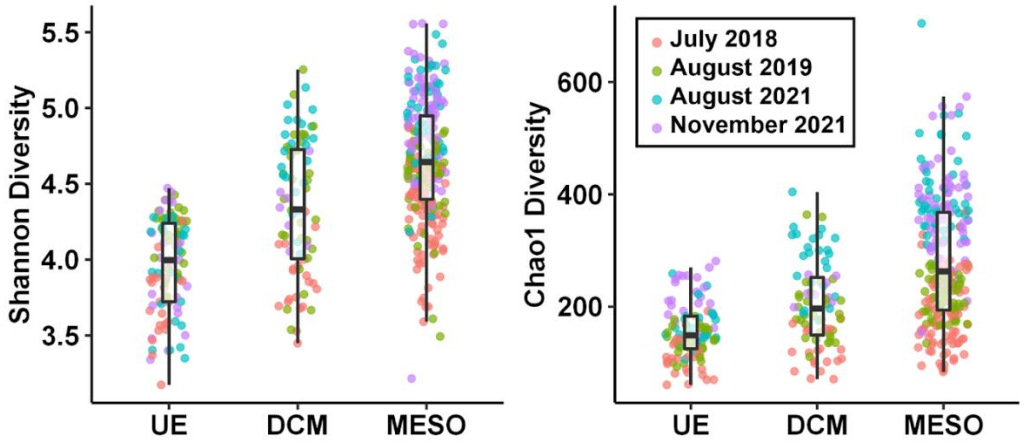
Appendix



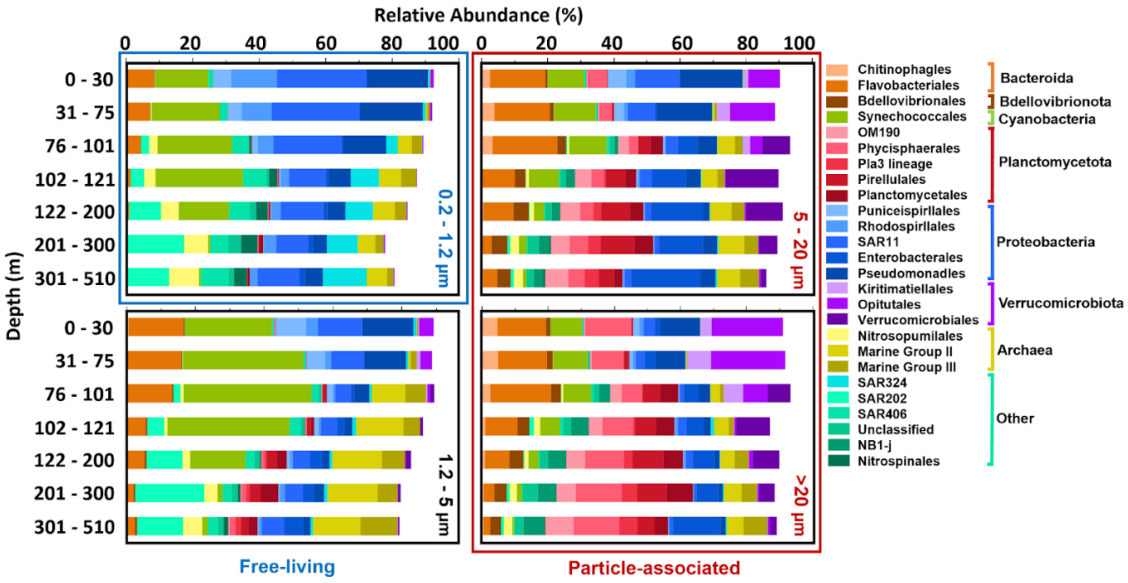
Appendix Figure 4.1: Depth profiles of the temperature, fluorescence, bacterial abundance, nitrate + nitrite, particulate organic carbon (POC) and dissolved organic carbon (DOC) across the upper 500 m of the water column during each of the time periods sampled.



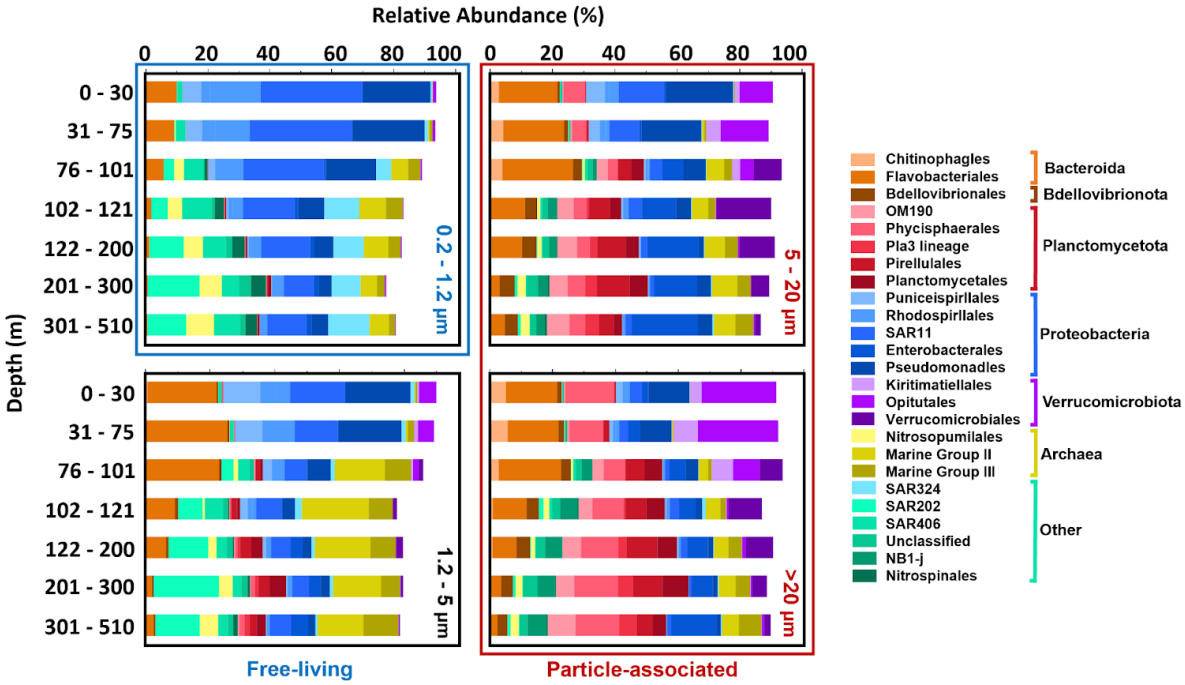
Appendix Figure 4.2: NMDS ordination of 16S community with cyanobacteria



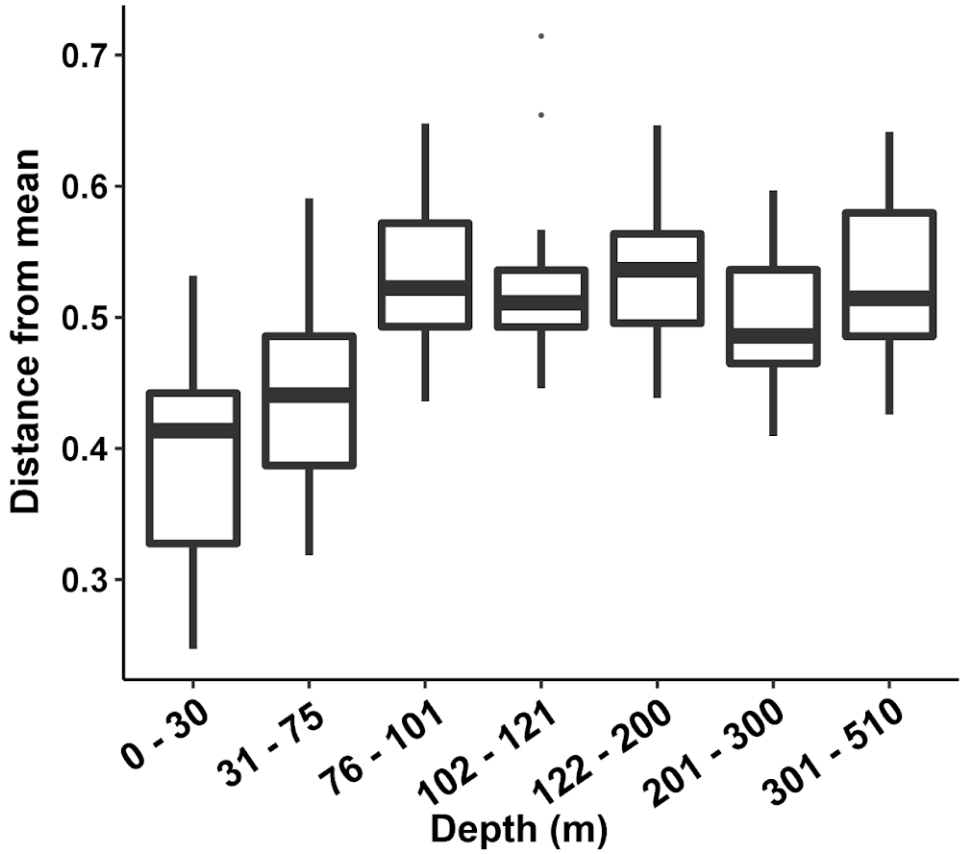
Appendix Figure 4.3: Alpha diversity of the 16S community of all size fractions with cyanobacteria



