AN ABSTRACT OF THE DISSERTATION OF

Mary Rose Gradoville for the degree of Doctor of Philosophy in Ocean, Earth and Atmospheric Sciences presented on March 10, 2017.

Title: Ecology and Environmental Controls of Two Keystone Groups of Oceanic Microorganisms: Diazotrophs and Pathogenic *Vibrio*

Abstract approved:

_____

Angelicque E. White

Marine bacteria play vital roles in every niche of the ocean, from small-scale symbioses to large-scale productivity and the regulation of Earth's climate. Recent advances in molecular tools now allow us to probe the genetic potential of entire microbial communities. The next step is linking these diverse communities to the critical functions they perform, in order to better understand how microbes regulate biogeochemical processes and predict how these processes may change as humans continue to alter the physical and chemical properties of the oceans. The three main body chapters of this dissertation use genomic tools and biogeochemical rate measurements to examine the ecology and environmental controls of two keystone groups of marine bacteria. Both Chapter II and Chapter III focus on marine diazotrophs, a group that converts dinitrogen gas into bioavailable nitrogen, thus helping to fuel productivity in oligotrophic, surface, open-ocean waters. Chapter II compares diazotrophic diversity and nitrogen fixation rates across three distinct ocean regions. This work indicates that diazotrophs are cosmopolitan in marine waters and that dominant diazotrophic groups have distinct biogeographical patterns, but that nitrogen fixation rates are restricted. Chapter III investigates the diversity and functional potential of microorganisms associated with colonies of the filamentous cyanobacterial diazotroph *Trichodesmium*. Molecular analyses revealed that the *Trichodesmium* colonies used in this study were inhabited by diverse assemblages of

microorganisms which were distinct from the surrounding bacterioplankton. Furthermore, epibionts included non-*Trichodesmium* diazotrophs and organisms with the genetic potential to influence colony nutrient acquisition. Chapter IV focuses on a separate group of marine bacterioplankton, *Vibrio* spp., which have significance for environmental health and marine disease due to the pathogenicity of several species. This chapter tracks the abundance of oyster-pathogenic *Vibrio* spp. in select Oregon estuaries and a shellfish hatchery. Results indicate that local growth drives elevated abundances of the oyster-pathogen *V. coralliilyticus* in Netarts Bay, and suggest that this species is particularly resistant to hatchery cleaning procedures.

Ecology and Environmental Controls of Two Keystone Groups of Oceanic Microorganisms: Diazotrophs and Pathogenic *Vibrio*

by
Mary Rose Gradoville

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented March 10, 2017
Commencement June 2017

Doctor of Philosophy dissertation of Mary Rose Gradoville
presented on March 10, 2017

APPROVED:

Major Professor, representing Ocean, Earth and Atmospheric Sciences

---

Dean of the College of Earth, Ocean and Atmospheric Sciences

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Mary Rose Gradoville, Author

ACKNOWLEDGEMENTS

Many people have helped make this dissertation possible, both through contributions to the research and through essential moral support through the process. I am extremely grateful for the many opportunities I have had in the College of Earth, Ocean, and Atmospheric Sciences, and for my community within the College and in Corvallis and beyond.

First and foremost, this dissertation was made possible through the guidance, patience, and wisdom of Angel White. Angel has pushed me to become a more capable and confident scientist and communicator. She has given me the constructive criticism needed to improve the science, the freedom to make my own mistakes and discoveries, and the resources to make it all happen. I am incredibly thankful to have been Angel's student for the past four years.

Many people have helped with my research projects. My committee members, Byron Crump, Ricardo Letelier, Burke Hales, Claudia Häse, and Staci Simonich have all been valuable resources; in particular, Byron Crump has been very gracious with his time and has welcomed me into his laboratory group. Outside of my committee, many others have helped to collect data, run samples, and provide feedback on data analysis, writing, and presentations: Deniz Bombar, Jon Zehr, Matthew Church, Sam Wilson, Katie Watkins-Brandt, Marnie-Jo Zirbel, Morgaine McKibbon, Jennifer McKay, Jennifer Bachman, Mike Verado, Ed Dever, Carla Schubiger, Alan Barton, Sue Cudd, Cale Miller, Connie Gradoville, Richard Feely, Dana Greely, Mark Dasenko, Anne-Marie Girard, Stefan Green, Kendra Turk-Kubo, Steve Pacella, Cameron Allen, Andrew Thurber, and Rick Colwell, just to name a few.

My friends and peers have been a huge source of support through graduate school. There are too many people to name, but I truly value the time I have spent with fellow graduate students, especially my OEB cohort ('Team Pteropod') and the GeoMicro group. I have also been blessed with amazing housemates who have supported and cared for me through the trials and tribulations of my PhD: Alejandra Sanchez, Jack Zunka, Maxine Davies, and Dan Coffman.

Finally, I am grateful for my loving, fun, and eccentric family: Connie, Denny, Heidi, Zack, Lisa, Bryce, Noelani, Rhone, and Bella. Most of all, I want to thank my parents, Connie and Denny, whose love is my biggest blessing.

CONTRIBUTION OF AUTHORS

Chapter 2:

This study was conceived by M.R. Gradoville and A.E. White. Data were collected by M.R. Gradoville, D. Bombar, and A.E. White and analyzed by M.R. Gradoville, D. Bombar, B.C. Crump, R.M. Letelier, and A.E. White. M.R. Gradoville wrote the first draft of the manuscript. All authors contributed substantial revisions through the drafting process.

Chapter 3:

This study was conceived by M.R. Gradoville. Data were collected by M.R. Gradoville and analyzed by M.R. Gradoville, B.C. Crump, M.J. Church, R.M. Letelier, and A.E. White. M.R. Gradoville wrote the first draft of the manuscript. All authors contributed substantial revisions through the drafting process.

Chapter 4:

This study was conceived by M.R. Gradoville, A.E. White, and C. Häse. Data were collected by M.R. Gradoville and analyzed by M.R. Gradoville, B.C. Crump, C. Häse, and A.E. White. M.R. Gradoville wrote the manuscript.

TABLE OF CONTENTS Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

1. PROBING THE UNSEEN WORLD OF MARINE BACTERIA

1.1 Introduction

Microorganisms (single-celled life forms) dominate life on earth, both in biomass and biogeochemical significance. Humans depend on microbes for enumerable critical processes, from aiding in crop production and food digestion to regulating the climate. Prokaryotes (archaea and bacteria) first appeared in the geologic record nearly four billion years ago, and since that time microorganisms have radiated to include species capable of inhabiting practically every known ecosystem on the planet (Woese, 1987). The metabolic diversity and extreme abundance of microorganisms have allowed them to shape the oceans and atmosphere through geologic history. For example, the evolution of photosynthetic cyanobacteria led to the oxygenation of the oceans and atmosphere, enabling the evolution of multicellular life (Lyons et al., 2014).

In marine environments, microbes, especially bacteria, drive nearly all processes of biological carbon transfer, energy flow, and nutrient cycling. In the pelagic oceans, photosynthetic cyanobacteria and unicellular algae form the base of the food web (Falkowski and Raven, 2013), heterotrophic bacteria and protozoa dominate the remineralization of organic matter (Azam et al., 1983), and specific microbial groups drive the cycling of essential nutrients such as nitrogen and sulfur (Paerl and Pinckney, 1996). Understanding the magnitude of these carbon and nutrient transformations is a paramount goal of microbial oceanography (Kirchman, 2010), primarily because of the strong ways in which these processes regulate Earth's climate (Falkowski, 1997;

Ducklow et al., 2001). For example, the balance of photosynthesis, respiration, and sinking organic matter in surface ocean waters partially controls the sequestration of carbon dioxide into the oceans (Volk and Hoffert, 1985). Since the climate is not in a steady state (Stocker et al., 2013), it is pertinent to understand not only the magnitude of microbial rate processes, but also the environmental controls of the individual organisms regulating those processes. Like all interesting problems, questions about the diversity and environmental drivers of marine communities have revealed complex answers, due in part to the extensive genetic diversity of microorganisms.

The diversity of marine microbes has long astonished and perplexed microbial ecologists. Hutchinson (1961) elegantly described the so-called "paradox of the plankton": how can so many species coexist in the same large, seemingly uniform environment of the sea? While some mechanisms have been proposed to help explain this paradox, including intermediate disturbance and microniches, neither the maintenance of biodiversity nor the full extent of microbial diversity are fully understood (Pedrós-Alió, 2012). However, the past several decades have increased our appreciation of the diversity, abundance, and biogeochemical importance of marine microorganisms (Kirchman, 2010). The discovery that the vast majority of marine bacteria could not be cultured using traditional plating techniques (Jannasch and Jones, 1959; Staley and Konopka, 1985) foreshadowed subsequent discoveries into the diversity and metabolic potential of these organisms enabled by DNA sequencing (Woese, 1987; Venter et al., 2004). The development and mass implementation of high-throughput environmental sequencing methods in the last decade represents yet another revolution, as estimates of microbial diversity within many ecosystems have increased by orders of magnitude

(Pedrós-Alió, 2012). Current phylogenetic analyses underscore the extreme diversity of bacteria, and show that the majority of bacterial lineages include no cultured representatives (Hug et al., 2016).

How can microbial ecologists derive meaningful information from a field with so many unknowns? In the face of such overwhelming diversity, it is sometimes useful to focus on groups of organisms rather than individual taxa. For instance, taxonomically heterogeneous organisms can be binned into functional groups or functional types, which together mediate the same process or share the same ecological niche. Functional types can sometimes be targeted without the need for DNA sequencing (e.g. photosynthetic pigments reflect functional groups of phytoplankton, (Wright, 1991)) and have been used to inform biogeochemical ocean models (Follows et al., 2007; Weber and Deutsch, 2012). However, detailed examinations of functional groups have revealed important species- or strain-specific differences in physiology and ecological controls.

This dissertation explores two groups of microorganisms for which understanding diversity at the "species" level (an indefinable yet unavoidable term for microorganisms) is critical. First, Chapters II and III center on the ecology of marine diazotrophs, a functional group of prokaryotes that converts atmospheric nitrogen gas into bioavailable ammonia. This functional group includes photoautotrophs, photoheterotrophs, and chemoheterotrophs, which differ strongly in their responses to environmental stimuli such as light, oxygen, nutrients, and organic matter (Zehr et al., 2003; Sohm et al., 2011; Bombar et al., 2016). Even within the diazotrophic genus *Trichodesmium*, laboratory isolates differ in their response to elevated carbon dioxide levels, indicating that species-level biogeographical information is prerequisite to predicting the climate response of

this group (Hutchins et al., 2013; Gradoville et al., 2014). Second, Chapter IV explores environmental controls of species from the *Vibrio* genus. *Vibrio* are a group of heterotrophic bacteria which includes both pathogenic and benign representatives. While members of this group share some common characteristics, including mechanisms for nutrient acquisition and surface attachment, environmental controls and life history strategies are often species-specific, which can have strong implications for human health and marine disease (Thompson et al., 2006). For example, the discovery that the pathogen *Vibrio cholera* (causative agent of the cholera disease in humans) associates with zooplankton led to a simple water filtration procedure which drastically reduced cholera instances in Bangladesh villages (Colwell et al., 2003). Investigations into the physical, chemical, and environmental controls of other, less-studied pathogenic *Vibrio* species may reveal environmental parameters which could similarly help to predict or prevent disease outbreaks in humans or animals.

A holistic understanding of microbial functioning requires knowing both the players and the processes. Linking community genomics to functions of interest, and ultimately to biogeochemical cycles, is a grand challenge for microbial ecologists. This problem is complicated by the difficulty of attributing rates to individual cells, but it is necessary due to species- and strain-specific physiology. In this dissertation, I worked both inside and outside of the microbial "black box," using genomics as a lens to view microbial communities as well as using oceanographic and chemical tools to measure community rates and examine biogeochemical features of the marine environment.

1.2 Thesis objectives

The three chapters in this dissertation explore the ecology and environmental drivers of two important groups of marine bacteria: diazotrophs (Chapters II and III) and pathogenic *Vibrio* spp. (Chapter IV). Chapter II examines the relationship between community nitrogen fixation rates and diazotrophic diversity across ocean biomes. Chapter III examines microbial communities inhabiting *Trichodesmium* colonies and the potential for community interactions that could affect colony community functioning. Chapter IV investigates the environmental controls of an oyster-pathogenic *Vibrio* species in Oregon estuaries and a local shellfish hatchery. These chapters add to our understanding of the ecology and environmental controls of these keystone groups of marine microorganisms.

1.3 References

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Marine Ecology Progress Series* 10(3)**,** 257-263.

Bombar, D., Paerl, R.W., and Riemann, L. (2016). Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. *Trends in Microbiology* 24(11)**,** 916-927.

Colwell, R.R., Huq, A., Islam, M.S., Aziz, K., Yunus, M., Khan, N.H., et al. (2003). Reduction of cholera in Bangladeshi villages by simple filtration. *Proceedings of the National Academy of Sciences* 100(3)**,** 1051-1055.

Ducklow, H.W., Steinberg, D.K., and Buesseler, K.O. (2001). Upper ocean carbon export and the biological pump. *Oceanography* 14(4)**,** 50-58.

Falkowski, P.G. (1997). Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* 387(6630)**,** 272-275.

Falkowski, P.G., and Raven, J.A. (2013). *Aquatic photosynthesis.* Princeton University Press.

Follows, M.J., Dutkiewicz, S., Grant, S., and Chisholm, S.W. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* 315(5820)**,** 1843-1846.

Gradoville, M.R., White, A.E., Böttjer, D., Church, M.J., and Letelier, R.M. (2014). Diversity trumps acidification: Lack of evidence for carbon dioxide enhancement of *Trichodesmium* community nitrogen or carbon fixation at Station ALOHA. *Limnology and Oceanography* 59(3)**,** 645-659.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., et al. (2016). A new view of the tree of life. *Nature Microbiology* 1**,** 16048.

Hutchins, D.A., Fu, F.-X., Webb, E.A., Walworth, N., and Tagliabue, A. (2013). Taxon-specific response of marine nitrogen fixers to elevated carbon dioxide concentrations. *Nature Geoscience* 6(9)**,** 790-795.

Hutchinson, G.E. (1961). The paradox of the plankton. *The American Naturalist* 95(882)**,** 137-145.

Jannasch, H.W., and Jones, G.E. (1959). Bacterial populations in sea water as determined by different methods of enumeration. *Limnology and Oceanography* 4(2)**,** 128-139.

Kirchman, D.L. (2010). *Microbial ecology of the oceans.* John Wiley & Sons.

Lyons, T.W., Reinhard, C.T., and Planavsky, N.J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506(7488)**,** 307-315.

Paerl, H.W., and Pinckney, J. (1996). A mini-review of microbial consortia: their roles in aquatic production and biogeochemical cycling. *Microbial Ecology* 31(3)**,** 225-247.

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Annual Review of Marine Science* 4**,** 449-466.

Sohm, J.A., Hilton, J.A., Noble, A.E., Zehr, J.P., Saito, M.A., and Webb, E.A. (2011). Nitrogen fixation in the South Atlantic Gyre and the Benguela upwelling system. *Geophysical Research Letters* 38(16)**,** L16608.

Staley, J.T., and Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Reviews in Microbiology* 39(1)**,** 321-346.

Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., et al. (2013). *IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.* Cambridge University Press.

Thompson, F.L., Austin, B., and Swings, J. (2006). *The biology of Vibrios.* ASM press.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304(5667)**,** 66-74.

Volk, T., and Hoffert, M.I. (1985). Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO2 changes. *The Carbon Cycle and Atmospheric CO: Natural Variations Archean to Present*: 99-110.

Weber, T., and Deutsch, C. (2012). Oceanic nitrogen reservoir regulated by plankton diversity and ocean circulation. *Nature* 489(7416)**,** 419-422.

Woese, C.R. (1987). Bacterial evolution. *Microbiological Reviews* 51(2)**,** 221.

Wright, S. (1991). Improved HPLC method for the analysis of chlorophylls and carotenoids from marine phytoplankton. *Marine Ecology Progress Series* 77**,** 183-196.

Zehr, J.P., Jenkins, B.D., Short, S.M., and Steward, G.F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environmental Microbiology* 5(7)**,** 539-554.

CHAPTER 2: DIVERSITY AND ACTIVITY OF NITROGEN–FIXING COMMUNITIES ACROSS OCEAN BASINS

Mary R. Gradoville[a], Deniz Bombar[b,*], Byron C. Crump[a], Ricardo M. Letelier[a], Jonathan P. Zehr[b], and Angelicque E. White[a]

[a]College of Earth, Ocean, and Atmospheric Sciences, 104 CEOAS Administration Building, Oregon State University, Corvallis, Oregon, USA

[b]Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, California, USA

[*]Current address: Marine Biological Section, Department of Biology, University of Copenhagen, Helsingør, Denmark.

In press, *Limnology and Oceanography*

2. DIVERSITY AND ACTIVITY OF NITROGEN-FIXING COMMUNITIES ACROSS OCEAN BASINS

2.1 Abstract

Recent observations of N_2 fixation rates and the presence of nitrogenase (*nifH*) genes from heterotrophic N_2-fixing (diazotrophic) prokaryotes in unusual habitats challenge the paradigm that pelagic marine N_2 fixation is constrained to cyanobacteria in warm, oligotrophic, surface waters. Here, we compare N_2 fixation rates and diazotrophic diversity (assessed via high-throughput *nifH* sequencing) from a region known to be dominated by cyanobacterial diazotrophs (the North Pacific Subtropical Gyre, NPSG) to two regions dominated by heterotrophic diazotrophs: the Eastern South Pacific (ESP, from the Chilean upwelling system to the subtropical gyre) and the Pacific Northwest coastal upwelling system (PNW). We observed distinct biogeographical patterns among the three regions. Diazotrophic community structure differed strongly between the NPSG, dominated by cyanobacterium UCYN-A, and the ESP, dominated by heterotrophic *nifH* group 1J/1K, yet surface N_2 fixation rates were similar in magnitude (up to 5.1 nmol N L^{-1} d^{-1}). However, while diverse, predominantly heterotrophic *nifH* genes were recovered from the PNW and the mesopelagic of the NPSG, N_2 fixation rates were undetectable in both of these environments (although glucose amendments stimulated low rates in the deep NPSG). Our work suggests that while diazotrophs may be nearly omnipresent in marine waters, the activity of this functional group is regionally restricted. Further, we show that the detection limits of the $^{15}N_2$ fixation assay suggest that many of the low N_2 fixation rates reported for the mesopelagic (often <0.1 nmol N L^{-1} d^{-1} in the literature) are

not indicative of active diazotrophy, highlighting the challenges of assessing the ecosystem significance of heterotrophic diazotrophs.

2.2 Introduction

Nitrogen (N) availability limits phytoplankton production in most marine ecosystems (Gruber 2004), and thus partially determines the drawdown of inorganic carbon from the atmosphere. The oceanic reservoir of bioavailable N is largely controlled by the balance of specific microbial processes: dinitrogen (N_2) fixation adds fixed N and denitrification and anaerobic ammonia oxidation remove fixed N. Whether the global oceanic N budget is balanced or not is a topic of longstanding controversy (see Galloway et al. 2004, Gruber 2004). Nevertheless, multiple lines of evidence suggest historical underestimations of both marine N_2 fixation rates (Codispoti 2007, Deutsch et al. 2007, Großkopf et al. 2012) and the diversity and abundance of marine N_2-fixers (diazotrophs) (Riemann et al. 2010). These findings have driven a reassessment of N_2-fixing habitats and diazotrophic diversity (reviewed in Bombar et al. 2016).

N_2 fixation was traditionally thought to be dominated by cyanobacteria inhabiting warm, oligotrophic, surface ocean waters (Karl et al. 2002), but this paradigm is being challenged. Sequencing the *nifH* gene, which encodes the iron-protein component of the enzyme nitrogenase that catalyzes N_2 fixation, has revealed potential marine diazotrophs spanning diverse prokaryotic lineages (Zehr et al. 1998), and *nifH* amplicons from putative heterotrophs have been reported to outnumber cyanobacterial amplicons in global surface ocean waters (Farnelid et al. 2011, Riemann et al. 2010). Furthermore, heterotrophic *nifH* genes and transcripts have been recovered from environments

previously unexplored for N_2 fixation, including coastal upwelling regimes, oxygen minimum zones, and the deep sea (Hewson et al. 2007, Jayakumar et al. 2012, Sohm et al. 2011). Active N_2 fixation in these environments, even at low rates, would increase overall marine N_2 fixation estimates and expand our understanding of the biogeographical patterns of an ecologically important functional group, the diazotrophs (e.g. Bonnet et al. 2013, Rahav et al. 2013).

The Eastern South Pacific (ESP), in particular, is a region that has garnered considerable attention as a potentially underappreciated habitat for marine N_2 fixation. Though historically understudied as an environment for diazotrophs (Luo et al. 2012), focus on the ESP increased after a numerical model predicted high N_2 fixation rates in the region (Deutsch et al. 2007). Despite model predictions, expeditions to the ESP have revealed low N_2 fixation rates in the oligotrophic gyre (0.08–0.88 nmol L^{-1} d^{-1}, Dekaezemacker et al. 2013, Halm et al. 2011), albeit highly variable rates in the Chilean upwelling system (0.1–127 nmol L^{-1} d^{-1}, Fernandez et al. 2011, Fernandez et al. 2015). The presence of inorganic N can suppress N_2 fixation by cyanobacterial diazotrophs (Knapp 2012, Knapp et al. 2012), so it is unsurprising that diazotrophic communities in the N-rich upwelling zone of the ESP appear to be heavily dominated by heterotrophic bacteria, particularly gamma-proteobacteria (Farnelid et al. 2011, Fernandez et al. 2011, Turk-Kubo et al. 2014). The quantitative Polymerase Chain Reaction (qPCR)-derived *nifH* abundances of select heterotrophic phylotypes, however, appear too low to account for even the low N_2 fixation rates observed to date (Turk-Kubo et al. 2014). This mismatch between proxies for cell abundances and rates implies an underestimation of diazotrophic abundance or cell-specific rates, and/or an overestimation of N_2 fixation

rates. These findings point to significant concerns regarding the inaccuracy of historical N_2 fixation rate measurements (Dabundo et al. 2014, Mohr et al. 2010) that have since been partially addressed (Wilson et al. 2012, Böttjer et al. 2016) and, more importantly, highlight our incomplete knowledge of the physiology, biogeography, and ecology of heterotrophic diazotrophs.

In order to understand the linkage between the assembly of diazotrophic communities and the magnitude of N_2 fixation rates, we have developed a cross-ecosystem study spanning oligotrophic gyres to upwelling regimes. Specifically, we present the first pairing of N_2 fixation rates with fine-scale diazotrophic diversity assessments (via high-throughput *nifH* sequencing) in three oceanic regions: the ESP (from the upwelling zone to the gyre), the North Pacific Subtropical Gyre (NPSG), and the Pacific Northwest coast of the United States (PNW). While previous observations in the ESP indicate low N_2 fixation rates and cryptic heterotrophic diazotrophs, the NPSG fosters moderate N_2 fixation rates presumably driven by highly abundant cyanobacterial diazotrophs (0.6–3.2 nmol N L^{-1} d^{-1}, Böttjer et al. 2016). Meanwhile, N_2 fixation has not been reported in the PNW, where the upwelling of cold, nitrate-rich waters would theoretically select against the growth of diazotrophs. Comparisons of these three potential habitats enable a global biogeographical perspective of marine diazotrophy. Here, we also address discontinuities between the genetic potential for diazotrophy and the presence of detectable N_2 fixation rates.

2.3 Methods

2.3.1 Study sites and sampling strategy

Diazotrophic diversity and N_2 fixation rates were assessed in (1) the ESP (Center for Microbial Oceanography: Research and Education cruise aboard the R/V Melville, Biogeochemical Gradients: Role in Arranging Planktonic Assemblages [BIG RAPA], November to December 2010), (2) the Pacific Northwest California Current system (PNW, West Coast Ocean Acidification cruise aboard NOAA *Fairweather* and R/V *Point Sur*, August 2012, see Feely et al. 2013), and (3) at Station ALOHA (A Long Term Habitat Assessment) in the NPSG (Hawaiian Ocean Experiment Budget of Energy cruise aboard R/V *Kilo Moana*, March 2014) (Fig. 2.1). A summary of oceanographic sampling locations and conditions is provided in Table 2.1.

Water samples from all cruises were collected using sampling bottles attached to a CTD (conductivity, temperature, depth) rosette. In the ESP, N_2 fixation was assessed in surface waters (5m), and diazotrophic diversity was assessed in surface and mesopelagic waters (5–420m). N_2 fixation rates and diazotrophic diversity were assessed at depths targeting the surface mixed layer, oxycline, and oxygen minimum zone in the PNW, and at standard depths of 25m (surface mixed layer), 75m (~base of mixed layer), and 200m (top of mesopelagic) in the NPSG.

2.3.2 Nitrogen fixation rates

N_2 fixation rates were assessed using a modification (Mohr et al. 2010, Wilson et al. 2012) of the original $^{15}N_2$ uptake method (Montoya et al. 1996) to avoid problems of bubble dissolution. $^{15}N_2$ gas (Cambridge Isotopes, 99%) was dissolved in degassed, filtered seawater. The $^{15}N_2$ content of this $^{15}N_2$-enriched seawater was validated via Membrane Inlet Mass Spectrometry (MIMS), and the $^{15}N_2$-enriched seawater was added

to $^{15}N_2$ incubations (2–10% by volume, see *Supplementary Information*). Purities of $^{15}N_2$ stocks were not assessed, although previous Cambridge Isotope $^{15}N_2$ batches have contained only trace contaminants of $^{15}NH_4^+$; at most this contamination is estimated to result in perceived N_2 fixation rates of <0.02 nmol N L^{-1} d^{-1} (Dabundo et al. 2014).

Before N_2 fixation rate sampling, all bottles were thoroughly acid-washed and milli-Q rinsed; bottles were also rinsed with water from the target collection depth three times before filling. These steps were intended to minimize potential trace metal contamination. For sampling, duplicate or triplicate polycarbonate bottles (1.1–4.5L) were filled to capacity, avoiding air contamination, and capped with septa closures. Samples were spiked with $^{15}N_2$ enriched seawater and incubated on deck for 24 h at approximately *in situ* temperature, either in the dark or with screening to mimic approximate *in situ* light conditions (Table S2.1, see *Supplementary Information* for a more detailed description of $^{15}N_2$ incubations). Incubations were terminated by gentle filtration of suspended material onto 25 mm diameter pre-combusted GF/F filters. Additionally, time-zero samples (the natural abundance of $\partial^{15}N$ of suspended particulate N) were collected at each time point and depth. Filters were frozen at 80°C and shipped to Oregon State University, where they were dried at 60°C overnight and packed into tin and silver capsules. Isotopic composition and masses of particulate N and C were measured with an isotope ratio mass spectrometer at the University of California Davis Stable Isotope Facility, or at Oregon State University. N_2 fixation rates were calculated according to Montoya et al. (1996) with the modification that rates are expressed as nmol N L^{-1} d^{-1} rather than nmol N_2 L^{-1} d^{-1}.

2.3.3 Diazotrophic Diversity

On all cruises, seawater for DNA was collected in 1–4 L acid-washed, milliQ-rinsed polycarbonate bottles and gently filtered onto 25 mm 0.2 μm Supor membranes (Pall Corporation). Filters from PNW and NPSG cruises were stored at –20°C or –80°C until extraction using the DNeasy Plant MiniKit (Qiagen) following manufacturer's instructions, with modifications to include freeze-thaw and proteinase K treatment for additional cell disruption. On the ESP cruise, water was size-fractionated, and DNA samples (0.2–10 μm fraction only) were transferred into sterile bead beating tubes containing a 1:1 mixture of 0.1 and 0.5 mm glass beads (BioSpec Products), flash-frozen, and stored at –80°C until extraction. All ESP DNA extractions, qPCR and PCR reactions were carried out as described by Turk-Kubo et al. (2012) (see *Supplementary Information*). DNA concentrations were measured with PicoGreen using a MicroMax 384 plate reading fluorometer, and extracts were stored at –20°C or –80°C.

The polymerase chain reaction (PCR) was used to amplify *nifH* genes for all samples using nested degenerate *nifH* primers (Zani et al. 2000, Zehr and McReynolds 1989). PCRs were performed using a Veriti thermocycler (Applied Biosystems) with 10 μL or 20 μL reaction volumes. The first round of PCR contained 1X PCR buffer, 200 μmol L^{-1} dNTPs, 3% BSA, 2.5 mmol L^{-1} Mg^{2+}, 0.1U Platinum High Fidelity *Taq* polymerase (Invitrogen), 1 μL DNA, and 1 μmol L^{-1} nifH1 and nifH2 primers (Zehr and McReynolds 1989). This reaction was cycled at 94°C for 7 min, then 30 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min, and finally 72°C for 7 min. The second round of *nifH* PCR used the same thermocycling conditions (except for a 5 min initial denaturation step) and components, but used 1 μL PCR product from the first reaction

and custom primers comprised of gene-specific sites, dual-indexed barcodes, Illumina linkers, and a sequencing primer binding region, similar to those described by Kozich et al. (2013) (Table S2.2). Triplicate PCR products were visualized by gel electrophoresis, pooled, and quantified as above. Only samples with three successful PCR reactions were included for sequencing, and negative controls and filter blanks were sequenced despite the absence of visual bands after amplification. Amplicons were pooled at equal concentrations, cleaned using the UltraClean PCR (MoBio) and AMPure XP Bead cleanup kits, and sequenced using MiSeq Standard v.3, 2×300 bp paired-end sequencing at Oregon State University.

Sequence reads were demultiplexed using the Illumina MiSeq Reporter (MSR) version 2.5.1. Forward and reverse barcodes were used for demultiplexing; however, due to poor quality of reverse reads, only forward reads were used for phylogenetic analyses. Quality control was performed using mothur (Schloss et al 2009), discarding reads with ambiguities, poor quality (average score <25 or any score <20), or homopolymers (>8 bp). Forward primers were removed, and all sequences were trimmed to 213 bp. Operational taxonomic units (OTUs) were clustered at 97% nucleotide sequence similarity, and a *de novo* chimera check was performed using USearch (Edgar 2010). OTUs containing chimeras, frameshifts, and non-*nifH* sequences were removed, and sequences were subsampled to 10,386 sequences per sample.

Translated OTUs were classified into canonical *nifH* gene clusters (Zehr et al. 2003b) using BLAST-p similarity to a reference database of *nifH* sequences previously assigned to *nifH* groups (http://www.jzehrlab.com/#!nifh-database/c1coj). Sequences with <90% amino acid similarity to any reference sequence or equal similarity to

sequences from multiple *nifH* groups are termed "undefined". Shannon diversity calculations and nonmetric multidimensional scaling analyses (Bray-Curtis similarity) were performed using QIIME (Caporaso et al. 2010). Sequences are available from NCBI (accession SRP078449).

2.3.4 Diazotrophic quantification

Droplet digital PCR (ddPCR) was used to quantify *nifH* copies from UCYN-A in NPSG samples and alpha-HQ586648 in ESP samples. Droplet digital PCR was performed using the Bio-Rad QX200 system. Diluted DNA extracts were used as templates in assays designed to amplify *nifH* genes from UCYN-A and alpha-HQ586648. Primer/probe sets were slightly modified from those described by Church et al. (2005a) (forward primer: GGCTATAACAACGTTTTATGCGTTGA, reverse primer: ACCACGACCAGCACATCCA, probe: TCCGGTGGTCCTGAGCCTGGA) and Zhang et al. (2011) (forward primer: ATCACCGCCATCAACTTCCT, reverse primer: AGACCACGTCGCCCAGAAC, probe: CGCCTACGATGACGTGGATTACGTGTCC) to minimize mismatches with sequences from OTU 1 (UCYN-A) and OTU 2 (alpha-HQ586648). Reactions consisted of 10 μL PCR mastermix (Bio-Rad), 900 nM primers, 250 nM probes, and ~1–50 ng DNA. Droplet generation, PCR, and scanning were conducted at the Oregon State University Center for Genome Research and Biocomputing according to manufacturer instructions (Bio-Rad), with an annealing temperature of 57°C. Standard curves were not used for these assays as ddPCR produces absolute concentrations without the need for standards. No template controls (NTC) were included in the ddPCR run; no UCYN-A or alpha-HQ586648 gene copies were detected

in any of the 6 NTC samples. Data were analyzed using the QuantaSoft analysis software package.

2.3.5 Statistical methods

Simple linear regressions and one-way ANOVA with subsequent Tukey Honest Significance Difference (HSD) tests of multiple comparisons were performed using the program R (http://www.r-project.org/). Significant differences in community structure among regions were tested using a one-way ANOSIM in PRIMER. A sensitivity analysis for N_2 fixation rate measurements was performed as described by Montoya et al. (1996). The minimum quantifiable rate for each water mass was calculated using standard propagation of errors via the observed variability between replicate samples (Table S2.3; region-averaged values provided in Table 2.2). We also used the approach of Montoya et al. (1996) to calculate an alternative detection limit by setting $[A_{\mathrm{PNf}} - A_{\mathrm{PN0}}]$ equal to 0.00146 atom %.

2.4 Results

2.4.1 Oceanographic Conditions

Samples used for this study came from diverse hydrographic and biogeochemical regions (Fig. 2.1, Table 2.1). The ESP cruise transited through a biogeochemical gradient from highly productive, upwelling-influenced waters on the Chilean shelf to warmer, stratified, oligotrophic stations in the South Pacific gyre (Rii et al. 2016). The PNW cruise transited through continental shelf and slope stations in the Northern California

Current system during the summer upwelling season. The NPSG cruise was conducted at Station ALOHA in oligotrophic conditions (Table 2.1).

2.4.2 Nitrogen fixation rates

N_2 fixation rates varied by region and by depth (Fig. 2.2). On the ESP transit cruise, surface N_2 fixation rates were highest near the coast (5.1 nmol N L^{-1} d^{-1} at Stn. 1), and decreased across the westward transit, with the lowest rates in the gyre (0.46 nmol N L^{-1} d^{-1} at Stn. 7) (Fig. 2.2). All rates exceeded the minimum quantifiable rate, defined here as the propagated error for each measurement, which ranged from 0.15–0.77 nmol N L^{-1} d^{-1} (Table S2.3). Subsurface N_2 fixation rates were not assessed in the ESP.

In the NPSG, we measured rates of 2.5–3.9 nmol N L^{-1} d^{-1} at 25m and 0.4–1.9 nmol N L^{-1} d^{-1} at 75m. N_2 fixation was undetectable at 200m under ambient conditions, though amendment with 3 μmol L^{-1} glucose stimulated rates of 0.1–0.3 nmol N L^{-1} d^{-1}; these rates exceeded the minimum quantifiable rate, which ranged from 0.04–1.08 nmol N L^{-1} d^{-1} for the region (Table S2.3).

N_2 fixation was not detected at any station or depth on the PNW cruise. Most measurements fell below the minimum quantifiable rate, which ranged from 0.06–0.69 nmol N L^{-1} d^{-1} (Fig. 2.2, Table S2.3). Two measurements were above the minimum quantifiable rate but below the alternate limit of detection suggested by Montoya et al. (1996) (calculated by setting $[A_{PNf} - A_{PN0}]$ equal to 0.00146 atom %, Table S2.3) and are thus regarded as not detected.

2.4.3 Diazotrophic diversity

Diazotrophic diversity was assessed via high-throughput sequencing of partial *nifH* gene sequences. Only samples with three successful PCR amplifications (achieved for 18/21 NPSG samples, 20/23 ESP samples, and 12/13 PNW samples) and sufficient reads to subsample down to 10,386 sequences per sample were retained (n = 49 samples). Rarefaction analyses show that our sequencing effort was sufficient to fully assess *nifH* diversity from most samples (Fig. S2.1). The subsampled datasets were used for all downstream analyses (excluding Fig. S2.1 and S2.2).

To test for the possibility of contaminant sequences from PCR reagents or other sample processing procedures (Zehr et al. 2003a), representative sequences from each OTU were aligned against known contaminant *nifH* sequences. Four OTUs had >95% nucleotide identity to known contaminants; the majority of these sequences were from ESP samples (Table S2.4). None of the negative control samples contained *nifH* sequences; however, two of four triplicate-pooled filter blank samples contained a small number of sequences (1574 and 634) (Fig. S2.2, see *Supplementary Information*).

Clustering sequences at 97% nucleotide similarity yielded 201 OTUs; 150 of these (>99% of total sequences) passed quality control procedures and were used for analyses. All OTUs clustered with *nifH* sequences from Clades I and III, and most were associated with group 1J/1K (alpha- and beta-proteobacteria) and 1B (cyanobacteria) (Fig. 2.3, 2.4). Diazotrophic community structure differed strongly between regions (Global ANOSIM R =0.87, $p = 0.001$) (Fig. 2.5). Likewise, the regions were dominated by different diazotrophic taxa (Fig. 2.4) and contained few shared OTUs (Fig. 2.3).

In the NPSG, most *nifH* sequences were classified as cyanobacteria (Fig. 2.4). 66.7% of NPSG sequences grouped with sequences from the cyanobacterium UCYN-A

(Fig. 2.3, 2.4, S2.3); lower limit concentrations of the most abundant UCYN-A phylotype (OTU 1, 55.8% of NPSG sequences) ranged from 3.3×10^4–4.4×10^5 *nifH* copies L^{-1} in 25m and 75m water (Table S2.5). Other prominent cyanobacterial phylotypes grouped with *Trichodesmium* (~4.8% of NPSG sequences), and <0.1% of NPSG sequences were closely related to *Richelia*, *Crocosphaera*, and unidentified cyanobacteria sharing <94% nucleotide similarity to cultured representatives. Other than cyanobacteria, 20.6% of NPSG sequences were assigned to the gamma-proteobacterial group 1G (Fig. 2.4). These sequences are dominated by OTU 3, which matches with qPCR primer/probe sets designed to amplify the group "Gamma A" (Church et al. 2005a, Langlois et al. 2015, Moisander et al. 2008). Together, cyanobacteria and group 1G represent >99% of sequences from 25m and 75m samples. While *nifH* genes were successfully amplified from all 25m and 75m samples, triplicate PCRs were successful for only 3/6 of the 200m samples; dominant taxa and OTUs were less consistent among these samples (Fig. 2.4, S2.3).

Heterotrophs dominated the diazotrophic community in the <10 μm size fraction of ESP samples (Fig. 2.4). Most sequences were assigned to group 1J/1K, and <0.2% of sequences were assigned to cyanobacteria. Five OTUs (2, 4, 5, 10, and 16), comprising >90% of total ESP sequences (Fig. S2.3), were tested *in silico* against published qPCR primer/probe sets targeting heterotrophic marine diazotrophs. OTU 2 matched the alpha-HQ586648 forward primer designed by Zhang et al. (2011) (probe and reverse primers were beyond the range of our partial *nifH* sequences, Table S2.6); lower-limit abundance estimates of alpha-HQ586648 ranged from 1.5×10^2–4.4×10^3 *nifH* copies L^{-1} (Table S2.5). OTUs 4, 5, 10, and 16 did not match existing primer/probe sets

(Table S2.6). Diazotrophic diversity in the ESP was not assessed for the >10 μm size fraction, which may have biased against the recovery of *nifH* genes from large or particle-associated diazotrophs.

Diazotrophic community composition from PNW samples was highly variable among stations and depths. The majority of PNW sequences clustered with heterotrophic taxa, though a sample from 150 m at Stn. 37 was dominated by UCYN-A sequences (Fig. 2.4). Overall, the most abundant *nifH* group was 1J/1K, which constituted 40.4% of total PNW sequences. However, a substantial fraction of PNW sequences were also assigned to groups 1A (delta-proteobacteria), 1G (gamma-proteobacteria), 1O/1P (beta-, gamma-proteobacteria) and group III (anaerobic delta-proteobacteria) (Fig. 2.4). Unlike the ESP and NPSG, very few OTUs were dominant in more than one sample (Fig. S2.3).

Alpha diversity metrics were calculated using the *nifH* sequence dataset. Shannon diversity was significantly higher in the PNW (Tukey HST $p = 0.004$) and the NPSG (Tukey HST $p = 0.015$) than in the ESP (Fig. 2.6). There was a negative relationship between Shannon diversity and N_2 fixation rates in the NPSG (linear regression, $R^2 = 0.316$, $p = 0.015$) but no significant relationship in the ESP ($p > 0.05$, Table S2.7). There were no significant relationships between Shannon diversity and temperature, oxygen concentration, or nitrate concentration (linear regression, $p > 0.05$, Table S2.7).

2.5 Discussion

2.5.1 Cosmopolitan diazotrophs with taxa-specific biogeography

Increasing reports of heterotrophic diazotrophs in “unusual” environments have challenged the common assumption that marine N_2 fixation is dominated by

cyanobacteria that inhabit warm, N-limiting waters (Bombar et al. 2016). Here, we probed for *nifH* genes in three regions: the NPSG, considered a classical habitat for diazotrophs, the ESP, where the functionally important diazotrophs appear cryptic (Turk-Kubo et al. 2014), and the PNW, where diazotrophy would not be expected. We observed the genetic potential for diazotrophy in all environments assessed, including the cold, N-rich waters of the PNW. Our data show strong differences in diazotrophic community structure between regions, implying distinct biogeography for these organisms.

The first global high-throughput *nifH* sequencing effort by Farnelid et al. (2011) reported heterotrophic diazotrophs outnumbering cyanobacteria in all regions assessed, including one sample from the NPSG. In contrast, cyanobacteria dominated our *nifH* sequences in all 14 NPSG euphotic zone samples (Fig. 2.4). In agreement with previous studies at Stn. ALOHA using qPCR and *nifH* clone libraries (Böttjer et al. 2014, Church et al. 2005a, Church et al. 2008), we found that the cyanobacterium UCYN-A was present at high concentrations (10^4–10^5 *nifH* copies L^{-1}) and dominated the diazotrophic community in the euphotic zone (Fig. 2.4, Table S2.5). We also observed a large proportion of euphotic zone *nifH* sequences associated with Gamma A, a globally distributed group of marine gamma-proteobacteria that has been previously detected at Stn. ALOHA (Church et al. 2005a, Langlois et al. 2015).

In contrast to the NPSG, the ESP diazotrophic community appears dominated by heterotrophs, a finding which agrees with previous studies (Halm et al. 2011, Turk-Kubo et al. 2014). However, while Halm et al. (2011) and Turk-Kubo et al. (2014) reported gamma-proteobacterial dominance, the majority of our sequences grouped with 1J/1K (alpha- and beta-proteobacteria), and the gamma-proteobacterial phylotypes previously

quantified in the ESP were not dominant in our dataset (Fig. 2.4, Table S2.6). It is possible that using only the <10 μm size fraction for ESP samples introduced bias by excluding particle-associated heterotrophs or large cyanobacterial cells. However, the absence of diatoms known to harbor diazotrophic cyanobacterial symbionts in a 16S rRNA plastid dataset obtained on the same cruise (Rii, pers. comm.) and our poor success in qPCR assays targeting gamma-proteobacteria using all DNA size fractions (data not shown) lead us to conclude that the 1J/1K group was likely dominant.

Perhaps the most striking feature in the ESP is the consistency of dominant diazotrophic taxa and phylotypes among stations and depths (Fig. 2.4, S2.3). This homogeneity is quite surprising considering the strong biogeochemical gradients in temperature, nutrients and chlorophyll across the Chilean upwelling zone into the South Pacific gyre. In particular, OTU 2 was the most abundant phylotype in both the gyre and coastal stations (Fig. S2.3), and also matches *nifH* sequences previously recovered from the Peruvian upwelling system, the Indian ocean, and the South China Sea (Fernandez et al. 2011, Shiozaki et al. 2014, Zhang et al. 2011). This finding counters previous observations of few heterotrophic diazotrophs with basin-wide distributions (Turk-Kubo et al. 2014). At present we do not have an explanation for the observed homogeneity of diazotrophic phylotypes along this oceanographic transect; further studies are needed to explore the stability of diazotrophic assemblages in this region in response to environmental perturbations.

While our deep-sequencing effort in the ESP and NPSG euphotic zone provides a higher resolution examination of diazotrophic diversity in previously explored regions, to our knowledge, our study is the first to recover *nifH* genes from PNW and mesopelagic

NPSG waters. PNW samples were dominated by diverse heterotrophic taxa, while mesopelagic NPSG samples contained relatively abundant cyanobacteria and heterotrophs, perhaps originating from sinking organic particles (Fig. 2.3, 2.4). Both of these regions had highly dissimilar diazotrophic communities among samples compared to samples from the NPSG euphotic zone and the ESP (Fig. 2.5). Our findings add to the growing literature reporting diazotrophic genes in deep, dark, and N-rich environments.

2.5.2 Mixed evidence for nitrogen fixation in "unusual" environments

Marine N_2 fixation is typically facultative and regulated by physiological and environmental constraints (Paerl et al. 1987); thus, the presence of heterotrophic diazotrophs does not indicate active N_2 fixation. Yet recent studies have reported not only the presence of *nifH* genes, but also active N_2 fixation in aphotic, N-rich waters, and it has been suggested that these depth-integrated rates could have a significant impact on the marine N budget (Bonnet et al. 2013). While mesopelagic N_2 fixation rates are low (typically $\leq$0.5 nmol N L^{-1} d^{-1} and often $\leq$0.1 nmol N L^{-1} d^{-1}, e.g. Bonnet et al. 2013, Rahav et al. 2013), the absence of any detectable N_2 fixation is rarely reported. Here, we tested for N_2 fixation at coastal PNW stations encompassing a range of temperature, nitrate, and oxygen concentrations (Table 2.1, Table S2.1). We observed no measurable N_2 fixation rates. Likewise, we observed no measurable N_2 fixation rates in 200m NPSG samples under ambient conditions, though 3 μmol L^{-1} glucose amendments stimulated low rates (Fig. 2.2).

The absence of detectable N_2 fixation rates in the PNW and mesopelagic NPSG samples is not surprising when considering the sensitivity of the $^{15}N_2$ assay. Our

minimum quantifiable rates ranged from 0.06–0.69 and 0.04–0.12 nmol N L^{-1} d^{-1} for PNW and 200m NPSG samples, respectively, and thus preclude the detection of N_2 fixation rates as low as those reported for some mesopelagic environments (e.g. Benavides et al. 2016). Our sensitivity analysis shows that the variability in the $\delta^{15}N$ and concentration of ambient particulate N were the dominant sources of error in N_2 fixation rate measurements (Table 2.2). An alternative detection limit, calculated as the N_2 fixation rates required to produce a change in ^{15}N of 0.00146 atom % (Montoya et al. 1996), produces values similar to our minimum quantifiable rates (Table 2.2). One reason for our relatively high detection limits is that we used observed particulate N concentrations in the calculations, as originally suggested by Montoya et al. (1996), rather than using the minimum N content required by the mass spectrometer (as in Fernandez et al. 2013). Using relatively low initial $^{15}N_2$ enrichments (~3-4 atom %, Table 2.2) further inflates detection limits; however, these low enrichments are a consequence of necessary improvements to the $^{15}N_2$ assay (Mohr. et al. 2010, Wilson et al. 2012). These sensitivity issues highlight the current need for standard reporting of accurate detection limits and a consensus on best practices in N_2 fixation research.

In contrast to the PNW and mesopelagic NPSG, we detected N_2 fixation rates in all surface NPSG and ESP samples (excluding one 75m NPSG sample; Fig. 2.2). N_2 fixation rates in the NPSG were within previously reported ranges (Böttjer et al. 2016). In the ESP, the relatively low N_2 fixation rates we report for the gyre (<1 nmol N L^{-1} d^{-1}) also agree with previous studies (Halm et al. 2011, Raimbault and Garcia 2008). N_2 fixation rates increased near the coast (Fig. 2.2), and stations 2–5 had rates >5 times those measured by Dekaezemacker et al. (2013) at nearby stations several months later (April–

May 2011). These differences could have arisen from underestimations by Dekaezemacker et al. (2013) due to delayed $^{15}N_2$ bubble dissolution (Mohr et al. 2010), and/or natural spatiotemporal variability. Previous N_2 fixation rate estimates in upwelling-influenced Chilean and Peruvian waters ranged from undetectable to extremely high (>100 nmol N L^{-1} d^{-1}) (Fernandez et al. 2011, Fernandez et al. 2015, Loescher et al. 2014, Raimbault and Garcia 2008) but due to the recent challenges in methodology (Mohr et al. 2010, Dabundo et al. 2014), such comparisons to earlier studies should be treated with caution. Overall, the low to moderate rates we measured add to the mounting evidence that ESP waters are unlikely to harbor the high N_2 fixation rates predicted by initial model projections (Deutsch et al. 2007), perhaps because iron limitation precludes extensive diazotrophy in this region (Knapp et al. 2016, Weber and Deutsch 2014).

Our findings in the ESP point to an emerging paradox: why would N_2 fixation be significant in surface waters where fixed N is available? Ammonium uptake (and to a lesser extent nitrate uptake) is less energetically expensive than N_2 fixation, and conventional wisdom suggests that diazotrophs should only have a competitive advantage over other phytoplankton when N is limiting (Knapp 2012). Yet numerous recent studies have reported N_2 fixation rates in high nitrate waters, often at depth (reviewed in Bombar et al. 2016). In our study, the highest N_2 fixation rates detected in the ESP were at Stn. 1, where nitrate concentrations were also highest (Table 2.1, Fig. 2.2), agreeing with other reports of enhanced N_2 fixation rates in upwelled waters (Sohm et al. 2011, Subramaniam et al. 2013). These results suggest that unlike cyanobacteria, heterotrophic marine N_2 fixation rates might not be reduced in the presence of fixed N (Knapp 2012). Recent

culture-based studies support this potential: Bentzon-Tilia et al. (2015) have shown that an alpha-proteobacterial isolate up-regulates N_2 fixation in response to NH_4^+ additions, possibly as a mechanism for internal redox regulation. Other studies report relationships between heterotrophic N_2 fixation and the supply of dissolved organic matter and hypothesize that organic matter-rich substrates (aggregates) or water masses are necessary to support heterotrophic diazotrophy (Benavides et al. 2015, Severin et al. 2015). Indeed, glucose amendments stimulated detectable N_2 fixation rates in our mesopelagic NPSG samples (Fig. 2.2). These collective findings help to explain heterotrophic diazotrophy in N-rich waters; however, they are largely correlative and cannot yet constrain the overall importance of heterotrophic diazotrophy to regional or oceanic N budgets.

2.5.3 Challenges in reconciling diazotrophic diversity, abundances, and nitrogen fixation rates

Our cross-ecosystem approach allows us to examine the relationship between diazotrophic genes and N_2 fixation rates in environments with and without measurable N_2 fixation, revealing an apparent negative relationship between diazotrophic diversity and function. The highest *nifH* diversity was observed in PNW samples where N_2 fixation was not detected (Fig. 2.6), and *nifH* diversity and N_2 fixation rates were inversely related in NPSG and ESP samples (Table S2.7). This finding is consistent with negative productivity/diversity relationships in many plant and animal communities (Rosenzweig and Abramsky 1993). Competition with successful taxa may lead to a decrease in overall diazotrophic species richness or evenness in environments favoring N_2 fixation.

While it is relatively unsurprising that overall *nifH* diversity does not positively correlate with N_2 fixation rates, we might expect a correspondence between total diazotrophic abundance and N_2 fixation rates based on the cell-specific rates of dominant organisms. We investigated this relationship in the ESP, where a mismatch between cell abundances and N_2 fixation rates has been previously described (Turk-Kubo et al. 2014). Though OTU 2 was the most abundant phylotype in the ESP dataset (Fig. S2.3), quantification of OTU 2 in our samples revealed low *nifH* concentrations (1.5×10^2–4.4×10^3 *nifH* copies L^{-1} assuming 100% extraction efficiency, Table S2.5). Using the approach of Turk-Kubo et al. (2014), and assuming the highest-reported cell-specific N_2 fixation rates by heterotrophic marine diazotrophs in culture, a concentration of 9.7×10^4 cells L^{-1} would be required to produce just the ESP-average minimum quantifiable rate (Table 2.2). Thus, OTU 2 appears insufficient to account for the measured N_2 fixation rates, a finding consistent with previous reports (Turk-Kubo et al. 2014). It is unlikely that this abundance/rate mismatch is driven by an overestimation of N_2 fixation rates due to ^{15}N contamination, since Cambridge $^{15}N_2$ gas has been observed to have minimal levels of $^{15}NH_4^+$ contamination equivalent to rates of <0.02 nmol N L^{-1} d^{-1} (Dabundo et al. 2014) that are well below our detection limits (Table 2.2, S2.3). Instead, we believe the abundance of active diazotrophs was underestimated. OTU 2 could have deceptively high relative abundances in the *nifH* sequence dataset due to biases in the degenerate *nifH* primers (Turk et al. 2011) or active diazotrophs in the >10 μm size fraction; if so, this phylotype might not represent the most abundant diazotroph in our samples. Other dominant phylotype(s) or a diverse community of diazotrophs may have contributed to measured N_2 fixation rates. Alternatively, OTU 2 could have extremely high cell-specific

N_2 fixation rates, perhaps due to a novel symbiosis. Quantitative PCR reactions performed on a limited number of ESP RNA samples indicate low but detectable *nifH* expression by a gamma-proteobacterial phylotype (ETSP-2, Turk-Kubo et al. 2014) but cannot resolve which diazotrophic taxa drive observed ESP N_2 fixation rates (data not shown).

While the specific organisms driving N_2 fixation rates in the ESP remain elusive, our finding of UCYN-A dominance in the NPSG was consistent among samples (Fig. 2.4, S2.3) and agrees with previous studies (Böttjer et al. 2014). Thus, since the mass ratio hypothesis predicts that the dominant species contributes most to community productivity (Grime 1998), we might expect a positive relationship between UCYN-A abundance and community N_2 fixation rates. However, ddPCR-derived UCYN-A *nifH* gene concentrations did not correlate to N_2 fixation rates in the NPSG ($R^2 = 0.12$, $p > 0.05$). In fact, we are not aware of any studies that have shown strong linkages between molecular estimates of diazotroph cell abundance and N_2 fixation rates in nature. The relationship between gene copies and rates is complicated by species-specific gene copy numbers (Sargent et al. 2016) and gene expression levels (Church et al. 2005b), and also by environmental controls such as light, temperature, nutrients, and dissolved organic matter (Benavides et al. 2015, Luo et al. 2014). Thus, the biogeochemical implications of gene-based diazotrophic abundances should be inferred with caution.

Reconciling the diversity and function of diazotrophic communities is further complicated by the different sensitivities for *nifH* diversity and N_2 fixation rate assays. Our $^{15}N_2$ tracer incubations produced average minimum quantifiable rates exceeding N_2 fixation rates previously reported in mesopelagic waters (often <1 nmol N L^{-1} d^{-1} , e.g.

Benavides et al. 2016, Dekaezemacker et al. 2013, Rahav et al. 2013). Detection limit concerns are further exacerbated by recent reports of $^{15}NH_4^+$ contaminants in $^{15}N_2$ gas, which can result in apparent N_2 fixation rates of up to 0.02 nmol N L^{-1} d^{-1} in low-contaminant Cambridge gas stocks and >100 nmol N L^{-1} d^{-1} in highly contaminated Campro gas stocks (Dabundo et al. 2014). It is not routine for authors to report detection limits for $^{15}N_2$ fixation rate measurements, yet if the limits we have calculated can be applied widely, then the biogeochemical significance of diazotrophy may have been overestimated by many studies, particularly in the ESP and the mesopelagic.

While low N_2 fixation rates can be difficult to detect, *nifH* PCR-based methods enable the detection of extremely rare diazotrophs, albeit with potential associated biases and contamination concerns. Since *nifH* sequence data are not quantitative, qPCR is often used to quantify specific diazotrophic phylotypes. This technique has the potential to underestimate total diazotrophic abundance, through mismatches in primer/probe binding sites (Lefever et al. 2013) and because samples are often only screened for specific diazotrophic phylotypes; for example, 4 out of 5 of our most abundant ESP phylotypes would not be targeted using published qPCR primer sets (Table S2.6). Yet interpretations of qPCR data can also overestimate the functional importance of heterotrophic phylotypes. These organisms are usually detected at concentrations of 10^2–10^3 (rarely 10^4) *nifH* copies L^{-1} (Langlois et al. 2015, Turk-Kubo et al. 2014), lower than the cell concentrations required to produce even our minimum quantifiable rates when assuming the highest reported cell-specific N_2 fixation rates from heterotrophic isolates recently cultured by Bentzon-Tilia et al. (2015) (Table 2.2). In short, it is easier to detect the

presence/absence of *nifH* genes than to accurately measure relatively low N_2 fixation rates (<1 nmol N L^{-1} d^{-1}), and the relationship between genes and rates remains elusive.

2.6 Conclusions and implications

Here, we illustrate a near-ubiquity of marine diazotrophs—we recovered *nifH* genes from every environment probed—but restricted regions of their functional significance to N cycling. The absence of detectable N_2 fixation rates despite the presence of *nifH* genes in the "unusual" environments of the coastal PNW and mesopelagic NPSG could be due to low diazotrophic abundances or low *nifH* expression by heterotrophic diazotrophs. Our data and others (e.g. Farnelid et al. 2011, Langlois et al. 2015) have begun to map out the biogeography of diazotrophic taxa, indicating a geographical constraint of cyanobacterial diazotrophs and the widespread distribution of putative heterotrophic diazotrophs. Future culture-based efforts can help to determine the ecological controls of these seemingly cosmopolitan heterotrophic diazotrophs and the mechanisms for variability in active N_2 fixation (Bentzon-Tilia et al. 2015). However, the absence of measurable N_2 fixation rates is also related to the inherent sensitivity of the $^{15}N_2$ fixation assay, a point that is rarely considered in marine N_2 fixation literature despite extensive detection limit calculations in the original paper describing the $^{15}N_2$ tracer method (Montoya et al. 1996). Our analyses show that this sensitivity issue precludes the detection of very low N_2 fixation rates. Thus, it seems improbable that mesopelagic N_2 fixation rates can accurately be extrapolated to significantly affect the global marine N budget.

Support for this project was provided primarily via the Center for Microbial Oceanography Research and Education (EF0424599 D.M. Karl, sub–awards to RML, AW, and JPZ), with additional contributions from the Simons Collaboration on Ocean Processes and Ecology (329108 to D.M. Karl with subaward to AW and JPZ). We thank S. Ferrón for MIMS analyses and K. Turk–Kubo for helpful discussions

2.7 Supplemental Information

Preparation of $^{15}N_2$ enriched seawater

For all ESP incubations, $^{15}N_2$-enriched seawater was prepared at sea. At each station, at least 3 h prior to dawn, surface seawater was collected, 0.2 μm-filtered, and bubbled with 99% ultrapure helium for 1 h. This filtered, degassed seawater was then transferred to a 1 L polycarbonate bottle capped with a septa closure, and 8 mL of $^{15}N_2$ gas was added using a gas-tight syringe. Enrichment bottles were shaken for ~10 min and stored at 20ºC for at least 1 h to encourage dissolution of the added gas.

For NPSG and PNW incubations, $^{15}N_2$-enriched seawater was prepared onshore ~1 week before departure using 0.2 μm-filtered seawater from Station ALOHA according to the methods of Wilson et al. (2012). Filtered seawater was bubbled with 99% ultrapure helium for ≥15 minutes, then heated, stirred, and vacuumed for ≥1 hour. The degassed seawater was dispensed into pre-vacuumed, gas-tight 3 L PTFE bags (Welch Fluorocarbon), which were subsequently weighed and cooled on ice. A gas-tight syringe was used to inject 12.9 mL^{-L} $^{15}N_2$ gas, and the bag was manually agitated for ~30 min. The small remaining $^{15}N_2$ gas bubble (~15% initial volume added) was removed, and the

enriched seawater was dispensed into glass serum bottles, which were capped, crimped, and stored at 4°C until use.

The $^{15}N_2$ content of enriched seawater was validated via Membrane Inlet Mass Spectrometry (MIMS). The $^{15}N_2$ enriched seawater used in NPSG incubations was analysed directly, while the content of $^{15}N_2$ enriched seawater used in ESP and PNW incubations was inferred by analysing subsequent batches of $^{15}N_2$ enriched seawater prepared using the same method for each respective cruise. NPSG (and inferred PNW) enriched seawater samples were analysed with MIMS according to the methods of Ferrón et al (2016). MIMS samples for the ESP were shipped from Chile to Oregon State post-cruise for validation of $^{15}N_2$ concentrations in enriched seawater, however shipping delays led to compromised samples. In lieu of having these values, we replicated the enriched seawater preparation methods at OSU (see above), and measured the resulting $^{15}N_2$ concentrations using MIMS in triplicate preparations. These values were used to calculate the initial A_{N2} and associated SD shown in Table 2 and Table S3.

Detailed methods for $^{15}N_2$ incubations

For ESP incubations, seawater from 5 m was collected prior to dawn for rate assays and natural abundance measurements ($^{15}N/^{14}N$). At coastal stations (1-3), triplicate 2.4 L polycarbonate bottles were filled to capacity; then, exactly 200 mL was removed and replaced with $^{15}N_2$-enriched seawater, and bottles were capped with septa closures. At gyre stations (4-7), incubations took place in 4.4 L polycarbonate bottles, and 350 mL of seawater was replaced with $^{15}N_2$-enriched seawater. Variable incubation volumes were used to ensure collection of sufficient particulate N across a transect of productivity.

Incubation bottles were transferred to on-deck incubators plumbed with surface seawater. These incubations were carried out in duplicate and incubated for 24 h (~dawn to dawn).

For PNW incubations, seawater was carefully sampled into 1.1 or 2.2 L polycarbonate bottles for $^{15}N_2$ assays and natural abundance samples, avoiding air contamination during CTD sampling. A syringe with tubing was used to dispense 130 mL of ice-cold $^{15}N_2$-enriched seawater into the bottom of incubation bottles, displacing an equal volume of sample while avoiding air contamination. Bottles were capped and incubated in the dark for 24 h at the average *in situ* temperature for the three depths (Table S2). Duplicate incubation bottles were used for each depth.

For NPSG incubations, seawater samples for incubations and natural abundance were collected before dawn and sampled into 4.4L polycarbonate bottles, avoiding air contamination. 130 mL $^{15}N_2$-enriched seawater was added to each sample as in PNW incubations. For 200 m samples, a 3 μM glucose treatment was included for some experiments. 25 m and 75 m samples were incubated in deckboard incubators cooled with surface seawater (~25°C) for 24 h (~dawn to dawn). Samples from 200 m were incubated in the dark at ~20.5°C.

ESP clone library sequencing and qPCR

A 359 base pair (bp) fragment of the *nifH* gene was amplified, cloned and sequenced from ESP surface waters using a previously described nested PCR approach with degenerate primers (Zehr and Turner, 2001). Amplified products were purified, cloned into pGEM-T vectors (Promega, Madison, WI, USA) and sequenced at the University of California Berkeley sequencing facility, using SP6 or T7 primers.

Previously designed qPCR primer and probe sets were used to enumerate *nifH* gene copy and transcript abundances of phylotypes “ETSP1” and “ETSP2” (Turk-Kubo et al 2014).

nifH negative control and filter blank samples

Four negative control samples (water substituted for DNA in PCR reactions) and four filter blank samples (Clean Supor filters which underwent DNA extraction and other sample processing procedures) were included for *nifH* Illumina sequencing. Triplicate PCR reactions were pooled for each of these samples, and the samples were sequenced despite the absence of gel electrophoresis bands corresponding to the partial *nifH* gene size. Of these 8 samples, only 2 filter blank samples retained any sequences after quality control procedures. One filter blank sample contained 1574 *nifH* sequences which all clustered to OTU 44, which was classified as 1J/K proteobacteria and contains 99% nucleic acid identity to *Methelocystus rosea* strain SV97. Only 1 OTU 44 sequence was found in any other sample in our dataset. The other filter blank sample contained 634 sequences comprised of 18 OTUs, and had a similar community structure to NPSG samples (Fig. S2). Though both of these filter blank samples contained far less sequences than the rarefication cutoff of 10 386, the presence of *nifH* sequences in blanks highlights the need for care in sample processing for the sensitive nested *nifH* PCR procedure.

2.8 References

Benavides M., P. H. Moisander, H. Berthelot, T. Dittmar, O. Grosso, and S. Bonnet. 2015. Mesopelagic N_2 fixation Related to Organic Matter Composition in the Solomon and Bismarck Seas (Southwest Pacific). PLoS One 10: e0143775, doi: 10.1371/journal.pone.0143775

Benavides, M. and others 2016. Basin. Basin-wide N_2 fixation in the deep waters of the Mediterranean Sea. Global Biogeochem. Cycles 30: 952–961.

Bentzon-Tilia, M., I. Severin, L. H. Hansen, and L. Riemann. 2015. Genomics and ecophysiology of heterotrophic nitrogen-fixing bacteria isolated from estuarine surface water. MBio 6: e00929–00915, doi: 10.1128/mBio.00929-15

Bonnet, S. and others. 2013. Aphotic N_2 Fixation in the Eastern Tropical South Pacific Ocean. PLoS One 8: e81265, doi:10.1371/journal.pone.0081265

Bombar, D., R. W. Pearl, and L. Riemann. 2016. Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. Trends Microbiol. 24: 916–927. doi: 10.1016/j.tim.2016.07.002

Böttjer, D. and others. 2016. Temporal variability of nitrogen fixation and particulate nitrogen export at Station ALOHA. Limnol. Oceanogr. 62: 200–216. doi: 10.1002/lno.10386

Böttjer, D., D. M. Karl, R. M. Letelier, D. A. Viviani, and M. J. Church. 2014. Experimental assessment of diazotroph responses to elevated seawater pCO_2 in the North Pacific Subtropical Gyre. Global Biogeochem. Cycles 28: 601–616.

Caporaso J. G. and others. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature methods 7: 335–336.

Church, M. J., K. M. Björkman, D. M. Karl, M. A. Saito, and J. P. Zehr. 2008. Regional distributions of nitrogen-fixing bacteria in the Pacific Ocean. Limnol. Oceanogr. 53: 63-77.

Church, M. J., B. D. Jenkins, D. M. Karl, and J. P. Zehr. 2005a. Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. Aquat. Microb. Ecol. 38: 3–14.

Church, M. J., C. M. Short, B. D. Jenkins, D. M. Karl, and J. P. Zehr. 2005b. Temporal patterns of nitrogenase gene (nifH) expression in the oligotrophic North Pacific Ocean. Appl. Environ. Microbiol. 71: 5362–5370.

Codispoti, L. An oceanic fixed nitrogen sink exceeding 400 Tg N a^{-1} vs the concept of homeostasis in the fixed-nitrogen inventory. Biogeosciences 4: 233–253.

Dabundo, R. and others 2014. The Contamination of Commercial $^{15}N_2$ Gas Stocks with ^{15}N-Labeled Nitrate and Ammonium and Consequences for Nitrogen Fixation Measurements. PLoS One 9: e110335, doi: 10.1371/journal.pone.0110335

Dekaezemacker, J., S. Bonnet, O. Grosso, T. Moutin, M. Bressac, and D. Capone. 2013. Evidence of active dinitrogen fixation in surface waters of the eastern tropical South Pacific during El Niño and La Niña events and evaluation of its potential nutrient controls. Global Biogeochem. Cycles 27: 768–779.

Deutsch, C., J. L. Sarmiento, D. M. Sigman, N. Gruber, and J. P. Dunne. 2007. Spatial coupling of nitrogen inputs and losses in the ocean. Nature 445: 163-167.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460-2461.

Farnelid, H. and others 2011. Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. PloS one 6: e19223, doi: 10.1371/journal.pone.0019223

Feely, R.A., S.R. Alin, B. Hales, G.C. Johnson, R.H. Byrne, W.T. Peterson and D. Greeley. Chemical and hydrographic profile measurements during the 2013 West Coast Ocean Acidification Cruise (WCOA2013, August 3-29, 2013). http://cdiac.ornl.gov/ftp/oceans/WCOA2013/. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee. doi: 10.3334/CDIAC/OTG.COAST_WCOA2013

Fernandez, C., L. Farías, and O. Ulloa. 2011. Nitrogen fixation in denitrified marine waters. PloS one 6: e20539, doi: 10.1371/journal.pone.0020539

Fernandez, C. L., and others. 2013 Community N_2 fixation and *Trichodesmium* spp. abundance along longitudinal gradients in the eastern subtropical North Atlantic. Ices J. Mar. Sci. 70: 223–231. fss142, doi: 10.1093/icesjms/fss142

Fernandez, C., M. L. González, C. Muñoz, V. Molina, and L. Farias. 2015. Temporal and spatial variability of biological nitrogen fixation off the upwelling system of central Chile (35–38.5° S). J. Geophys. Res.: Oceans 120: 3330-3349.

Ferrón S, del Valle D. A., Björkman K. M., Quay P. D., Church M. J., and D. M. Karl. 2016. Application of membrane inlet mass spectrometry to measure aquatic gross primary production by the ^{18}O in vitro method. Limnol. Oceanogr. Methods; e-pub ahead of print 15 July 2016, doi: 10.1002/lom3.10116

Galloway, J. N. and others 2004. Nitrogen cycles: past, present, and future. Biogeochemistry 70: 153-226.

Grime, J. 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. J. Ecol. 86: 902-910.

Großkopf, T. and others 2012. Doubling of marine dinitrogen-fixation rates based on direct measurements. Nature 488: 361-364.

Gruber, N. 2004. The dynamics of the marine nitrogen cycle and its influence on atmospheric CO_2 variations, p. 97-148. In Follows, M. and T. Oguz [eds.], The ocean carbon cycle and climate. Springer.

Halm, H. and others 2011. Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. ISME J 6: 1238-1249.

Hewson, I. and others 2007. Characteristics of diazotrophs in surface to abyssopelagic waters of the Sargasso Sea. Aquat. Microb. Ecol. 46: 15-30.

Jayakumar, A., M. Al-Rshaidat, B. B. Ward, and M. R. Mulholland. 2012. Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea. FEMS Microbiol. Ecol. 82: 597-606.

Karl, D. and others. 2002. Dinitrogen fixation in the world's oceans. Biogeochemistry 57: 47-98.

Knapp, A. N. 2012. The sensitivity of marine N_2 fixation to dissolved inorganic nitrogen. Front. Microbiol. 3: 374, doi:10.3389/fmicb.2012.00374.

Knapp, A. N., J. Dekaezemacker, S. Bonnet, J. A. Sohm, and D. G. Capone. 2012. Sensitivity of *Trichodesmium erythraeum* and *Crocosphaera watsonii* abundance and N_2 fixation rates to varying NO_3^- and PO_4^{3-} concentrations in batch cultures. Aquat. Microb. Ecol. 66: 223-236.

Knapp, A. N., K. L. Casciotti, W. M. Berelson, M. G. Prokopenko, and D. G. Capone. 2016. Low rates of nitrogen fixation in eastern tropical South Pacific surface waters. Proc. Natl. Acad. Sci. U.S.A.: 113: 4398–4403.

Kozich, J. J., S. L. Westcott, N. T. Baxter, S. K. Highlander, and P. D. Schloss. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl. Environ. Microbiol. 79: 5112–5120.

Langlois, R., T. Großkopf, M. Mills, S. Takeda, and J. Laroche. 2015. Widespread distribution and expression of Gamma A (UMB), an uncultured, diazotrophic, γ-

proteobacterial *nifH* phylotype. PloS one 10: e0128912. doi: 10.1371/journal.pone.0128912

Lefever S., F. Pattyn, J. Hellemans, and J. Vandesompele. 2013. Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. Clin. Chem. 59: 1470–1480, doi:10.1373/clinchem.2013.203653

Letunik, I. and P. Bork. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 44:W242–W245, doi: 10.1093/nar/gkw290

Loescher, C. R. and others 2014. Facets of diazotrophy in the oxygen minimum zone waters off Peru. ISME J 8: 2180–2192.

Luo, Y. -W. and others 2012. Database of diazotrophs in global ocean: abundances, biomass and nitrogen fixation rates. Earth System Science Data 4: 47-73.

Luo, Y. -W., I. Lima, D. M. Karl, C. A. Deutsch, and S. C. Doney. 2014. Data-based assessment of environmental controls on global marine nitrogen fixation. Biogeosciences 11: 691–708.

Mohr, W., T. Großkopf, D. W. R. Wallace, and J. Laroche. 2010. Methodological underestimation of oceanic nitrogen fixation rates. PloS one 5: e12583. doi: 10.1371/journal.pone.0012583

Moisander, P. H., R. A. Beinart, M. Voss, and J. P. Zehr. 2008. Diversity and abundance of diazotrophic microorganisms in the South China Sea during intermonsoon. ISME J 2: 954–967.

Montoya, J. P., M. Voss, P. Kahler, and D. G. Capone. 1996. A simple, high-precision, high-sensitivity tracer assay for N_2 fixation. Appl. Environ. Microbiol. 62: 986–993.

Paerl, H. W., K. M. Crocker, and L. E. Prufert. 1987. Limitation of N_2 fixation in coastal marine waters: Relative importance of molybdenum, iron, phosphorus, and organic matter availability. Limnol. Oceanogr. 32: 525-536.

Rahav, E. and others. 2013. Dinitrogen fixation in aphotic oxygenated marine environments. Front. Microbiol. 4: 227.

Raimbault, P., and N. Garcia. 2008. Evidence for efficient regenerated production and dinitrogen fixation in nitrogen-deficient waters of the South Pacific Ocean: impact on new and export production estimates. Biogeosciences 5: 323–338.

Riemann, L., H. Farnelid, and G. F. Steward. 2010. Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters. Aquat. Microb. Ecol. 61: 235–247.

Rii, Y. M., S. Duhamel, R. R. Bidigare, D. M. Karl, D. J. Repeta, and M. J. Church. 2016. Diversity and productivity of photosynthetic picoeukaryotes in biogeochemically distinct regions of the South East Pacific Ocean. Limnol. Oceanogr. 61: 806-824.

Rosenzweig, M. L., and Z. Abramsky. 1993. How are diversity and productivity related?, p 52–65. In Ricklefs, R. E. and Schluter D. [eds.], Species diversity in ecological communities. University of Chicago Press.

Sargent, E. C., A. Hitchcock, S. A. Johansson, R. Langlois, C. M. Moore, J. LaRoche, A. J. Poulton, and T. S. Bibby. 2016. Evidence for polyploidy in the globally important
diazotroph *Trichodesmium*. FEMS Microbiol. Lett. 362. doi: 10.1093/femsle/fnw244

Schloss, P. D. and others 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75: 7537-7541.

Severin I, Bentzon-Tilia M, Moisander PH, Riemann L. 2015. Nitrogenase expression in estuarine bacterioplankton influenced by organic carbon and availability of oxygen. FEMS Microbiol. Lett. 362: fnv105, doi: 10.1093/femsle/fnv105

Shiozaki, T., M. Ijichi, T. Kodama, S. Takeda, and K. Furuya. 2014. Heterotrophic bacteria are major nitrogen fixers in the euphotic zone of the Indian Ocean. Global Biogeochem. Cycles 28: 1096–1110.

Sohm, J. A., J. A. Hilton, A. E. Noble, J. P. Zehr, M. A. Saito, and E. A. Webb. 2011. Nitrogen fixation in the South Atlantic Gyre and the Benguela upwelling system. Geophys. Res. Lett. 38: L16608, doi: 10.1029/2011GL048315

Subramaniam, A., C. Mahaffey, W. Johns, and N. Mahowald. 2013. Equatorial upwelling enhances nitrogen fixation in the Atlantic Ocean. Geophys. Res. Lett. 40: 1766-1771.

Turk K.A., and others. 2011. Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the eastern North Atlantic. ISME J 5: 1201–1212.

Turk-Kubo K. A., K. M. Achilles, T. R. Serros, M. Ochiai, J. P. Montoya, and J. P. Zehr. 2012. Nitrogenase (*nifH*) gene expression in diazotrophic cyanobacteria in the Tropical North Atlantic in response to nutrient amendments. Front. Microbiol. 3: 386, doi: 10.3389/fmicb.2012.00386

Turk-Kubo, K. A., M. Karamchandani, D. G. Capone, and J. P. Zehr. 2014. The paradox of marine heterotrophic nitrogen fixation: Abundances of heterotrophic diazotrophs do not account for nitrogen fixation rates in the Eastern Tropical South Pacific. Environ. Microbiol. 16: 3095–3114.

Weber, T., and C. Deutsch. 2014. Local versus basin-scale limitation of marine nitrogen fixation. Proc. Natl. Acad. Sci. U.S.A. 111: 8741–8746.

Wilson, S. T., D. Böttjer, M. J. Church, and D. M. Karl. 2012. Comparative assessment of nitrogen fixation methodologies, conducted in the oligotrophic North Pacific Ocean. Appl. Environ. Microbiol .78: 6516–6523.

Zani, S., M. T. Mellon, J. L. Collier, and J. P. Zehr. 2000. Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. Appl. Environ. Microbiol. 66: 3119–3124.

Zehr, J. P., L. L. Crumbliss, M. J. Church, E. O. Omoregie, and B. D. Jenkins. 2003a. Nitrogenase genes in PCR and RT-PCR reagents: implications for studies of diversity of functional genes. Biotechniques 35: 996–1013.

Zehr, J. P., B. D. Jenkins, S. M. Short, and G. F. Steward. 2003b. Nitrogenase gene diversity and microbial community structure: a cross‐system comparison. Environ. Microbiol. 5: 539–554.

Zehr, J. P., and L. A. Mcreynolds. 1989. Use of degenerate oligonucleotides for amplification of the nifH gene from the marine cyanobacterium *Trichodesmium thiebautii*. Appl. Environ. Microbiol. 55: 2522–2526.

Zehr, J. P., M. T. Mellon, and S. Zani. 1998. New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. Appl. Environ. Microbiol. 64: 3444–3450.

Zehr J. and P. Turner. 2001. Nitrogen fixation: Nitrogenase genes and gene expression. Methods Microbiol. 30: 271-286.

Zhang, Y., Z. Zhao, J. Sun, and N. Jiao. 2011. Diversity and distribution of diazotrophic communities in the South China Sea deep basin with mesoscale cyclonic eddy perturbations. FEMS Microbiol Ecol 78: 417–427.