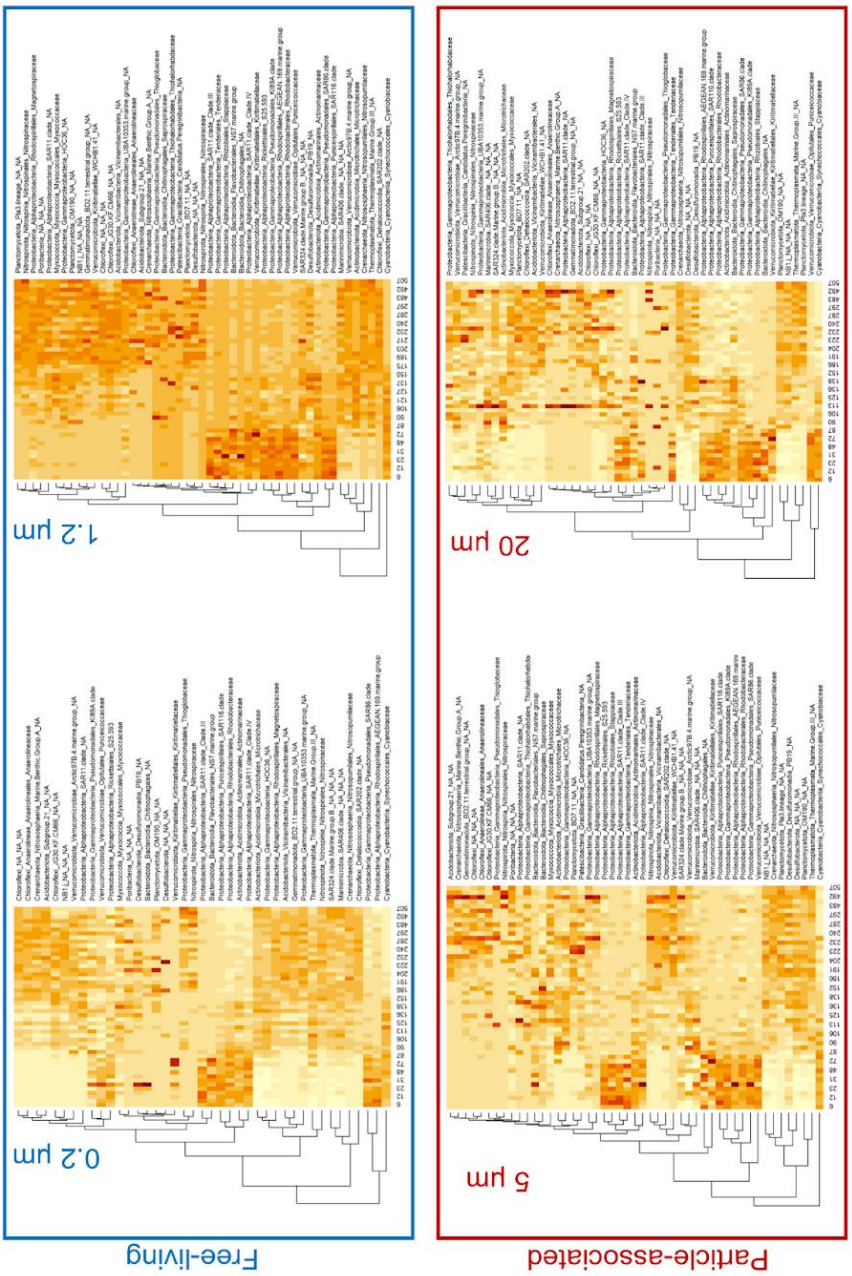
Appendix Figure 4.4: Microbial taxa (including cyanobacteria) grouped by order showing change in community across fraction and depth bins



Appendix Figure 4.5: Microbial taxa (excluding cyanobacteria) grouped by order showing change in community across fraction and depth bins



Appendix Figure 4.6: Distances from the mean community structure were calculated for each sample within each depth and values were visualized within depth bins. Communities within the upper water column (<75 m depth) displayed significantly (Tukey HSD $p < 0.05$) reduced distance to the mean than samples below 75 m.



Appendix Figure 4.7: Heatmap with cosmopolitan (present in >5 depths) families significantly enhanced in specific fractions & their relative abundance changes by depth

References

Abramson, L., Lee, C., Liu, Z., Wakeham, S. G., & Szlosek, J. (2010). Exchange between suspended and sinking particles in the northwest Mediterranean as inferred from the organic composition of in situ pump and sediment trap samples. *Limnology and Oceanography*, *55*(2), 725–739.

Adam, T. C., Burkepile, D. E., Holbrook, S. J., Carpenter, R. C., Claudet, J., Loiseau, C., Thiault, L., Brooks, A. J., Washburn, L., & Schmitt, R. J. (2020). Landscape-scale patterns of nutrient enrichment in a coral reef ecosystem: implications for coral to algae phase shifts. *Ecological Applications*, *31*(1). https://doi.org/10.1002/eap.2227

Alldredge, A. L. A., & Silver, M. W. (1988). Characteristics, Dynamics and Significance of Marine Snow. *Progress in Oceanography*, *20*(1), 41–82.

Alldredge, A. L., Cole, J. J., & Caron, D. A. (1986). Production of heterotrophic bacteria inhabiting macroscopic organic aggregates (marine snow) from surface waters. *Limnology and Oceanography*, *31*(1), 68–78. https://doi.org/10.4319/lo.1986.31.1.0068

Alldredge, A. L., Passow, U., & Logant, B. E. (1993). The abundance and significance of a class of large, transparent organic particles in the ocean. *Deep.Sea Research*, *1*(6), 1131–1140.

Altabet, M. A. (1990). Organic C, N, and stable isotopic composition of particulate matter collected on glass-fiber and aluminum oxide filters. In *Limnology and Oceanography* (Vol. 35, Issue 4, pp. 902–909). https://doi.org/10.4319/lo.1990.35.4.0902

Amacher, J., Neuer, S., & Lomas, M. (2013). DNA-based molecular fingerprinting of eukaryotic protists and cyanobacteria contributing to sinking particle flux at the Bermuda Atlantic time-series study. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *93*, 71–83. https://doi.org/10.1016/j.dsr2.2013.01.001

Andrade, I., Sangrà, P., Hormazabal, S., & Correa-Ramirez, M. (2014). Island mass effect in the Juan Fernández Archipelago (33°S), Southeastern Pacific. *Deep-Sea Research Part I: Oceanographic Research Papers*, *84*, 86–99. https://doi.org/10.1016/j.dsr.2013.10.009

Apprill, A., Holm, H., Santoro, A. E., Becker, C., Neave, M., Hughen, K., Donà, A. R., Aeby, G., Work, T., Weber, L., & McNally, S. (2021). Microbial ecology of coral-

dominated reefs in the Federated States of Micronesia. *Aquatic Microbial Ecology*, *86*, 115–136. https://doi.org/10.3354/AME01961

Apprill, A., Mcnally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Arnosti, C. (2011). Microbial extracellular enzymes and the marine carbon cycle. *Annual Review of Marine Science*, *3*, 401–425. https://doi.org/10.1146/annurev-marine-120709-142731

Arnosti, C., Wietz, M., Brinkhoff, T., Hehemann, J.-H., Probandt, D., Zeugner, L., & Amann, R. (2021). The Biogeochemistry of Marine Polysaccharides: Sources, Inventories, and Bacterial Drivers of the Carbohydrate Cycle. *Annual Reviews of Marine Science*, *13*, 81–108. https://doi.org/10.1146/annurev-marine-032020

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., & Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Marine ecology progress series*, 257-263.