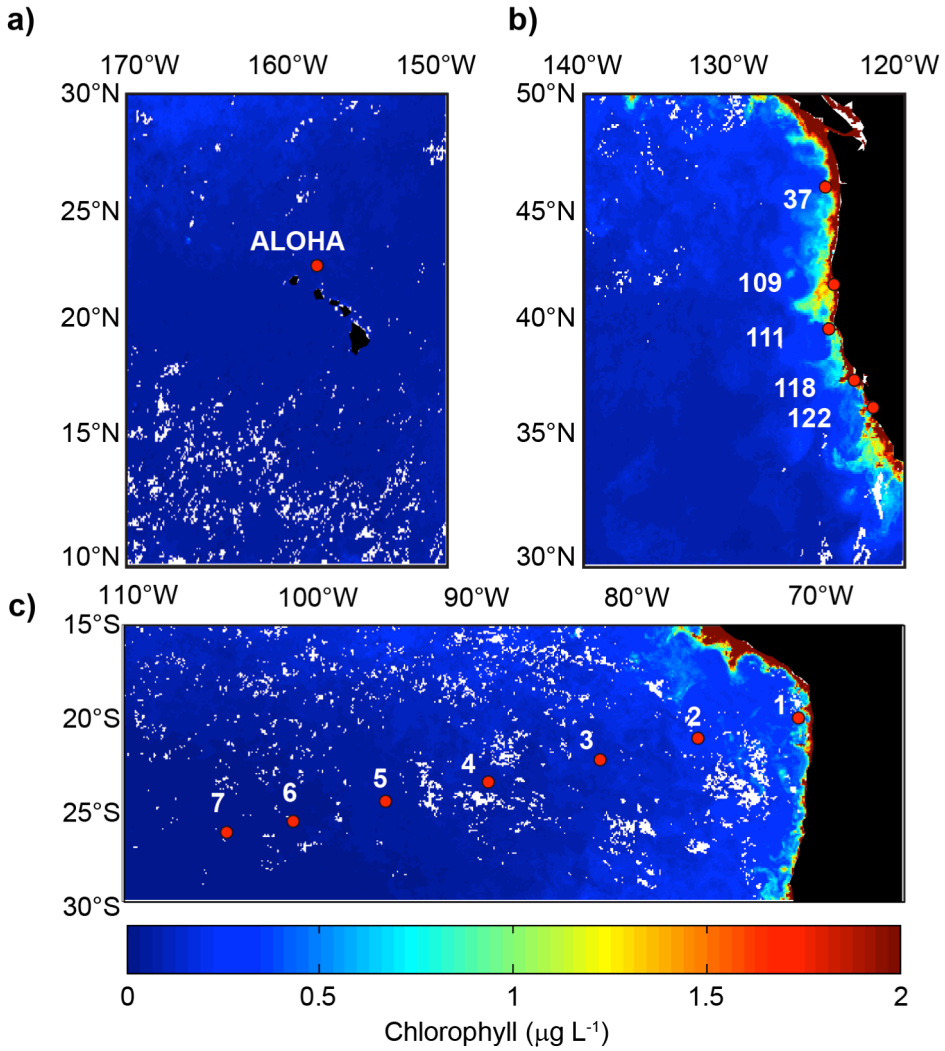

Figure 2.1: Location of sampling stations from the North Pacific Subtropical Gyre (NPSG, A), Pacific Northwest (PNW, B), and Eastern South Pacific (ESP, C) cruises. Stations are superimposed onto surface chlorophyll concentrations obtained from 9 km-binned MODIS satellite data (http://oceandata.sci.gsfc.nasa.gov/) from the month sampled (March 2014 for NPSG, August 2013 for PNW, and November 2010 for ESP).

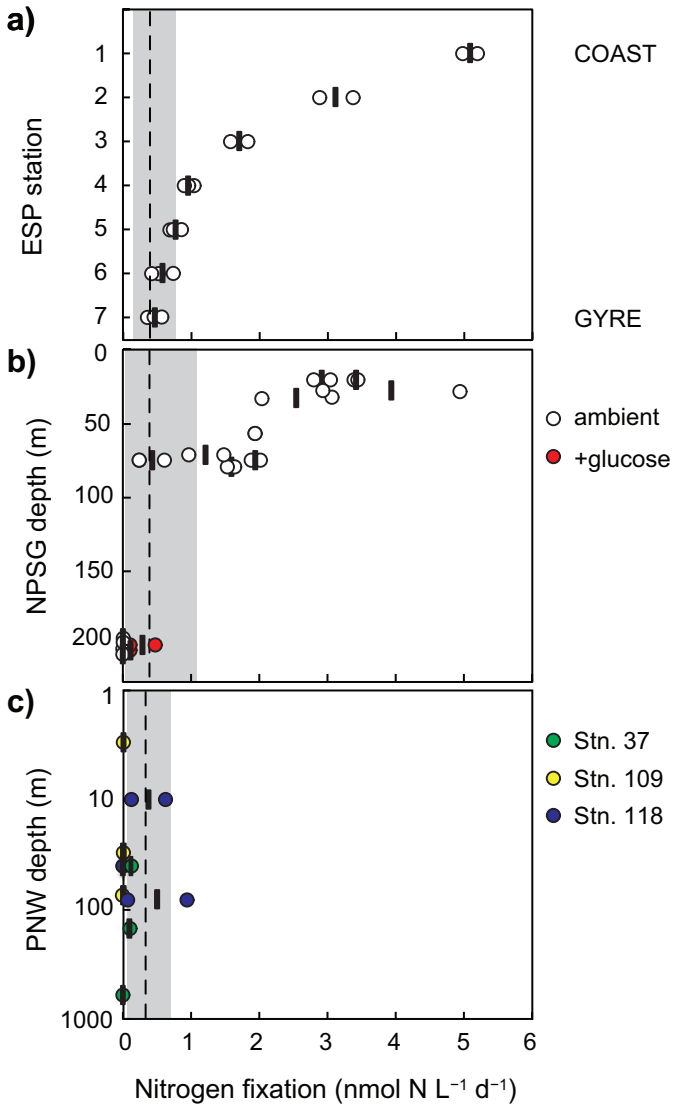

Figure 2.2: Nitrogen fixation rates on Eastern South Pacific (ESP, A), North Pacific Subtropical Gyre (NPSG, B), and Pacific Northwest (PNW, C) cruises, measured via the $^{15}N_2$ assimilation assay. Note that N_2 fixation rates at ESP stations were only measured in surface waters. Circles represent individual measurements and vertical dashes represent averages. Colors denote glucose amendment (B) or station (C). Black dashed lines represent cruise-averaged minimum quantifiable rates, with the full range of minimum quantifiable rates observed on each cruise (Table S2.3) indicated by grey shading. N_2 fixation rates below the cruise-minimum minimum quantifiable rate were moved to zero.

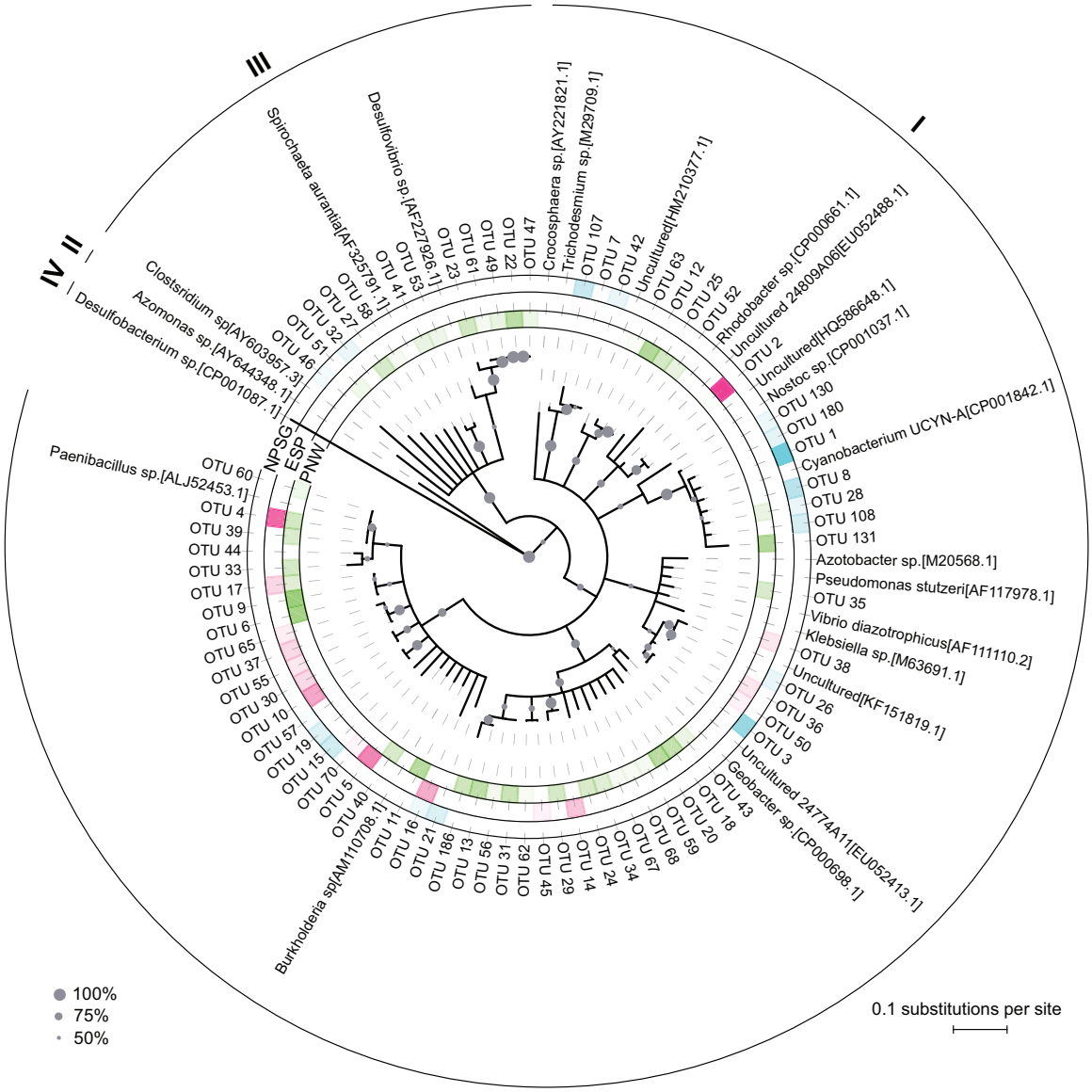

Figure 2.3: Neighbor-joining phylogenetic tree derived from partial *nifH* amino acid sequences from North Pacific Subtropical Gyre (NPSG), Eastern South Pacific (ESP), and Pacific Northwest (PNW) cruises. Representative sequences from each operational taxonomic unit (OTU) containing >10 sequences in the rarefied dataset (71 OTUs representing >99% of total sequences) are displayed on the tree. Colored bars represent the log-transformed relative abundances of OTUs from the NPSG (blue), PNW (green), and ESP (red) samples, with darker shading indicating higher relative abundance. Bootstrap values (1000 replicates) of >50% are represented in the tree with size-proportional grey circles. The taxonomic affiliations of OTUs with canonical *nifH* clades (Zehr et al. 2003b) are displayed in roman numerals. An alignment of the sequences represented in this tree is provided in Fig. S2.4. Tree was produced using the Interactive Tree of Life (http://itol.embl.de/, Letunic and Bork 2016).

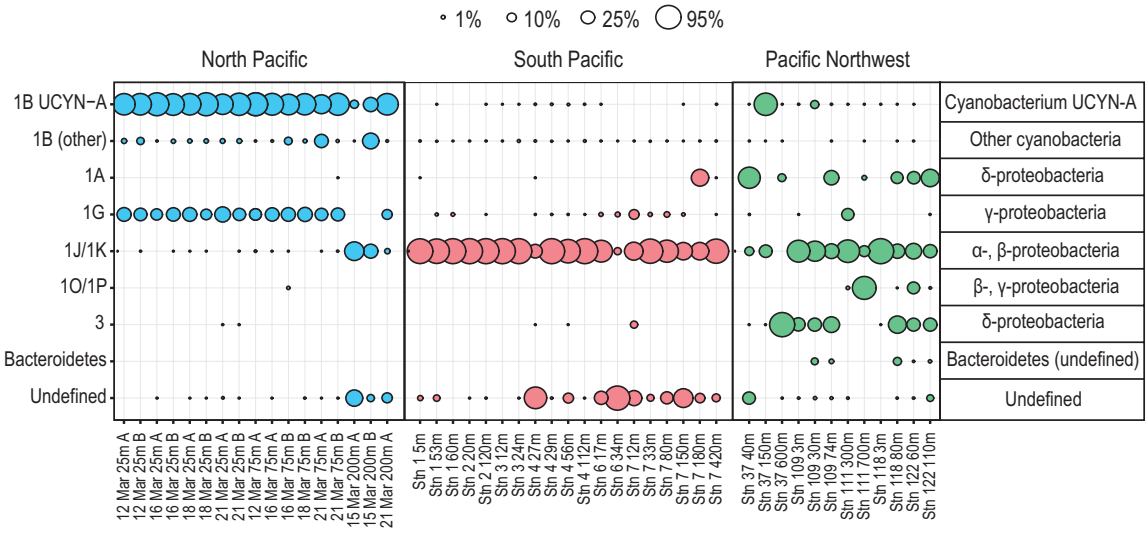

Figure 2.4: Percentages of sequences from North Pacific Subtropical Gyre (blue), Eastern South Pacific (red), and Pacific Northwest (green) samples assigned to *nifH* gene cluster groups. Representative taxa within each *nifH* group are provided to the right. Group 1B was divided into "UCYN-A" and "other cyanobacteria." Sequences with <90% amino acid similarity to reference sequences or equal similarity to multiple reference sequences defined to different *nifH* clusters are termed "Undefined."

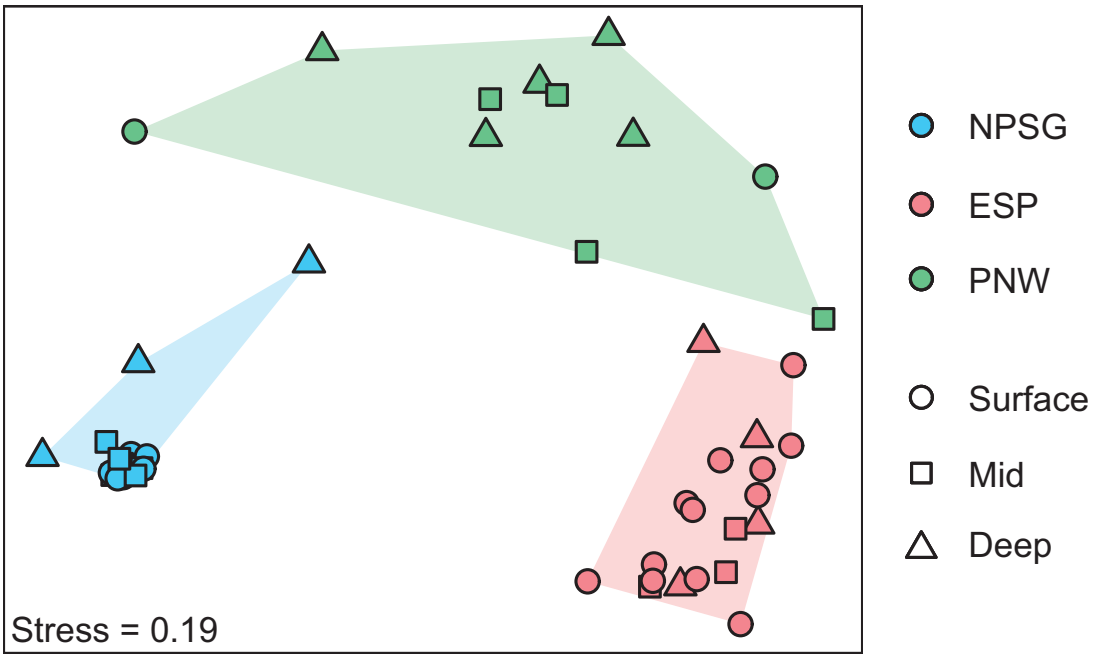

Figure 2.5: Non-metric multi-dimensional scaling (NMDS) plot derived from the Bray-Curtis dissimilarity matrix of *nifH* OTUs from North Pacific Subtropical Gyre (NPSG), Eastern South Pacific (ESP), and Pacific Northwest (PNW) cruises. Each point represents an individual sample. Definition of depth classes vary by cruise: 'Surface' indicates 25m (NPSG), above the deep chlorophyll maxima (DCM, ESP), or the oxygen maxima (PNW), 'Mid' indicates 75m (NPSG), ~DCM (ESP), or the oxycline (PNW), and 'Deep' indicates 200m (NPSG), below the DCM (ESP), or the oxygen minima (PNW).

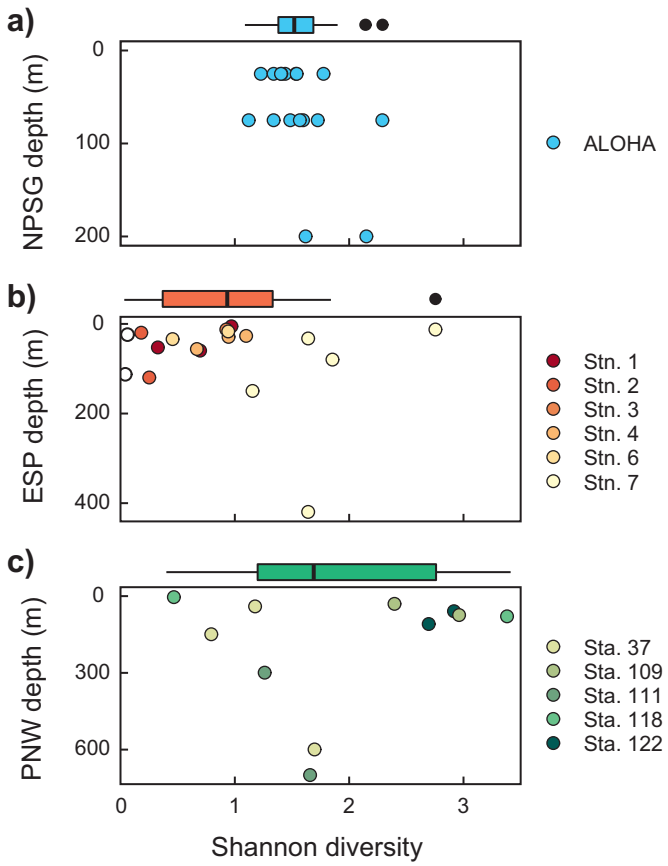

Figure 2.6: Estimates of Shannon diversity for North Pacific Subtropical Gyre (A), Eastern South Pacific (B), and Pacific Northwest (C) cruises. Boxplots above each panel represent the depth- and station-pooled Shannon diversity from each cruise. Values are derived from partial *nifH* DNA sequences (clustered into 97% OTUs).

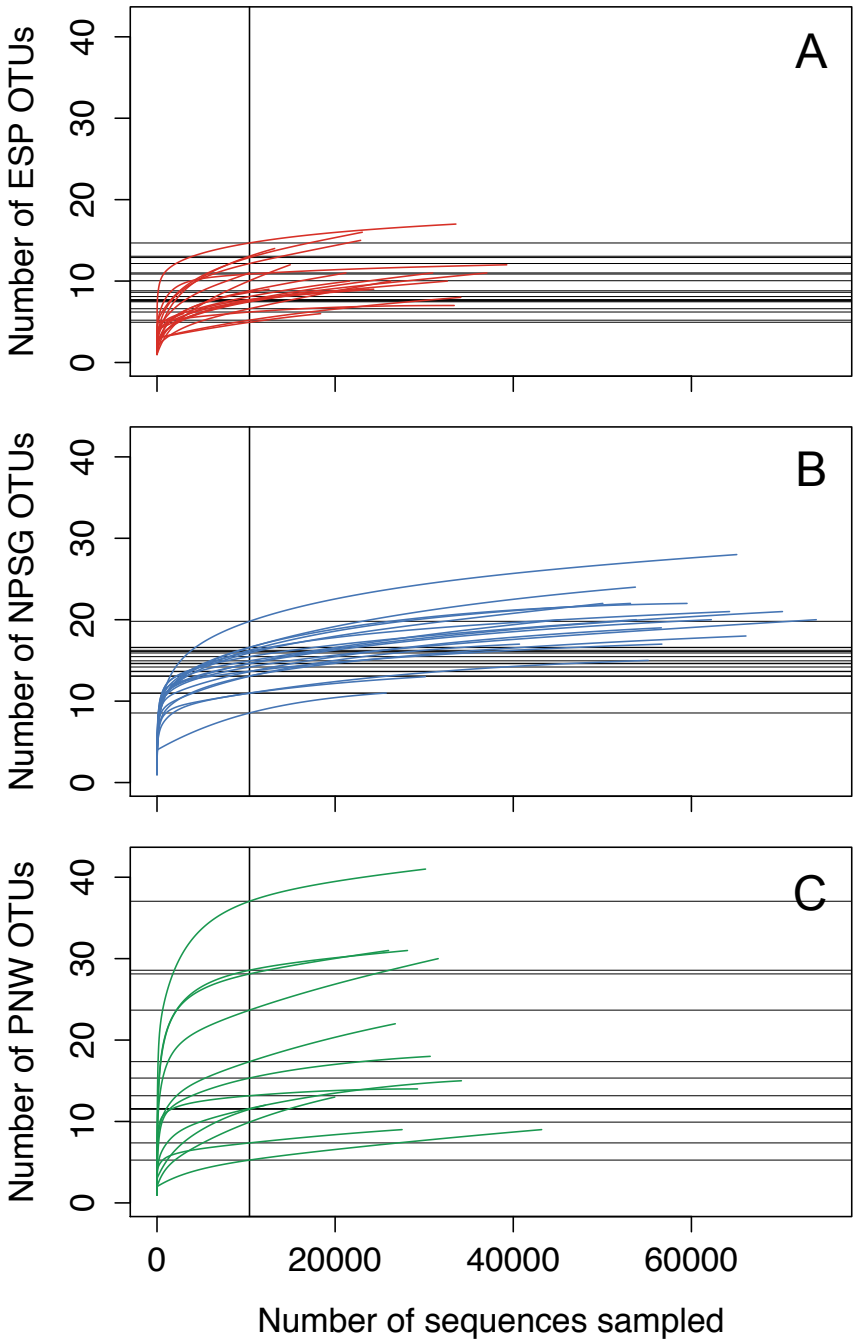

Figure S2.1: Rarefaction curves for ESP (A), NPSG (B), and PNW (C) samples. Curves were produced using the vegan rarecurve function (http://CRAN.R-project.org/package=vegan) for OTUs clustered at 97% nucleotide identity. All samples were subsampled to 10,368 sequences (shown in vertical lines) prior to phylogenetic analyses.

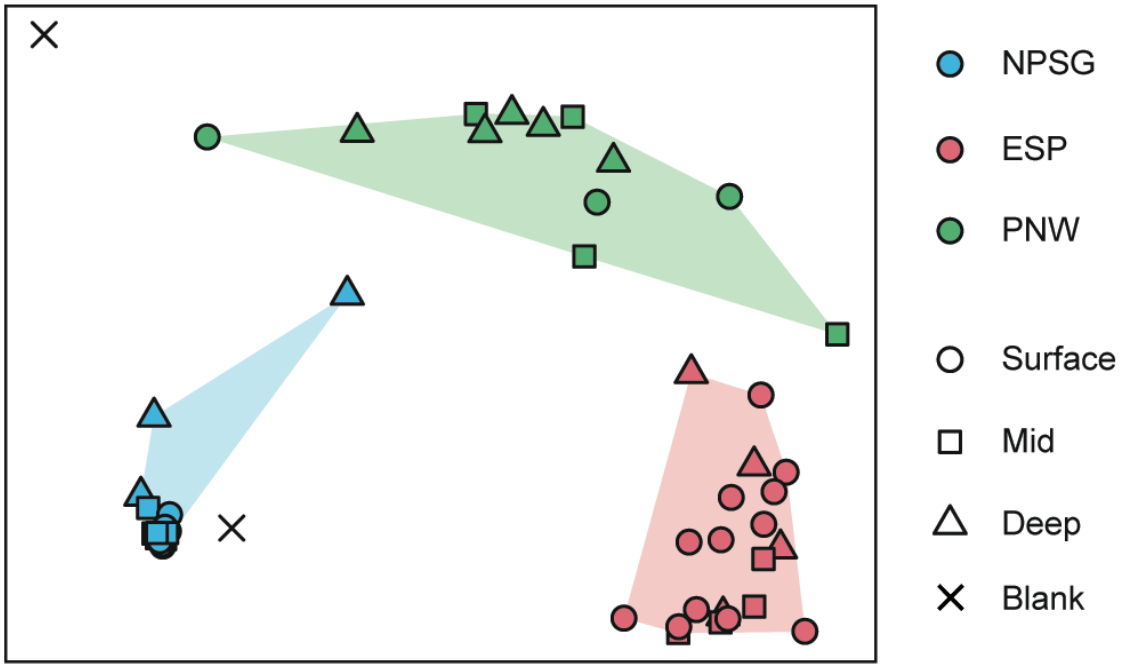

Figure S2.2: Non-metric multi-dimensional scaling (NMDS) plot derived from the Bray–Curtis dissimilarity matrix of *nifH* OTUs, constructed from dataset subsampled down to 634 sequences per sample in order to include filter blank samples. Each point represents an individual sample. Definition of depth classes vary by cruise: 'Surface' indicates 25m (NPSG), above the deep chlorophyll maxima (DCM, ESP), or the oxygen maxima (PNW), 'Mid' indicates 75m (NPSG), ~DCM (ESP), or oxycline (PNW), and 'Deep' indicates 200m (NPSG), below the DCM (ESP), or oxygen minima (PNW).

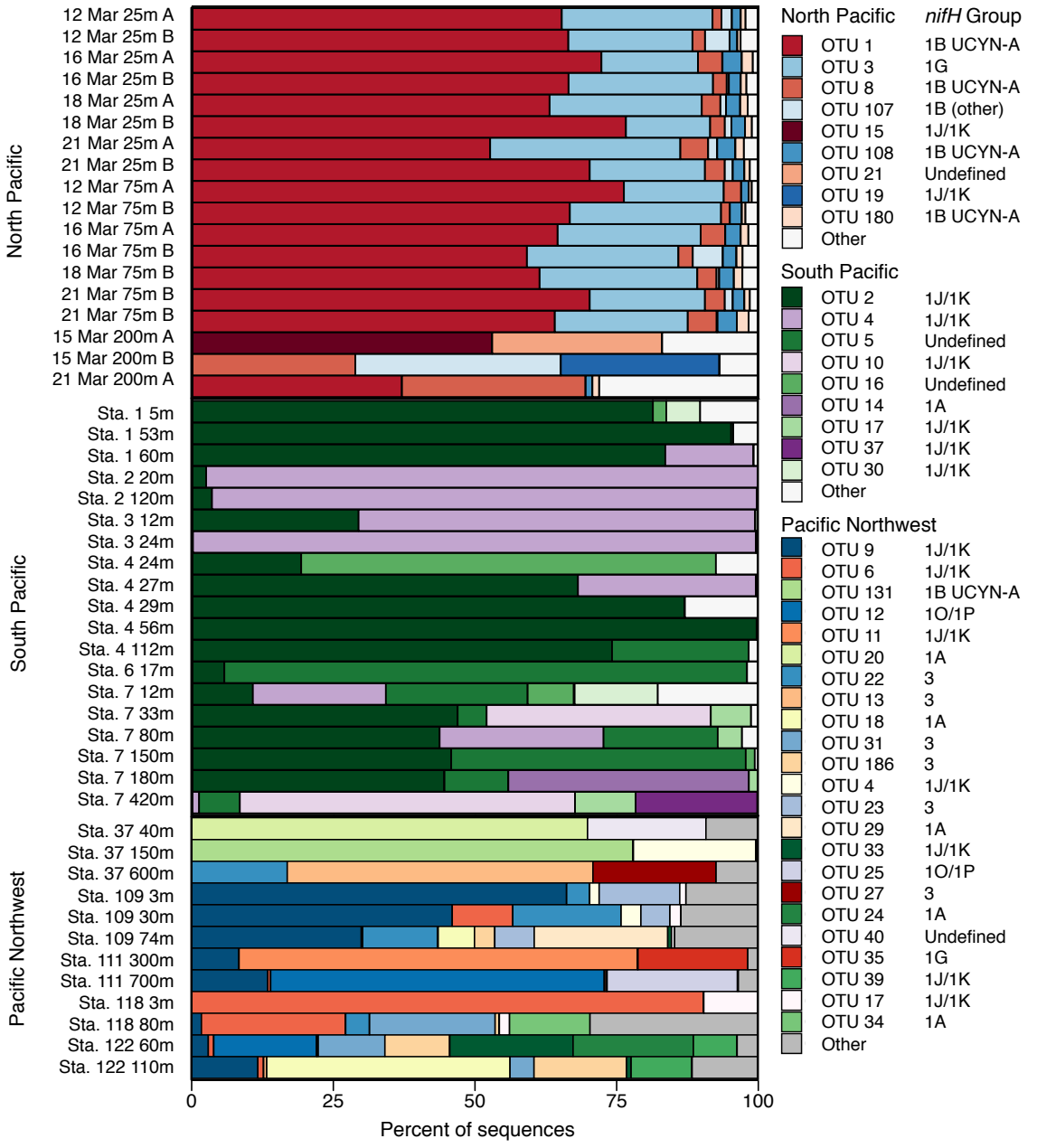

Figure S2.3: Relative abundances of *nifH* OTUs from each sample, based on partial *nifH* sequence dataset. OTUs representing >1% of total sequences from each cruise are shown; all other OTUs are bracketed into "Other." Taxonomical assignment of each OTU into *nifH* gene cluster groups is shown to the right.

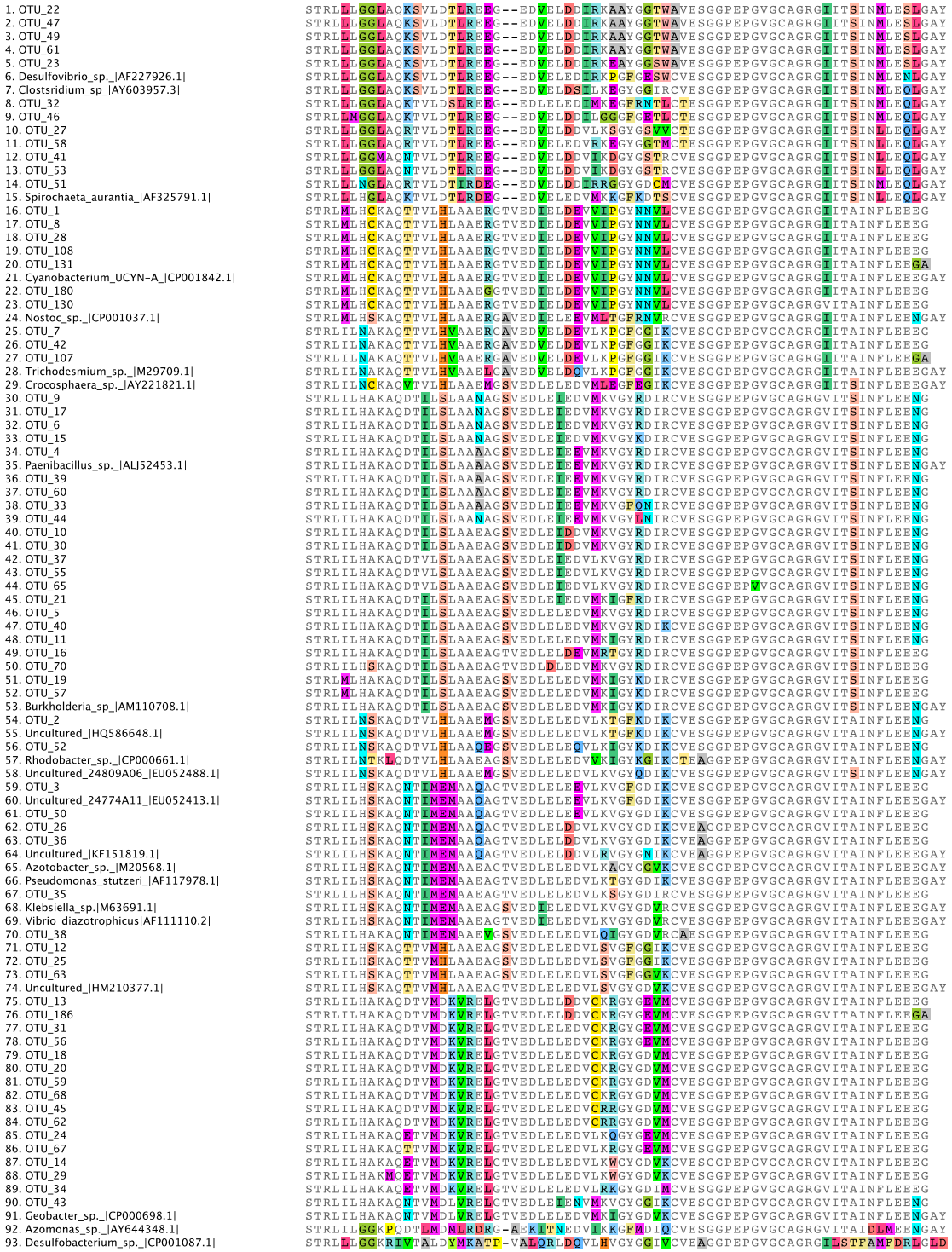

Figure S2.4: Alignment of the 93 sequences displayed in Fig. 2.3. Sequences were aligned in Geneious (v9.0.3) using a ClustalW alignment with the BLOSUM cost matrix and default gap penalties with free end gaps. Only the overlapping regions of reference sequences (72 amino acids) are shown. Colors show disagreements with the consensus.

Table 2.1: Regional summary and environmental conditions of sampling stations on North Pacific Subtropical Gyre (NPSG), Eastern South Pacific (ESP), and Pacific Northwest (PNW) cruises. Sea surface temperature (SST, °C) and chlorophyll fluorescence (Chl, μg L^{-1}) were measured using conductivity-temperature-depth sensors (excluding PNW Chl, which were measured fluorometrically). Nitrate + nitrite (N+N, μmol kg^{-1}) and phosphate (PO_4, μmol kg^{-1}) concentrations were measured with discrete bottle samples taken at ~5 m depth. ND indicates no data available.

Region	Summary	Station	Date	Location	SST	N+N	PO_4	Chl
NPSG	Occupation at Stn. ALOHA in the NPSG. N_2 fixation rates and *nifH* diversity assessed at 25m, 75m, and 200m.	ALOHA	3/12/14	22°46.1'N, 157°59.7'W	24.2	ND	ND	0.16
		ALOHA	3/15/14	22°46.8'N, 158°5.1'W	24.3	0.01	0.09	0.17
		ALOHA	3/17/14	22°44.5'N, 158°5.9'W	23.9	0.02	0.08	0.28
		ALOHA	3/18/14	22°46.5'N, 158°4.9'W	23.8	ND	ND	0.25
		ALOHA	3/19/14	22°46.7'N, 158°4.9'W	23.9	0.01	0.11	0.23
		ALOHA	3/21/14	22°46.7'N, 158°5.0'W	23.8	0.02	0.11	0.21
ESP	Transit from the nutrient-rich, productive Chilean upwelling system (Stn. 1) to the low-nutrient, low-chlorophyll gyre (Stn. 7). N_2 fixation measured at 5m only; *nifH* diversity measured from 5-420m.	1	11/21/10	20° 4.8'S, 70°48.0'W	18.9	0.44	0.88	0.83
		2	11/25/10	21°10.8'S, 76°57.0'W	17.9	0.25	0.48	0.53
		3	11/28/10	22°15.6'S, 82°21.0'W	17.9	0.23	0.45	0.20
		4	12/1/10	23°27.6' S, 88°45.6'W	18.7	ND	ND	0.24
		5	12/5/10	24°33.6'S, 94°43.2'W	19.7	ND	ND	0.11
		6	12/7/10	25°33.0'S, 100°8.4'W	20.7	ND	ND	0.06
		7	12/9/10	26°15.0'S, 103°57.6'W	21.5	0.27	0.21	0.07
PNW	Coastal cruise sampling continental slope and shelf stations in the productive California Current upwelling system. N_2 fixation (Stn. 37, 109, 118 only) and *nifH* diversity assessed from 3-700m at depths targeting surface, oxycline, and oxygen minima.	37	8/9/13	46°7.2'N, -124°54.6'W	15.8	0.07	0.31	0.19
		109	8/23/13	41°58.2'N, -124°24.0'W	10.8	16.86	1.39	ND
		111	8/24/13	40°6.0'N, -124°42.6'W	11.9	9.23	1.00	ND
		118	8/27/13	37°52.2'N, -123°3.6'W	14.8	0.50	0.32	ND
		122	8/28/13	36°43.8'N, -121°58.2'W	14.9	0.06	0.36	ND

Table 2.2: Cruise-averaged sensitivity analysis for N_2 fixation rate (NFR) measurements[a]

Cruise	Parameter (X)	Value	SD	δNFR/δX	Error contribution (SD × [δNFR / δX])^2	% Total error	Summary
ESP	Δt	1.00	1.94×10^{-17}	-1.17×10^{-1}	1.76×10^{-33}	0.00	
	A_{N2}	4.35%	5.10×10^{-3}	-4.21×10^{2}	7.58×10^{-2}	27.3	Mean.....1.81 nmol N L^{-1} d^{-1}
	A_{PN0}	0.366%	3.24×10^{-6}	-3.22×10^{4}	3.59×10^{-2}	11.3	MQR.....0.40 nmol N L^{-1} d^{-1}
	A_{PNf}	0.371%	7.18×10^{-6}	3.22×10^{4}	2.61×10^{-2}	19.9	LOD.....0.47 nmol N L^{-1} d^{-1}
	$[PN]_f$	1.25×10^{3}	1.71×10^{2}	1.35×10^{-3}	5.60×10^{-2}	41.5	HET-MQ....9.7 × 10^4 cells L^{-1}
NPSG	Δt	1.06	5.58×10^{-3}	−1.23	5.75×10^{-5}	0.02	
	A_{N2}	2.57%	1.16×10^{-3}	−57.9	1.37×10^{-2}	5.31	Mean.....1.32 nmol N L^{-1} d^{-1}
	A_{PN0}	0.372%	1.68×10^{-5}	1.40×10^{4}	1.13×10^{-1}	60.5	MQR.....0.39 nmol N L^{-1} d^{-1}
	A_{PNf}	0.378%	1.28×10^{-5}	1.41×10^{4}	1.10×10^{-1}	30.6	LOD.....0.21 nmol N L^{-1} d^{-1}
	$[PN]_f$	3.24×10^{2}	23.1	2.88×10^{-3}	1.14×10^{-2}	3.65	HET-MQ....9.5 × 10^4 cells L^{-1}
PNW	Δt	1.05	8.02×10^{-3}	4.10×10^{-2}	3.55×10^{-5}	0.01	
	A_{N2}	3.26%	3.82×10^{-3}	1.18×10^{-2}	2.50×10^{-3}	3.68	Mean.....0 nmol N L^{-1} d^{-1}
	A_{PN0}	0.370%	1.89×10^{-6}	-7.44×10^{2}	3.50×10^{-2}	38.6	MQR.....0.27 nmol N L^{-1} d^{-1}
	A_{PNf}	0.370%	3.24×10^{-6}	7.44×10^{2}	8.81×10^{-2}	44.9	LOD.....1.09 nmol N L^{-1} d^{-1}
	$[PN]_f$	2.30×10^{3}	2.00×10^{2}	-7.10×10^{-6}	2.65×10^{-3}	12.8	HET-MQR..6.6 × 10^4 cells L^{-1}

[a]The contribution of each source of error to the total uncertainty is provided, as performed by Montoya et al. (1996) using standard methods for error propagation. A_{N2}, A_{PN0}, and A_{PNf} represent the initial atom % $^{15}N_2$ gas, initial atom % ^{15}N of PN, and final atom % ^{15}N of PN, respectively. $[PN]_f$ represents the final concentration of N (nmol N L^{-1}), and Δt represents the incubation time (days). The third and fourth columns represent the observed average and standard deviation (SD) for each measured parameter from duplicate or triplicate incubation bottles. The fifth column represents the partial derivative of the NFR with respect to each parameter, evaluated using the provided average and SD. The sixth and seventh columns represent the absolute and relative error associated with each parameter. The total uncertainty associated with each measurement is termed the "Minimal Quantifiable Rate" (MQR). An alternative limit of detection (LOD) was calculated by setting $[A_{PNf} - A_{PN0}]$ equal to 0.00146 atom %, as described by Montoya et al. (1996). HET-MQR represents a minimal heterotrophic diazotrophic cell concentration required to produce the MQR, calculated using the highest cell-specific NFR reported by Bentzon-Tilia et al. (2015) for the alpha-proteobacterial isolate *R. palustris* BAL398 (0.52 fmol C_2H_4 $cell^{-1}$ h^{-1}; ~4.1 fmol N $cell^{-1}$ d^{-1}). Note that all values here represent regional averages; see Table S2.3 for the full sensitivity analysis.

Table S2.1: Incubation conditions and ancillary measurements for all $^{15}N_2$ experiments. Temperature (T), oxygen (O_2) and chloropigment fluorescence (Chl) were measured using conductivity-temperature-depth sensors. ND indicates no data available.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S2.2: Dual-index barcoded, sample-specific forward and reverse primers (5' → 3') used in this study. Sample barcodes are shown in bold. Forward and reverse degenerate PCR *nifH* primers are indicated by nifH3 (TGYGAYCCNAARGCNGA) and nifH4 (ADNGCCATCATYTCNCC; note the misprint of this primer in the original manuscript by Zani et al (2000)). NNNN indicate Illumina linker regions: AATGATACGGCGACCACCGAGATCTACAC for forward primers and CAAGCAGAAGACGGCATACGAGAT for reverse primers. Blue text indicates the binding site for Illumina sequencing primers, which were designed to optimize melting temperature during sequencing, as in Kozich et al (2013).

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S2.3: Sensitivity analysis for all nitrogen fixation rate (NFR) measurements. The contribution of each source of error to the total uncertainty is provided, as performed by Montoya et al. (1996). Average and SD (columns 13, 14, 18, 19, 23, 24,28, 29, 33, and 34) represent the observed average and standard deviation for each measured parameter from duplicate or triplicate incubation bottles. δNFR/ δX (columns 15, 20, 25, 30, and 35) represent the partial derivative of the NFR with respect to each parameter, evaluated using the provided average and standard deviation. Error Contribution and % Total Error (columns 16, 17, 21, 22, 26, 27, 31, 32, 36, and 37) represent the absolute and relative error associated with each parameter. The total uncertainty associated with each measurement is termed the "Minimal Quantifiable Rate" (column 10) and calculated using standard propagation of error. An alternative limit of detection was performed by setting the change in $^{15}N_2$ atom % of PN equal to 0.00146 atom % (column 12). NFR units are nmol N L^{-1} d^{-1}.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S2.4: List of OTUs with >95% similarity to known *nifH* PCR contaminants.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S2.5: Quantifications of UCYN-A *nifH* gene copies (in NPSG samples) and α-HQ586648 *nifH* gene copies **(**in ESP samples) assessed via droplet digital PCR. Values represent the average from duplicate DNA samples, with standard deviations in parentheses, when available. Abundances are calculated assuming 100% DNA extraction efficiency and 10% DNA extraction efficiency; we have previously assessed our DNA extraction protocol to extract ~10% of total DNA. ND represents no copies detected.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S2.6: *In silico* qPCR mismatch tests for the five most abundant OTUs from the ESP (OTU 2, 4, 5, 10, and 16). Show are the percentages of sequences within each OTU which have 0 mismatches, ≤1 mismatch, and ≤2 mismatches with forward primers (F), reverse primers (R) and probes from existing Taqman style qPCR assays targeting heterotrophic marine nitrogen fixing bacteria. Grey text indicates primers/probes outside of the range of the partial *nifH* DNA for most sequences generated in our study.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S2.7: Results from linear regression models predicting Shannon diversity (via partial *nifH* sequences clustered into 97% OTUs) as a function of N_2 fixation rate (NFR), temperature (Temp, °C), oxygen concentration (O_2, μmol kg^{-1}) and nitrate + nitrite concentration (N+N, μmol kg^{-1}).

See GradovilleMaryR2017_SupplementaryTables.xlsx

CHAPTER 3: MICROBIOME OF TRICHODESMIUM COLONIES FROM THE NORTH PACIFIC SUBTROPICAL GYRE

Mary R. Gradoville[a,†], Byron C. Crump[a], Ricardo M. Letelier[a], Matthew J. Church[b], and Angelicque E. White[a]

[a]College of Earth, Ocean and Atmospheric Sciences, 104 COAS Administration Building, Oregon State University, Corvallis, Oregon, USA

[b]Flathead Lake Biological Station, 32125 Bio Station Ln., Univeristy of Montana, Montana, USA

In review, *Frontiers in Microbiology*

3. MICROBIOME OF TRICHODESMIUM COLONIES FROM THE NORTH PACIFIC SUBTROPICAL GYRE

3.1 Abstract

Filamentous diazotrophic Cyanobacteria of the genus *Trichodesmium*, often found in colonial form, provide an important source of new nitrogen to tropical and subtropical marine ecosystems. Colonies are composed of several clades of *Trichodesmium* in association with a diverse community of bacterial and eukaryotic epibionts. We used high-throughput 16S rRNA and *nifH* gene sequencing, carbon and dinitrogen (N_2) fixation assays, and metagenomics to describe the diversity and functional potential of the microbiome associated with *Trichodesmium* colonies collected from the North Pacific Subtropical Gyre (NPSG). The 16S rRNA and *nifH* gene sequences from hand-picked colonies were predominantly (>99%) from *Trichodesmium* Clade I (i.e., *T. thiebautii*), which is phylogenetically and ecologically distinct from the Clade III IMS101 isolate used in most laboratory studies. The bacterial epibiont communities were dominated by Bacteroidetes, Alphaproteobacteria, and Gammaproteobacteria, including several taxa with a known preference for surface attachment, and were relatively depleted in the unicellular Cyanobacteria and small photoheterotrophic bacteria that dominate NPSG surface waters. Sequencing the *nifH* gene (encoding a subcomponent of the nitrogenase enzyme) identified non-*Trichodesmium* diazotrophs that clustered predominantly among the Cluster III *nifH* sequence-types that includes putative anaerobic diazotrophs. *Trichodesmium* colonies may represent an important habitat for these Cluster III diazotrophs, which were relatively rare in the surrounding seawater. Sequence analyses

of *nifH* gene transcripts revealed several cyanobacterial groups, including heterocystous *Richelia*, associated with the colonies. Both the 16S rRNA and *nifH* datasets indicated strong differences between *Trichodesmium* epibionts and picoplankton in the surrounding seawater, and also between the epibionts inhabiting *Trichodesmium* puff and tuft colony morphologies. Metagenomic sequence analyses suggested that lineages typically associated with a copiotrophic lifestyle comprised a large fraction of colony-associated epibionts, in contrast to the streamlined genomes typical of bacterioplankton in these oligotrophic waters. Epibiont metagenomes were enriched in genes involved in phosphate and iron acquisition and denitrification pathways relative to surface seawater metagenomes. We propose that the unique microbial consortium inhabiting colonies has a significant impact on the biogeochemical functioning of *Trichodesmium* colonies in pelagic environments.

3.2 Introduction

The filamentous, dinitrogen (N_2)-fixing (diazotrophic) cyanobacterium *Trichodesmium* provides a major source of bioavailable nitrogen (N) to the oligotrophic subtropical and tropical oceans (Karl et al., 2002; Capone et al., 2005). *Trichodesmium* abundances and N_2 fixation rates have been integral components of global N_2 fixation estimates and models (e.g. Coles et al., 2004; Mahaffey et al., 2005); thus, an accurate understanding of the physiology and ecology of this genus is crucial. Most *Trichodesmium* laboratory studies have used a single isolate, *T. erythraeum* strain IMS101, grown in culture with minimal heterotrophic bacteria. In contrast, natural *Trichodesmium* populations are composed of species from four phylogenetically distinct

clades (Hynes et al., 2012), which can vary in physiological traits such as carbon (C) affinity and phosphonate biosynthesis (Dyhrman et al., 2009; Hutchins et al., 2013). Furthermore, in nature they are commonly found associated with attached microorganisms (Borstad and Borstad, 1977). The diversity of this complex community likely affects the overall functioning of colonies (Gradoville et al., 2014), yet few studies have examined the ecology of *Trichodesmium* species and associated epibionts (although see Hmelo et al., 2012; Rouco et al., 2016).

Trichodesmium cells exist as free filaments or aggregate colonies (Letelier and Karl, 1996) with varying morphologies, namely spherical "puffs" and fusiform "tufts." These colonies have been reported to maintain an active and diverse assemblage of attached organisms, including bacteria, eukaryotic phytoplankton, protozoa, fungi, and copepods (Borstad and Borstad, 1977; Sheridan et al., 2002). *Trichodesmium* colonies constitute a favorable environment for associated epibionts by providing buoyancy (Walsby, 1992), elevated concentrations of dissolved organic N (Capone et al., 1994), and a substrate for attachment (O'Neil, 1998). Recent studies using 16S rRNA gene sequencing have shown that *Trichodesmium*-associated bacterial epibionts include surface-associated taxa (Hmelo et al., 2012) and that selective processes appear to drive epibiont community structure (Rouco et al., 2016).

Less is known about how associated microorganisms affect the functioning of the *Trichodesmium* holobiont. *Trichodesmium* colonies appear to be hotspots for microbial activity: hydrolytic enzymes activities are elevated within colonies (Stihl et al., 2001; Sheridan et al., 2002) and a metatranscriptome from *Trichodesmium* bloom material recovered more transcripts from associated organisms than from *Trichodesmium* cells

(Hewson et al., 2009). Microbial processes carried out by associated microorganisms have the potential to influence rates of N_2 or C fixation. For instance, quorum sensing by associated bacteria can increase alkaline phosphatase activity within colonies (Van Mooy et al., 2011), which could stimulate *Trichodesmium* dissolved organic phosphorus utilization, thereby increasing N_2 fixation rates when phosphate is limiting. Likewise, specific epibiont bacteria may secrete siderophores, chelating iron which could subsequently become bioavailable to *Trichodesmium* after photodegradation (Roe et al., 2012). Associated microorganisms may also directly contribute to the fixation of C and/or N_2. Phototrophs including filamentous Cyanobacteria (Siddiqui et al., 1992) and diatoms (Borstad and Borstad 1977) have historically been observed within *Trichodesmium* colonies. More recently, heterocystous cyanobacterial diazotrophs have been observed within *Trichodesmium* colonies (Momper et al., 2015) and *nifH* genes (encoding a subcomponent of the nitrogenase enzyme) phylogenetically clustering among facultative anaerobes and aerobic heterotrophic bacteria have been retrieved from *Trichodesmium* colonies (Gradoville et al., 2014). The degree to which these associated diazotrophs contribute to bulk colony N_2 fixation rates is unknown.

Here, we examine the microbiome associated with *Trichodesmium* colonies collected from the North Pacific Subtropical Gyre (NPSG). We used a combined approach of high-throughput 16S rRNA and *nifH* gene sequencing, metagenomics, and ^{13}C and $^{15}N_2$ fixation assays to survey the diversity of the *Trichodesmium* holobiont, test for the presence and activity of non-*Trichodesmium* colony-associated diazotrophs, and explore the functional potential of the colonies. We compare the colony-associated microbiome to the microbial community structure and metagenomic composition of

surrounding seawater, revealing diverse and unique microbial structure and functional potential associated with *Trichodesmium* colonies.

3.3 Methods

3.3.1 Sample collection

Samples were collected in March 2014 aboard the R/V *Kilo Moana* at Stn. ALOHA (A Long-term Oligotrophic Habitat Assessment; 22.45°N, 158°W), an open-ocean field site ~100 km north of Oahu (Table 3.1)). *Trichodesmium* colonies were collected using a 202 μm plankton net which was hand-towed at <2 km h^{-1} through near-surface waters (<10 m depth) for 10–15 min. Once recovered, colonies were isolated using an inoculating loop and rinsed twice with 0.2 μm-filtered surface seawater prior to all analyses. Colonies were sorted into morphological classes of spherical 'puffs' (further divided into 'radial puffs' and 'non-radial puffs' on 23 Mar), fusiform 'tufts,' and 'mixed' morphologies (Fig. 3.1), and filtered for subsequent extraction of DNA and RNA, or used for C and N_2 fixation measurements. Additionally, bulk seawater from 25 m depth was collected for comparison with *Trichodesmium* colony DNA. Seawater samples were collected using sampling bottles attached to a CTD (conductivity, temperature, depth) rosette, and subsampled into 4L acid-washed, MilliQ-rinsed polycarbonate bottles prior to filtration.

3.3.2 Carbon and nitrogen fixation rates

C and N_2 fixation rates were measured using the ^{13}C method of Legendre and Gosselin (1997) and a modification of the $^{15}N_2$ uptake method of Montoya et al. (1996) to

avoid delayed bubble dissolution (Mohr et al., 2010; Wilson et al., 2012). $^{15}N_2$ was added to incubations via $^{15}N_2$-enriched seawater, which was prepared onshore ~1 week prior to departure using Stn. ALOHA surface seawater according to the methods of Wilson et al. (2012). Briefly, seawater was 0.2 μm-filtered and degassed, then dispensed into gas-tight 3 L PTFE bags (Welch Fluorocarbon); 12.9 mL^{-L} $^{15}N_2$ gas (Cambridge Isotopes, 99%) was injected into the bag, which was manually agitated to facilitate dissolution. This $^{15}N_2$-enriched seawater was dispensed into glass serum bottles, which were capped, crimped, and stored at 4°C until use. The $^{15}N_2$ content of enriched seawater was validated via Membrane Inlet Mass Spectrometry according to the methods of Böttjer et al. (2016).

For the incubations, 20-30 colonies were transferred into 37 mL glass serum bottles filled with 0.2 μm-filtered surface seawater. Samples were spiked with 4 mL $^{15}N_2$-enriched seawater and 0.5 mL of 48 mmol L^{-1} ^{13}C bicarbonate stock, and bottles were topped off with filtered seawater, capped with Viton septa and aluminum caps, and crimp-sealed. Samples were incubated from dawn to dusk (~12 hr) in flow-through deckboard incubators with blue acrylic shading used to simulate ~60% of the sea-surface irradiance. Incubations were terminated by gentle filtration onto 25 mm diameter pre-combusted glass fiber filters (Whatman GF/F). Additionally, 20–30 colonies were preserved for $\delta^{15}N$ natural abundance (time-zero samples) after each net tow. Filters were flash-frozen and shipped to Oregon State University, where they were dried at 60°C overnight and packed into tin and silver capsules. Isotopic composition and masses of particulate N and C were measured with an isotope ratio mass spectrometer at Oregon State University. Fixation rates were calculated according to Montoya et al. (1996) and

normalized to particulate C concentrations; thus, N_2 fixation rates are expressed as nmol N μmol C^{-1} d^{-1} rather than nmol N_2 L^{-1} d^{-1}.

3.3.3 Nucleic acid extraction, amplification, and sequencing

For samples used for subsequent extraction of DNA and RNA, 20–30 *Trichodesmium* colonies of each morphotype were transferred into filtered seawater and gently filtered onto 25 mm diameter, 0.2 μm polyethersulfone Supor filters (Pall Corporation). Samples for subsequent extraction of planktonic DNA from 25 m seawater were filtered onto 0.2 μm Supor filters using a peristaltic pump. Filters were placed into empty microcentrifuge tubes (DNA) or microcentrifuge tubes containing 0.5 mL RNAlater (RNA), flash-frozen, transported in liquid nitrogen to Oregon State University, and stored at –80°C until analysis. DNA was extracted using the DNeasy Plant MiniKit (Qiagen), with a modified protocol to include a freeze-fracture step and Proteinase K treatment. RNA was extracted using the RNeasy MiniKit (Qiagen) according to manufacturer instructions, with additional steps for cell disruption through flash-freezing and bead-beating filters in mixtures of 500 µL RLT buffer, 5 µL β-mercaptoethanol, and 200 µL of mixed 0.1 mm and 0.5 mm glass beads (Biospec products). Possible carry-forward DNA contamination was minimized from RNA extracts by using the Turbo DNA-free kit (Ambion), and extracts were quantified using a Qubit RNA HS Assay kit (Invitrogen). Complimentary DNA (cDNA) was synthesized using the SuperScript III First-Strand kit (Invitrogen) according to the manufacturer's instructions, using the nifH3 gene-specific primer (Zani et al., 2000). DNA and cDNA were quantified with the Quant-

iT PicoGreen dsDNA Assay Kit (Invitrogen) using a MicroMax 384 plate reading fluorometer, and extracts were stored at –20°C or –80°C.

The polymerase chain reaction (PCR) was used to amplify a portion of the 16S rRNA gene (DNA samples only), targeting the entire prokaryotic community, and the *nifH* gene, targeting diazotrophs (for both DNA and cDNA). All PCR reactions were performed using a Veriti (Applied Biosystems) or DNAEngine (Bio-Rad) thermocycler and 10 or 20 μL reaction volumes. 16S PCR consisted of 1X HotMasterMix (5 PRIME), 1 μL DNA extract, and 5 pmol 515f (GTGCCAGCMGCCGCGGTAA) and 806r (GGACTACHVGGGTWTCTAAT) primers (Caporaso et al., 2010) which were modified to include Illumina adapters and dual-index barcodes as described by Kozich et al. (2013). Thermal cycling conditions for 16S gene amplifications were: 94°C for 3 min, followed by 30 cycles of 94°C for 45 s, 50°C for 60 s, and 72°C for 90 s, with a final 72°C extension for 10 min.

The *nifH* gene was amplified using nested degenerate *nifH* primers (Zani et al. 2000, Zehr and McReynolds 1989). The first round contained 1X PCR buffer, 0.1U Platinum High Fidelity *Taq* polymerase (Invitrogen), 200 μmol L^{-1} dNTPs, 3% BSA, 2.5 mmol L^{-1} Mg^{2+}, 1 μL DNA or cDNA, and 1 μmol L^{-1} nifH1 and nifH2 primers (Zehr and McReynolds 1989). Reaction conditions were: 94°C for 7 min, followed by 30 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min, and a final 72°C extension for 7 min. The second round of *nifH* PCR used the same components and thermocycling conditions as the first round, except the DNA extract was replaced with 1 μL of the amplified product generated during the first round PCR reaction, and custom primers were used, consisting of gene-specific sites (nifH3 and nifH4), dual-indexed barcodes,

Illumina linkers, and a sequencing primer binding region, similar to those described by Kozich et al. (2013; Table S3.1). PCR negative controls and filter blank samples were included in PCR reactions.

Triplicate PCR reactions were visualized by gel electrophoresis, then pooled and quantified as above. Samples were only sequenced if they had three successful PCR reactions, except for PCR negative controls and filter blanks, which were sequenced despite the absence of visual gel bands after amplification. 16S and *nifH* gene amplicons were pooled to equimolar concentrations, cleaned using both the UltraClean PCR (MoBio) and AMPure XP Bead cleanup kits, and sequenced at Oregon State University using MiSeq Standard v.3, 2×300 bp paired-end sequencing.

Metagenomes were constructed from two *Trichodesmium* puff DNA samples (Fig. 3.1). Libraries were constructed using an Illumina Nextera XT library prep kit, and cleaned using the AMPure XP Bead cleanup kits. Samples were sequenced on an Illumina MiSeq using a v.3 MiSeq Reagent Kit and a 2×300 bp paired-end protocol. Metagenome library preparation, cleaning, and sequencing were carried out by the Oregon State University Center for Genome Research and Biocomputing Center.

3.3.4 Bioinformatic analyses

Sequence reads from 16S amplicons, *nifH* amplicons, and metagenomes were demultiplexed using the Illumina MiSeq Reporter (MSR) version 2.5.1. For 16S rRNA gene sequences, primers were removed using MSR. The majority of 16S rRNA gene paired-end reads were merged and screened for quality, retaining sequences between 245-254 bp with no ambiguities using mothur (Schloss et al 2009). For a subset of 16S rRNA

gene samples, only forward reads were used for phylogenetic analyses due to the poor quality of reverse reads. The reverse primer was trimmed from forward reads, and reads with ambiguities, homopolymers (>8 bp) or poor quality (average score <25 or any score <20) were removed using mothur. Finally, forward reads with lengths between 245-254 bp were retained and combined with the paired-end 16S sequences for subsequent analyses. Singletons were removed, operational taxonomic units (OTUs) were clustered at 97% nucleotide sequence similarity, and a chimera check was performed with the Gold ChimeraSlayer reference database using USearch (Edgar 2010). Taxonomy was assigned in QIIME using the Silva v123 reference database, and sequences classified as chloroplasts, mitochondria, Archaea, Eukaryota, or an unknown domain were removed. Sequences were subsampled to 7,011 sequences per sample, resulting in near-saturation for most rarefaction curves (Fig. S3.1). Nonmetric multidimensional scaling analyses (NMDS, via Bray-Curtis similarity) and alpha diversity metric calculations were performed using QIIME (Caporaso et al. 2010). This same procedure was performed on reduced datasets containing *Trichodesmium* OTUs only (excluding 25 m seawater samples; rarefied to 2,248 sequences per sample) and containing all non-*Trichodesmium* OTUs (rarefied to 3,128 sequences per sample).

For *nifH* amplicons, though both forward and reverse barcodes were used for demultiplexing, only forward reads were used for phylogenetic analyses due to the poor quality of reverse reads. Reads with ambiguities, poor quality, or homopolymers were discarded. Forward primers were removed, sequences were trimmed to 244 bp, and OTUs were clustered at 97% nucleotide sequence similarity using USearch with a *de novo* chimera checker (Edgar 2010). OTUs containing chimeras, frameshifts, and non-

nifH sequences were removed. Sequences were subsampled to 9,651 sequences per sample, saturating most rarefaction curves (Fig. S3.1). The *nifH* OTUs were translated and phylogenetically classified into *nifH* gene clusters (Zehr et al. 2003b) via BLAST-p similarity to a reference database of *nifH* gene sequences (http://www.jzehrlab.com/#!nifh-database/c1coj). Sequences were termed "undefined" if they had equal amino acid similarity to sequences from multiple *nifH* gene sequence-types. BLASTn searches of the National Center for Biotechnology Information (https://blast.ncbi.nlm.nih.gov) were also performed for select *nifH* and 16S rRNA gene OTUs.

Metagenome sequences were demultiplexed using the Illumina MiSeq Reporter (MSR) version 2.5.1. All further processing steps were performed for the two *Trichodesmium* colony metagenomes and a metagenome previously constructed from Stn. ALOHA surface seawater DNA (15 m depth, 0.2 μm pore-size filter) on 30 July 2015 (Wilson et al., submitted; NCBI BioProject accession PRJNA358725 (BioSample S37C001). Raw reads were assembled separately for each sample using MEGAHIT (Li et al., 2015). Assemblies were uploaded to the Joint Genome Institute Genomes Online Database (https://gold.jgi.doe.gov/), where coding sequences (CDS) were predicted and annotated to the Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa and Goto, 2000; Huntemann et al., 2015). Metagenome sequences were processed according to the methods of Nalven (2016). Sequence reads were then trimmed for quality using seqtk (https://github.com/lh3/seqtk) and mapped back to CDS using Bowtie 2 (Langmead and Salzberg, 2012). Counts (one for single reads and two for paired reads mapped), CDS lengths, and alignment lengths were extracted using SAMtools (Li et al., 2009), and

counts were normalized to account for length of reads and length of CDS (Wagner et al. 2012). Counts within KEGG ortholog groups (KO) were summed and normalized as counts per million mapped to KO-annotated contigs (Genes Per Million [GPM], Wagner et al., 2012) and as counts per million mapped to KO-annotated contigs of known function (designated GPMK). GPM counts were used to analyze overall taxonomy, while GPMK were used for functional analyses. Counts from each KO were also divided into categories assigned to Cyanobacteria (assumed to be predominantly *Trichodesmium*) and non-Cyanobacteria. Details on the assembly and annotation of each sample are provided in Table 2.

All sequences are available from NCBI (accession SRP078449).

3.3.5 Statistical analyses

Two-way ANOVA with subsequent Tukey Honest Significant Difference (HSD) post-hoc tests were used to test the effect of day and sample type on N_2 fixation rates and alpha diversity metrics. The Welch Two Sample t-test was used to test for differences in the relative proportion of puff and tuft sequences in dominant OTUs, using the Bonferroni correction for multiple comparisons. Both ANOVA and t-tests were performed using the program R (http://www.r-project.org/). Detection limits for N_2 fixation rate measurements were calculated using standard propagation of errors via the observed variability between replicate samples as described by Gradoville et al. (*in press*) (Table S3.2).

3.4 Results

3.4.1 Carbon and nitrogen fixation rates

Shipboard incubation experiments showed that *Trichodesmium* colonies were actively fixing N_2 and C. $^{15}N_2$ fixation rates ranged from 0.24 to 4.16 nmol N μmol C^{-1} d^{-1} (Table 3.1)); all rates exceeded detection limits (Table S3.2). ^{13}C fixation rates ranged from 173 to 243 nmol C μmol C^{-1} d^{-1} (Table 3.1)). Both $^{15}N_2$ and ^{13}C rates were normalized to C content rather than colony number due to the known variability in the size of *Trichodesmium* colonies (Letelier and Karl, 1996). These ranges are similar to previously reported *Trichodesmium* colony-specific (Lomas et al., 2012) and C-specific (Gradoville et al., 2014) N_2 and C fixation rates. N_2 fixation rates varied by day of sampling (two-way ANOVA, $p < 0.01$) but not by morphology ($p > 0.05$). C fixation rates did not vary by either day or morphology (two-way ANOVA, $p > 0.05$).

3.4.2 Trichodesmium species diversity

The *Trichodesmium* species diversity within our samples was assessed via PCR amplification and sequencing of 16S rRNA and *nifH* genes. Sequences from both genes indicate that *Trichodesmium* Clade I (e.g., *T. thiebautii*) dominated our samples, with Clade III (e.g., *T. erythraeum*) representing <1% of *Trichodesmium* sequences (Fig. 3.2). The 16S rRNA gene dataset contained 2 OTUs classified as *Trichodesmium*, with the most abundant OTU classified as Clade I (16S OTU 2, 99.5% of *Trichodesmium* 16S rRNA sequences). Likewise, 2 of the 3 *Trichodesmium nifH* gene OTUs (*nifH* OTU 1 and *nifH* OTU 27) were classified as Clade I and together comprised 99.9% of the *Trichodesmium nifH* gene sequences (Fig. 3.2). Puff, tuft, and mixed morphology

samples from 16S rRNA and *nifH* genes all contained >99% Clade I *Trichodesmium* sequences (Table S3.3), and the *Trichodesmium* community structure did not vary by morphology (Fig. 3.3C). No sequences from the 16S rRNA or *nifH* gene datasets were classified as *Trichodesmium* Clade II or Clade IV.

3.4.3 Microbial diversity via 16S rRNA gene amplicons

The microbial diversity of the *Trichodesmium* microbiome was assessed using high-throughput sequencing of partial 16S rRNA genes from 17 *Trichodesmium* colony samples and 4 surface seawater samples for comparison. *Trichodesmium* sequences represented 24–75% of rRNA amplicons from colony samples; the remaining 25–76% of sequences corresponded to associated bacteria, termed epibionts (though it is possible that a subset of these organisms were endobionts). The most abundant epibiotic taxa belonged to Bacteroidetes (Cytophagia, Sphingobacteriales, and Flavobacteriales), Alphaproteobacteria (predominantly Rhodobacteriales, Rhodospirillales, and Rhizobiales), and Gammaproteobacteria (e.g. *Marinicella sp.*, *Alteromonas sp.*, Oceanospirillales) (Fig. 3.4, S3.2, Table S3.3). Even at broad phylum- and class-level taxonomic groupings, the *Trichodesmium* epibiont community differed from the bacterial community in the surrounding seawater: all colonies were relatively enriched in Bacteroidetes, and puff and mixed colony samples were enriched in Acidobacteria and Deltaproteobacteria, compared to the surrounding seawater (Fig. 3.4). Additionally, some of the most abundant taxa in NPSG near-surface seawater samples, including the Cyanobacteria *Prochlorococcus sp.* and *Synechococcus sp.,* the Actinobacteria *Actionomarina sp.*, and marine groups AEGEAN-169, SAR11, SAR86, and SAR116,

were relatively depleted or absent in *Trichodesmium* colony samples (Fig. 3.5, S3.2). At the 97% identity level, *Trichodesmium* colonies and seawater samples had few dominant OTUs in common (Fig. S3.2). NMDS analyses provide further evidence that the community structure of the epibionts was distinct from that of the surrounding seawater and also illustrated greater dissimilarity among *Trichodesmium* samples than among surface seawater samples (Fig. 3.3).

The *Trichodesmium* epibiont community varied with colony morphology. *Trichodesmium* colonies with puff morphology (n = 8 samples) contained a smaller fraction of *Trichodesmium* sequences (24–51% *Trichodesmium* 16S rRNA), and thus a larger fraction of epibiont sequences, than tuft morphologies (n = 5 samples; 57–75% *Trichodesmium*16S rRNA) (Fig. 3.4). The epibiont communities of puff colonies contained a larger fraction of Bacteroidetes (including Cytophagia and Saprospiraceae) and Deltaproteobacteria (including Desulfuromonadales) than tuft colonies (Fig. 3.4, S3.2). Tuft colonies contained a larger fraction of non-*Trichodesmium* Cyanobacteria (predominantly *Limnothrix*) and Gammaproteobacteria (including Alteromonadaceae, Oleiphilaceae, and Piscirickettsiaceae) than puff colonies. There were differences between puff and tuft colony epibionts at the OTU level: over half of the most abundant *Trichodesmium* OTUs had significantly different relative abundances between the two morphotypes (Fig S3.2). NMDS analyses demonstrated that the overall epibiont community structure varied by morphology, with puff colonies clustering separately from tuft colonies (Fig. 3.3).

Alpha diversity metrics were calculated from 16S rRNA gene OTUs at 97% identity (Table 3.3). Both diversity (Shannon) and species richness (Chao1) varied by

sample type (i.e., seawater or morphology) and by day of sampling ($p < 0.05$, two-way ANOVA). Species richness did not vary among *Trichodesmium* morphologies (Tukey HSD $p > 0.05$), but all morphotypes had significantly lower (by a factor of ~2) species richness than surface seawater samples (Tukey HSD $p \leq 0.001$). Diversity was higher in *Trichodesmium* puff samples and mixed morphology samples than in tuft samples (Tukey HSD $p < 0.001$). *Trichodesmium* samples of all morphotypes had lower diversity than seawater samples (Tukey HSD $p < 0.05$); however, when excluding *Trichodesmium* OTUs, diversity in samples of all *Trichodesmium* morphotypes were not significantly different from seawater (Tukey HSD $p > 0.05$). Thus, the *Trichodesmium* epibiont community had lower species richness, but insignificant differences in evenness, compared to seawater.

3.4.4 Diazotroph diversity via nifH amplicons

We sequenced partial *nifH* genes and transcripts from *Trichodesmium* colonies, and from surface seawater DNA samples for comparison, to test for the presence and transcriptional activities of non-*Trichodesmium* diazotrophs associated with the colonies. While sequences belonging to *Trichodesmium* dominated the *nifH* genes, we also recovered non-*Trichodesmium nifH* genes and transcripts (Fig. 3.5). In the DNA samples, *Trichodesmium* represented 64–99% of *nifH* sequences, with an average of 7% of sequences corresponding to non-*Trichodesmium* diazotrophs. Most non-*Trichodesmium nifH* DNA sequences were classified as *nifH* Cluster III, a group that includes anaerobic microorganisms such as *Desulfovibrio* and *Clostridium* (Zehr et al., 2003). Non-*Trichodesmium* groups other than Cluster III represented 1.5% of *nifH* gene sequences,

and included previously identified *nifH* groups such as 1G (presumed Gammaproteobacteria), 1J/1K (presumed Alpha- and Betaproteobacteria), and a very small percentage of sequences belonging to the Cyanobacteria UCYN-A. The non-*Trichodesmium* diazotrophs associated with the colonies were distinct from diazotrophic taxa in the surrounding surface seawater, where *nifH* gene sequences were dominated by UCYN-A and presumed Gammaproteobacteria and contained <0.01% *nifH* Cluster III.

A much smaller fraction of the *nifH* transcript sequences belonged to non-*Trichodesmium* diazotrophs (Fig. 3.5). Sequences phylogenetically related to the *nifH* Cluster III, 1G, and 1J/1K, which constituted a modest proportion of *nifH* gene sequences, were conspicuously absent from the *nifH* transcript sequences. The small fraction of non-*Trichodesmium nifH* transcripts (0–3.5%) belonged to Cyanobacteria, predominantly cyanobacterium UCYN-A and *Richelia/Calothrix* (with the exception of one sample containing 0.01% of the 1J/1K Alpha- and Betaproteobacteria *nifH* transcripts).

The *nifH* DNA and RNA sequences show that *Trichodesmium* puff and tuft colonies harbored different communities of non-*Trichodesmium* diazotrophs (Fig. 3.5). Puff colonies harbored a larger fraction of Cluster III (average 9.6% of *nifH* gene sequences) than tuft colonies (average 2.1% of *nifH* DNA sequences), while tuft colonies harbored a larger fraction of 1G (presumed Gammaproteobacteria, average 2.9% of *nifH* gene sequences) than tufts (average 0.2% of *nifH* gene sequences). Additionally, *nifH* transcripts from puff and tuft colonies morphologies included different phylotypes of heterocystous cyanobacteria (*Richelia*/*Calothrix*). One puff RNA sample (collected 21 Mar) contained transcripts derived from the *Calothrix* SC01/HET-3 group (Foster and

Zehr, 2006; Foster et al., 2010). No tuft samples contained *Calothrix* SC01 sequences, but all 4 tuft RNA samples contained transcripts derived from *Richelia*/HET-1 group (Church et al., 2005b). Neither heterocystous phylotype matched qPCR primer sets developed by Momper et al. (2015) to target heterocystous cyanobiont hetDA (≥4 mismatches with forward primer for both phylotypes; reverse primer was out of our sequencing region).

3.4.5 Metagenomic taxonomy and functional potential

We sequenced metagenomes from two *Trichodesmium* puff samples collected on 23 Mar 2014 ('radial puff' and 'non-radial puff") and assembled and annotated these sequences along with sequences from a publically available Stn. ALOHA surface seawater metagenome collected in July 2015. Colony metagenomes were dominated by bacteria (>99% of total counts), with ~70% of counts assigned to Cyanobacteria (Table 3.4). Cyanobacteria accounted for 64% and 77% of *rpoB* genes from radial and non-radial puff colonies, respectively. Thus, assuming that all genomes contain one copy of *rpoB* (Mollet et al., 1997) and that the majority of cyanobacterial counts are *Trichodesmium*, both total metagenome and *rpoB* gene counts produce conservative estimates of ~1 epibiont cell for every 2-3 *Trichodesmium* cells within colonies. Less than 1% of *Trichodesmium* colony counts were assigned to Eukarya, Archaea, or viruses, compared to 3.6% of surface seawater counts (Table 3.4). Eukaryotes represented 0.5% and 0.3% of non-radial and radial puff colony counts, respectfully, with dominant groups including green algae (Streptophyta and Chlorophyta), chordates, heterotrophic flagellates (Choanoflagellida), arthropods, diatoms, ciliates, and fungi (Table 3.4, S3.4).

The relative abundances of bacterial taxa mirrored trends observed in the 16S rRNA gene dataset, with the majority of colony sequences belonging to Cyanobacteria (primarily *Trichodesmium*), Alphaproteobacteria, Bacteroidetes, and Gammaproteobacteria (Table 3.4, Fig. 3.4).

Metagenome counts were annotated to KO and normalized to GPM in order to compare the relative abundance of genes and pathways between samples. However, ~75% of assembled contigs from *Trichodesmium* colonies failed KO annotation, far exceeding the ~25% of failed contig annotations observed in the surface seawater sample (Table 3.2). Furthermore, of the sequences that were successfully mapped to annotated contigs, *Trichodesmium* samples contained a larger fraction of KO with unknown function than the surface seawater sample (Table 3.2). This resulted in smaller GPM values from *Trichodesmium* metagenomes than the surface seawater metagenome for most KEGG gene categories (Fig. S3.3). Hence, we chose to use a normalization of counts per million mapped to a KO of known function (GPMK) in order to compare the functional potential of *Trichodesmium* colonies and surface seawater.

The gene contents of the *Trichodesmium* colony samples were distinct from those observed in the near-surface seawater. Colonies contained ~40% fewer *rpoB* GPM than the seawater samples (both in Cyanobacteria and non-Cyanobacteria fractions, Table 3.2), suggesting larger average genome sizes for *Trichodesmium* and epibiont cells. Summing KOs from KEGG gene groups revealed broad functional differences between colonies and surface seawater (Fig. 3.6). Seawater samples were relatively enriched in KEGG groups including nucleotide and amino acid metabolism, transcription, translation,

and replication and repair, while the colony samples were relatively enriched in energy metabolism, metabolism of terpenoids and polyketides, and cell motility.

Trichodesmium colony and surface seawater metagenomes also differed in the abundances of specific genes and pathways involved in nutrient cycling (Fig. 3.7). Colonies were enriched in genes encoding alkaline phosphatase and transporters for phosphate, phosphonates, and Fe(II), but depleted in Fe(III) transporter genes, compared to seawater. There were similar abundances of phosphate starvation response and Fe complex (siderophore) transport genes in colonies and seawater; however, the majority of these genes in the colonies belonged to non-Cyanobacteria (epibionts), which only represented ~30% of total colony metagenome counts. Thus, phosphate starvation response and Fe complex transport genes were enriched in epibionts compared to the surrounding plankton.

Trichodesmium colonies were also enriched in N cycling genes. Compared to seawater, the colony metagenomes contained higher total N metabolism gene abundances (34% and 46% higher abundances in radial and non-radial puffs, respectively, Table S3.5), and were strongly enriched in genes involved in N transformation pathways (Fig. 3.7). Nitrogenase genes were ~2000X more abundant in colonies than seawater, and included a large fraction assigned to non-Cyanobacteria (11% and 20% of nitrogenase genes in radial and non-radial puff colonies, respectively). Assimilatory nitrate reduction genes were present in both colony and seawater samples, but were ~5X more abundant in colonies, where the majority of genes corresponded to Cyanobacteria. Dissimilatory nitrate reduction and denitrification genes were absent in seawater samples but present in both colony samples; genes in these pathways were nearly exclusively assigned to non-

Cyanobacteria (Fig. 3.7). Genes involved in nitrification pathways were not observed in colony or seawater metagenomes.

3.5 Discussion

Several decades of research have documented the presence of bacterial and eukaryotic epibionts inhabiting *Trichodesmium* colonies (Borstad and Borstad, 1977; Siddiqui et al., 1992; Rouco et al. 2016), but the taxonomic composition and functional potential of these associated communities are not well-understood. Here, we used a variety of molecular tools to probe the diversity of *Trichodesmium* and associated epibionts in colonies from the NPSG. We found that the colonies were dominated by a single clade of *Trichodesmium*, but harbored a diverse community of associated microorganisms. These microbial assemblages were distinct from the surrounding seawater, differed by colony morphology, and included bacteria with a known preference for surface attachment, as well as putative anaerobic diazotrophs. Colony metagenomes include genes and pathways not present in *Trichodesmium* genomes, including siderophore transport and denitrification genes, which likely affects the biogeochemical functioning of *Trichodesmium* colonies.

3.5.1 Trichodesmium species diversity

The abundance and distribution of *Trichodesmium* have been studied extensively, but most work has focused on *Trichodesmium* at the genus-level, using techniques including microscopy, video plankton recording, and satellite imaging (e.g. Dugdale, 1961; Subramaniam et al., 2001; Davis and McGillicuddy, 2006). In the laboratory,

Trichodesmium isolates have been phylogenetically classified into four major clades (based on the hetR and ITS genes), with the majority of isolates falling into Clade I (e.g. *T. thiebautii*) and Clade III (e.g. *T. erythraeum*) (Orcutt et al., 2002; Hynes et al., 2012) but the geographical distributions of these clades in field populations has only begun to be investigated. Our finding of Clade I dominance is in agreement with recent surveys in the N. Pacific, N. Atlantic, and S. Pacific (Hmelo et al., 2012; Gradoville et al., 2014; Rouco et al., 2014; Rouco et al., 2016), which all observed the majority of *Trichodesmium* sequences belonging to Clade I. However, most physiological studies of *Trichodesmium* use the cultivated Clade III laboratory isolate *T. erythraeum* IMS101. Isolates from Clade I and Clade III appear to respond differently to environmental stimuli; for example, elevating pCO_2 enhances rates of N_2 and C fixation by Clade III isolates IMS101 and GBRTRLI101 but not the Clade I isolate H9-4 (Hutchins et al., 2007; Hutchins et al., 2013). While more work is needed to resolve the spatial and temporal variability of *Trichodesmium* species biogeography, current evidence suggests that at a global scale *Trichodesmium* Clade I may be more abundant than Clade III. Hence, modeling studies using the response of isolate IMS101 to predict the pCO_2 response of natural *Trichodesmium* populations should be viewed with caution. In addition, our findings agree with previous reports that colony morphology is not an accurate proxy for *Trichodesmium* clade (Hynes et al., 2012), as both puff and tuft colony samples were composed of >99% Clade I (Fig. 3.2).