Azam, F., & Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology*, *5*(10), 782–791. https://doi.org/10.1038/nrmicro1747

Azam, F., Smith, D. C., Steward, G. F., & Hagstrom, ~k. (1993). Sources of Carbon for the Microbial Loop Bacteria-Organic Matter Coupling and Its Significance for Oceanic Carbon Cycling. *Microb Ecol*, *28*, 167–179.

Baumas, C. M. J., Le Moigne, F. A. C., Garel, M., Bhairy, N., Guasco, S., Riou, V., Armougom, F., Grossart, H. P., & Tamburini, C. (2021). Mesopelagic microbial carbon production correlates with diversity across different marine particle fractions. *ISME Journal*, *15*(6), 1695–1708. https://doi.org/10.1038/s41396-020-00880-z

Bishop, J. K. B., Edmond, J. M., Ketten, D. R., Bacon, M. P., & Silker, W. B. (1977). The chemistry, biology, and vertical flux of particulate matter from the upper 400 m of the equatorial Atlantic Ocean. *Deep Sea Research*, *24*, 511–548.

Blanco-Bercial, L., Parsons, R., Bolaños, L. M., Johnson, R., Giovannoni, S. J., & Curry, R. (2022). The protist community traces seasonality and mesoscale hydrographic features in the oligotrophic Sargasso Sea. *Frontiers in Marine Science*, *9*. https://doi.org/10.3389/fmars.2022.897140

Boeuf, D., Edwards, B. R., Eppley, J. M., Hu, S. K., Poff, K. E., Romano, A. E., Caron, D. A., Karl, D. M., & DeLong, E. F. (2019). Biological composition and microbial dynamics of sinking particulate organic matter at abyssal depths in the oligotrophic open ocean. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(24), 11824–11832. https://doi.org/10.1073/pnas.1903080116

Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., & Weber, T. (2019). Multi-faceted particle pumps drive carbon sequestration in the ocean. In *Nature* (Vol. 568, Issue 7752, pp. 327–335). Nature Publishing Group. https://doi.org/10.1038/s41586-019-1098-2

Buesseler, K. O., Boyd, P. W., Black, E. E., Siegel, D. A., Designed, D. A. S., & Performed, E. E. B. (2020). Metrics that matter for assessing the ocean biological carbon pump. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(18), 9679–9687. https://doi.org/10.1073/pnas.1918114117/-/DCSupplemental

Buesseler, K. O., Lamborg, C. H., Boyd, P. W., Lam, P. J., Trull, T. W., Bidigare, R. R., Bishop, J. K. B., Casciotti, K. L., Dehairs, F., Elskens, M., Honda, M., Karl, D. M., Siegel, D. A., Silver, M. W., Steinberg, D. K., Valdes, J., Mooy, B. V., & Wilson, S. (2007). Revisiting Carbon Flux Through the Ocean's Twilight Zone. *Science, 316*, 567–570. https://doi.org/10.1016/j.pocean.2006.1010.1007

Burd, A. B., & Jackson, G. A. (2009). Particle aggregation. In *Annual Review of Marine Science* (Vol. 1, pp. 65–90). Annual Reviews Inc. https://doi.org/10.1146/annurev.marine.010908.163904

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Carlson, C. A., Ducklow, H. W., & Michaels, A. F. (1994). Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. *Nature, 371*, 405–408.

Carlson, C. A., Ducklowt, H. W., & Sleeters, T. D. (1996). Stocks and dynamics of bacterioplankton in the northwestern Sargasso Sea. *DeepSea Research II*, *43*(3), 491–515.

Carlson, Craig A., & Hansell, D. A. (2015). DOM Sources, Sinks, Reactivity, and Budgets. In *Biogeochemistry of Marine Dissolved Organic Matter* (pp. 65–126). Elsevier. https://doi.org/10.1016/B978-0-12-405940-5.00003-0

Carlson, C. A., Morris, R., Parsons, R., Treusch, A. H., Giovannoni, S. J., & Vergin, K. (2009). Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones

of the northwestern Sargasso Sea. *ISME Journal*, *3*(3), 283–295. https://doi.org/10.1038/ismej.2008.117

Casciotti, K. L., Trull, T. W., Glover, D. M., & Davies, D. (2008). Constraints on nitrogen cycling at the subtropical North Pacific Station ALOHA from isotopic measurements of nitrate and particulate nitrogen. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *55*(14–15), 1661–1672. https://doi.org/10.1016/j.dsr2.2008.04.017

Chavez, F., and J. Toggweiler (1995), Physical estimates of global new production: the upwelling contribution, in *Upwelling in the Ocean: Modern Processes and Ancient Records*, edited by C. Summerhayes, K. Emeis, M. Angel, R. Smith and B. Zeitzschel, pp. 313-320, John Wiley and Sons, New York.

Chikaraishi, Y., Ogawa, N. O., Kashiyama, Y., Takano, Y., Suga, H., Tomitani, A., Miyashita, H., Kitazato, H., & Ohkouchi, N. (2009). Determination of aquatic food-web structure based on compound-specific nitrogen isotopic composition of amino acids. *Limnology and Oceanography: Methods*, *7*(NOV), 740–750. https://doi.org/10.4319/lom.2009.7.740

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

Cho, B. C., & Azam, F. (1990). Biogeochemical significance of bacterial biomass in the ocean's euphotic zone. *Marine Ecology Progress Series*, *63*, 253–259.

Chow, J. S., Lee, C., & Engel, A. (2015). The influence of extracellular polysaccharides, growth rate, and free coccoliths on the coagulation efficiency of Emiliania huxleyi. *Marine Chemistry*, *175*, 5–17. https://doi.org/10.1016/j.marchem.2015.04.010

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

Cottrell, M. T., & Kirchman, D. L. (2000). Natural Assemblages of Marine Proteobacteria and Members of the Cytophaga-Flavobacter Cluster Consuming Low-and High-Molecular-Weight Dissolved Organic Matter. *Applied and Environmental Microbiology*, *66*, 1692–1697. https://journals.asm.org/journal/aem

Cruz, B. N., Brozak, S., & Neuer, S. (2021). Microscopy and DNA-based characterization of sinking particles at the Bermuda Atlantic Time-series Study station point to zooplankton mediation of particle flux. *Limnology and Oceanography*, *66*(10), 3697–3713. https://doi.org/10.1002/lno.11910

D'Elia C. F., Webb K. L. (1977) The dissolved nitrogen flux of reef corals. In: Proceedings of the third international coral reef symposium, pp 325–331

Dadon-Pilosof, A., Conley, K. R., Jacobi, Y., Haber, M., Lombard, F., Sutherland, K. R., Steindler, L., Tikochinski, Y., Richter, M., Glöckner, F. O., Suzuki, M. T., West, N. J., Genin, A., & Yahel, G. (2017). Surface properties of SAR11 bacteria facilitate grazing avoidance. *Nature Microbiology*, *2*(12), 1608–1615. https://doi.org/10.1038/s41564-017-0030-5

Dam, G. D., & Drapeau D T. (1995). Coagulation efficiency, organic-matter glues and the dynamics of particles during a phytoplankton bloom in a mesocosm study. *Deep-Sea Research II*, *42*(1), 111–123.

Dandonneau, Y., & Charpyt, L. (1985). An empirical approach to the island mass effect in the south tropical Pacific based on sea surface chlorophyll concentrations. *Deep-Sea Research*, *32*(6), 707–721.

Datta, M. S., Sliwerska, E., Gore, J., Polz, M. F., & Cordero, O. X. (2016). Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms11965

DeLong E F, Preston C M, Mincer T, Rich V, Hallan S J, Frigaard N, Martinez A, Sullivan M B, Edwards R, Brito B R, Chisholm S W, & Karl D M. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science*, *311*, 496–503. https://doi.org/10.1126/science.1122876

DeLong, E. F., Franks, D. G., & Alldredge, A. L. (1993). Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnology and Oceanography*, *38*(5), 924–934. https://doi.org/10.4319/lo.1993.38.5.0924

DeVries, T., & Weber, T. (2017). The export and fate of organic matter in the ocean: New constraints from combining satellite and oceanographic tracer observations. *Global Biogeochemical Cycles*, *31*(3), 535–555. https://doi.org/10.1002/2016GB005551

Dilling, L., & Alldredge, A. L. (2000). Fragmentation of marine snow by swimming macrozooplankton: A new process impacting carbon cycling in the sea. *Deep-Sea Research I*, *47*, 1227–1245.

Doherty, S. C., Maas, A. E., Steinberg, D. K., Popp, B. N., & Close, H. G. (2021). Distinguishing zooplankton fecal pellets as a component of the biological pump using compound-specific isotope analysis of amino acids. *Limnology and Oceanography*, *66*(7), 2827–2841. https://doi.org/10.1002/lno.11793

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

Duarte, C. M., & Cebrián, J. (1996). The fate of marine autotrophic production. *Limnology and Oceanography*, *41*(8), 1758–1766. https://doi.org/10.4319/lo.1996.41.8.1758

Ducklow, H. W., Steinberg, D. K., & Buesseler, K. O. (2001). Upper Ocean Carbon Export and the Biological Pump. *Oceanography*, *14*(4), 50–58.

Duret, M. T., Lampitt, R. S., & Lam, P. (2019). Prokaryotic niche partitioning between suspended and sinking marine particles. *Environmental Microbiology Reports*, *11*(3), 386–400. https://doi.org/10.1111/1758-2229.12692

Ebersbach, F., Assmy, P., Martin, P., Schulz, I., Wolzenburg, S., & Nöthig, E. M. (2014). Particle flux characterisation and sedimentation patterns of protistan plankton during the iron fertilisation experiment LOHAFEX in the Southern Ocean. *Deep-Sea Research Part I: Oceanographic Research Papers*, *89*, 94–103. https://doi.org/10.1016/j.dsr.2014.04.007

Emeis, K. C., Mara, P., Schlarbaum, T., Möbius, J., Dähnke, K., Struck, U., Mihalopoulos, N., & Krom, M. (2010). External N inputs and internal N cycling traced by isotope ratios of nitrate, dissolved reduced nitrogen, and particulate nitrogen in the eastern Mediterranean Sea. *Journal of Geophysical Research: Biogeosciences*, *115*(4). https://doi.org/10.1029/2009JG001214

Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S. G., González, J. M., & Pedrós-Alió, C. (2013). Ecology of marine bacteroidetes: A comparative genomics approach. *ISME Journal*, *7*(5), 1026–1037. https://doi.org/10.1038/ismej.2012.169

Frade, P. R., Glasl, B., Matthews, S. A., Mellin, C., Serrão, E. A., Wolfe, K., Mumby, P. J., Webster, N. S., & Bourne, D. G. (2020). Spatial patterns of microbial communities across surface waters of the Great Barrier Reef. *Communications Biology*, *3*(1). https://doi.org/10.1038/s42003-020-01166-y

Francis, T. Ben, Bartosik, D., Sura, T., Sichert, A., Hehemann, J. H., Markert, S., Schweder, T., Fuchs, B. M., Teeling, H., Amann, R. I., & Becher, D. (2021). Changing expression patterns of TonB-dependent transporters suggest shifts in polysaccharide consumption over the course of a spring phytoplankton bloom. *ISME Journal*, *15*(8), 2336–2350. https://doi.org/10.1038/s41396-021-00928-8

Francois, R., Honjo, S., Krishfield, R., & Manganini, S. (2002). Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean. *Global Biogeochemical Cycles*, *16*(4), 34-1-34–20. https://doi.org/10.1029/2001gb001722

Fuchsman, C. A., Staley, J. T., Oakley, B. B., Kirkpatrick, J. B., & Murray, J. W. (2012). Free-living and aggregate-associated Planctomycetes in the Black Sea. *FEMS Microbiology Ecology*, *80*(2), 402–416. https://doi.org/10.1111/j.1574-6941.2012.01306.x

Fuhrman, J. A., Hewson, I., Schwalbach, M. S., Steele, J. A., Brown, M. V, & Naeem, S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(35), 13104–13109. www.pnas.orgcgidoi10.1073pnas.0602399103

Ganesh, S., Parris, D. J., Delong, E. F., & Stewart, F. J. (2014). Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME Journal*, *8*(1), 187–211. https://doi.org/10.1038/ismej.2013.144

Gast G J, Wiegman, S., Wieringa, E., van Duyll, F. C., & Bak R P. (1998). Bacteria in coral reef water types: removal of cells, stimulation of growth and mineralization. *Marine Ecology Progress Series*, *167*, 37–45.

Giering, S. L. C., Sanders, R., Lampitt, R. S., Anderson, T. R., Tamburini, C., Boutrif, M., Zubkov, M. V., Marsay, C. M., Henson, S. A., Saw, K., Cook, K., & Mayor, D. J. (2014). Reconciliation of the carbon budget in the ocean's twilight zone. *Nature*, *507*(7493), 480–483. https://doi.org/10.1038/nature13123

Giovannoni S J, Britschgi T B, Moyer C L, & Field K G. (1990). Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, *345*, 60–63.

Giovannoni, S. J., & Vergin, K. L. (2012). Seasonality in Ocean Microbial Communities. *Science*, *335*, 671–676. https://www.science.org

Giovannoni, S. J., Rappei, M. S., Vergin, K. L., & Adair, N. L. (1996). 6S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria (molecular ecology/phylogeny/thermophily). *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 7979–7984. https://www.pnas.org

Goldberg, S. J., Carlson, C. A., Hansell, D. A., Nelson, N. B., & Siegel, D. A. (2009). Temporal dynamics of dissolved combined neutral sugars and the quality of dissolved organic matter in the Northwestern Sargasso Sea. *Deep-Sea Research Part I: Oceanographic Research Papers*, *56*(5), 672–685. https://doi.org/10.1016/j.dsr.2008.12.013

Goldthwait, S., Yen, J., Brown, J., & Alldredge, A. (2004). Quantification of marine snow fragmentation by swimming euphausiids. *Limnology and Oceanography*, *49*(4 I), 940–952. https://doi.org/10.4319/lo.2004.49.4.0940

Gordon, D. A., & Giovannoni, S. J. (1996). Detection of Stratified Microbial Populations Related to Chlorobium and Fibrobacter Species in the Atlantic and Pacific Oceans. *Applied and Environmental Microbiology*, *62*(4), 1171–1177. https://journals.asm.org/journal/aem

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*. https://doi.org/10.1038/ncomms10581

Gowing, M. M., & Wishner, K. E. (1992). Feeding ecology of benthopelagic zooplankton on an eastern tropical Pacific seamount. *Marine Biology*, *112*, 451–467.