3.5.2 Diversity of associated microbiome

Our samples contained diverse bacterial and eukaryotic taxa associated with *Trichodesmium* colonies. While relative abundances derived from 16S rRNA gene and metagenomic sequences can be biased by taxa-specific gene copy numbers and genome sizes, respectively, the relatively large fraction of non-*Trichodesmium* sequences in both datasets suggest epibionts are numerically abundant within the colonies. Though colony metagenome sequences were dominated by bacteria, we also observed sequences from viruses, Archaea, and many eukaryotic taxa previously observed associated with *Trichodesmium* colonies (Borstad and Borstad, 1977; Sheridan et al., 2002). Bacterial species richness within colonies was ~10-fold higher than the richness previously assessed for Atlantic colonies using clone libraries (Hmelo et al. 2012), and ~half of the richness in surrounding seawater, reaffirming that colonies harbor a diverse epibiont community (Sheridan et al., 2002; Rouco et al., 2016).

In our study, the first to directly compare high-throughput 16S rRNA gene sequences from *Trichodesmium* colonies and the surrounding surface seawater, we found that epibiont communities were distinct from the surrounding bacterioplankton. The warm, oligotrophic waters of the NPSG are known to be dominated by the Cyanobacteria *Prochlorococcus* (Campbell et al., 1994) and photo- and chemoheterotrophs including SAR11 and Rhodobacteraceae (DeLong et al., 2006). Indeed, the most abundant taxa in our near-surface seawater samples were those clustering among *Prochlorococcus*, *Synechococcus* and the small photoheterotroph Actinomarina (SAR11 represented only 3.2% of seawater sequences, likely due to a known bias in the 16S rRNA gene primer set used, (Apprill et al., 2015)); however, these taxa were all conspicuously absent from *Trichodesmium* colony samples. The relative absence of typically oligotrophic bacteria

with streamlined genomes (e.g. *Prochloroccus*, *Actinomarina*, SAR11, and AEGEAN-169) in colonies could be due to elevated nutrient concentrations within colonies favoring copiotrophic taxa (Lauro et al., 2009; Giovannoni et al., 2014). Instead, colony epibionts were dominated by Bacteroidetes, Alphaproteobacteria, and Gammaproteobacteria, which is consistent with previous 16S rRNA gene surveys of microbial communities associated with *Trichodesmium* (Hmelo et al., 2012; Rouco et al., 2016). Several dominant epibiont taxa have been previously observed associated with marine particulates, including the Bacteroidetes classes Cytophagia and Flavobacteriia (DeLong et al., 1993; Crump et al., 1999; Bryant et al., 2016), Alteromonadales (Fontanez et al., 2015), and Planctomycetes (DeLong et al., 1993). Though epibiont communities had several abundant taxa in common with the surrounding seawater at the order-level (e.g. Rhodobacterales, Rhodospirillales, Oceanospirillales), there were few commonalities with surface seawater phylotypes at the 97% OTU-level. The distinct community structure and lower species richness of epibionts compared to surrounding bacterioplankton, and the commonalities between epibiont taxa from our samples and previous *Trichodesmium* studies (Hmelo et al., 2012; Rouco et al., 2016) together suggest that *Trichodesmium* colonies provide a niche favoring select bacterial taxa.

In addition, we observed distinct epibiont communities associated with puff and tuft colonies, in agreement with Rouco et al. (2016), as well as evidence that certain bacterial species may consistently associate with specific morphotypes. Tuft colonies contained a larger fraction of *Trichodesmium* 16S rRNA gene sequences than puff colonies, possibly due to less colonizable surface area in this morphotype, which likely drives the lower diversity values observed for tufts (Table 3.3). This finding contrasts

with the microscopic observations of Sheridan et al. (2002), who reported tuft colonies harboring higher bacterial densities than puff colonies. Furthermore, the epibiont composition differed between the two morphologies, both in terms of phyla-level taxonomy (e.g. puffs contained more Bacteroidetes and tufts contained more non-*Trichodesmium* Cyanobacteria, Fig. 3.4) and, even more strikingly, in the relative abundance of specific phylotypes (Fig. S3.2). For example, a phylotype clustering among the filamentous Cyanobacteria *Limnothrix* represented 11.7% of non-*Trichodesmium* tuft sequences but only 0.1% of non-*Trichodesmium* puff sequences. Filamentous Cyanobacteria have been observed in close association with *Trichodesmium* filaments from tuft colonies (e.g. Paerl et al., 1989a; Siddiqui et al., 1992; Hewson et al., 2009), and *Limnothrix*-like sequences represented 31% of 16S rRNA gene clone library sequences in tuft (but not puff) colonies from the N. Atlantic (Hmelo et al., 2012). Thus, this *Limnothrix* phylotype may be a common associate of *Trichodesmium* tufts. Likewise, *Microscilla* represented 7.7% of tufts but only 0.01% of non-*Trichodesmium* puff sequences, and this genus has been previously recovered from *Trichodesmium* tufts in the N. Pacific, N. Atlantic, and Caribbean Sea (Janson et al., 1999; Rouco et al., 2016). Puff colonies also contained abundant phylotypes relatively absent from tufts, including Alphaproteobacteria and Bacteroidetes phylotypes, the cyanobacterium *Rivularia*, and a *Marinicella* phylotype which shares 100% nucleotide sequence identity to a sequence previously recovered from *Trichodesmium* colonies (accession GU726121). It is remarkable that so many phylotypes had significantly different relative abundances between the two morphotypes (Fig. S3.2), and also that many of the most abundant genera from our samples have also been dominant in previous surveys of *Trichodesmium*

epibionts (Hmelo et al., 2012; Rouco et al., 2016). Since the species composition of *Trichodesmium* did not vary by colony morphology, physical (i.e. filament compactness, colonizable surface area) or chemical properties of puff and tuft colonies likely drive the observed differences in epibiont community structure.

3.5.3 Colony-associated diazotrophs

There have been several reports of cyanobacterial and heterotrophic diazotrophs associated with *Trichodesmium* colonies (Paerl et al., 1989b; Gradoville et al., 2014; Momper et al., 2015), but the community composition and metabolic activity of these organisms have been largely unexplored. Here, we used high-throughput sequencing of partial *nifH* genes and transcripts explore the diversity of colony-associated diazotrophs. We observed non-*Trichodesmium nifH* genes (including genes from putative heterotrophs) in all *Trichodesmium* DNA samples, representing 1–35% of the colony *nifH* sequences (Fig. 3.5).

The ecological importance of non-cyanobacterial marine diazotrophs is a current enigma in N_2 fixation research: non-cyanobacterial *nifH* genes have been recovered from numerous marine environments (Bombar et al., 2016), but rates of N_2 fixation in marine environments dominated by non-cyanobacterial diazotrophs are often low or undetectable (e.g. Knapp et al. 2016, Gradoville et al. in review). Here, we found robust evidence that *Trichodesmium* colonies comprise yet another habitat for these seemingly cosmopolitan organisms. The majority of our non-*Trichodesmium nifH* gene sequences phylogenetically grouped among Cluster III *nifH* genes, which includes diverse anaerobic microorganisms (Zehr et al., 2003). The possibility of anaerobic bacteria inhabiting

Trichodesmium colonies appears plausible since colonies have been reported to contain anoxic microzones (Paerl and Bebout, 1988); indeed, we also found denitrification and Fe(II) transporter genes enriched in colony metagenomes (see *Functional potential within Trichodesmium colonies*). However, *nifH* Cluster III contains diverse lineages (Zehr et al., 2003), and the physiology and ecology of these organisms are not well-understood. In our study, the three most abundant Cluster III OTUs each share <85% nucleotide identity with any cultured representative in the BLASTn database. One of these OTUs matches a qPCR primer/probe set designed by Church et al. (2005a) to quantify a specific group of Cluster III *nifH* sequence-types in the NPSG, while all three OTUs share >99% nucleotide identity with sequences previously obtained from *Trichodesmium* colonies at Stn. ALOHA (Gradoville et al., 2014). Such results suggest *Trichodesmium* colonies may selectively harbor members of the Cluster III *nifH* phylotypes, including organisms not currently captured by existing Cluster III qPCR primers and probes (Church et al. 2005).

It is interesting to note that both Cluster III and the 1J/1K (presumed Alpha- and Betaproteobacteria) group had higher relative abundances in our *Trichodesmium* colony samples than in the surrounding seawater, where *nifH* sequences were dominated by the unicellular cyanobacterium UCYN-A, the Gammaproteobacterial *nifH* group 1G, and other Cyanobacteria including *Trichodesmium* (Fig. 3.5). This suggests that *Trichodesmium* colonies may represent a niche for Cluster III and 1J/1K diazotrophs. It is possible that the relative enrichment of these groups in *Trichodesmium* colonies could reflect a preference for marine particulates—for example, Bryant et al. (2016) observed marine plastic particles to be enriched in *nifH* genes—rather than a unique property of the colonies themselves. Marine particles may be favorable environments for heterotrophic

diazotrophs (Bombar et al., 2016), especially putative anaerobic Cluster III taxa, which could inhabit anoxic microzones of particles (Benavides et al., 2015). Future research is needed to determine whether *Trichodesmium* colonies represent an important niche for *nifH* Cluster III diazotrophs in the NPSG and other oceanic regions.

Though non-cyanobacterial taxa were present in *Trichodesmium nifH* genes, the absence of non-cyanobacterial *nifH* transcripts from RNA samples suggests that these taxa were not actively fixing N_2 at the time of sampling (Fig. 3.5). However, the *Trichodesmium nifH* RNA samples included non-*Trichodesmium* Cyanobacterial transcripts, with the majority belonging to two *Calothrix*/*Richelia* OTUs. Neither of these OTUs matched qPCR primer sets designed by Momper et al. (2015) to quantify *Calothrix*-like heterocystous cyanobionts recently observed inhabiting *Trichodesmium* puff colonies from the NPSG. All tuft RNA samples contained *Calothrix*/*Richelia* sequences matching primer/probe sets for group HET-1 (Church et al., 2005b), while *Calothrix*/*Richelia* in one puff RNA sample match primer/probe sets for the SC01/HET-3 group (Foster and Zehr, 2006; Foster et al., 2010). This suggests that there may be morphotype-specific associations between heterocystous Cyanobacteria and *Trichodesmium* other than the cohabitation described by Momper et al. (2015). Our observation of *Calothrix*/*Richelia* sequences in *nifH* transcripts but not *nifH* genes likely results from poor amplification of this group by the *nifH* primers used (Turk-Kubo et al., 2015) and high cell-specific transcription rates. This group appears to be much more transcriptionally active than the non-cyanobacterial diazotrophs associated with colonies (observed in *nifH* genes but not *nifH* transcripts), a finding which agrees with previous

nifH gene expression surveys using bulk seawater from Stn. ALOHA (Church et al. 2005b).

3.5.4 Functional potential within Trichodesmium colonies

Our metagenomic data suggest that *Trichodesmium* epibionts may benefit from a colony-associated lifestyle and influence nutrient cycling within colonies. Epibionts appeared to possess larger average genome sizes than bulk plankton, suggesting non-streamlined genomes, consistent with the relative absence of oligotrophic taxa (*Prochlorococcus, Actinomarina*, etc.) observed in colonies. Furthermore, epibionts were depleted in genes encoding basic metabolic functions and cellular machinery needed for replication compared to seawater metagenomes, again consistent with a lack of streamlined genomes (Giovannoni et al., 2014). Instead, colony samples were enriched in genes involved in motility, which could be useful in a colony-associated lifestyle, and in metabolic pathways not present in the seawater metagenome (Fig. 3.6, 3.7, Table S3.5). Additionally, a large fraction of *Trichodesmium* colony contigs failed annotation. This could be due to the large fraction of non-coding DNA in the *Trichodesmium* genome (Walworth et al., 2015), but could also arise from a larger fraction of uncultivated microorganisms in *Trichodesmium* microbiome than in the surrounding seawater.

The NPSG is a chronically oligotrophic system, with production rates limited by the availability of N (Karl et al., 1997) and sometimes P (Karl et al., 1995). Since diazotrophs such as *Trichodesmium* circumvent N limitation through N_2 fixation, their growth and N_2 fixation rates are typically limited by the availability of P and/or Fe (as well as light and temperature, Luo et al., 2014). Hence, there is considerable interest in

understanding the mechanisms of P and Fe acquisition by *Trichodesmium*. We observed enriched alkaline phosphatase, phosphate transport, and phosphonate transport genes in colonies, which agrees with previous demonstrations of efficient organic phosphorus scavenging and utilization by *Trichodesmium* (Dyhrman et al., 2006). Furthermore, colony epibionts contained genes encoding the synthase for acyl homoserine lactoses (Table S3.5), quorum sensing molecules which have been shown to stimulate alkaline phosphatase activity by *Trichodesmium* cells in culture (Van Mooy et al., 2011). We also found that phosphate starvation response genes were enriched in epibionts, which could reflect P-limitation due to the release of inorganic and organic N compounds by *Trichodesmium* cells (Capone et al., 1994; Mulholland et al., 2004).

The genes involved in Fe transport also differed between colony and seawater metagenomes. Fe(II) transporters were enriched in colonies, consistent with previous observations of these genes in *Trichodesmium* isolates (Chappell and Webb, 2010), but were nearly absent in seawater. In well-oxygenated seawater, most Fe exists as Fe(III), hence our observation of epibionts enriched in Fe(II) transport genes suggests anoxic microzones within colonies (Paerl and Bebout, 1988) may result in reduction of Fe(III). Additionally, we found low abundances of Cyanobacterial siderophore transport genes, reflecting the inability of *Trichodesmium* to use highly chelated Fe sources (Chappell and Webb, 2010), but these genes were enriched in non-Cyanobacterial epibionts. Our observations of abundant P and Fe acquisition genes in *Trichodesmium* and epibionts could reflect competition for these resources in the colony community. However, we also found metagenomic evidence for previously described potential mutualisms, as epibionts

could facilitate *Trichodesmium* nutrient uptake through quorum sensing (Van Mooy et al., 2011) and siderophore production (Roe et al., 2012).

Finally, we observed the genetic capacity for denitrification within *Trichodesmium* colonies. Both colony samples contained all necessary genes for the denitrification and dissimilatory nitrate reduction pathways, while no genes from either pathway were observed in the seawater sample. Furthermore, 16S OTU 18, comprising 9.9% of non-*Trichodesmium* 16S rRNA gene sequences from tuft colonies, was classified as the denitrifier *Nisaea sp*. These results agree with Wyman et al. (2013), who reported *nosZ* amplicons isolated from *Trichodesmium* colonies in the Arabian Sea. It is possible that low-oxygen microzones within the colonies and nitrate supplied through diurnal migration (Walsby, 1978) could allow for active denitrification, producing a tight spatial coupling between N_2 fixation and denitrification within colonies and reducing apparent colony N_2 fixation rates.

3.6 Conclusions

Our multifaceted high-throughput sequencing approach enabled a detailed view of the *Trichodesmium* colony microbiome. While the species composition of *Trichodesmium* was dominated by a single clade and uniform in all of our samples, the community structure of bacterial epibionts differed between puff and tuft colony morphologies, suggesting that differences in biogeochemical rates among colony morphologies may be driven by processes carried out by the associated microbiome. Epibionts appear copiotrophic, with the genetic capacity to influence colony nutrient cycling. Additionally, we found that colonies contained active Cyanobacterial

diazotrophs, and presumed heterotrophic and anaerobic diazotrophs, suggesting that *Trichodesmium* colonies harbor a unique microbial community with the potential to influence rate processes classically attributed to *Trichodesmium spp.*

Support for this project was provided by the National Science foundation through the Center for Microbial Oceanography: Research and Education (C-MORE; EF0424599) and the Simons Collaboration on Ocean Processes and Ecology (SCOPE award ID 329108). We are grateful to Ed Delong and Frank Aylward for supplying Stn. ALOHA metagenome data.

3.7 References

Apprill, A., McNally, S., Parsons, R., Webe, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137.

Benavides, M., Moisander, P.H., Berthelot, H., Dittmar, T., Grosso, O., Bonnet, S. (2015). Mesopelagic N_2 fixation related to organic matter composition in the Solomon and Bismarck Seas (Southwest Pacific). *PloS one* 10: e0143775.

Bombar, D., Paerl, R.W., Riemann, L. (2016). Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. *Trends Microbiol.* 24, 916–927.

Borstad, G., Borstad, L. (1977). "The *Oscillatoria erythraea* (Cyanophyta) community of associates," in *Cooperative Investigations of the Caribbean and Adjacent Regions-II*, *FAO Fish. Rep., 200*, ed. HB Stewart, 51–57.

Böttjer, D., Dore, J.E., Karl, D.M., Letelier, R.M., Mahaffey, C., Wilson, S.T., et al. (2016). Temporal variability of nitrogen fixation and particulate nitrogen export at Station ALOHA. *Limnol. Oceanogr.* doi: 10.1002/lno.10386

Bryant, J.A., Clemente, T.M., Viviani, D.A., Fong, A.A., Thomas, K.A., Kemp, P., et al. (2016). Diversity and activity of communities inhabiting plastic debris in the North Pacific Gyre. *mSystems* 1: e00024–00016.

Campbell, L., Nolla, H., Vaulot, D. (1994). The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. *Limnol. Oceanogr.* 39, 954–961.

Capone, D.G., Burns, J.A., Montoya, J.P., Subramaniam, A., Mahaffey, C., Gunderson, T., et al. (2005). Nitrogen fixation by *Trichodesmium* spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Glob. Biogeochem. Cycles* 19, GB2024.

Capone, D.G., Ferrier, M.D., Carpenter, E.J. (1994). Amino acid cycling in colonies of the planktonic marine cyanobacterium *Trichodesmium thiebautii*. *Appl. Env. Microbiol.* 60, 3989–3995.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7, 335–336.

Chappell, P.D., Webb, E.A. (2010). A molecular assessment of the iron stress response in the two phylogenetic clades of *Trichodesmium*. *Env. Microbiol.* 12, 13–27.

Church, M.J., Jenkins, B.D., Karl, D.M., Zehr, J.P. (2005a). Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. *Aquat. Microb. Ecol.* 38, 3–14.

Church, M.J., Short, C.M., Jenkins, B.D., Karl, D.M., Zehr, J.P. (2005b). Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. *Appl. Env. Microbiol.* 71, 5362–5370.

Coles, V.J., Hood, R.R., Pascual, M., Capone, D.G. (2004). Modeling the impact of *Trichodesmium* and nitrogen fixation in the Atlantic Ocean. *J. Geophys. Res.* 109, C06007.

Crump, B.C., Armbrust, E.V., Baross, J.A. (1999). Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl. Env. Microbiol.* 65, 3192–3204.

Davis, C.S., McGillicuddy, D.J. (2006). Transatlantic abundance of the N_2-fixing colonial cyanobacterium *Trichodesmium*. *Science* 312, 1517–1520.

DeLong, E.F., Franks, D.G., Alldredge, A.L. (1993). Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* 38, 924–934.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.U., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311, 496–503.

Dugdale, R. (1961). Nitrogen fixation in the Sargasso Sea. *Deep Sea Res. Pt. II:* 7, 297–300.

Dyhrman, S., Chappell, P., Haley, S., Moffett, J., Orchard, E., Waterbury, J., *et al*. (2006). Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* 439, 68–71.

Dyhrman, S.T., Benitez-Nelson, C.R., Orchard, E.D., Haley, S.T., Pellechia, P.J. (2009). A microbial source of phosphonates in oligotrophic marine systems. *Nat. Geosci.* 2, 696–699.

Fontanez, K.M., Eppley, J.M., Samo, T.J., Karl, D.M., DeLong, E.F. (2015). Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Front. Microbiol.* 6, 469.

Foster, R.A., Goebel, N.L., Zehr, J.P. (2010). Isolation of *Calothrix rhizosoleniae* (cyanobacteria) strain SC01 from *Chaetoceros* (bacillariophyta) spp. diatoms of the Subtropical North Pacific ocean. *J. Phycol.* 46, 1028–1037.

Foster, R.A., Zehr, J.P. (2006). Characterization of diatom–cyanobacteria symbioses on the basis of *nifH*, *hetR* and 16S rRNA sequences. *Env. Microbiol.* 8, 1913–1925.

Giovannoni, S.J., Cameron Thrash, J., Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565.

Gradoville M.R., Bombar, D., Crump, B.C., Letelier, R.M., White, A.E., (2014). Diversity and activity of nitrogen-fixing communities across ocean basin. *Limnol. Oceanogr.* in press.

Gradoville, M.R., White, A.E., Böttjer, D., Church, M.J., Letelier, R.M. (2014). Diversity trumps acidification: Lack of evidence for carbon dioxide enhancement of *Trichodesmium* community nitrogen or carbon fixation at Station ALOHA. *Limnol. Oceanogr* 59, 645–659.

Hewson, I., Poretsky, R.S., Dyhrman, S.T., Zielinski, B., White, A.E., Tripp, H.J., et al. (2009). Microbial community gene expression within colonies of the diazotroph, *Trichodesmium*, from the Southwest Pacific Ocean. *ISME J.* 3, 1286–1300.

Hmelo, L.R., Van Mooy, B.A.S., Mincer, T.J. (2012). Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. *Aquat. Microb. Ecol.* 67, 1–14.

Huntemann, M., Ivanova, N.N., Mavromatis, K., Tripp, H.J., Paez-Espino, D., Palaniappan, K., et al. (2015). The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v. 4). *Stand. Genomic Sci.* 10, 86.

Hutchins, D, Fu, FX, Zhang, Y, Warner, M, Feng, Y, Portune, K, et al. (2007). CO_2 control of *Trichodesmium* N_2 fixation, photosynthesis, growth rates, and elemental ratios: Implications for past, present, and future ocean biogeochemistry. *Limnol. Oceanogr.* 52, 1293–1304.

Hutchins, D.A., Fu, F-X., Webb, E.A., Walworth, N., Tagliabue, A. (2013). Taxon-specific response of marine nitrogen fixers to elevated carbon dioxide concentrations. *Nat. Geosci.* 6, 790–795.

Hynes, A.M., Webb, E.A., Doney, S.C., Waterbury, J.B. (2012). Comparison of cultured *Trichodesmium* (Cyanophyceae) with species characterized from the field. *J. Phycol.* 48, 196–210.

Janson, S., Bergman, B., Carpenter, E.J., Giovannoni, S.J., Vergin, K. (1999). Genetic analysis of natural populations of the marine diazotrophic cyanobacterium *Trichodesmium*. *FEMS Microbiol. Ecol.* 30, 57–65.

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

Karl, D., Letelier, R., Hebel, D., Tupas, L., Dore, J., Christian, J., et al. (1995). Ecosystem changes in the North Pacific subtropical gyre attributed to the 1991–92 El Nino. *Nature* 373, 230–234.

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J., Hebel, D. (1997). The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* 388, 533–538.

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., et al. (2002). Dinitrogen fixation in the world's oceans. *Biogeochemistry* 57, 47–98.

Knapp, A.N., Casciotti, K.L., Berelson, W.M., Prokopenko, M.G., Capone, D.G. (2016). Low rates of nitrogen fixation in eastern tropical South Pacific surface waters. *PNAS* 113, 4398–4403.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Env. Microbiol.* 79, 5112–5120.

Langmead, B., Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods* 9, 357–359.

Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., et al. (2009). The genomic basis of trophic strategy in marine bacteria. *PNAS* 106, 15527–15533.

Legendre, L., Gosselin, M. (1997). Estimation of N or C uptake rates by phytoplankton using ^{15}N or ^{13}C: revisiting the usual computation formulae. *J. Plankton. Res.* 19, 263–271.

Letelier, R.M., Karl, D.M. (1996). Role of *Trichodesmium* spp. in the productivity of the subtropical North Pacific Ocean. *Mar. Ecol. Prog. Ser.* 133, 263–273.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Lomas, M., Hopkinson, B., Losh, J., Ryan, D., Shi, D., Xu, Y., et al. (2012). Effect of ocean acidification on cyanobacteria in the subtropical North Atlantic. *Aquat. Microb. Ecol.* 66, 211–222.

Lundgren, P., Janson, S., Jonasson, S., Singer, A., Bergman, B. (2005). Unveiling of novel radiations within *Trichodesmium* cluster by *hetR* gene sequence analysis. *Appl. Env. Microbiol.* 71, 190–196.

Luo, Y-W., Lima, I., Karl, D., Deutsch, C., Doney, S. (2014). Data-based assessment of environmental controls on global marine nitrogen fixation. *Biogeosciences* 11, 691–708.

Mahaffey, C., Michaels, A.F., Capone, D.G. (2005). The conundrum of marine N_2 fixation. *Am. J. Sci.* 305, 546–595.

Mohr, W., Großkopf, T., Wallace, D.W.R., LaRoche, J. (2010). Methodological underestimation of oceanic nitrogen fixation rates. *PloS one* 5, e12583.

Mollet, C., Drancourt, M., Raoult, D. (1997). *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol. Microbiol.* 26, 1005–1011.

Momper, L.M., Reese, B.K., Carvalho, G., Lee, P., Webb, E.A. (2015). A novel cohabitation between two diazotrophic cyanobacteria in the oligotrophic ocean. *ISME J.* 9, 882–893.

Montoya, J.P., Voss, M., Kahler, P., Capone, D.G. (1996). A simple, high-precision, high-sensitivity tracer assay for N_2 fixation. *Appl. Env. Microbiol.* 62, 986–993.

Mulholland, M.R., Bronk, D.A., Capone, D.G. (2004). Dinitrogen fixation and release of ammonium and dissolved organic nitrogen by *Trichodesmium* IMS101. *Aquat. Microb. Ecol.* 37, 85–94.

Nalven, S.G. (2016). Decoding DOM Degradation with metatranscriptomics: How do sunlight and microbial communities interact to degrade dissolved organic matter in Arctic freshwaters? [dissertation/master's thesis]. [Corvallis (OR)]: Oregon State University.

O'Neil, J.M. (1998). The colonial cyanobacterium *Trichodesmium* as a physical and nutritional substrate for the harpacticoid copepod *Macrosetella gracilis*. *J. Plankton. Res.* 20, 43–59.

Orcutt, K., Rasmussen, U., Webb, E.A., Waterbury, J.B., Gundersen, K., Bergman, B. (2002). Characterization of *Trichodesmium* spp. by genetic techniques. *Appl. Env. Microbiol.* 68, 2236–2245.

Paerl, H.W., Bebout, B.M. (1988). Direct measurement of O_2-depleted microzones in marine *Oscillatoria*: relation to N_2 fixation. *Science* 241, 442–445.

Paerl, H.W., Bebout, B.M., Prufert, L.E. (1989a). Bacterial associations with marine *Oscillatoria* sp. (*Trichodesmium* sp.) populations: Ecological implications. *J. Phycol.* 25, 773–784.

Paerl, H.W., Priscu, J.C., Brawner, D.L. (1989b). Immunochemical localization of nitrogenase in marine *Trichodesmium* aggregates: Relationship to N_2 fixation potential. *Appl. Env. Microbiol.* 55, 2965–2975.

Roe, K.L., Barbeau, K., Mann, E.L., Haygood, M.G. (2012). Acquisition of iron by *Trichodesmium* and associated bacteria in culture. *Env. Microbiol.* 14, 1681–1695.

Rouco, M., Haley, S.T., Dyhrman, S.T. (2016). Microbial diversity within the *Trichodesmium* holobiont. *Env. Microbiol.* 18, 5151–5160.

Rouco, M., Warren, H.J., McGillicuddy, D.J., Waterbury, J.B., Dyhrman, S.T. (2014). *Trichodesmium* sp. clade distributions in the western North Atlantic Ocean. *Limnol. Oceanogr.* 59, 1899–1909.

Sheridan, C., Steinberg, D., Kling, G. (2002). The microbial and metazoan community associated with colonies of *Trichodesmium* spp.: a quantitative survey. *J. Plankton. Res.* 24, 913–922.

Siddiqui, P., Bergman, B., Carpenter, E. (1992). Filamentous cyanobacterial associates of the marine planktonic cyanobacterium *Trichodesmium*. *Phycologia* 31, 326–337.

Stihl, A., Sommer, U., Post, A.F. (2001). Alkaline phosphatase activities among populations of the colony-forming diazotrophic cyanobacterium *Trichodesmium* spp. (cyanobacteria) in the Red Sea. *J. Phycol.* 37, 310–317.

Subramaniam, A., Brown, C.W., Hood, R.R., Carpenter, E.J., Capone, D.G. (2001). Detecting *Trichodesmium* blooms in SeaWiFS imagery. *Deep Sea Res. Pt. II* 49, 107–121.

Turk-Kubo, K.A., Frank, I.E., Hogan, M.E., Desnues, A., Bonnet, S., Zehr, J.P. (2015). Diazotroph community succession during the VAHINE mesocosm experiment (New Caledonia lagoon). *Biogeosciences* 12, 7435–7452.

Van Mooy, B.A., Hmelo, L.R., Sofen, L.E., Campagna, S.R., May, A.L., Dyhrman, S.T., et al. (2011). Quorum sensing control of phosphorus acquisition in *Trichodesmium* consortia. *ISME J.* 6, 422–429.

Wagner, G.P., Kin, K., Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* 131, 281–285.

Walsby, A. (1978). The properties and buoyancy-providing role of gas vacuoles in *Trichodesmium* Ehrenberg. *Brit. Phycol. J.* 13, 103–116.

Walsby, A. (1992). "The gas vesicles and buoyancy of *Trichodesmium*," in *Marine pelagic cyanobacteria: Trichodesmium and other diazotrophs*, eds. E.J. Carpenter, D.G. Capone, J.G. Rueter. (Springer Netherlands), 141–161.

Walworth, N., Pfreundt, U., Nelson, W.C., Mincer, T., Heidelberg, J.F., Fu, F., et al. (2015). *Trichodesmium* genome maintains abundant, widespread noncoding DNA in situ, despite oligotrophic lifestyle. *PNAS* 112, 4251–4256.

Wilson, S.T., Böttjer, D., Church, M.J., Karl, D.M. (2012). Comparative assessment of nitrogen fixation methodologies, conducted in the oligotrophic North Pacific Ocean. *Appl. Env. Microbiol.* 78: 6516–6523.

Wyman, M., Hodgson, S., Bird, C. (2013). Denitrifying Alphaproteobacteria from the Arabian Sea that express *nosZ*, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. *Appl. Env. Microbiol.* 79, 2670–2681.

Zani, S., Mellon, M.T., Collier, J.L., Zehr, J.P. (2000). Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. *Appl. Env. Microbiol.* 66, 3119–3124.

Zehr, J.P., Jenkins, B.D., Short, S.M., Steward, G.F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Env. Microbiol.* 5, 539–554.

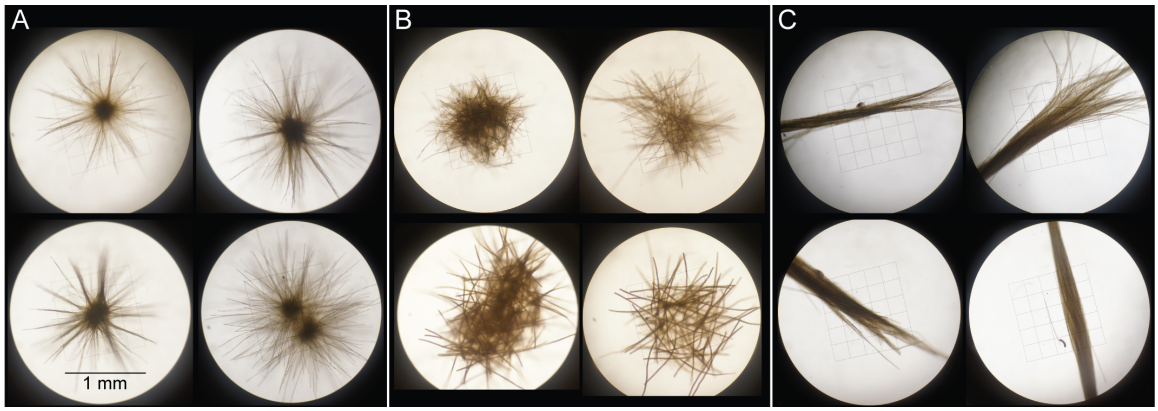

Figure 3.1: *Trichodesmium* colonies sorted into morphological classes of "radial puffs" (A), "non-radial puffs" (B), and "tufts" (C) on 23 Mar 2014. On all other collection days, "puffs" designate mixtures of morphotypes (A) and (B), "tufts" designate morphotype (C), and "mixed" designates mixtures of all three morphotypes.

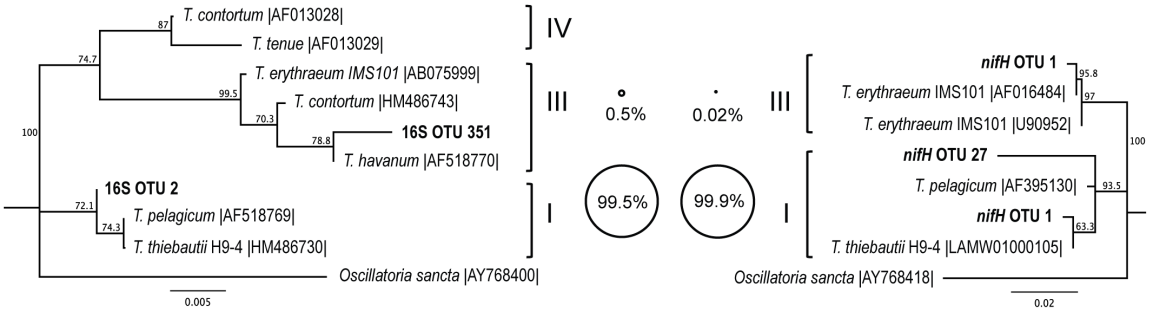

Figure 3.2: Neighbor joining phylogenetic trees depicting the relationships between *Trichodesmium* OTUs (97% nucleotide similarity) from partial 16S rRNA (left) and partial *nifH* (right) gene sequences, together with reference sequences from cultivated representatives (accession numbers given). Major *Trichodesmium* clades (Lundgren et al., 2005) are shown in Roman numerals. Bubble plots depict the percentage of *Trichodesmium* DNA sequences from this study which group with each clade, according to partial 16S (left) and partial *nifH* (right) amplicon datasets. Bootstrap values (1000 replicates) of >50% are provided. Scale bars represent nucleotide substitutions per site.

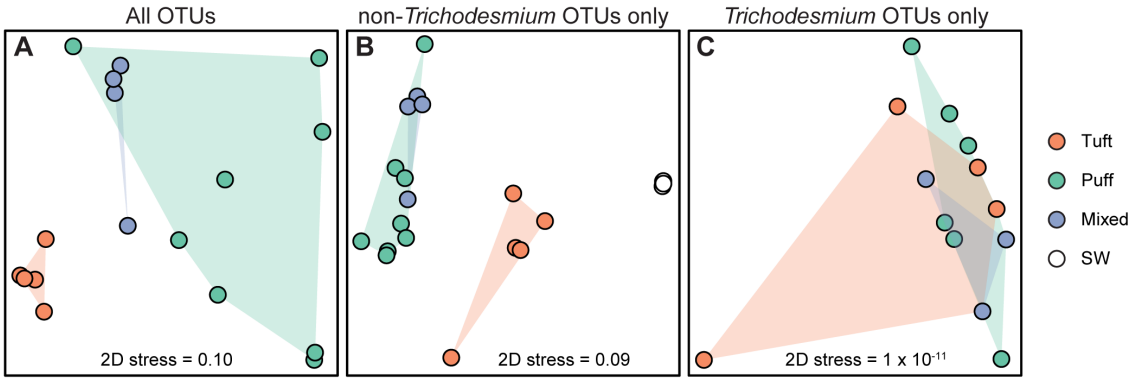

Figure 3.3: Non-metric multi-dimensional scaling (NMDS) plots derived from the Bray–Curtis dissimilarity matrix of 16S OTUs from (A) all *Trichodesmium* colony sample OTUs (7011 sequences per sample), (B) *Trichodesmium* colony and surface seawater samples excluding *Trichodesmium* OTUs (3128 sequences per sample), and (C) *Trichodesmium* colony samples excluding non-*Trichodesmium* OTUs (2248 sequences per sample). Each point represents an individual sample. Colors represent sample type (*Trichodesmium* tuft colonies, *Trichodesmium* puff colonies, mixed *Trichodesmium* colonies, and bulk seawater (SW) from 25m).

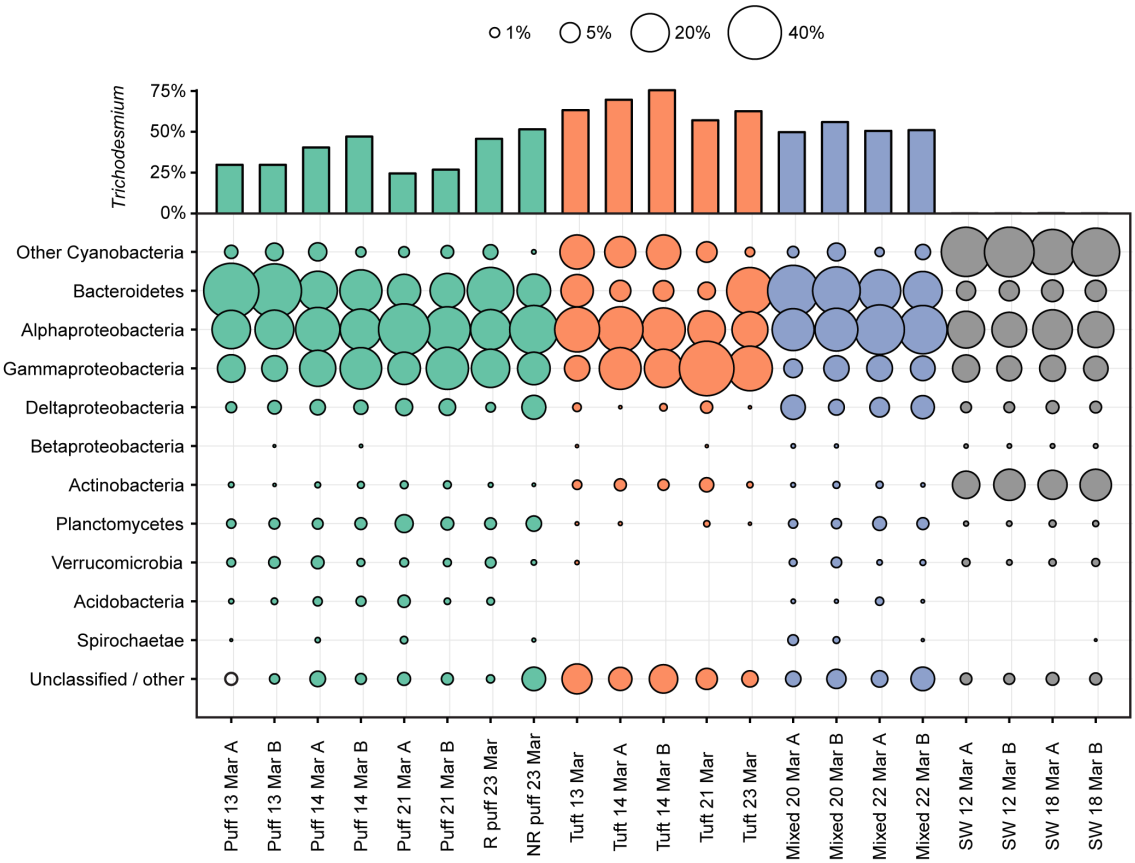

Figure 3.4: Percentages of partial 16S rRNA gene sequences from each sample assigned to bacterial taxa. *Trichodesmium* percentages are displayed in the upper bar plot while percentages of non-*Trichodesmium* taxa are displayed in area plots. Color indicates the sample type/morphology (green, red, blue, and grey for puff colonies, tuft colonies, mixed colonies, and bulk 25m seawater (SW), respectively). R denotes radial and NR denotes non-radial puff morphologies (see Fig. 3.1).