Grabowski, E., Letelier, R. M., Laws, E. A., & Karl, D. M. (2019). Coupling carbon and energy fluxes in the North Pacific Subtropical Gyre. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09772-z

Grossart, H. P. (2010). Ecological consequences of bacterioplankton lifestyles: Changes in concepts are needed. *Environmental Microbiology Reports*, *2*(6), 706–714. https://doi.org/10.1111/j.1758-2229.2010.00179.x

Grossart, H. P., Tang, K. W., Kiørboe, T., & Ploug, H. (2007). Comparison of cell-specific activity between free-living and attached bacteria using isolates and natural assemblages. *FEMS Microbiology Letters*, *266*(2), 194–200. https://doi.org/10.1111/j.1574-6968.2006.00520.x

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Darzi, Y., Audic, S., Berline, L., Brum, J. R., Coelho, L. P., Espinoza, J. C. I., Malviya, S., Sunagawa, S., Dimier, C., Kandels-Lewis, S., Picheral, M., Poulain, J., Searson, S., … Gorsky, G. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, *532*(7600), 465–470. https://doi.org/10.1038/nature16942

Guidi, L., Legendre, L., Reygondeau, G., Uitz, J., Stemmann, L., & Henson, S. A. (2015). A new look at ocean carbon remineralization for estimating deepwater sequestration. *Global Biogeochemical Cycles*, *29*(7), 1044–1059. https://doi.org/10.1002/2014GB005063

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., Picheral, M., & Gorsky, G. (2009). Effects of phytoplankton community on production, size and export of large aggregates: A world-ocean analysis. *Limnology and Oceanography*, *54*(6), 1951–1963. https://doi.org/10.4319/lo.2009.54.6.1951

Haas, A. F., Nelson, C. E., Kelly, L. W., Carlson, C. A., Rohwer, F., Leichter, J. J., Wyatt, A., & Smith, J. E. (2011). Effects of coral reef benthic primary producers on dissolved organic carbon and microbial activity. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027973

Hansell, D. A. (2013). Recalcitrant Dissolved Organic Carbon Fractions. *Annual Review of Marine Science*, *5*(1), 421–445. https://doi.org/10.1146/annurev-marine-120710-100757

Hansell, D., Carlson, C., Repeta, D., & Schlitzer, R. (2009). Dissolved Organic Matter in the Ocean: A Controversy Stimulates New Insights. *Oceanography*, *22*(4), 202–211. https://doi.org/10.5670/oceanog.2009.109

Halewood, E., Opalk, K., Custals, L., Carey, M., Hansell, D., & Carlson, C. A. (2022). GO-SHIP Repeat Hydrography: Determination of dissolved organic carbon (DOC) and total dissolved nitrogen (TDN) in seawater using High Temperature Combustion Analysis.[GOOS ENDORSED PRACTICE].

Hannides, C. C. S., Popp, B. N., Anela Choy, C., & Drazen, J. C. (2013). Midwater zooplankton and suspended particle dynamics in the North Pacific Subtropical Gyre: A stable isotope perspective. *Limnology and Oceanography*, *58*(6), 1931–1946. https://doi.org/10.4319/lo.2013.58.6.1931

Hansell, D. A., Carlson, C. A., & Suzuki, Y. (2002). Dissolved organic carbon export with North Pacific Intermediate Water formation. *Global Biogeochemical Cycles*, *16*(1), 7-1-7–8. https://doi.org/10.1029/2000gb001361

Heins, A., Reintjes, G., Amann, R. I., & Harder, J. (2021). Particle Collection in Imhoff Sedimentation Cones Enriches Both Motile Chemotactic and Particle-Attached Bacteria. *Frontiers in Microbiology, 12*. https://doi.org/10.3389/fmicb.2021.643730

Hench, J. L., Leichter, J. J., & Monismith, S. G. (2008). Episodic circulation and exchange in a wave-driven coral reef and lagoon system. *Limnology and Oceanography*, *53*(6), 2681–2694. https://doi.org/10.4319/lo.2008.53.6.2681

Henson, S. A., Sanders, R., & Madsen, E. (2012a). Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean. *Global Biogeochemical Cycles*, *26*(1). https://doi.org/10.1029/2011GB004099

Henson, S. A., Sanders, R., & Madsen, E. (2012b). Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean. *Global Biogeochemical Cycles*, *26*(1). https://doi.org/10.1029/2011GB004099

Ivanova, E. P., López-Pérez, M., Zabalos, M., Nguyen, S. H., Webb, H. K., Ryan, J., Lagutin, K., Vyssotski, M., Crawford, R. J., & Rodriguez-Valera, F. (2015). Ecophysiological diversity of a novel member of the genus Alteromonas, and description of Alteromonas mediterranea sp. nov. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *107*(1), 119–132. https://doi.org/10.1007/s10482-014-0309-y

Ivars-Martinez, E., Martin-Cuadrado, A. B., D'Auria, G., Mira, A., Ferriera, S., Johnson, J., Friedman, R., & Rodriguez-Valera, F. (2008). Comparative genomics of two ecotypes of the marine planktonic copiotroph Alteromonas macleodii suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME Journal*, *2*(12), 1194–1212. https://doi.org/10.1038/ismej.2008.74

James, A. K., Washburn, L., Gotschalk, C., Maritorena, S., Alldredge, A., Nelson, C. E., Hench, J. L., Leichter, J. J., Wyatt, A. S. J., & Carlson, C. A. (2020). An Island Mass Effect Resolved Near Mo'orea, French Polynesia. *Frontiers in Marine Science, 7*. https://doi.org/10.3389/fmars.2020.00016

Jarman, C. L., Larsen, T., Hunt, T., Lipo, C., Solsvik, R., Wallsgrove, N., Ka'apu-Lyons, C., Close, H. G., & Popp, B. N. (2017). Diet of the prehistoric population of Rapa Nui (Easter Island, Chile) shows environmental adaptation and resilience. *American Journal of Physical Anthropology*, *164*(2), 343–361. https://doi.org/10.1002/ajpa.23273

Johannes, R. E., Alberts, J., Kinzie, R. A., Pomeroy, L. R., Sottile, W. , & et al. (1972). The Metabolism of Some Coral Reef Communities: A Team Study of Nutrient and Energy

Flux at Eniwetok. *Bioscience*, 541–543. https://academic.oup.com/bioscience/article/22/9/541/264573

Jones, E. C. (1962). Evidence of an Island Effect upon the Standing Crop of Zooplankton near the Marquesas Islands, Central Pacific. *ICES Journal of Marine Science*, *27*, 223–231. https://academic.oup.com/icesjms/article/27/3/223/738905

Kaiser, K., & Benner, R. (2008). Major bacterial contribution to the ocean reservoir of detrital organic carbon and nitrogen. *Limnology and Oceanography*, *53*(1), 99–112. https://doi.org/10.4319/lo.2008.53.1.0099

Kaiser, K., & Benner, R. (2009). Biochemical composition and size distribution of organic matter at the Pacific and Atlantic time-series stations. *Marine Chemistry*, *113*(1–2), 63–77. https://doi.org/10.1016/j.marchem.2008.12.004

Kaiser, K., & Benner, R. (2012). Characterization of lignin by gas chromatography and mass spectrometry using a simplified CuO oxidation method. *Analytical Chemistry*, *84*(1), 459–464. https://doi.org/10.1021/ac202004r

Kappelmann, L., Krüger, K., Hehemann, J. H., Harder, J., Markert, S., Unfried, F., Becher, D., Shapiro, N., Schweder, T., Amann, R. I., & Teeling, H. (2019). Polysaccharide utilization loci of North Sea Flavobacteriia as basis for using SusC/D-protein expression for predicting major phytoplankton glycans. *ISME Journal*, *13*(1), 76–91. https://doi.org/10.1038/s41396-018-0242-6

Karl, D. M., Church, M. J., Dore, J. E., Letelier, R. M., & Mahaffey, C. (2012). Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(6), 1842–1849. https://doi.org/10.1073/pnas.1120312109

Kelly, L. W., Nelson, C. E., Haas, A. F., Naliboff, D. S., Calhoun, S., Carlson, C. A., Edwards, R. A., Fox, M. D., Hatay, M., Johnson, M. D., Kelly, E. L. A., Lim, Y. W., Macherla, S., Quinlan, Z. A., Silva, G. G. Z., Vermeij, M. J. A., Zgliczynski, B., Sandin, S. A., Smith, J. E., & Rohwer, F. (2019). Diel population and functional synchrony of microbial communities on coral reefs. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-09419-z

Kelly, L. W., Nelson, C. E., Petras, D., Koester, I., Quinlan, Z. A., Arts, M. G. I., Nothias, L.-F., Comstock, J., White, B. M., Hopmans, E. C., Van Duyl, F. C., Carlson, C. A., Aluwihare, L. I., Dorrestein, P. C., Haas, A. F., Designed, A. F. H., & Performed, A. F. H. (2022). Distinguishing the molecular diversity, nutrient content, and energetic

potential of exometabolomes produced by macroalgae and reef-building corals. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(5). https://doi.org/10.1073/pnas.2110283119/-/DCSupplemental

Kelly, L. W., Williams, G. J., Barott, K. L., Carlson, C. A., Dinsdale, E. A., Edwards, R. A., Haas, A. F., Haynes, M., Lim, Y. W., McDole, T., Nelson, C. E., Sala, E., Sandin, S. A., Smith, J. E., Vermeij, M. J. A., Youle, M., & Rohwer, F. (2014). Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), 10227–10232. https://doi.org/10.1073/pnas.1403319111

Knee, K. L., Crook, E. D., Hench, J. L., Leichter, J. J., & Paytan, A. (2016). Assessment of Submarine Groundwater Discharge (SGD) as a Source of Dissolved Radium and Nutrients to Moorea (French Polynesia) Coastal Waters. *Estuaries and Coasts*, *39*(6), 1651–1668. https://doi.org/10.1007/s12237-016-0108-y

Kujawinski, E. B. (2011). The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annual Review of Marine Science*, *3*(1), 567–599. https://doi.org/10.1146/annurev-marine-120308-081003

Legendre, L., Demers, S., Delesalle, B., & Harnois, C. (1988). Biomass and photosynthetic activity of phototrophic picoplankton in coral reef waters (Moorea Island, French Polynesia). *Marine Ecology Progress Series*, *47*, 153–160.