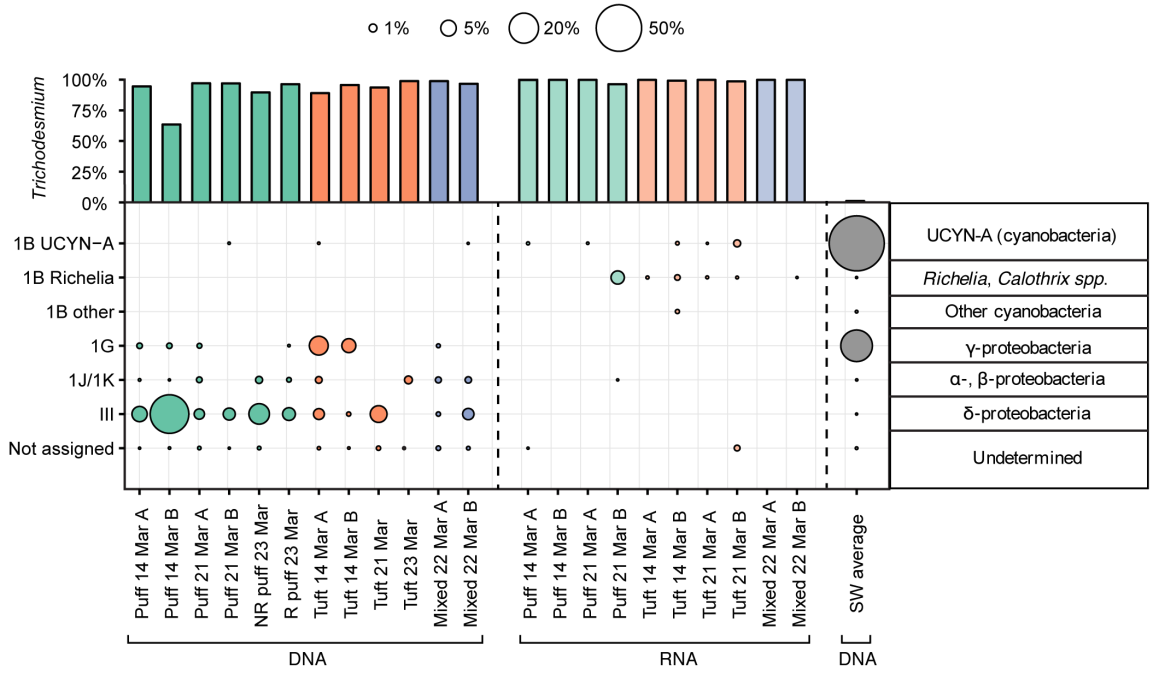

Figure 3.5: Percentages of *nifH* gene (left, right) and transcript (center) sequences from each sample assigned to *nifH* cluster groups. *Trichodesmium* are displayed in the upper bar plot while other taxa are displayed in the lower area plot. Representative taxa from canonical *nifH* clusters (Zehr et al., 2003) are shown to the right. Color indicates the sample type/morphology (green, red, blue, and grey for puff colonies, tuft colonies, mixed colonies, and bulk 25m seawater (SW), respectively). R denotes radial and NR denotes non-radial puff morphologies (see Fig. 3.1).

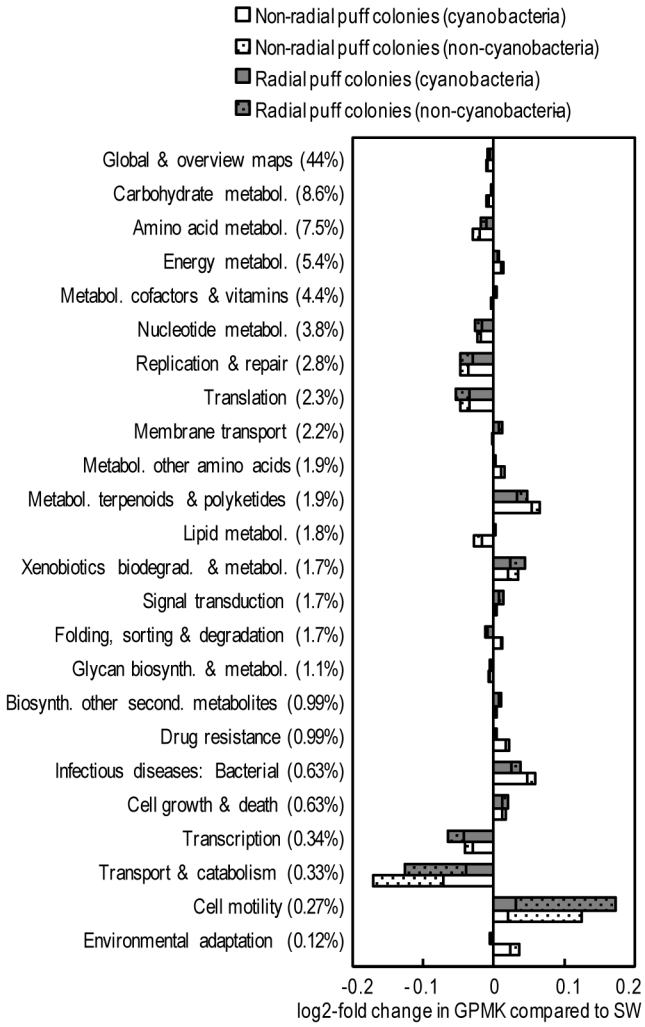

Figure 3.6: Relative abundance of KEGG gene groups in *Trichodesmium* colony samples (collected 23 Mar 2014) compared to a surface seawater sample from Stn. ALOHA (collected 30 July 2015). The total percentage of total counts from three summed samples are provided in parentheses. Pathways involved with organ systems, human disease, and/or representing <0.1% of total GPMK were excluded. Pathways displayed represent >97% of total GPMK.

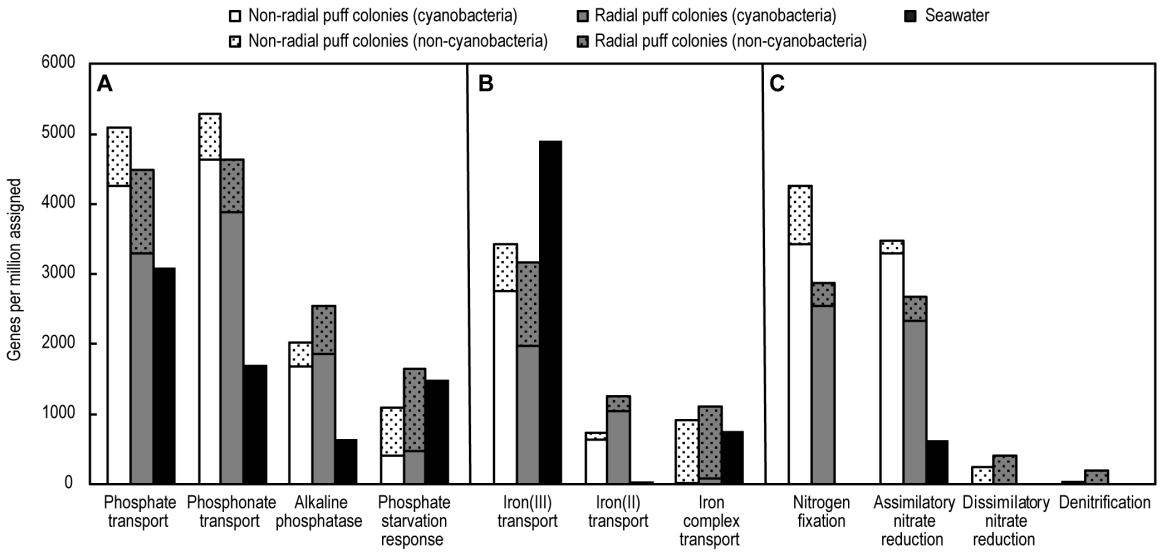

Figure 3.7: Abundances (GPMK) of select genes and pathways involved phosphorus (A), iron (B) and nitrogen (C) cycling pathways from *Trichodesmium* colonies and surface seawater metagenomes. For *Trichodesmium* colonies, solid areas represent cyanobacterial counts and dotted areas represent non-cyanobacterial counts. See Table S3.5 for a list of KO included in each pathway.

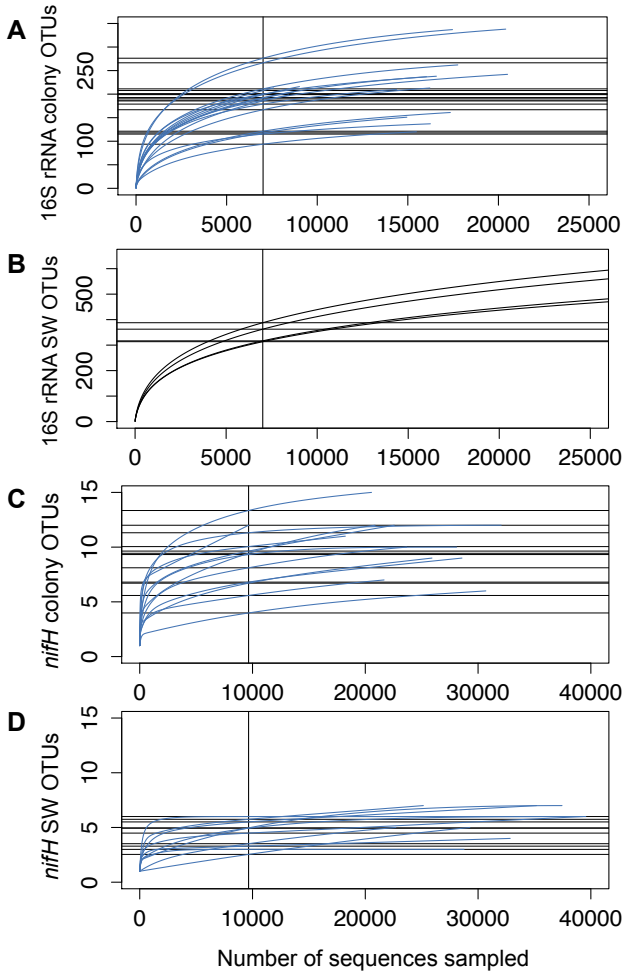

Figure S3.1: Rarefaction curves for *Trichodesmium* colony and 25 m seawater wamples from 16S rRNA (A, B) and *nifH* (C, D) genes. Curves were produced using the vegan rarecurve function (http://CRAN.R-project.org/package =vegan).

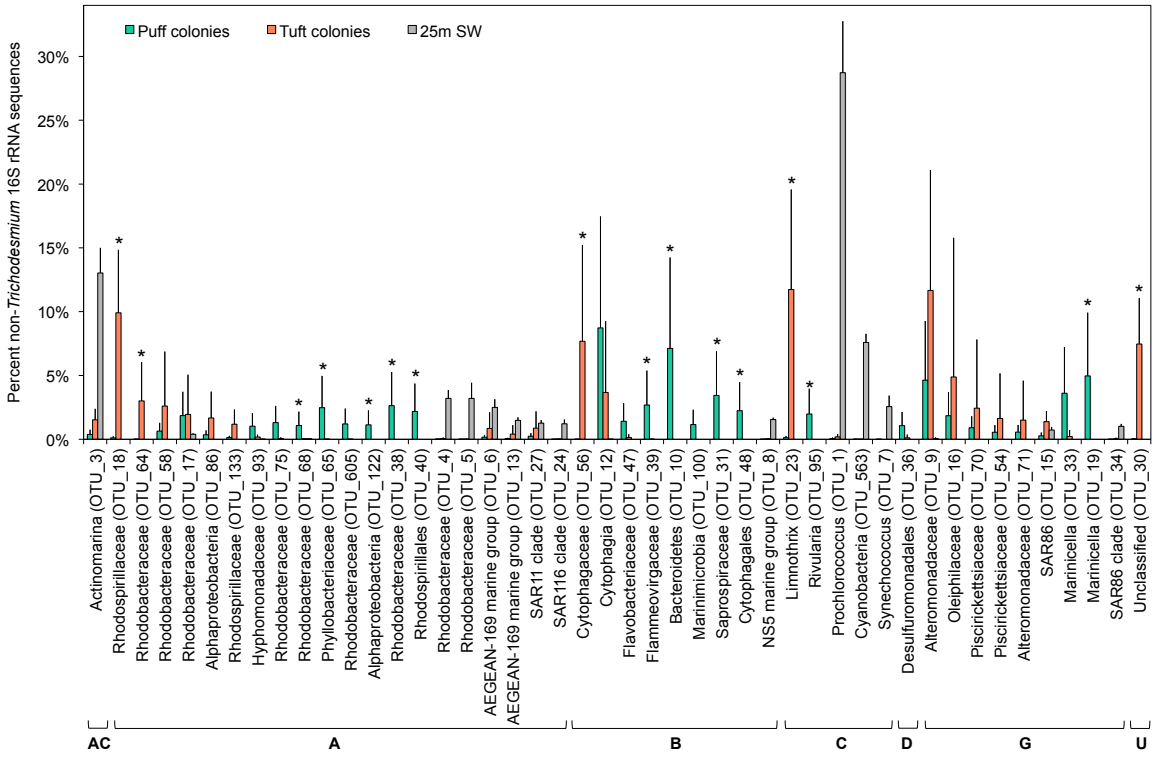

Figure S3.2: Dominant OTUs from the partial 16S rRNA dataset. The percent of non-*Trichodesmium* 16S rRNA gene sequences from puff, tuft, and 25m SW samples for OTUs containing an average of >1% of sequences for puff or tuft samples are displayed; this represents 64%, 77%, and 14% of non-*Trichodesmium* 16S rRNA sequences from puff colonies, tuft colonies, and 25m SW samples, respectively. Brackets show OTUs classified as actinobacteria (AC), alpha-proteobacteria (A), beta-proteobacteria (B), Cyanobacteria (C), delta-proteobacteria (D), gamma-proteobacteria (G), and unclassified taxonomy (U). Error bars represent SD from puff samples (n=7), tuft samples (n=5), and 25m SW samples (n=4). Asterisks denote a significant difference between puff and tuft samples ($p < 0.001$, Welch Two Sample t-test with Bonnferroni correction for multiple comparisons).

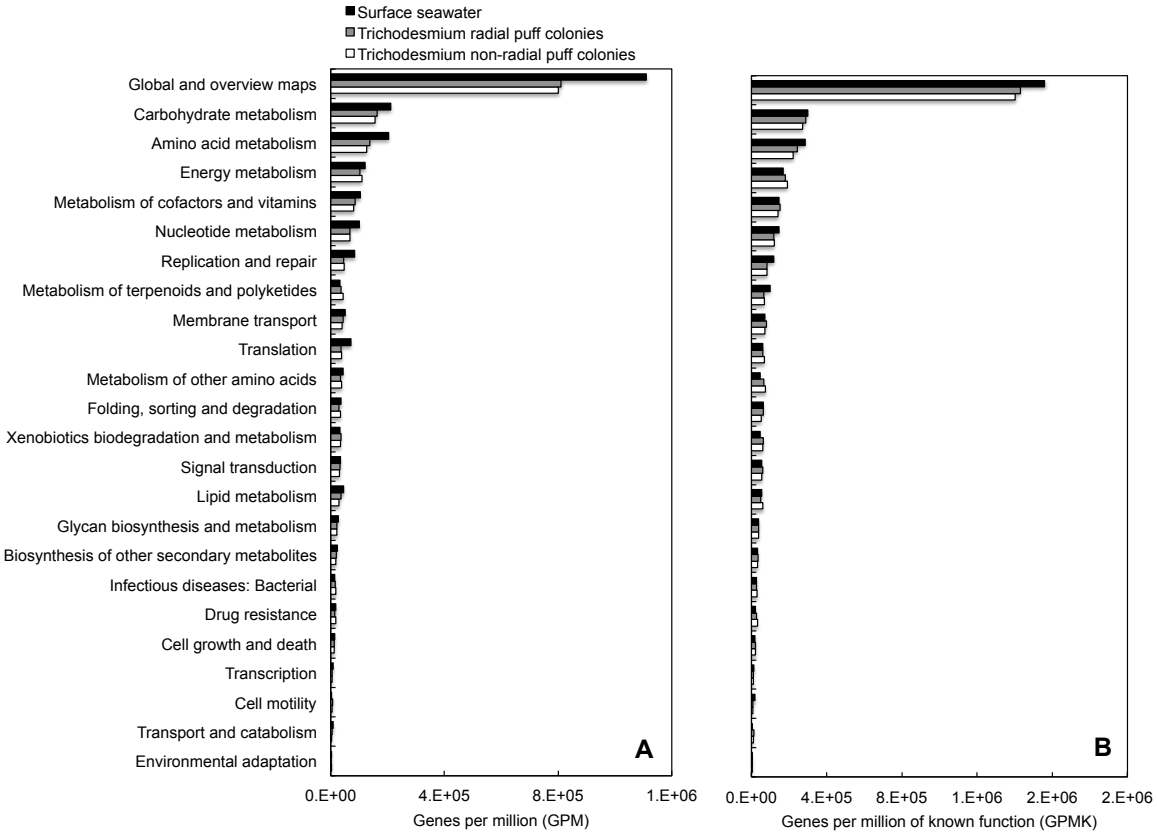

Figure S3.3: KEGG categories including summed KO from *Trichodesmium* colony and surface seawater samples using GPM (A) and GPMK (B) normalization methods.

Table 3.1: Summary and environmental conditions and samples collected in a March 2014 cruise at Stn. ALOHA. All samples were collected pre-dawn, with the exception of 23 Mar 2014, when samples were collected mid-afternoon. Sea surface temperature (SST) and surface chlorophyll fluorescence (Chl) were measured at 25 m depth using conductivity-temperature-depth sensors. Rates represent averages of duplicate incubation bottles, with standard deviations in parentheses. R denotes radial; NR denotes non-radial (see Fig. 3.1).

Date	SST (°C)	Chl (μg L^{-1})	Morphologies used	Measurements	C fixation rate (nmol C μmol C^{-1} d^{-1})	N_2 fixation rate (nmol N μmol C^{-1} d^{-1})
12 Mar	24.2	0.16	25m seawater only	DNA	ND	ND
13 Mar	24.2	0.15	puff, tuft	DNA	ND	ND
14 Mar	24.1	0.16	puff, tuft	DNA, RNA, rates	Puff: 7.9 (1.2) Tuft: 7.2 (0.6)	Puff: 0.02 (0.003) Tuft: 0.01 (0.004)
18 Mar	23.8	0.25	25m seawater only	DNA	ND	ND
20 Mar	23.8	0.22	mixed	DNA, rates	9.1 (1.8)	0.09 (0.05)
21 Mar	23.8	0.21	puff, tuft	DNA, RNA	ND	ND
22 Mar	23.8	0.20	mixed	DNA, RNA, rates	10.1 (1.9)	0.14 (0.08)
23 Mar	23.8	0.18	mixed	rates	9.7 (1.5)	0.17 (0.05)
23 Mar	23.9	0.11	R puff, NR puff, tuft	DNA, microscopy	ND	ND

Table 3.2: Summary of assembly, annotation, and mapping. Parenthetic values represent only those KO assigned to non-Cyanobacteria.

	Trichodesmium **non-radial puff colonies (23 Mar 2014)**	***Trichodesmium*** **radial puff colonies (23 Mar 2014)**	**Stn. ALOHA surface seawater (30 July 2015)**
Illumina paired-end reads	13,294,194	8,629,462	14,035,332
Contigs assembled	1,771,587	1,341,086	444,296
Contigs annotated to KO	454,684	290,117	330,104
Contigs annotated (%)	25.7%	21.6%	74.3%
Counts mapped to KO	6,446,495	3,743,920	3,669,469
Counts mapped to KO of known function[1]	3,664,674	2,116,212	2,550,585
Genomes per million genes[1,2]	1606 (1504)	1756 (1843)	2838
KO of known funtion (%)	56.8 (57.9)	56.5 (59.6)	69.5%

[1]Length-corrected counts
[2]*rpoB* GPM

Table 3.3: Diversity and species richness estimates. Estimates are derived from partial 16S rRNA gene sequences using all OTUs (7011 sequences per sample) and OTUs excluding *Trichodesmium* sequences (3128 sequences per sample). Data are presented as averages within sample type (n = 8 puff, 5 tuft, 4 mixed morphology, and 4 25m bulk seawater (SW) samples), with standard deviations in parentheses.

	Sample Type	Diversity (Shannon)	Species Richness (Chao1)
All OTUs	Puff	4.5 (0.8)	268 (70)
	Tuft	2.5 (0.5)	208 (58)
	Mixed	3.8 (0.2)	263 (15)
	SW	5.1 (0.3)	513 (55)
non-*Trichodesmium* OTUs only	Puff	5.6 (0.6)	250 (66)
	Tuft	4.4 (0.6)	212 (57)
	Mixed	5.8 (0.3)	269 (38)
	SW	5 (0.3)	426 (27)

Table 3.4: Taxonomic assignments from three metagenome samples. Data represent percentages of length-corrected reads mapped to KO-annotated assemblies.

	Trichodesmium **non-radial puff colonies (23 Mar 2014)**	***Trichodesmium*** **radial puff colonies (23 Mar 2014)**	**Stn. ALOHA surface seawater (30 July 2015)**
Cyanobacteria	75.9%	65.5%	31.8%
α-proteobacteria	11.1%	12.5%	41.6%
γ-proteobacteria	3.2%	5.3%	11.2%
δ-proteobacteria	1.2%	0.9%	0.8%
β-proteobacteria	0.5%	1.1%	0.7%
Bacteroidetes	4.1%	8.6%	5.4%
Firmicutes	1.1%	1.3%	1.3%
Planctomycetes	0.8%	1.9%	0.3%
Actinobacteria	0.5%	0.6%	1.0%
Verrucomicrobia	0.1%	0.5%	0.6%
Chloroflexi	0.1%	0.2%	0.1%
Eukaryota	0.5%	0.3%	2.6%
Archaea	0.1%	0.2%	0.5%
Viruses	0.1%	0.1%	0.5%
Other bacteria	0.7%	1.0%	1.6%

Table S3.1: Dual-index barcoded, forward and reverse nifH primers (5' -> 3') used in this study. Sample barcodes are shown in bold. Forward and reverse *nifH* PCR primers are indicated by nif3 (TGYGAYCCNAARGCNGA) and nif4 (ADNGCCATCATYTCNCC; note the misprint of this primer in the original manuscript by Zani et al (2000)). NNNN indicate Illumina linker regions: AATGATACGGCGACCACCGAGATCTACAC (forward) and CAAGCAGAAGACGGCATACGAGAT (reverse). Blue text indicates the binding site for sequencing primers, which were designed to optimize melting temperature during sequencing, as described by Kozich et al (2013).

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S3.2: Sensitivity analysis showing detection limits for all nitrogen fixation rate (NFR) measurements. The contribution of each source of error to the total uncertainty is provided, as performed by Montoya et al. (1996). Average and SD columns represent the observed average and standard deviation for each measured parameter. δNFR/ δX columns represent the partial derivative of the nitrogen fixation rate (NFR) with respect to each parameter, evaluated using the provided average and standard deviation. Error Contribution and % Total Error columns represent the absolute and relative error associated with each parameter. The total uncertainty associated with each measurement is considered the detection limit, and was calculated using standard propagation of error. NFR units are nmol N μmol C^{-1} d^{-1}.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S3.3 Taxonomic assignment and number of reads from each sample assigned to each 16S rRNA OTU.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S3.4 Metagenome taxonomy counts (GPM) at the phylum level.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S3.5 Metagenome counts (GPMK) from each sample belonging KO from select genes and pathways (see Fig. 3.7), along with the percent of *Trichodesmium* colony sample reads taxonomically assigned to Cyanobacteria.

See GradovilleMaryR2017_SupplementaryTables.xlsx

CHAPTER 4: ENVIRONMENTAL CONTROLS OF OYSTER-PATHOGENIC *VIBRIO* SPP. IN OREGON ESTUARIES AND A SHELLFISH HATCHERY

Mary R. Gradoville[a], Byron C. Crump[a], Claudia Häse[b], and Angelicque E. White[a]

[a]College of Earth, Ocean and Atmospheric Sciences, 104 COAS Administration Building, Oregon State University, Corvallis, Oregon, USA

[b]Department of Biomedical Sciences, College of Veterinary Medicine, Oregon State University, Corvallis, Oregon, USA

In preparation for submission to *Applied and Environmental Microbiology*

1752 N Street N.W., Washington, D.C., 20036

4. ENVIRONMENTAL CONTROLS OF OYSTER-PATHOGENIC *VIBRIO* SPP. IN OREGON ESTUARIES AND A SHELLFISH HATCHERY

4.1 Abstract

Vibrio spp. have been a persistent concern for Pacific Northwest larval bivalve hatcheries, which are vulnerable to environmental pathogens in seawater resources used for rearing larvae, yet the biogeochemical drivers of oyster-pathogenic *Vibrio* spp. in their planktonic state are poorly understood. Here, we present data tracking shellfish-toxigenic *Vibrio* bacteria in Netarts Bay and Yaquina Bay, Oregon, as well as in coastal waters and a local shellfish hatchery through the 2015 upwelling season. *Vibrio* populations were quantified using a culture-independent approach of high-throughput *Vibrio*-specific 16S rRNA amplicon sequencing paired with droplet digital PCR, and abundances were compared to local chemical and biological conditions. The most abundant putative pathogenic *Vibrio* spp. in our samples was *V. coralliilyticus*. Environmental concentrations of total *Vibrio* spp. and *V. coralliilyticus* were highest in Netarts Bay sediment samples, and higher in seawater from Netarts Bay than seawater from Yaquina Bay or nearshore coastal waters. In Netarts Bay seawater, the highest concentrations of *V. coralliilyticus* were observed during low tide, and abundances increased throughout the summer. We suggest that planktonic *Vibrio* populations in Netarts Bay seawater are seeded from the underlying estuarine sediment, and that warm, shallow waters in estuarine mudflats facilitate the local growth of the *V. coralliilyticus* pathogen. Samples from larval oyster tanks in Whiskey Creek Hatchery, which uses seawater pumped directly from Netarts Bay, contained significantly lower total *Vibrio*

spp. than did the bay-water, while *V. coralliilyticus* abundances were roughly similar to those in bay-water, resulting in a 50-fold increase in the relative abundance of the *V. coralliilyticus* pathogen in hatchery tanks. This suggests that hatchery disinfection practices less effectively remove this pathogen than other *Vibrio* spp.

4.2 Importance

Vibrio spp. are a serious concern for economically important aquaculture efforts. It has been argued that oyster-pathogenic *Vibrio* spp. contributed to recent mortality events in Pacific Northwest shellfish hatcheries (Elston et al. 2008); however, these events are often sporadic and unpredictable. The success of hatcheries is critically linked to the chemical and biological composition of inflowing seawater resources; thus, it is pertinent to understand the biogeochemical drivers of oyster-pathogenic *Vibrio* spp. in their free-living state. Here, we show that Netarts Bay, the location of a local hatchery, is enriched in oyster-pathogenic *V. coralliilyticus* compared to coastal seawater, and hypothesize that favorable condition promote the local growth of this pathogen. Furthermore, *V. coralliilyticus* appears to persist within hatchery conditions. These results improve our understanding of the ecology and environmental controls of the *V. coralliilyticus* pathogen, and could be used to improve future aquaculture efforts.

4.3 Introduction

The *Vibrionaceae* (*Vibrio*) are a genetically and ecologically diverse group of pathogenic and benign heterotrophic, gram-negative bacteria present in most, if not all, marine ecosystems (Urakawa and Rivera, 2006; Wietz et al. 2010). *Vibrionaceae*

include >130 described species (Gomez-Gil et al. 2014) with diverse life histories. Populations of free-living marine *Vibrio* can survive in a dormant state during unfavorable conditions (Colwell, 1993) but grow rapidly in response to temperature and nutrient pulses (Eilers et al. 2000), aided by numerous ribosome genes (Heidelberg et al. 2000). Additionally, many *Vibrio* species are associated with marine particles and/or living hosts, where they can act as mutual symbionts or disease agents. *Vibrio* species include several human pathogens, including *V. cholerae* (Koch, 1884; Prouty and Klose, 2006), *V. parahaemolyticus*, and *V. vulnificus*, and numerous pathogens of marine mammals, fish, and shellfish (Takemura et al. 2014).

Pathogenic *Vibrio* have been a historical concern for the aquaculture industry (Elston et al. 1984; Estes et al. 2004), and it has been argued that infections of *V. coralliilyticus* (formerly misclassified as *V. tubiashii;* Ben-Haim and Rosenberg, 2002; Wilson et al. 2013) currently threaten production in Pacific Northwest oyster hatcheries (Elston et al. 2008). These hatcheries use seawater pumped directly from coastal or estuarine waters with minimal treatment steps (i.e., sand filters and heating); thus, larval rearing success is critically linked to the chemical and biological composition of seawater resources. Indeed, hatchery production has been depressed in recent years, and while pathogenic *Vibrio* spp. were initially implicated (Elston et al. 2008), Barton et al. (2012, 2015) offered strong evidence that unfavorable carbonate chemistry (ocean acidification) was the more significant cause. Wind-driven coastal upwelling in summer months delivers waters to hatcheries which have a low aragonite saturation state, a carbonate system parameter that is mechanistically linked to larval growth and fitness under controlled conditions (Waldbusser et al. 2014) and has been correlated to larval

production in an Oregon hatchery (Barton et al. 2012). However, hatchery mortality events are sporadic and unpredictable, and the determinations of the underlying mechanisms—acidification, pathogens, and/or other stressors—are often speculative. Thus, the possibility that *V. coralliilyticus* infections may have played a role in hatchery mortality events cannot be excluded. Furthermore, while much recent attention has focused on monitoring the intrusion of acidified waters into coastal and estuarine environments (Feely et al. 2008; Waldbusser and Salisbury, 2014), it is not known whether the spatiotemporal distributions of oyster-pathogenic *Vibrio* spp. are linked to particular offshore water masses or local chemical or biological conditions.

The abundance and community structure of coastal and estuarine *Vibrio* populations are shaped by environmental drivers (Thompson et al. 2004; Tout et al. 2015; Siboni et al. 2016). However, these patterns can be complex, as individual *Vibrio* species display different relationships with physical/chemical conditions (e.g. temperature, salinity) and with local biology, due in part to species-specific associations with hosts including phytoplankton and zooplankton (Takemura et al. 2014). Previous studies of *V. coralliilyticus* in coral reef systems have observed increased abundances associated with elevated temperatures (Tout et al. 2015; Amin et al. 2016), but the ecology of this organism has rarely been assessed in temperate ecosystems. In the only previous study of *V. coralliilyticus* in the Pacific Northwest coastal upwelling system, Elston et al. (2008) hypothesized that *V. coralliilyticus* populations in Netarts Bay, OR were seeded via upwelling of nearshore waters, with subsequent local growth following relaxation of upwelling-favorable winds and warming. However, this study was qualitative in nature and relied on plating and culturing methods which bias *Vibrio* community structure

(Colwell, 1993). Thus, the ecological controls of oyster-pathogenic *Vibrio* spp. in coastal upwelling-influenced systems, including estuaries containing larval hatcheries, remains poorly understood.

To this end, we designed a study to track the diversity and abundance of total and oyster-pathogenic *Vibrio* spp. in Oregon estuaries, coastal waters, and a larval oyster hatchery. Many previous studies of *Vibrio* ecology have involved plating and culturing isolates from environmental samples in order to sequence multiple gene loci (Sawabe et al. 2013), an approach which limits the number of cells that can be sequenced (Pedrós-Alió, 2012) and which can bias community structure (Colwell, 1993). We instead used high-throughput *Vibrio*-specific 16S rRNA amplicon sequencing and droplet digital PCR to quantify *Vibrio* populations. Here, we investigate the spatiotemporal distributions of the diversity and abundance of *Vibrio* populations in Oregon estuaries through the 2015 upwelling season as well as the environmental controls of the most dominant pathogen, *V. coralliilyticus*.

4.4 Results

4.4.1 Environmental conditions

A wide range of biological and chemical conditions were encountered in Netarts Bay and Yaquina Bay over the course of the summer 2015 sampling period, and in vertical profiles of the water column at offshore stations. The range of temperature, salinity, chlorophyll, nutrients, the partial pressure of carbon dioxide (P_{CO2}), and daily wind stress are presented in Table S4.2 for the estuarine and offshore sampling stations. Sampling spanned upwelling- and downwelling-favorable conditions (Table S4.2).

4.4.2 Diversity of Vibrio spp. populations and putative oyster-pathogenic Vibrio

Vibrio spp. diversity was assessed via high throughput sequencing of a 491 bp region of the 16S rRNA gene using *Vibrio*-specific primers. DNA from *V. coralliilyticus* strain RE22 was also sequenced as a positive control. Only samples with three successful PCR amplifications were used for sequencing (23/23 coastal seawater samples, 21/22 Yaquina Bay seawater samples, 30/30 Netarts tidal flat seawater samples, 18/28 Netarts tidal flat sediment samples, 58/62 Netarts WCSH inflow seawater samples, and 9/19 WCSH larval tank samples). Subsampling down to 7145 sequences per sample resulted in near-saturation for most rarefaction curves (data not shown). Clustering sequences at 97% nucleotide identity resulted in 1950 *Vibrio* spp. OTUs.

The community structure and species composition of *Vibrio* populations varied with sample type. All estuary samples displayed similar *Vibrio* community structure, which was distinct from the structure of coastal seawater and WCSH larval tank samples (Fig. 4.2). 8/9 WCSH larval tank samples clustered separately from environmental samples and close to DNA from cultured *V. coralliilyticus* strain RE22. The taxonomic composition of *Vibrio* spp. also varied by sample type (Fig. S4.2). For example, coastal seawater samples were relatively enriched in OTUs clustering with *Photobacterium* spp. and *V. caribbeanicus* compared to estuary samples. The most dominant OTU in the dataset clustered with *V. lentus* (Fig. S4.2).

Putative oyster-pathogenic *Vibrio* spp. were classified based on inferred phylogenetic relationships between *Vibrio* OTUs and a representative set of 134 cultured *Vibrio* species (Gomez-Gil et al. 2014) (Fig. S1). Our samples contained sequences

clustering with the bivalve larval pathogens *V. coralliilyticus*, *V. tubiashii, V. parahaemolyticus, V. pectenicida*, and *V. cholerae*, but did not contain sequences clustering with *V. alginolyticus*, *V. splendidus*, or *V. vulnificus* (Fig. S4.3). The most abundant putative-pathogenic species was *V. coralliilyticus* in all sample types, with the exception of Yaquina Bay seawater samples, where the most abundant putative pathogen was *V. cholerae*. We focused remaining analyses on *V. coralliilyticus*.

4.4.3 Abundances of Vibrio spp. and V. coralliilyticus

The fraction of total *Vibrio* sequences classified as *V. coralliilyticus* was higher in WCSH larval rearing tanks than in all other sample types (Fig. 4.3A, S4.3). *V. coralliilyticus* represented a median of 22% of *Vibrio* sequences in tank samples, a 60-fold higher percentage than coastal seawater samples and a 39-fold higher percentage than Netarts WCSH intake seawater samples (Tukey HSD $p < 0.0001$ for both). Among environmental samples, the median relative proportion of *V. coralliilyticus* sequences was higher in Netarts seawater overlying tidal flats than in all other sample types (Tukey HSD $p < 0.05$ for all). *V. coralliilyticus* represented the smallest fraction of total *Vibrio* spp. in Netarts sediment samples, where fractions were significantly lower than coastal seawater (Tukey HSD $p < 0.01$), Netarts WCSH intake seawater (Tukey HSD $p < 0.0001$), and Netarts tidal flat seawater (Tukey HSD $p < 0.0001$).

Relative abundances of *V. coralliilyticus* from 16S rRNA sequence data were combined with ddPCR-derived concentrations of total *Vibrio* spp. 16S rRNA gene copies to produce estimates for *V. coralliilyticus* cell abundances (assuming 10% DNA extraction efficiency and 12 16S rRNA genes per genome) (Fig. 4.3). Total *Vibrio* spp.

abundance estimates ranged from $1.1 \times 10^1 – 3.2 \times 10^5$ cells mL^{-1}, averaging ~0.3% of total heterotrophic bacteria (assessed via flow cytometry) across sample types. When normalizing to mL $seawater^{-1}$ or g $sediment^{-1}$, the highest concentrations of *V. coralliilyticus* were observed in Netarts Bay tidal flat station sediment and seawater samples, and in WCSH larval tank samples (Fig. 4.3). Among seawater samples, *V. coralliilyticus* concentrations were lowest in coastal seawater and highest in the Netarts tidal flats. Total *Vibrio* spp. abundance estimates followed similar trends among sample types, excluding larval tank samples, which contained the lowest concentrations of total *Vibrio* spp. Normalizing abundances to DNA content resulted in similar trends across sample types, with the exception of Netarts tidal flat sediment samples; these samples had the highest abundances of all sample types when normalized to sediment mass but the lowest abundances when normalized to DNA content, which was ~three orders magnitude higher in sediment (g^{-1}) than in seawater (mL^{-1}) (Fig. S4.4). At coastal OR stations, the abundances of *Vibrio* spp., *V. coralliilyticus*, and total heterotrophic bacteria all decreased with depth (Fig. 4.4).

The abundances of *V. coralliilyticus* and total *Vibrio* spp. were highly variable. In Netarts Bay tidal flat seawater samples, the variability within biological replicates of *Vibrio* spp. and *V. coralliilyticus* were an order of magnitude greater than the biological variability observed in total heterotrophic bacterial abundance (assessed via flow cytometry) (Table S4.2). Likewise, variability between biological replicates of *Vibrio* spp. and *V. coralliilyticus* in Netarts WCSH intake seawater was an order of magnitude higher than that of total heterotrophic bacteria, and was nearly equal in magnitude to the total daily variability (Table S4.2). Despite this large biological heterogeneity, there were

significant differences in *V. coralliilyticus* concentrations among days for the Netarts WCSH intake time series (One-way ANOVA, $p < 0.001$; Fig. 4.5) and marginally significant differences among Netarts tidal flat stations (One-way ANOVA $p = 0.08$, Fig. 4.6). Additionally, *V. coralliilyticus* had a larger total range (3 orders of magnitude in Netarts WCSH intake water, Fig. 4.5) and larger day-to-day variability than concentrations of total heterotrophic bacteria.

4.4.4 Environmental predictors of Vibrio spp. and V. coralliilyticus

Despite the high variability observed within biological replicates, regression models showed evidence for environmental predictors of *V. coralliilyticus* and *Vibrio* spp. concentrations. Concentrations of both groups were positively correlated to the day of year (Fig. 4.5, linear regression p-value < 0.001, $R^2 = 0.26$ and 0.22 for *Vibrio* spp. and *V. coralliilyticus*, respectively), suggesting a seasonal cycle in *Vibrio* growth. Models testing environmental explanatory variables indicate that both *V. coralliilyticus* and total *Vibrio* spp. were negatively correlated to northward, downwelling-favorable wind stress and positively correlated to dissolved inorganic phosphorus concentrations (Table 4.3). Additionally, total *Vibrio* spp. concentrations were positively correlated to temperature, while *V. coralliilyticus* concentrations were significantly higher at low tide than high tide. Furthermore, tide, wind stress, and phosphate were also significant explanatory variables predicting the ratio of *V. coralliilyticus* to heterotrophic bacteria, indicating that these variables help to explain how patterns of *V. coralliilyticus* diverge from the average bacterioplankton, whose abundances were negatively related to N+N and positively related to temperature, nitrate, and phosphate (Table 4.3).

4.5 Discussion

Environmental pathogens such as *Vibrio* spp. act as disease agents for a variety of hosts, including shellfish, but can also persist and thrive in marine environments independently (Cangelosi et al. 2004). Understanding the ecology of these organisms in their free-living state can help determine environmental controls of marine disease and improve aquaculture efforts. In this study, we used a culture-independent molecular approach to investigate the diversity and spatiotemporal abundance patterns of oyster-pathogenic *Vibrio* spp. in Oregon estuaries, coastal seawater, and a larval oyster hatchery. The most abundant putative pathogen in our samples clustered with *V. coralliilyticus*, a known oyster pathogen (Richards et al. 2015). We report elevated *V. coralliilyticus* concentrations in Netarts Bay compared to coastal seawater and Yaquina Bay seawater, where favorable conditions appear to drive the local growth of this organism, especially in warm, late summer months. Furthermore, *V. coralliilyticus* represented a large fraction of total *Vibrio* populations in WCSH larval rearing tanks, implying the persistence of this organism in hatchery conditions.

4.5.1 Local growth of V. coralliilyticus in Netarts Bay

There is an economic incentive to understand the environmental drivers of *V. coralliilyticus* abundances in Netarts Bay, where a local shellfish hatchery is vulnerable to aquatic pathogens in seawater resources. A study by Elston et al. (2008) suggested that *V. coralliilyticus* abundances in Netarts Bay might be related to oceanographic conditions, postulating that the upwelling of deep waters seed high concentrations of this pathogen into the bay, and that subsequent relaxation and warming could fuel explosive

growth. We did not find evidence to support this hypothesis. Concentrations of both *Vibrio* spp. and *V. coralliilyticus* decreased with depth at coastal stations (Fig. 4.4), indicating that upwelling conditions would supply lower seed concentrations than downwelling conditions. Additionally, high concentrations of *Vibrio* spp. and *V. coralliilyticus* in Netarts WCSH intake water were not associated with physical or chemical characteristics of the cold, salty, nutrient-rich upwelling source water (Fig. 4.7); in fact, the highest concentrations in this time series occurred during a period of low salinity and warm temperatures (Fig. 4.5). Thus, we found no evidence linking the abundances of *V. coralliilyticus* in Netarts Bay to coastal upwelling, suggesting that the threat of *Vibrio* pathogens to bivalves may be temporally uncoupled from ocean acidification stress in this environment.

We postulate that the moderate abundances of *V. coralliilyticus* we observed in Netarts Bay were not advected from nearshore waters, but rather that favorable estuarine conditions promote the local growth of this organism. In this study, *V. coralliilyticus* were the most abundant in tidal flat samples; specifically, seawater sampled directly above shallow seagrass/macroalgae beds contained the highest concentrations of *V. coralliilyticus* (10^2–10^3 cells mL^{-1}) (Fig. 4.1, 4.6). During low tide (when our samples were collected), these areas are partially isolated, and connected to the larger bay by small tidal channels. Thus, these shallow tidal flat pools may act as environmental incubators, where higher residence times may minimize the dilution of estuarine *V. coralliilyticus* populations with coastal seawater. It may also be possible that these shallow, stagnant waters are heated more efficiently than deeper channels, promoting the growth of total *Vibrio* spp. and *V. coralliilyticus* (Tout et al. 2015). Furthermore,

concentrations of *V. coralliilyticus* in Netarts WCSH intake pipe samples were higher during low tide than high tide, suggesting that the tidal input of nearshore oceanic waters dilutes local *Vibrio* populations in Netarts Bay. This contrasts with total heterotrophic bacteria, which were not strongly related to tide conditions, and suggests a unique ecology of *V. coralliilyticus*.

It is also likely that sediment reservoirs seed *Vibrio* populations into Netarts Bay seawater. Median concentrations of *Vibrio* spp. were an order of magnitude higher in tidal flat sediments than in tidal flat seawater, and two orders of magnitude higher than in coastal OR seawater (Fig. 4.3), consistent with previous reports of sediments as reservoirs for *Vibrio* spp. in estuaries (Vezzulli et al. 2009; Chase et al. 2015). The overall *Vibrio* community structure in Netarts Bay seawater was similar to Netarts sediment (Fig. 4.2), suggesting an interaction between these two reservoirs, which may help explain the high concentrations of *Vibrio* spp. in shallow tidal flat pools. However, the tidal flat sediment *Vibrio* communities contained a smaller relative proportion of *V. coralliilyticus* than was observed in tidal flat seawater samples (Fig. S4.3). This species may be seeded from alternate estuarine sources, such as seagrass, macroalgae. The advection of coastal seawater may also transport *V. coralliilyticus* into the bay, but this would require high local growth rates to explain the elevated abundances in Netarts seawater (Fig. 4.3).

Abundances of total *Vibrio* and *V. coralliilyticus* increased throughout the summer in Netarts WCSH inflow samples, suggesting a seasonality of these organisms. This could result from higher temperatures in the second half of the summer (Fig. 4.5) promoting increased growth rates. Indeed, *Vibrio* spp. concentrations were positively

related to temperature in Netarts WCSH inflow samples (Table 4.3). Elevated seawater temperatures have been linked to global increases in *Vibrio* spp. concentrations and to incidences of pathogenic *Vibrio* infections worldwide (Vezzulli et al. 2013); thus, *Vibrio* spp. concentrations in Netarts Bay could be expected to increase in future warming oceans. It should also be noted that our sampling year (2015) was characterized by anomalously high sea surface temperatures (~2-3°C above climatological averages) in the Northeast Pacific (McKibben et al. 2017) which may have further increased *Vibrio* growth rates and/or shifted community interactions (e.g. Peterson et al. 2016).

One challenge to inferring environmental controls from this dataset is the high degree of small-scale variability in *Vibrio* abundance. The coefficient of variation we observed among biological replicates was an order of magnitude larger for *Vibrio* spp. and *V. coralliilyticus* concentrations than for total heterotrophic bacteria (Table S4.2). This high variability likely reflects the stochastic collection of *Vibrio* spp. associated with large suspended particles or spatial heterogeneity within the estuary. Metabolic flexibility and the ability to produce extracellular enzymes including chitinase (Hunt et al. 2008) allow *Vibrio* spp. to grow on a wide range of substrates; *Vibrio* spp. have been observed associated with zooplankton (Heidelberg et al. 2002), phytoplankton (Islam et al. 1990), and marine detritus (Lyons et al. 2007). The fraction of particle-associated *Vibrio* has been found to be substantial: plankton-associated *Vibrio* concentrations (g biomass^{-1}) can be several orders of magnitude higher than free-living concentrations (mL seawater^{-1}) (Turner et al. 2009). Since we filtered seawater onto 0.2 μm filters without a pre-filtration step, our samples reflect both the free-living and particle-associated fractions. Further

studies are needed to determine whether *V. coralliilyticus* is consistently associated with specific living (e.g. a planktonic or benthic organism) or detrital reservoirs.

4.5.2 Hatchery conditions favor the growth or persistence of V. coralliilyticus

Vibriosis is a serious disease for hatchery-reared oyster larvae, and it has been argued that *V. coralliilyticus* infections may have contributed to severe mortality events observed in Pacific Northwest hatcheries over the past two decades (Elston et al. 2008). Hatchery outbreaks could in theory be a direct consequence of high pathogenic *Vibrio* concentrations in seawater resources; alternatively, hatcheries could become contaminated with toxigenic *Vibrio* spp. and/or environmental conditions could trigger increased virulence or larval susceptibility (Kimes et al. 2012; Asplund et al. 2014). Here, we report lower total concentrations of *Vibrio* spp. in larval tanks than in inflowing seawater, but a shift in the *Vibrio* community structure, with *V. coralliilyticus* representing a >50-fold larger fraction of total *Vibrio* spp. in tanks relative to that found in inflowing seawater (Fig. 4.3, S4.3). The absolute concentrations of *V. coralliilyticus* in larval tanks were well below those required to induce mortality under laboratory conditions (Richards et al. 2015), and no samples used in this study were collected during severe mortality events (A. Barton, *pers. comm.*). However, the striking dominance of *V. coralliilyticus* over other *Vibrio* species in larval rearing tanks compared to inflowing water implies that this pathogen is particularly resistant to hatchery disinfection methods, successful in hatchery conditions, and/or implicates contamination within the hatchery.

Following the initial seed-stock collapses and the work of Elston et al. (2008), hatchery personnel across the industry undertook extensive measures to reduce *Vibrio*

abundance in hatchery waters (Barton et al. 2015). During summer 2015, these practices included sand-filtering inflowing seawater, bubbling seawater with ozone prior to larval inoculation, disinfecting tanks between larval batches (every 2–3 days), and using sterile techniques when handling phytoplankton cultures. The significantly lower overall abundance of *Vibrio* spp. in WCSH waters than in bay-waters suggests that these measures have been generally effective at total *Vibrio* removal, but the persistence of *V. coralliilyticus* at near, or even slightly enriched above, bay-water levels suggests that the measures are less effective at removal of the most likely pathogenic form. We view this as a simplest explanation of our observations.

Alternatively, it is possible that the higher relative proportion of *V. coralliilyticus* in hatchery tanks could be due to contamination or elevated growth rates of this pathogen. Previous studies have documented substantial levels of *Vibrio* spp. in phytoplankton cultures, oyster broodstocks, thiosulfate, and air within hatcheries using similar disinfection methods (Sainz-Hernández and Maeda-Martínez, 2005; Elston et al. 2008). Here, we observed moderate *V. coralliilyticus* concentrations in larval tanks containing phytoplankton and several size classes of larvae, fertilized eggs prior to the addition of food, and also in tanks containing sand-filtered, heated seawater prior to the addition of any larvae or phytoplankton (Table 4.2). Thus, it is unlikely that the observed *V. coralliilyticus* concentrations were introduced through phytoplankton or broodstock contamination. Additionally, it is possible that the sand-filtration and ozone treatment reduced the concentration of total *Vibrio* spp., but that heating seawater to ~25°C favored the growth of *V. coralliilyticus*. Elevating seawater temperature has been previously demonstrated to increase the abundance of *V. coralliilyticus* relative to other *Vibrio*

species (Tout et al. 2015). While our study demonstrated a clear shift in *Vibrio* community structure from the WCSH intake water to the WCSH larval tank water, the underlying mechanisms for this shift remain speculative.

4.5.3 Implications for virulence and larval disease

The frequent non-linearity between pathogen abundance, toxin production, and host mortality is a challenge in disease ecology (e.g. McKibben et al. 2015). This is especially problematic for the management of vibriosis in bivalves, for which the mechanism of virulence is somewhat unclear (de O Santos et al. 2011). Initial work attributed the pathogenicity of *V. coralliilyticus* in both bivalve larvae and coral hosts to the extracellular zinc metalloprotease *vcpA* (Ben-Haim et al. 2003b; Hasegawa et al. 2008; Sussman et al. 2009). However, recent work has demonstrated that this metalloprotease is not required for *V. coralliilyticus* pathogenesis in multiple animal hosts (de O Santos et al. 2011), or for *V. tubiashii* pathogenesis in oyster larvae (Mersni-Achour et al. 2015). Sequencing the genomes of *V. coralliilyticus* isolates have revealed a diverse repertoire of virulence factors (de O Santos et al. 2011; Mersni-Achour et al. 2015) including metalloproteases, hemolysins, and membrane proteins, which may function independently or in concert to induce pathogenicity (Kimes et al. 2012).

Furthermore, the pathogenicity and infection potential of *V. coralliilyticus* are likely regulated by environmental cues. Temperature can regulate *V. coralliilyticus* virulence factors (Kimes et al. 2012) and elevating temperatures >27°C increases pathogenicity in corals (Ben-Haim et al. 2003a). This temperature effect may contribute to the success of *V. coralliilyticus* in larval hatcheries, where seawater is typically heated

to improve larval growth. Seawater chemistry may also regulate the infection potential of *V. coralliilyticus*: Asplund et al. (2014) reported that the closely related *V. tubiashii* was more successful at infecting blue mussels under conditions of ocean acidification. A similar ocean acidification effect on interactions between *V. coralliilyticus* and the *C. gigas* larvae could have severe consequences for the Pacific Northwest shellfish industry, which is already threatened by the intrusion of acidified upwelled waters and bracing for future changes to carbonate chemistry (Barton et al. 2012; Barton et al. 2015).

The virulence of *V. coralliilyticus* is also likely affected by community interactions with other microorganisms, especially other *Vibrio* spp. Host organisms are often inhabited by diverse *Vibrio* communities rather than clonal populations (Wendling et al. 2014; Tout et al. 2015). Recent work has shown that the virulence of multiple pathogenic *Vibrio* species can be synergistic (Gay et al. 2004). Furthermore, the presence of benign strains can contribute to the virulence of pathogenic *Vibrio* spp. (Lemire et al. 2015), possibly through the density-dependent production of quorum sensing molecules, which include regulators of virulence factors (Zhu et al. 2002). Adding to this complexity, specific *Vibrio* virulence factors can be strain-dependent (Kimes et al. 2012) and exchanged through lateral gene transfer. In the current study, larval tank samples included large fractions of sequences classified as *V. coralliilyticus*, as well as the *V. tubiashii* pathogen and putative-benign species including *V. penaeicida*, *V. lentos*, and *Allivibrio fisheri* (Fig. S4.3, S3). This diverse community may collaborate to facilitate *V. coralliilyticus* infections. It should also be noted that due to phylogenetic incongruities among genes used for taxonomic identification of *Vibrionaceae* (Sawabe et al. 2013), our culture-independent approach can only classify sequences as “putative” pathogens. Thus,

is possible that some sequences were misclassified as *V. coralliilyticus*. However, the fact that dominant OTUs classified as *V. coralliilyticus* shared high similarity to sequences derived from cultured *V. coralliilyticus* RE22, which were processed using the same laboratory and *in silico* methods (data not shown) gives us confidence in our results. Likewise, it is possible that sequences classified as benign species are actually pathogenic due to either misclassification or our incomplete knowledge of pathogenic potential within the *Vibrionaceae.*

While more work is needed to elucidate how physical, chemical, and biological cues regulate *V. coralliilyticus* pathogenicity in seawater and hatchery conditions, this study provides a critical first step toward understanding the ecology of this pathogen in temperate estuary systems. Netarts Bay appears to be a favorable environment for *V. coralliilyticus*, with sediment reservoirs likely seeding *Vibrio* populations and shallow tidal pools allowing for local growth of this pathogen. Our findings of higher *V. coralliilyticus* abundances at low tide and in late summer could be used to inform hatchery practices, as the threat of *Vibrio* pathogens is expected to worsen as temperatures increase in the coming decades. Furthermore, the stark community shift in *Vibrio* populations from Netarts WCSH intake water to larval tank water suggests that *V. coralliilyticus* infections could be a concern even when concentrations in the bay are low. Future work exploring mechanisms and environmental controls of the toxicity of *V. coralliilyticus* are needed in order to predict the risks imposed by this pathogen.

4.6 Methods

4.6.1 Sample collection

Biological (DNA, flow cytometry, chlorophyll *a*) and chemical (temperature (T), salinity (S), carbonate system parameters, nitrate + nitrite (N+N), and phosphate) samples were collected from Oregon estuaries and coastal seawater during summer 2015 (Table 4.1). Estuary samples were collected from Netarts Bay, a shallow, tidally dominated bay located on the northern Oregon coast, and from Yaquina Bay, a drowned river estuary in central Oregon. Estuarine sampling was designed to encompass a range of tidal height, time of day, offshore wind stress, and local chemistry. Additionally, coastal seawater samples were collected on research cruises in October 2014 and September 2015.

In Netarts Bay, samples were collected from tidal flat stations and from inflowing seawater and larval rearing tanks at the Whiskey Creek Shellfish Hatchery (WCSH), located on the eastern edge of the bay (Fig. 4.1). Seawater is continuously pumped from the bay into WCSH through a pipe located ~0.5m above a seagrass bed where the water depth is on average ~2m. Seawater samples were collected from the WCSH outflow of this pipe (~5 second pipe residence time), which also flows through an analytical system measuring T, S, the partial pressure of carbon dioxide and total carbon dioxide (P_{CO2} and TCO_2 respectively). Additionally, water samples were collected from WCSH larval rearing tanks (Table 4.2). These tanks are filled once every ~3 days with seawater which is collected via the inflow pipe and sand filtered, UV-treated, and heated to 25°C; subsequently, tanks are inoculated with larvae or sperm and eggs. Thus, larval tank samples reflect Netarts Bay seawater that has been chemically and biologically altered. On three occasions during low tide, seawater and sediment samples were also collected from stations in the Netarts Bay tidal flats. Sampling locations included tidal channels, isolated seawater pools, seagrass beds, mudflats, and sand flats (Fig. 4.1).

In Yaquina Bay, samples were collected from the Oregon State University pumphouse dock, located ~2.5 km from the mouth of the Yaquina river (44.62°N, –124.04°W). A hand-held niskin sampling bottle was used to collect samples ~1m from above the bottom, where a YSI 6600 series sonde equipped with T and S sensors was moored. The average sea level at this station was ~3m.

Coastal seawater samples were collected at continental shelf and slope stations on and near the Newport Hydroline on cruises of opportunity in September 2015 (*R/V Elahka*) and October 2015 (*R/V Oceanus*) (Table 4.1). On both cruises, samples were collected with Nisken sampling bottles attached to a CTD (conductivity, temperature, depth) rosette. Sampling depths targeted the surface mixed layer, bottom water, and the oxycline.

4.6.2 DNA preservation, extraction, amplification, and sequencing

Seawater and tank water samples used for subsequent DNA extraction were sampled in duplicate into triple-rinsed 1–2 L dark polycarbonate bottles, and 300–1700 mL was immediately filtered onto 25 or 47 mm diameter, 0.2 μm polyethersulfone Supor filters (Pall Corporation) using a peristaltic pump. Filters were placed into microcentrifuge tubes, flash-frozen, and transported in liquid nitrogen to Oregon State University, where they were stored at –80°C until analysis. DNA was extracted from filters using the DNeasy Plant MiniKit (Qiagen), with a modified protocol to include additional steps for cell disruption through flash-freezing, bead-beating with 200 μL of mixed 0.1 mm and 0.5 mm glass beads (Biospec products), and proteinase K treatment. For Netarts Bay sediment samples, acid-washed plastic syringe corers were used to

collect duplicate samples of the top 1 mm of sediment from each station. Sediment samples were transferred into Whirl-Pak bags (Nasco) and transported on dry ice to Oregon State University where they were stored at –80°C. DNA was extracted from 0.25 g of sediment using the DNeasy PowerSoil DNA isolation kit (Qiagen) according to manufacturer instructions for wet soil samples. DNA extracts were quantified with the Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen) using a MicroMax 384 plate reading fluorometer and stored at –20°C or –80°C.

The polymerase chain reaction (PCR) was used to amplify a 491 bp region of the 16S rRNA gene using a *Vibrionaceae*-specific primer set. DNA was amplified using a two-stage "targeted amplicon sequencing" approach (Bybee et al. 2011; Green et al. 2015). The first-stage primers contained gene-specific regions VF169 (Yong et al. 2006) and 680R (Thompson et al. 2004) and common sequence tags, as described previously (Moonsamy et al. 2013). These reactions were performed using DNAEngine (BioRad) thermocyclers and 15 μL reaction volumes consisting of 1X HotStarTaq *Plus* Master Mix (Qiagen), 1 μL DNA extract (diluted 1:10 in PCR-clean water), and 0.2 μM forward and reverse primers. Reactions were cycled at 95°C for 2 min, followed by 35 cycles of 95°C for 15 s, 53°C for 30 s, and 72°C for 30 s, with a final 72°C extension for 10 min. For each sample, PCR reactions were run in triplicate, visualized by gel electrophoresis, pooled, and quantified as above. Samples were only sequenced if they had three successful PCR reactions, excluding PCR negative controls and filter blank samples, which were sequenced despite the absence of visual gel bands after amplification.

PCR amplicons were shipped on dry ice to the DNA Services Facility at the University of Illinois at Chicago for further processing. Here, a second PCR amplification

was performed using Access Array Barcode Library primers (Fluidigm) containing common sequence linkers, unique barcodes (reverse primer only) and Illumina adapters. These reactions were performed in 10 μL reaction volume using MyTaq HS 2X mastermix (Bioline), and were cycled at 95°C for 5 min, followed by 8 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec. PCR products were purified and normalized using SequalPrep plates (Life Technologies), quantified using a Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher) with a GENios Pro Fluorescence Microplate Reader (Tecan), and pooled using an epMotion5075 liquid handling workstation (Eppendorf). Pooled libraries were spiked with 15% phiX sequenced using MiSeq Standard v.3, 2×300 bp paired-end sequencing. The sequencing reaction was initiated using the Fluidigm sequencing primers targeting the gene-specific primer and common sequence tag regions. De-multiplexing of reads was performed on instrument. Sequencing was performed at the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign (UIUC).

4.6.3 Bioinformatic analyses

Sequence reads from *Vibrio* spp. 16S rRNA amplicons were demultiplexed using the Illumina MiSeq Reporter (MSR) version 2.5.1. Paired-end reads were merged via the make.contigs command in mothur (Schloss et al. 2009) using a deltaq value of 20 as an additional quality control measure because of the relatively short (~54 bp) overlapping region for forward and reverse reads. This option discards reads through the merging process when forward and reverse reads have different bases at the same position and the difference in quality score is <20. Merged reads were screened for quality using

screen.seqs in mothur, retaining sequences between 480–540 bp in length with no ambiguities and or homopolymers of >8 bp. Singletons were removed, operational taxonomic units (OTUs) were clustered at 97% nucleotide sequence similarity, and a chimera check was performed against the Gold ChimeraSlayer reference database using USearch (Edgar, 2010). Sequences were subsampled to 7145 sequences per sample.

Vibrio spp. OTUs with >100 sequences in the rarefied dataset were aligned against sequences from 134 cultured *Vibrionaceae* species, and a maximum likelihood phylogenetic tree was constructed using PhyXML (Guindon et al. 2010). Nonmetric multidimensional scaling (NMDS) analyses and alpha diversity calculations were performed using QIIME (Caporaso et al. 2010).

4.6.4 Vibrio spp. and heterotrophic bacteria quantification

Droplet digital PCR (ddPCR) was used to quantify total *Vibrio* spp. 16S rRNA copies using a Bio-Rad QX200 system. Reactions consisted of 10 μL Evagreen PCR mastermix (Bio-Rad), 200 nM 567F and 680R primers (Thompson et al. 2004), and ~2 ng DNA, with a total reaction volume of 20 μL. Droplet generation, PCR, and scanning were conducted at the Oregon State University Center for Genome Research and Biocomputing, according to manufacturer instructions (BioRad) but using an annealing temperature of 57°C. Unlike qPCR, ddPCR produces absolute concentrations without the need for standards. Data were analysed using the QuantaSoft analysis software package. Filter blank samples (clean filters which underwent extraction and processing steps) were included on the run, and a detection limit was calculated as the average plus three standard deviations of triplicate filter blanks. Droplet digital PCR was also used to detect

V. coralliilyticus using a primer set designed to quantify the *vcpA* metalloprotease gene of *V. coralliilyticus* (formerly *V. tubiashii*) strain RE22 (Gharaibeh et al. 2009) for a subset of samples. However, *vcpA* concentrations were all below or near our calculated detection limit; these data were not further analysed.

Heterotrophic bacterial cell densities in seawater samples were measured using flow cytometry. Duplicate 3 mL subsamples were pipetted into cryovials and fixed with fresh paraformaldehyde at a final concentration of 1% (volume volume^{-1}). Fixed samples were inverted and incubated at room temperature in the dark for 10 min, then flash-frozen and transported in liquid nitrogen to Oregon State University and stored at –80°C. For analysis, samples were thawed on ice in the dark then spiked with Fluoresbrite 1 mm beads, stained with SYBR Green I (Marie et al. 1997), and run on a Becton-Dickinson FASCaliber flow cytometer with a 488 nm laser. Bacterial cells and beads were distinguished from other particulate matter by their side light scatter and green fluorescence.

4.6.5 Ancillary data

Discrete samples were preserved for Chl *a*, nitrate, and phosphate. For Chl *a*, 50 mL seawater was filtered onto 25 mm GF/F filters (Whatman), which were placed in snap-cap tubes, wrapped in foil and flash-frozen. For nutrient samples, 25 mL seawater was frozen in high-density polyethylene bottles. Samples for Chl *a* and nutrients were transported in liquid nitrogen to OSU and stored at –80°C until analysis. Chl *a* was extracted in acetone at –20°C for 48 h, then analyzed with a Turner Model 10-AU fluorometer using the methods of Welschmeyer (1994). Nutrient samples were thawed,

filtered through 25 mm GF/F filters (Whatman), and analyzed via phosphomolybdic acid reduction for phosphate and cadmium reduction (Armstrong et al. 1967) for N + N using a Technicon Auto Analyzer II.

At WCSH, the P_{CO2} and TCO_2 of hatchery inflow water were measured real-time using the Burke-O-Lator 3000 (see Barton et al. 2015). Tidal heights were estimated for the Yaquina Bay and Netarts Bay WCSH intake time points using the program xtide (available at http://www.flaterco.com/xtide/) using "South Beach, Yaquina Bay, Oregon" and "Netarts, Netarts Bay, Oregon" sites, respectively. Daily wind stress was derived from winds observed at Newport, Oregon (http://damp.coas.oregonstate.edu/windstress/).

4.6.6 Statistical analyses

Concentrations of bacteria, *Vibrio* spp., and *V. coralliilyticus* were log-transformed prior to all regressions in order to improve model assumptions of normality and equal variance. One-way ANOVA with subsequent Tukey Honest Significance Difference (HSD) tests of multiple comparisons were performed to test for differences in concentrations among sample types, and to compare tests for differences in concentrations between days and stations. All statistical tests were performed using the program R (http://www.r-project.org/).

Four separate type II linear regression models were used to test for environmental predictors of *V. coralliilyticus*, *Vibrio* spp., and total heterotrophic bacteria (volumetric concentrations), as well as for the ratio of *V. coralliilyticus* to total heterotrophic bacteria. Graphical analysis did not indicate strong co-linearity between explanatory variables tested (temperature, salinity, daily wind stress, discrete tide, P_{CO2}, nitrate, and phosphate).

Initial models incorporated all explanatory variables, then subsequent models were reduced to only incorporate variables with statistically significant predictive power ($p < 0.05$); summary statistics are reported from the reduced models. The relative contribution of each explanatory variable to the total R^2 value of each model was calculated using the R package relaimpo (Grömping, 2006). Analysis of residuals using the acf function in R indicated no problems with temporal autocorrelation.

Support for this project was provided by the Environmental Protection Agency Science to Achieve Results (Award #F13E20903). We thank Alan Barton and Sue Cudd from Whiskey Creek Shellfish Hatchery for help with sampling efforts and D. Reinert for producing Fig. 4.1.

4.7 References

Amin, A., Feng, G., Al-saari, N., Meirelles, P.M., Yamazaki, Y., Mino, S., et al. (2016). The first temporal and spatial assessment of *Vibrio* diversity of the surrounding seawater of coral reefs in Ishigaki, Japan. *Frontiers in Microbiology* 7, 1185.

Armstrong, F., Stearns, C., and Strickland, J. (1967). The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. *Deep Sea Research and Oceanographic Abstracts* 14(3), 381-389.

Asplund, M.E., Baden, S.P., Russ, S., Ellis, R.P., Gong, N., and Hernroth, B.E. (2014). Ocean acidification and host–pathogen interactions: blue mussels, *Mytilus edulis*, encountering *Vibrio tubiashii*. *Environmental microbiology* 16(4), 1029-1039.

Barton, A., Hales, B., Waldbusser, G.G., Langdon, C., and Feely, R.A. (2012). The Pacific oyster, *Crassostrea gigas*, shows negative correlation to naturally elevated carbon dioxide levels: Implications for near-term ocean acidification effects. *Limnology and Oceanography* 57(3), 698-710.

Barton, A., Waldbusser, G.G., Feely, R.A., Hales, B., and Langdon, C.J. (2015). Impacts of coastal acidification on the Pacific Northwest shellfish industry and adaptation strategies implemented in response. *Oceanography* 28(2), 146-159.

Ben-Haim, Y., and Rosenberg, E. (2002). A novel *Vibrio* sp. pathogen of the coral *Pocillopora damicornis*. *Marine Biology* 141(1), 47-55.

Ben-Haim, Y., Thompson, F., Thompson, C., Cnockaert, M., Hoste, B., Swings, J., et al. (2003a). *Vibrio coralliilyticus* sp. nov., a temperature-dependent pathogen of the coral *Pocillopora damicornis*. *International Journal of Systematic and Evolutionary Microbiology* 53(1), 309-315.

Ben-Haim, Y., Zicherman-Keren, M., and Rosenberg, E. (2003b). Temperature-regulated bleaching and lysis of the coral *Pocillopora damicornis* by the novel pathogen *Vibrio coralliilyticus*. *Applied and Environmental Microbiology* 69(7), 4236-4242.

Bybee, S.M., Bracken-Grissom, H., Haynes, B.D., Hermansen, R.A., Byers, R.L., Clement, M.J., et al. (2011). Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. *Genome Biology and Evolution* 3, 1312-1323.

Cangelosi, G.A., Freitag, N.E., and Riley-Buckley, M.S. (2004). *From outside to inside: environmental microorganisms as human pathogens.* Washington, D.C. : American Society for Microbiology.

Chase, E., Young, S., and Harwood, V.J. (2015). Sediment and vegetation as reservoirs of *Vibrio vulnificus* in the Tampa Bay Estuary and Gulf of Mexico. *Applied and Environmental Microbiology* 81(7), 2489-2494.

Colwell, R.R. (1993). Nonculturable but still viable and potentially pathogenic. *Zentralblatt für Bakteriologie* 279(2), 154-156.

de O Santos, E., Alves, N., Dias, G.M., Mazotto, A.M., Vermelho, A., Vora, G.J., et al. (2011). Genomic and proteomic analyses of the coral pathogen *Vibrio coralliilyticus* reveal a diverse virulence repertoire. *The ISME Journal* 5(9), 1471-1483.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19), 2460-2461.

Eilers, H., Pernthaler, J., and Amann, R. (2000). Succession of pelagic marine bacteria during enrichment: a close look at cultivation-induced shifts. *Applied and Environmental Microbiology* 66(11), 4634-4640.

Elston, R.A., Hasegawa, H., Humphrey, K.L., Polyak, I.K., and Hase, C. (2008). Re-emergence of *Vibrio tubiashii* in bivalve shellfish aquaculture: severity, environmental drivers, geographic extent and management. *Diseases of Aquatic Organisms* 82(2), 119.

Estes, R.M., Friedman, C.S., Elston, R.A., and Herwig, R.P. (2004). Pathogenicity testing of shellfish hatchery bacterial isolates on Pacific oyster *Crassostrea gigas* larvae. *Diseases of Aquatic Organisms* 58(2), 223-230.

Feely, R.A., Sabine, C.L., Hernandez-Ayon, J.M., Ianson, D., and Hales, B. (2008). Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science* 320(5882), 1490-1492.

Gay, M., Berthe, F., and Leroux, F. (2004). Screening of *Vibrio* isolates to develop an experimental infection model in the Pacific oyster *Crassostrea gigas*. *Diseases of Aquatic Organisms* 59(1), 49-56.

Gharaibeh, D.N., Hasegawa, H., and Häse, C.C. (2009). Development of a quantitative real-time PCR assay for detection of *Vibrio tubiashii* targeting the metalloprotease gene. *Journal of Microbiological Methods* 76(3), 262-268.

Gomez-Gil, B., Thompson, C.C., Matsumura, Y., Sawabe, T., Iida, T., Christen, R., et al. (2014). "The Family Vibrionaceae," in *The Prokaryotes – Gammaproteobacteria,* ed. E. Rosenberg. (Berlin: Springer-Verlag), 659-747.

Green, S.J., Venkatramanan, R., and Naqib, A. (2015). Deconstructing the polymerase chain reaction: understanding and correcting bias associated with primer degeneracies and primer-template mismatches. *PLoS One* 10(5), e0128122.

Grömping, U. (2006). Relative importance for linear regression in R: the package Relaimpo. *Journal of Statistical Software* 17(1), 1-27.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59(3), 307-321.

Hasegawa, H., Lind, E.J., Boin, M.A., and Häse, C.C. (2008). The extracellular metalloprotease of *Vibrio tubiashii* is a major virulence factor for Pacific oyster (*Crassostrea gigas*) larvae. *Applied and Environmental Microbiology* 74(13), 4101-4110.

Heidelberg, J., Heidelberg, K., and Colwell, R. (2002). Bacteria of the γ-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. *Applied and Environmental Microbiology* 68(11), 5498-5507.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406(6795), 477-483.

Hill, J., and Wheeler, P. (2002). Organic carbon and nitrogen in the northern California current system: comparison of offshore, river plume, and coastally upwelled waters. *Progress in Oceanography* 53(2), 369-387.

Hunt, D.E., Gevers, D., Vahora, N.M., and Polz, M.F. (2008). Conservation of the chitin utilization pathway in the Vibrionaceae. *Applied and Environmental Microbiology* 74(1), 44-51.

Islam, M., Drasar, B., and Bradley, D. (1990). Long-term persistence of toxigenic *Vibrio cholerae* 01 in the mucilaginous sheath of a blue-green alga, *Anabaena variabilis*. *The Journal of Tropical Medicine and Hygiene* 93(2), 133-139.

Kimes, N.E., Grim, C.J., Johnson, W.R., Hasan, N.A., Tall, B.D., Kothary, M.H., et al. (2012). Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *The ISME Journal* 6(4), 835-846.

Koch, R. (1884). An address on cholera and its bacillus. *British Medical Journal* 2(1236), 453.

Lemire, A., Goudenège, D., Versigny, T., Petton, B., Calteau, A., Labreuche, Y., et al. (2015). Populations, not clones, are the unit of *Vibrio* pathogenesis in naturally infected oysters. *The ISME Journal* 9(7), 1523-1531.

Lyons, M.M., Lau, Y.-T., Carden, W.E., Ward, J.E., Roberts, S.B., Smolowitz, R., et al. (2007). Characteristics of marine aggregates in shallow-water ecosystems: implications for disease ecology. *EcoHealth* 4(4), 406-420.

Marie, D., Partensky, F., Jacquet, S., and Vaulot, D. (1997). Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Applied and Environmental Microbiology* 63(1), 186-193.

McKibben, S.M., Peterson, W., Wood, A.M., Trainer, V.L., Hunter, M., and White, A.E. (2017). Climatic regulation of the neurotoxin domoic acid. *Proceedings of the National Academy of Sciences* 114(2), 239-244.

McKibben, S.M., Watkins-Brandt, K.S., Wood, A.M., Hunter, M., Forster, Z., Hopkins, A., et al. (2015). Monitoring Oregon Coastal Harmful Algae: Observations and implications of a harmful algal bloom-monitoring project. *Harmful Algae* 50, 32-44.

Mersni-Achour, R., Cheikh, Y.B., Pichereau, V., Doghri, I., Etien, C., Dégremont, L., et al. (2015). Factors other than metalloprotease are required for full virulence of French *Vibrio tubiashii* isolates in oyster larvae. *Microbiology* 161, 997-1007.

Moonsamy, P., Williams, T., Bonella, P., Holcomb, C., Höglund, B., Hillman, G., et al. (2013). High throughput HLA genotyping using 454 sequencing and the Fluidigm Access Array™ system for simplified amplicon library preparation. *Tissue Antigens* 81(3), 141-149.

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Annual Review of Marine Science* 4, 449-466.

Peterson, W., Bond, N., and Robert, M. (2016). The Blob (Part Three): Going, going, gone? *PICES Press* 24(1), 46.

Prouty, M.G., and Klose, K.E. (2006). *Vibrio cholerae*: the genetics of pathogenesis and environmental persistence. *The Biology of Vibrios*, 311-339.

Richards, G.P., Watson, M.A., Needleman, D.S., Church, K.M., and Häse, C.C. (2015). Mortalities of Eastern and Pacific oyster larvae caused by the pathogens *Vibrio*

coralliilyticus and *Vibrio tubiashii*. *Applied and Environmental Microbiology* 81(1), 292-297.

Sainz-Hernández, J.C., and Maeda-Martínez, A.N. (2005). Sources of *Vibrio* bacteria in mollusc hatcheries and control methods: a case study. *Aquaculture Research* 36(16), 1611-1618.

Sawabe, T., Ogura, Y., Matsumura, Y., Feng, G., Amin, A.R., Mino, S., et al. (2013). Updating the *Vibrio* clades defined by multilocus sequence phylogeny: proposal of eight new clades, and the description of *Vibrio tritonius* sp. nov.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* 75(23), 7537-7541.

Siboni, N., Balaraju, V., Carney, R., Labbate, M., and Seymour, J.R. (2016). Spatiotemporal dynamics of *Vibrio* spp. within the Sydney Harbour Estuary. *Frontiers in Microbiology* 7.

Sussman, M., Mieog, J.C., Doyle, J., Victor, S., Willis, B.L., and Bourne, D.G. (2009). *Vibrio* zinc-metalloprotease causes photoinactivation of coral endosymbionts and coral tissue lesions. *PLoS One* 4(2), e4511.

Sutton, D., and Garrick, R. (1993). Bacterial disease of cultured giant clam *Tridacna gigas* larvae. *Diseases of Aquatic Organisms* 16(1), 47-53.

Takemura, A.F., Chien, D.M., and Polz, M.F. (2014). Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Frontiers in Microbiology* 5. doi: 10.3389/fmicb.2014.00038.

Thompson, J.R., Randa, M.A., Marcelino, L.A., Tomita-Mitchell, A., Lim, E., and Polz, M.F. (2004). Diversity and dynamics of a North Atlantic coastal *Vibrio* community. *Applied and Environmental Microbiology* 70(7), 4103-4110.

Tout, J., Siboni, N., Messer, L.F., Garren, M., Stocker, R., Webster, N.S., et al. (2015). Increased seawater temperature increases the abundance and alters the structure of natural *Vibrio* populations associated with the coral *Pocillopora damicornis*. *Frontiers in Microbiology* 6.

Turner, J.W., Good, B., Cole, D., and Lipp, E.K. (2009). Plankton composition and environmental factors contribute to *Vibrio* seasonality. *The ISME Journal* 3(9), 1082-1092.

Urakawa, H., and Rivera, I.N.G. (2006). Aquatic Environment. *The Biology of Vibrios*, 175-189.

Vezzulli, L., Colwell, R.R., and Pruzzo, C. (2013). Ocean warming and spread of pathogenic vibrios in the aquatic environment. *Microbial Ecology* 65, 817-825.

Vezzulli, L., Pezzati, E., Moreno, M., Fabiano, M., Pane, L., Pruzzo, C., et al. (2009). Benthic ecology of Vibrio spp. and pathogenic Vibrio species in a coastal Mediterranean environment (La Spezia Gulf, Italy). *Microbial ecology* 58, 808-818.

Waldbusser, G.G., Hales, B., Langdon, C.J., Haley, B.A., Schrader, P., Brunner, E.L., et al. (2014). Saturation-state sensitivity of marine bivalve larvae to ocean acidification. *Nature Climate Change* 5, 273–280.

Waldbusser, G.G., and Salisbury, J.E. (2014). Ocean acidification in the coastal zone from an organism's perspective: multiple system parameters, frequency domains, and habitats. *Ann. Rev. Mar. Sci.* 6, 221-247.