Leichter, J. J., Alldredge, A. L., Bernardi, G., Brooks, A. J., Carlson, C. A., Carpenter, R. C., Edmunds, P. J., Fewings, M. R., Hanson, K. M., Hench, J. L., Holbrook, S. J., Nelson, C. E., Schmitt, R. J., Toonen, R. J., Washburn, L., & Wyatt, A. S. J. (2013). Biological and physical interactions on a tropical island coral reef: Transport and retention processes on moorea, French Polynesia. *Oceanography*, *26*(3), 52–63. https://doi.org/10.5670/oceanog.2013.45

Liu, S., Longnecker, K., Kujawinski, E. B., Vergin, K., Bolaños, L. M., Giovannoni, S. J., Parsons, R., Opalk, K., Halewood, E., Hansell, D. A., Johnson, R., Curry, R., & Carlson, C. A. (2022). Linkages Among Dissolved Organic Matter Export, Dissolved Metabolites, and Associated Microbial Community Structure Response in the Northwestern Sargasso Sea on a Seasonal Scale. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.833252

Lomas, M. W., Bates, N. R., Johnson, R. J., Knap, A. H., Steinberg, D. K., & Carlson, C. A. (2013). Two decades and counting: 24-years of sustained open ocean biogeochemical measurements in the Sargasso Sea. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *93*, 16–32. https://doi.org/10.1016/j.dsr2.2013.01.008

Lomas, M. W., Steinberg, D. K., Dickey, T., Carlson, C. A., Nelson, N. B., Condon, R. H., & Bates, N. R. (2010). Increased ocean carbon export in the Sargasso Sea linked to climate variability is countered by its enhanced mesopelagic attenuation. *Biogeosciences*, *7*, 57–70. www.biogeosciences.net/7/57/2010/

Longhurst, A. R., & Harrison, W. G. (1988). Vertical nitrogen flux from the oceanic photic zone by diel migrant zooplankton and nekton. *Deep-Sea Research*, *35*(6), 881–889.

Longhurst, A. R., Bedo, A. W., Harrison, W. G., Head, E. J. H., & Sameoto, D. D. (1990). Vertical flux of respiratory carbon by oceanic diel migrant biota. *Deep-Sea Research*, *37*(4), 685–694.

Longhurst, A., Sathyendranath, S., Platt, T., & Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research*, *17*(6), 1245–1271. https://academic.oup.com/plankt/article-abstract/17/6/1245/1441236

López-López, A., Bartual, S. G., Stal, L., Onyshchenko, O., & Rodríguez-Valera, F. (2005). Genetic analysis of housekeeping genes reveals a deep-sea ecotype of Alteromonas macleodii in the Mediterranean Sea. *Environmental Microbiology*, *7*(5), 649–659. https://doi.org/10.1111/j.1462-2920.2004.00733.x

López-Pérez, M., & Rodriguez-Valera, F. (2016). Pangenome evolution in themarine bacterium alteromonas. *Genome Biology and Evolution*, *8*(5), 1556–1570. https://doi.org/10.1093/gbe/evw098

López-Pérez, M., Gonzaga, A., & Rodriguez-Valera, F. (2013). Genomic diversity of "deep ecotype" Alteromonas macleodii isolates: Evidence for pan-mediterranean clonal frames. *Genome Biology and Evolution*, *5*(6), 1220–1232. https://doi.org/10.1093/gbe/evt089

López-Pérez, M., Gonzaga, A., Ivanova, E. P., & Rodriguez-Valera, F. (2014). Genomes of Alteromonas australica, a world apart. *BMC Genomics*, *15*(1). https://doi.org/10.1186/1471-2164-15-483

Marín, I. (2014)."Proteobacteria," in Encyclopedia of astrobiology. Eds. R. Amils, M.Gargaud, J. Cernicharo Quintanilla, H. J. Cleaves, W. M. Irvine, D. Pinti and M. Viso(Berlin, Heidelberg: Springer Berlin Heidelberg), 1–2.

Marsay, C. M., Sanders, R. J., Henson, S. A., Pabortsava, K., Achterberg, E. P., & Lampitt, R. S. (2015). Attenuation of sinking particulate organic carbon flux through the mesopelagic ocean. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(4), 1089–1094. https://doi.org/10.1073/pnas.1415311112

Martin, J. H., Knauer, G. A., Karl, K. M., & Broenkow, W. W. (1987). VERTEX: carbon cycling in the northeast Pacific. *Deep-Sea Research*, *34*(2), 267–285.

Martinez, E., & Maamaatuaiahutapu, K. (2004). Island mass effect in the Marquesas Islands: Time variation. *Geophysical Research Letters*, *31*(18). https://doi.org/10.1029/2004GL020682

McCave, I. N. (1975). Vertical flux of particles in the ocean. *Deep Sea Research and Oceanographic Abstracts*, *22*(7), 491–502.

McClelland, J. W., & Montoya, J. P. (2002). Trophic relationships and the nitrogen isotopic composition of amino acids in plankton. *Ecology*, *83*(8), 2173–2180. https://doi.org/10.1890/0012-9658(2002)083[2173:TRATNI]2.0.CO;2

McCliment, E. A., Nelson, C. E., Carlson, C. A., Alldredge, A. L., Witting, J., & Amaral-Zettler, L. A. (2012). An all-taxon microbial inventory of the Moorea coral reef ecosystem. *ISME Journal*, *6*(2), 309–319. https://doi.org/10.1038/ismej.2011.108

McNally, S. P., Parsons, R. J., Santoro, A. E., & Apprill, A. (2017). Multifaceted impacts of the stony coral Porites astreoides on picoplankton abundance and community composition. *Limnology and Oceanography*, *62*(1), 217–234. https://doi.org/10.1002/lno.10389

Mestre, M., Ruiz-González, C., Logares, R., Duarte, C. M., Gasol, J. M., & Sala, M. M. (2018). Sinking particles promote vertical connectivity in the ocean microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(29), E6799–E6807. https://doi.org/10.1073/pnas.1802470115

Michael, V., Frank, O., Bartling, P., Scheuner, C., Göker, M., Brinkmann, H., & Petersen, J. (2016). Biofilm plasmids with a rhamnose operon are widely distributed determinants of the “swim-or-stick” lifestyle in roseobacters. *ISME Journal*, *10*(10), 2498–2513. https://doi.org/10.1038/ismej.2016.30

Michaels, A. F., & Silver, M. W. (1988). Primary production, sinking fluxes and the microbial food web. *Deep-Sea Research*, *35*(4), 473–490.

Moorea Coral Reef LTER and Washburn, L. 2019. MCR LTER: Coral Reef: Temperature and Salinity subset of moorings FOR01, FOR04 and FOR05 from 2005-2014 ver 5. Environmental Data Initiative. https://doi.org/10.6073/pasta/6a506c7e46d737304468e960ddbc1303

Moormann, M., Zähringer, U., Moll, H., Kaufmann, R., Schmid, R., & Altendorf, K. (1997). A new glycosylated lipopeptide incorporated into the cell wall of a smooth variant of Gordona hydrophobica. *Journal of Biological Chemistry*, *272*(16), 10729–10738. https://doi.org/10.1074/jbc.272.16.10729

Moriarty, D. J. W. (1979). Biomass of suspended bacteria over coral reefs. *Marine Biology*, *53*, 193–195.

Moriarty, D. J. W., Pollard, P. C., & Hunt, W. G. (1985). Temporal and spatial variation in bacterial production in the water column over a coral reef. *Marine Biology*, *85*, 285–292.

Morris, R. M., Vergin, K. L., Cho, J. C., Rappé, M. S., Carlson, C. A., & Giovannoni, S. J. (2005). Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. *Limnology and Oceanography*, *50*(5), 1687–1696. https://doi.org/10.4319/lo.2005.50.5.1687

Neall, V. E., & Trewick, S. A. (2008). The age and origin of the Pacific islands: A geological overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1508), 3293–3308. https://doi.org/10.1098/rstb.2008.0119

Nelson, C. E., & Carlson, C. A. (2012). Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environmental Microbiology*, *14*(6), 1500–1516. https://doi.org/10.1111/j.1462-2920.2012.02738.x

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. *ISME Journal*, *5*(8), 1374–1387. https://doi.org/10.1038/ismej.2011.12

Nelson, C. E., Goldberg, S. J., Wegley Kelly, L., Haas, A. F., Smith, J. E., Rohwer, F., & Carlson, C. A. (2013). Coral and macroalgal exudates vary in neutral sugar composition and differentially enrich reef bacterioplankton lineages. *ISME Journal*, *7*(5), 962–979. https://doi.org/10.1038/ismej.2012.161

Nelson, C. E., Kelly, L. W., & Haas, A. F. (2023). Microbial Interactions with Dissolved Organic Matter Are Central to Coral Reef Ecosystem Function and Resilience. *Annual Reviews of Marine Science*, *15*, 431–460. https://doi.org/10.1146/annurev-marine-042121

Nelson, C. E., & Wear, E. K. (2014). Microbial diversity and the lability of dissolved organic carbon. *Proceedings of the National Academy of Sciences*, *111*(20), 7166–7167. https://doi.org/10.1073/pnas.1405751111

Nowicki, M., DeVries, T., & Siegel, D. A. (2022). Quantifying the Carbon Export and Sequestration Pathways of the Ocean's Biological Carbon Pump. *Global Biogeochemical Cycles*, *36*(3). https://doi.org/10.1029/2021GB007083

Odum, H. T., & Odum, E. P. (1995). Trophic Structure and Productivity of a Windward Coral Reef Community on Eniwetok Atoll. *Ecological Monograph*, *25*(3), 291–320.

Oksanen, J., Blanchet, F. G., Friendly, M., & Kindt, R. (2020). *Community Ecology Package*. https://github.com/vegandevs/vegan

Ortega-Retuerta, E., Joux, F., Jeffrey, W. H., & Ghiglione, J. F. (2013). Spatial variability of particle-attached and free-living bacterial diversity in surface waters from the Mackenzie River to the Beaufort Sea (Canadian Arctic). *Biogeosciences*, *10*(4), 2747–2759. https://doi.org/10.5194/bg-10-2747-2013

Ortega-Retuerta, E., Mazuecos, I. P., Reche, I., Gasol, J. M., Álvarez-Salgado, X. A., Álvarez, M., Montero, M. F., & Arístegui, J. (2019). Transparent exopolymer particle (TEP) distribution and in situ prokaryotic generation across the deep Mediterranean Sea and nearby North East Atlantic Ocean. *Progress in Oceanography*, *173*, 180–191. https://doi.org/10.1016/j.pocean.2019.03.002

Palacios, D. M. (2002). Factors influencing the island-mass effect of the Galápagos Archipelago. *Geophysical Research Letters*, *29*(23). https://doi.org/10.1029/2002GL016232

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Parsons, R. J., Breitbart, M., Lomas, M. W., & Carlson, C. A. (2012). Ocean time-series reveals recurring seasonal patterns of virioplankton dynamics in the northwestern Sargasso Sea. *ISME Journal*, *6*(2), 273–284. https://doi.org/10.1038/ismej.2011.101

Passow, U. (2002). Transparent exopolymer particles (TEP) in aquatic environments. *Progress in Oceanography*, *55*, 287–333. www.elsevier.com/locate/pocean

Ploug, H., Grossart, H. P., Azam, F., & Jørgensen, B. B. (1999). Photosynthesis, respiration, and carbon turnover in sinking marine snow from surface waters of Southern California Bight: Implications for the carbon cycle in the ocean. *Marine Ecology Progress Series*, *179*, 1–11. https://doi.org/10.3354/meps179001

Poff, K. E., Leu, A. O., Eppley, J. M., Karl, D. M., Delong, E. F., Cordero, O. X., Crump, B. C., Designed, E. F. D., & Performed, E. F. D. (2021). Microbial dynamics of elevated carbon flux in the open ocean's abyss. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(4). https://doi.org/10.1073/pnas.2018269118/-/DCSupplemental

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., ... & de Vargas, C. (2012). CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS biology*, *10*(11), e1001419.

Pomeroy, L. R. (1974). The ocean's food web, a changing paradigm. *Bioscience*, *24*(9), 499-504.

Popp, B. N., B. S. Graham, R. J. Olson, C. C. S. Hannides, M.Lott, G. López-Ibarra, and F. Galván-Magaña. 2007. Insightinto the trophic ecology of yellowfin tuna, Thunnusalbacares, from compound-specific nitrogen isotope analy-sis of proteinaceous amino acids, p. 173-190. InT. E. Daw-son and R. T. W. Siegwolf [eds.], Stable isotopes as indica-tors of ecological change. Academic.

Porter, K. G., & Feig, Y. S. (1980). The use of DAPI for identifying and counting aquatic microflora. *Limnology and Oceanography*, *25*(5), 943–948. https://doi.org/10.4319/lo.1980.25.5.0943

Preston, C. M., Durkin, C. A., & Yamahara, K. M. (2020). DNA metabarcoding reveals organisms contributing to particulate matter flux to abyssal depths in the North East Pacific ocean. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *173*. https://doi.org/10.1016/j.dsr2.2019.104708

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved

data processing and web-based tools. *Nucleic Acids Research*, *41*(D1). https://doi.org/10.1093/nar/gks1219

Raapoto, H., Martinez, E., Petrenko, A., Doglioli, A., Gorgues, T., Sauzède, R., Maamaatuaiahutapu, K., Maes, C., Menkes, C., & Lefèvre, J. (2019). Role of Iron in the Marquesas Island Mass Effect. *Journal of Geophysical Research: Oceans*, *124*(11), 7781–7796. https://doi.org/10.1029/2019JC015275

Reintjes, G., Arnosti, C., Fuchs, B., & Amann, R. (2019). Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation. *ISME Journal*, *13*(5), 1119–1132. https://doi.org/10.1038/s41396-018-0326-3

Reintjes, G., Heins, A., Wang, C., & Amann, R. (2023). Abundance and composition of particles and their attached microbiomes along an Atlantic Meridional Transect. *Frontiers in Marine Science*, *10*. https://doi.org/10.3389/fmars.2023.1051510

Richardson T L, & Jackson G A. (2007). Small phytoplankton and carbon export from the surface ocean. *Science*, *315*(5813), 835–838. https://doi.org/10.1126/science.1133417

Rieck, A., Herlemann, D. P. R., Jürgens, K., & Grossart, H. P. (2015). Particle-associated differ from free-living bacteria in surface waters of the baltic sea. *Frontiers in Microbiology*, *6*(DEC). https://doi.org/10.3389/fmicb.2015.01297

Riley, J. S., Sanders, R., Marsay, C., Le Moigne, F. A. C., Achterberg, E. P., & Poulton, A. J. (2012). The relative contribution of fast and slow sinking particles to ocean carbon export. *Global Biogeochemical Cycles*, *26*(1). https://doi.org/10.1029/2011GB004085

Saino T, & Hattori A. (1980). 15N natural abundance in oceanic suspended particulate matter. *Nature*, *283*, 752–754.

Salazar, G., Cornejo-Castillo, F. M., Borrull, E., Díez-Vives, C., Lara, E., Vaqué, D., Arrieta, J. M., Duarte, C. M., Gasol, J. M., & Acinas, S. G. (2015). Particle-association lifestyle is a phylogenetically conserved trait in bathypelagic prokaryotes. *Molecular Ecology*, *24*(22), 5692–5706. https://doi.org/10.1111/mec.13419

Sarmiento, J. L., and N. Gruber (2006), *Ocean biogeochemical dynamics*, Princeton University Press Princeton, New Jersey, USA.

Schultz, D., Zühlke, D., Bernhardt, J., Francis, T. Ben, Albrecht, D., Hirschfeld, C., Markert, S., & Riedel, K. (2020). An optimized metaproteomics protocol for a holistic taxonomic and functional characterization of microbial communities from marine particles. *Environmental Microbiology Reports*, *12*(4), 367–376. https://doi.org/10.1111/1758-2229.12842

Siegenthaler, U., & Sarmiento, J. L. (1993). Atmospheric carbon dioxide and the ocean. *Nature*, *365*(6442), 119–125. https://doi.org/10.1038/365119a0

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

Simon, M., Grossart, H.-P., Schweitzer, B., & Ploug, H. (2002). Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology*, *28*, 175–211. www.int-res.com

Skoog, A., & Benner, R. (1997). Aldoses in various size fractions of marine organic matter: Implications for carbon cycling. *Limnology and Oceanography*, *42*(8), 1803–1813. https://doi.org/10.4319/lo.1997.42.8.1803

Smith, D. C., Simon, M., Alldredge, A. L., & Azam, F. (1992). Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature*, *359*, 139–142.

Smith, M. W., Allen, L. Z., Allen, A. E., Herfort, L., & Simon, H. M. (2013). Contrasting genomic properties of free-living and particle-attached microbial assemblages within a coastal ecosystem. *Frontiers in Microbiology*, *4*(MAY). https://doi.org/10.3389/fmicb.2013.00120

Sorokin, P.D. (1993). Coral Reef Ecology. *Ecological Studies*.

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., Knap, A. H., & Steinberg, D. K. (2001). Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep-Sea Research II*, *48*, 1405–1447.

Steinberg, D. K., Steinberg, D. K., Carlson, C. A., Bates, N. R., Goldthwait, S. A., Madin, L. P., & Michaels, A. F. (2000). Zooplankton vertical migration and the active transport of dissolved organic and inorganic carbon in the Sargasso Sea. *Deep-Sea Research I*, *47*, 137–158.

Steinberg, D. K., Van Mooy, B. A. S., Buesseler, K. O., Boyd, P. W., Kobari, T., & Karl, D. M. (2008). Bacterial vs. zooplankton control of sinking particle flux in the ocean's twilight zone. *Limnology and Oceanography*, *53*(4), 1327–1338. https://doi.org/10.4319/lo.2008.53.4.1327

Tanaka Y, Nakajima R. 2018. Dissolved organic matter in coral reefs: distribution, production, and bacterial consumption. In Coral Reef Studies of Japan, ed. A Iguchi, C Hongo, pp. 7–27. Singapore: Springer

Thompson J R, Pacocha S, Pharino C, Klepac-Ceraj V, Hunt D E, Benoit J, Sarma-Rupavtarm R, Distel D L, & Polz M F. (2005). Genotypic diversity within a natural coastal bacterioplankton population. *Science*, *307*, 1307–1311. https://doi.org/10.1126/science.1105750

Torreton, J.-P., & Dufour, P. (1996). Temporal and spatial stability of bacterioplankton biomass and productivity in an atoll lagoon. *Aquatic Microbial Ecology*, *11*, 251–261.

Tout, J., Jeffries, T. C., Webster, N. S., Stocker, R., Ralph, P. J., & Seymour, J. R. (2014). Variability in microbial community composition and function between different niches within a coral reef. *Microbial Ecology*, *67*(3), 540–552. https://doi.org/10.1007/s00248-013-0362-5

Treusch, A. H., Demir-Hilton, E., Vergin, K. L., Worden, A. Z., Carlson, C. A., Donatz, M. G., Burton, R. M., & Giovannoni, S. J. (2012). Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME Journal*, *6*(3), 481–492. https://doi.org/10.1038/ismej.2011.117

Treusch, A. H., Vergin, K. L., Finlay, L. A., Donatz, M. G., Burton, R. M., Carlson, C. A., & Giovannoni, S. J. (2009). Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME Journal*, *3*(10), 1148–1163. https://doi.org/10.1038/ismej.2009.60

Turner, J. T. (2015). Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump. In *Progress in Oceanography* (Vol. 130, pp. 205–248). Elsevier Ltd. https://doi.org/10.1016/j.pocean.2014.08.005

Valencia, B., Stukel, M. R., Allen, A. E., McCrow, J. P., Rabines, A., Palenik, B., & Landry, M. R. (2021). Relating sinking and suspended microbial communities in the California Current Ecosystem: digestion resistance and the contributions of phytoplankton taxa to export. *Environmental Microbiology*, *23*(11), 6734–6748. https://doi.org/10.1111/1462-2920.15736

Vollbrecht, C., Moehlenkamp, P., Gove, J. M., Neuheimer, A. B., & McManus, M. A. (2021). Long-Term Presence of the Island Mass Effect at Rangiroa Atoll, French Polynesia. *Frontiers in Marine Science*, *7*(January), 1–15. https://doi.org/10.3389/fmars.2020.595294

Wakeham, S. G., & Canuel, E. A. (1988). Organic geochemistry of particulate matter in the eastern tropical North Pacific Ocean: implications for particle dynamics. *Journal of Marine Research*, *46*(1), 183–213. https://doi.org/10.1357/002224088785113748

Wear, E. K., Wilbanks, E. G., Nelson, C. E., & Carlson, C. A. (2018). Primer selection impacts specific population abundances but not community dynamics in a monthly time-series 16S rRNA gene amplicon analysis of coastal marine bacterioplankton. *Environmental Microbiology*, *20*(8), 2709–2726. https://doi.org/10.1111/1462-2920.14091

Weber, L., Gonzalez-Díaz, P., Armenteros, M., & Apprill, A. (2019). The coral ecosphere: A unique coral reef habitat that fosters coral–microbial interactions. *Limnology and Oceanography*, *64*(6), 2373–2388. https://doi.org/10.1002/lno.11190

Weber, L., González-Díaz, P., Armenteros, M., Ferrer, V. M., Bretos, F., Bartels, E., Santoro, A. E., & Apprill, A. (2020). Microbial signatures of protected and impacted Northern Caribbean reefs: changes from Cuba to the Florida Keys. *Environmental Microbiology*, *22*(1), 499–519. https://doi.org/10.1111/1462-2920.14870

Weber, T., Cram, J. A., Leung, S. W., DeVries, T., & Deutsch, C. (2016). Deep ocean nutrients imply large latitudinal variation in particle transfer efficiency. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), 8606–8611. https://doi.org/10.1073/pnas.1604414113

Wegner, C. E., Richter-Heitmann, T., Klindworth, A., Klockow, C., Richter, M., Achstetter, T., Glöckner, F. O., & Harder, J. (2013). Expression of sulfatases in Rhodopirellula baltica and the diversity of sulfatases in the genus Rhodopirellula. *Marine Genomics*, *9*, 51–61. https://doi.org/10.1016/j.margen.2012.12.001

Williams, P. J. (1984). A review of measurements of respiration rates of marine plankton populations. In: Hobbie, J. E. & Williams, P. J. (eds.) Heterotrophic activity in the Sea. New York: Plenum Press

Wilson, S. E., & Steinberg, D. K. (2010). Autotrophic picoplankton in mesozooplankton guts: Evidence of aggregate feeding in the mesopelagic zone and export of small

phytoplankton. *Marine Ecology Progress Series, 412*, 11–27. https://doi.org/10.3354/meps08648

Zamanillo, M., Ortega-Retuerta, E., Nunes, S., Rodríguez-Ros, P., Dall'osto, M., Estrada, M., Sala, M. M., & Simó, R. (2019). Main drivers of transparent exopolymer particle distribution across the surface Atlantic Ocean. *Biogeosciences, 16*(3), 733–749. https://doi.org/10.5194/bg-16-733-2019

Ziervogel, K., Steen, A. D., & Arnosti, C. (2010). Changes in the spectrum and rates of extracellular enzyme activities in seawater following aggregate formation. *Biogeosciences, 7*, 1007–1015. www.biogeosciences.net/7/1007/2010/