Welschmeyer, N.A. (1994). Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. *Limnol. Oceanogr.* 39, 1985-1992.

Wendling, C.C., Batista, F.M., and Wegner, K.M. (2014). Persistence, seasonal dynamics and pathogenic potential of *Vibrio* communities from Pacific oyster hemolymph. *PLoS One* 9(4), e94256.

Wietz, M., Gram, L., Jørgensen, B., and Schramm, A. (2010). Latitudinal patterns in the abundance of major marine bacterioplankton groups. *Aquatic Microbial Ecology* 61(2), 179-189.

Wilson, B., Muirhead, A., Bazanella, M., Huete-Stauffer, C., Vezzulli, L., and Bourne, D.G. (2013). An improved detection and quantification method for the coral pathogen *Vibrio coralliilyticus*. *PloS one* 8(12), e81800. doi: 10.1371/journal.pone.0081800.

Yong, L., Guanpin, Y., Hualei, W., Jixiang, C., Xianming, S., Guiwei, Z., et al. (2006). Design of *Vibrio* 16S rRNA gene specific primers and their application in the analysis of seawater *Vibrio* community. *Journal of Ocean University of China* 5(2), 157-164.

Zhu, J., Miller, M.B., Vance, R.E., Dziejman, M., Bassler, B.L., and Mekalanos, J.J. (2002). Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* 99(5), 3129-3134.

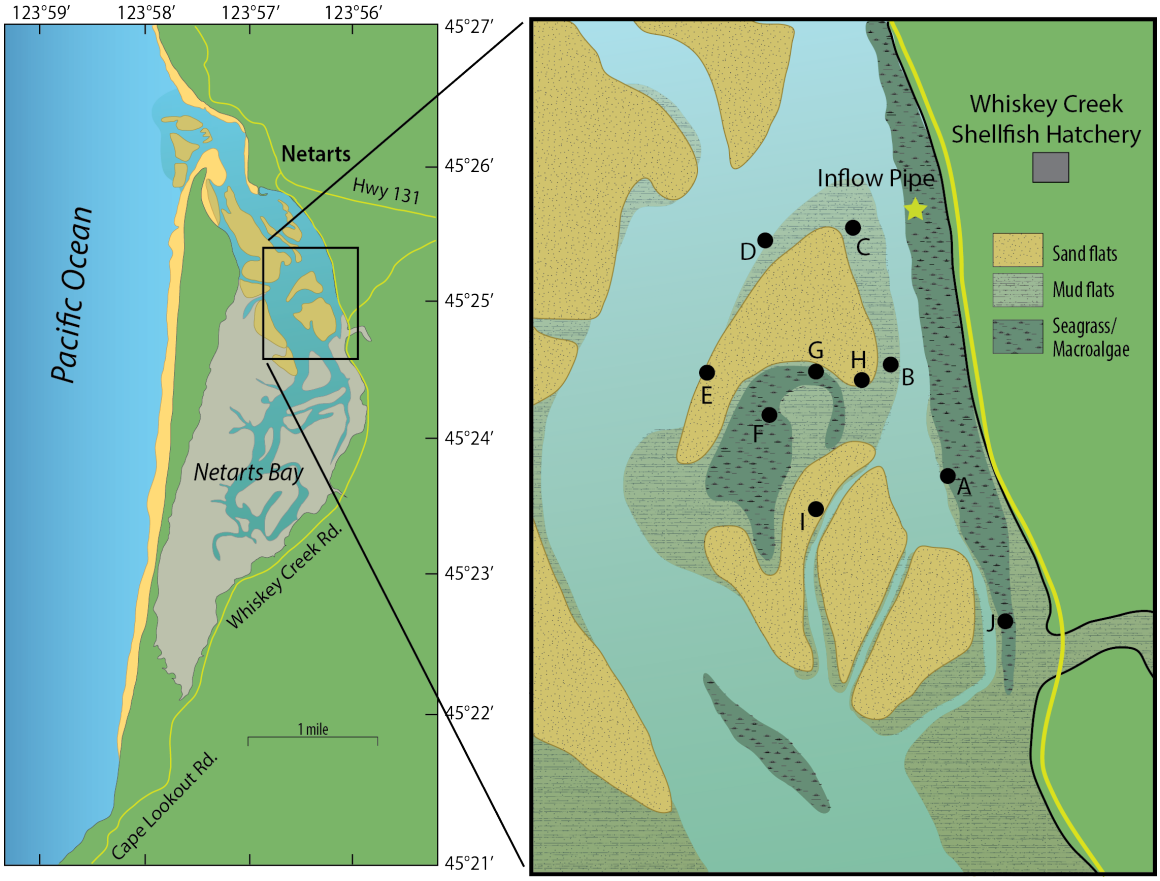

Figure 4.1: Map of Netarts Bay showing locations of Whiskey Creek Shellfish Hatchery (WCSH), the WCSH pipe inlet, and tidal flat sampling stations.

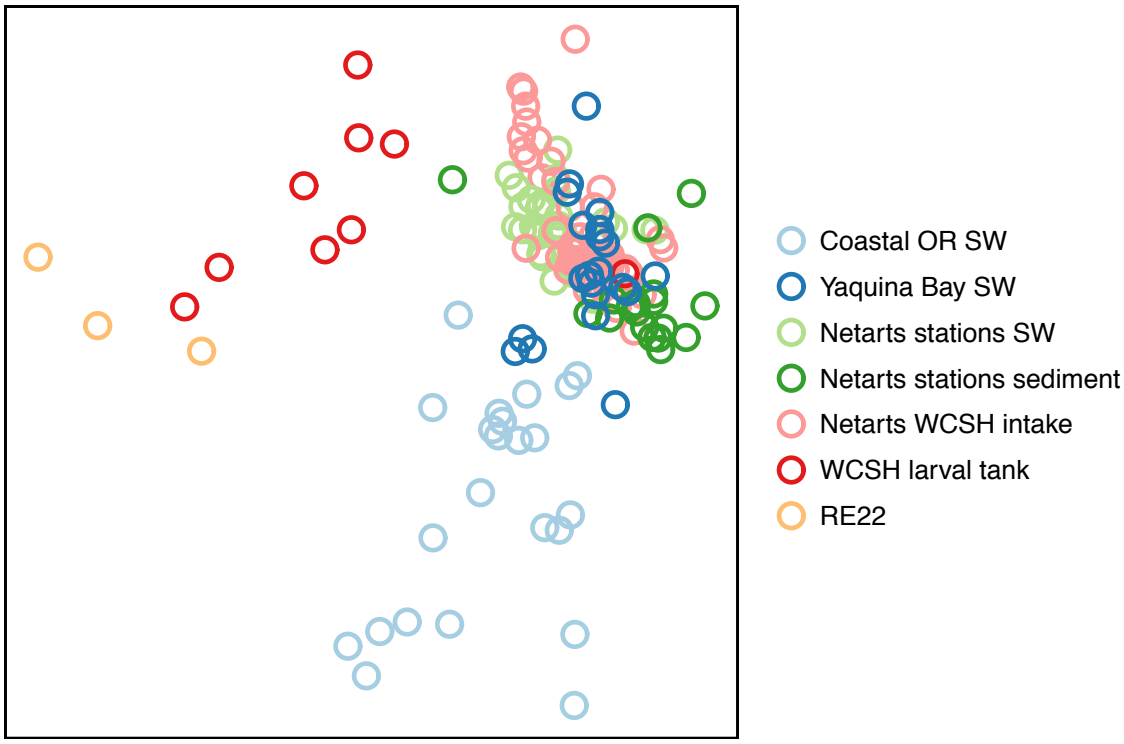

Figure 4.2: Non-metric multi-dimensional scaling (NMDS) plot derived from the Bray–Curtis dissimilarity matrix of *Vibrio* spp. 16S OTUs clustered at 97% identity. Each point represents an individual sample. Colors represent sample type. Three replicate DNA samples from a *V. coralliilyticus* strain RE22 culture were sequenced and included for comparison.

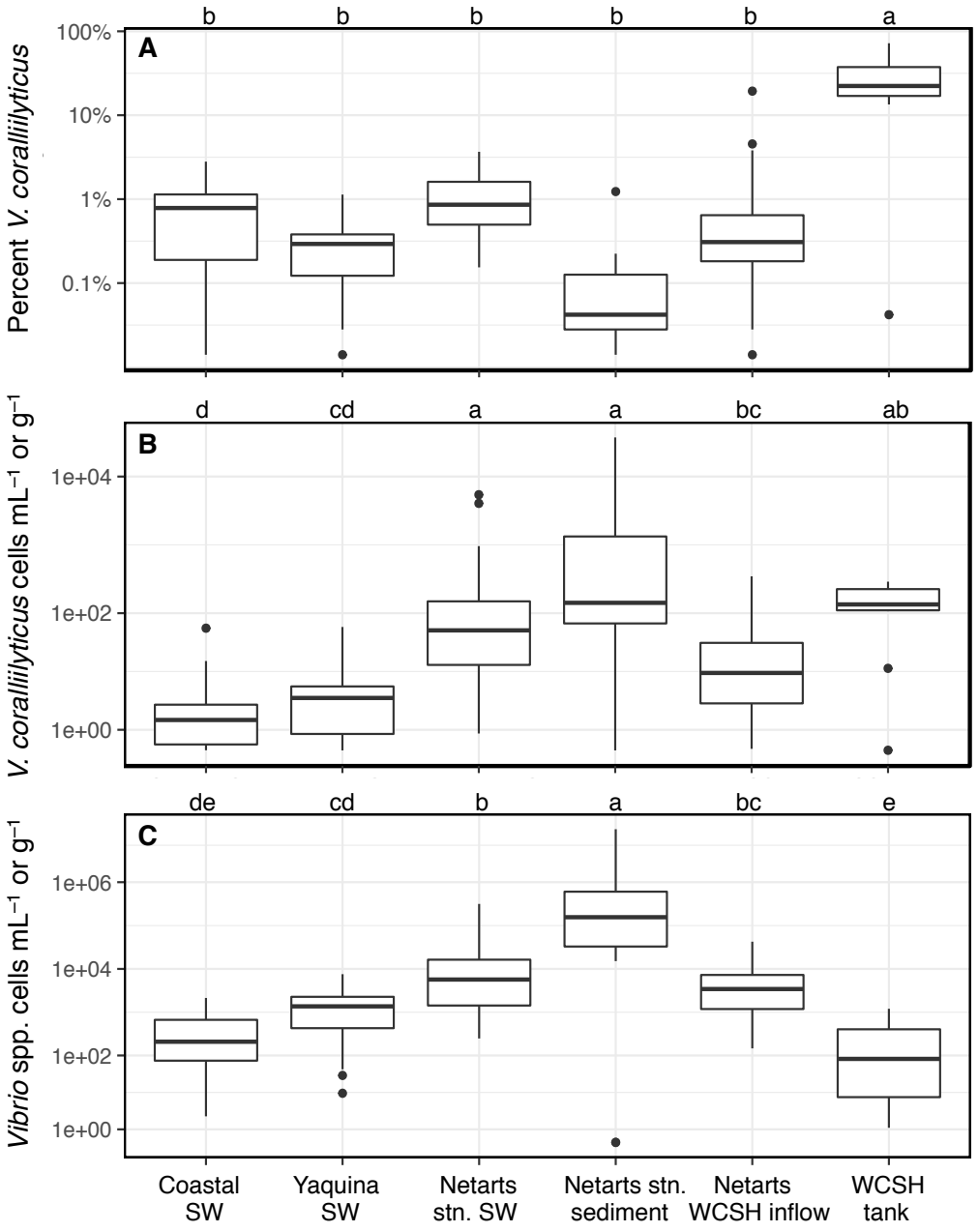

Figure 4.3: The percentage of total *Vibrio* spp. classified as *V. coralliilyticus* (A) and abundance estimates for *V. coralliilyticus* (B) and total *Vibrio spp.* (C). Concentrations are normalized to mL seawater or to gram sediment (Netarts stn. sediment samples only). Letters above each panel note statistical significance, where different letters signify significant differences in median values within a panel (Tukey HSD $p < 0.05$), and categories with the same letter are not statistically different from one another. Boxplots represent medians as thick horizontal lines, 25-75% quantiles as boxes, 1.5 times the inter-quartile range as whiskers, and outliers as dots.

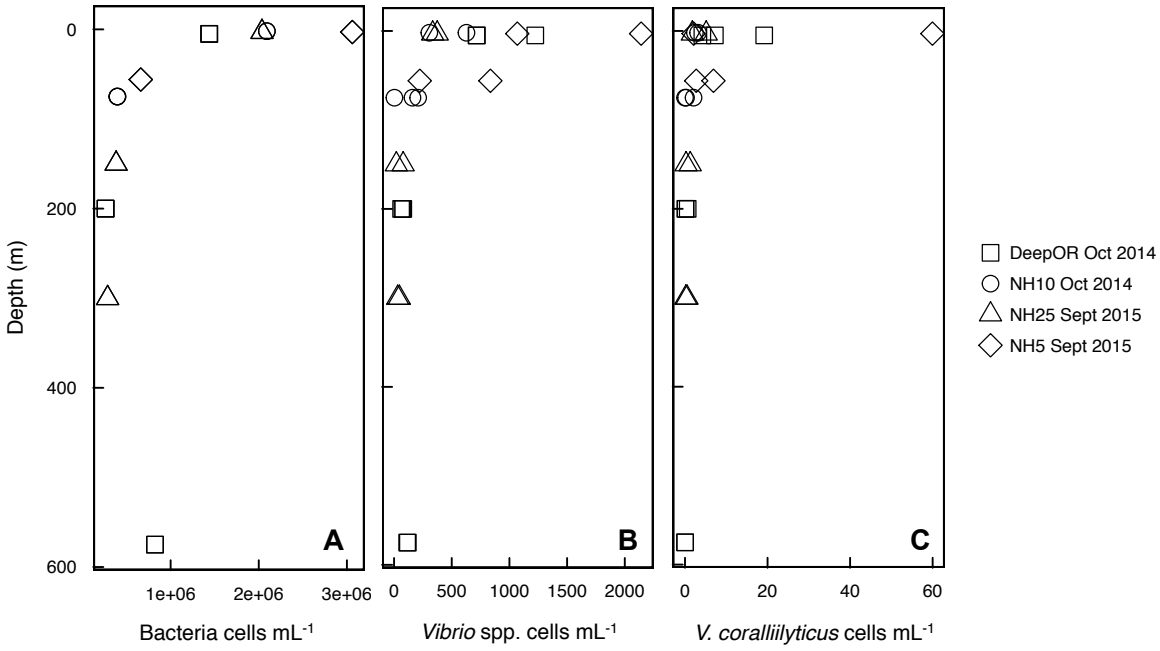

Figure 4.4: Abundances of heterotrophic bacteria (A), *Vibrio* spp. (B) and *V. coralliilyticus* (C) at shelf-break Oregon stations. See Table 4.1 for a description of sampling sites. "DeepOR" signifies Stn. CE0405.

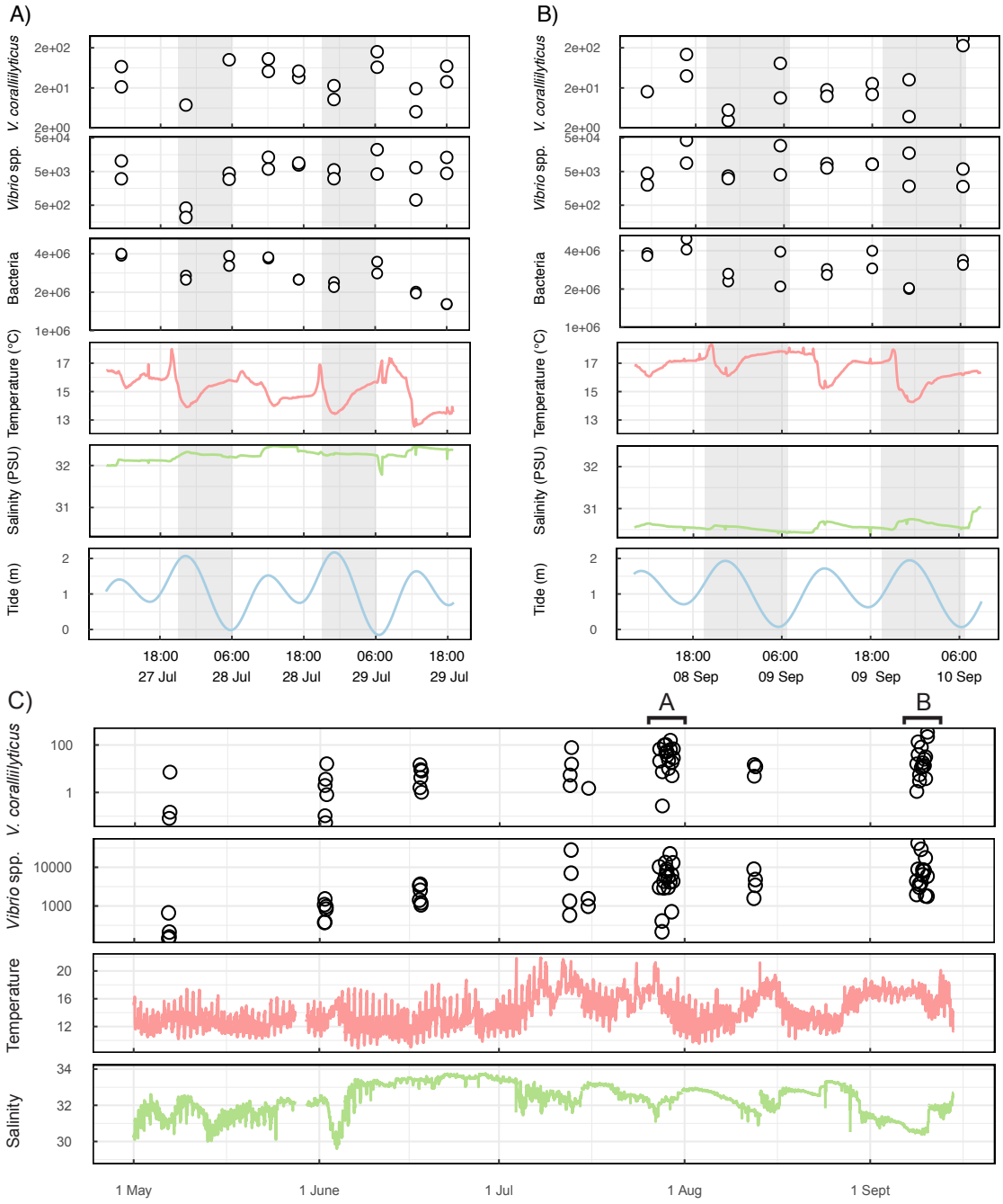

Figure 4.5: Time series from the Netarts WCSH inlet in summer 2015. Concentrations of *Vibrio* spp. and *V. coralliilyticus* (cells mL^{-1}), temperature (T, °C) and salinity (S, PSU) are presented for the entire time series (C), as well as total heterotrophic bacterial counts (cells mL^{-1}) and tidal height (m) for two intensive sampling periods (A, B).

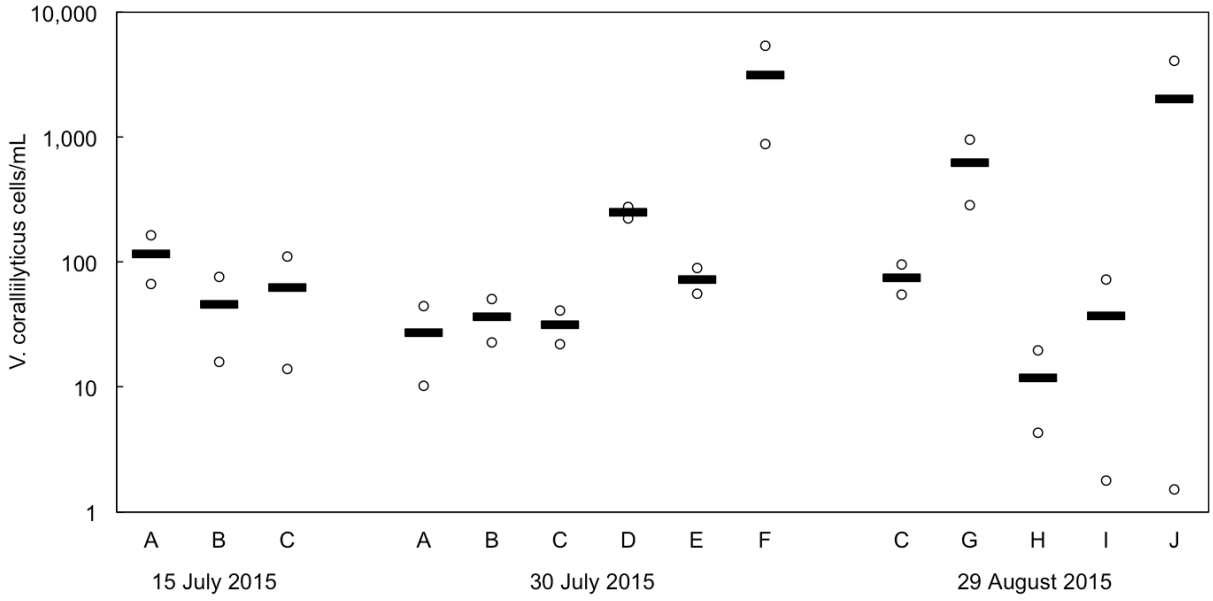

Figure 4.6 *V. coralliilyticus* concentrations in seawater from Netarts tidal flat stations. Stations were sampled during low tide from ~07:00–09:00) on 15 July, 30 July, and 29 August 2015. Circles represent individual duplicate samples; dark bars represent averages. See Fig. 4.1 for locations of sampling stations.

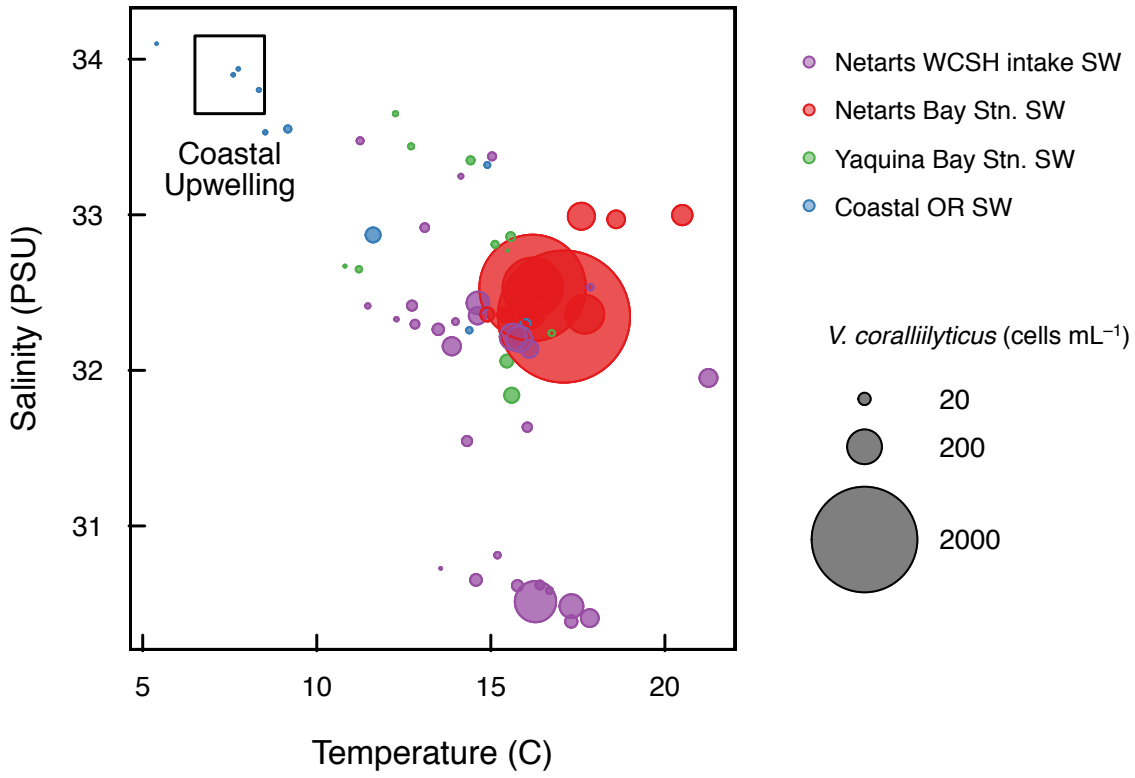

Figure 4.7: Concentrations *V. coralliilyticus* in all seawater samples plotted over the temperature and salinity of seawater. The box represents contains samples from 200m water at Coastal OR Stn. CE0405, thus reflecting the approximate temperature and salinity expected for upwelling source water.

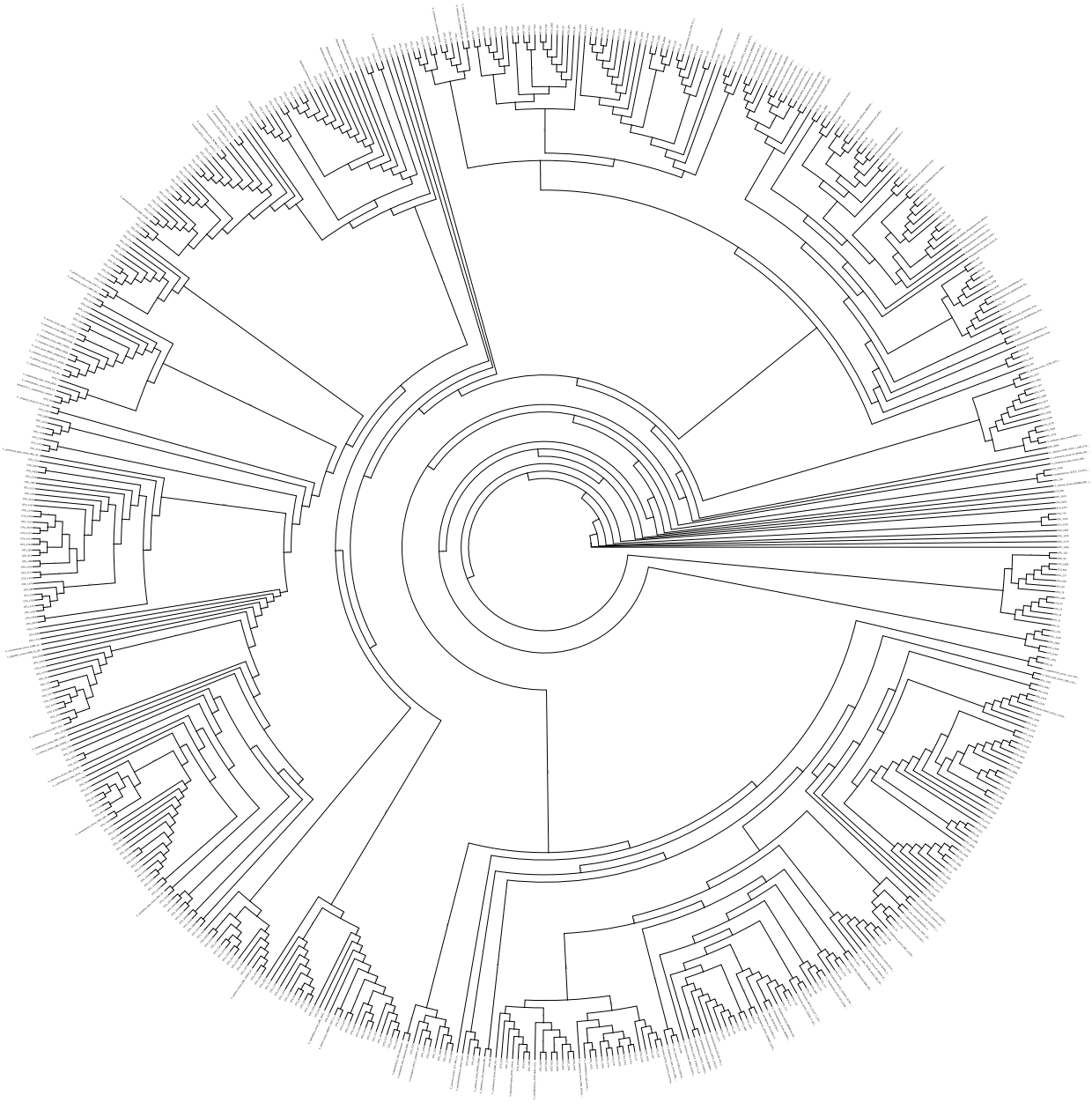

Figure S4.1: Maximum likelihood phylogenetic tree depicting *Vibrio* 16S rRNA gene sequences from this study. A representative sequence from each operational taxonomic unit (OTU) containing >100 sequences in the rarefied dataset (>98% of total rarefied sequences) and sequences from 134 *Vibrionaceae* isolates are displayed on the tree.

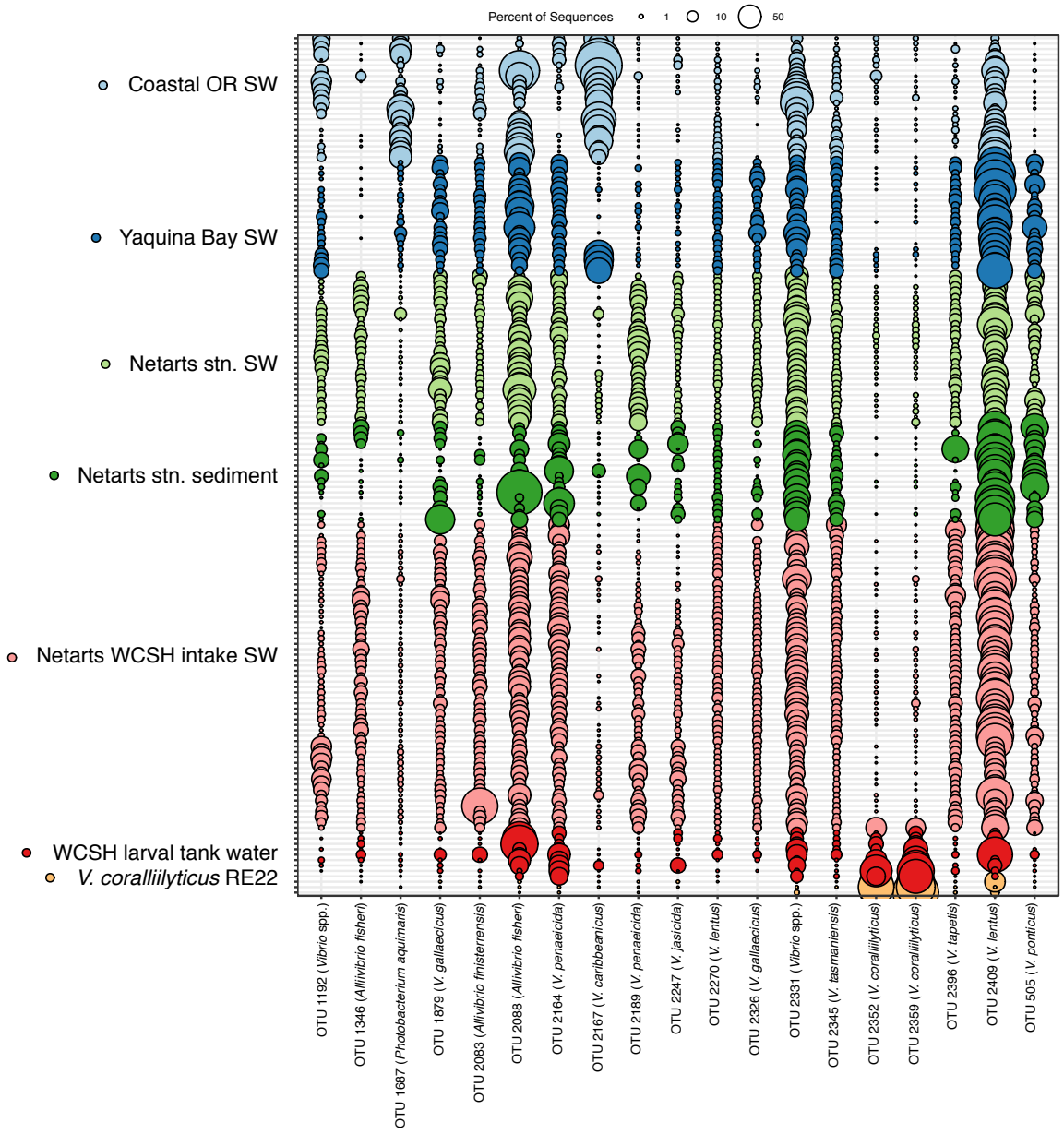

Figure S4.2: Relative abundances of dominant *Vibrio* spp. operational taxonomic units (OTUs) across all DNA samples. The 19 most abundant OTUs are depicted, representing >65% of total sequences from the rarefied dataset. Note that OTUs identified as *V. coralliilyticus* also clustered with *V. neptunius* and *V. neresis*, and that OTUs identified as *V. ponticus* also clustered with *V. alfacsensis*.

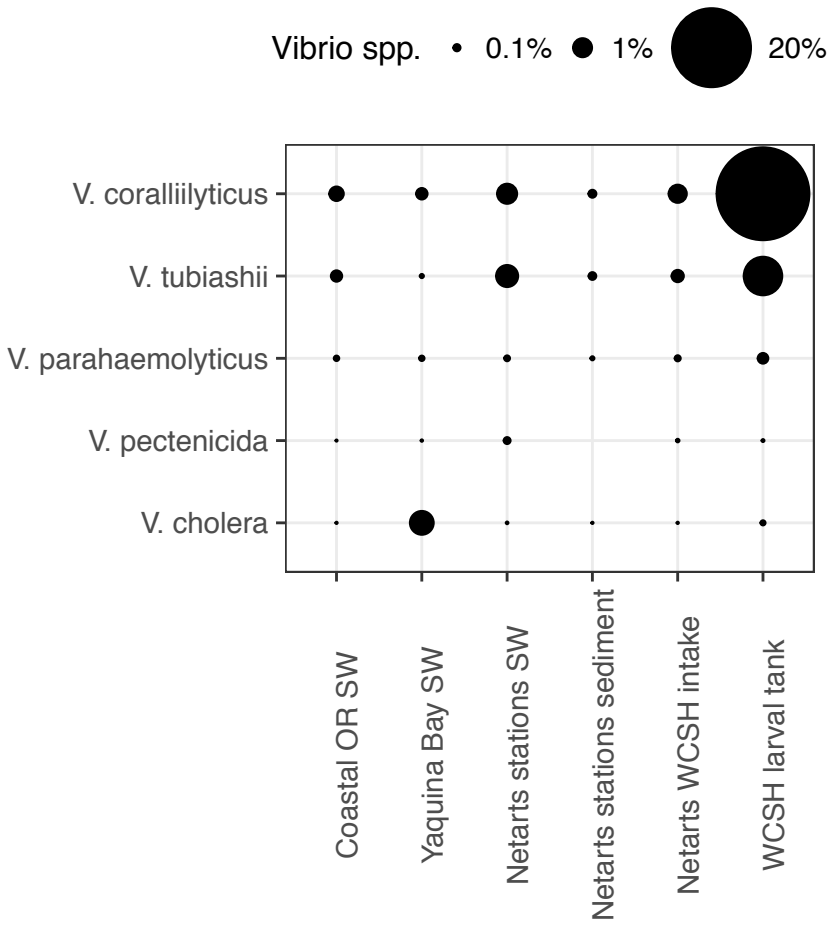

Figure S4.3: Average percentage of total *Vibrio* spp. classified as putative pathogens from the six sample types

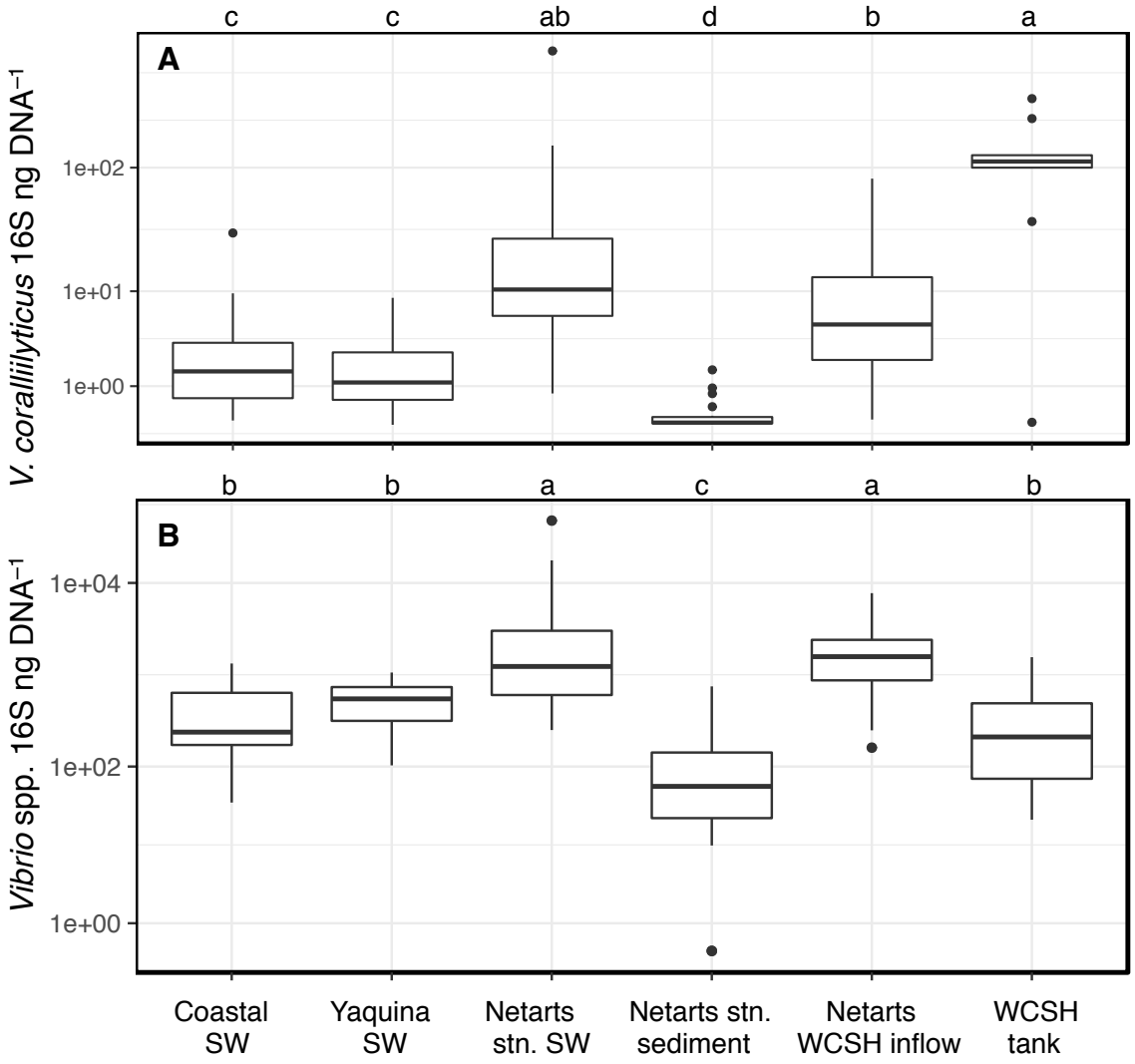

Figure S4.4: DNA-normalized concentrations of *V. coralliilyticus* (A) and total *Vibrio spp.* (B). Letters above each panel note statistical significance, where different letters signify significant differences in median values within a panel (Tukey HSD $p < 0.05$), and categories with the same letter are not statistically different from one another. Boxplots represent medians as thick horizontal lines, 25-75% quantiles as boxes, 1.5 times the inter-quartile range as whiskers, and outliers as dots.

Table 4.1: Description of sampling stations used in this study.

Location	Sample Type(s)	Description	Sampling time period	Days sampled	DNA samples
Netarts WCSH inflow	DNA, FCM, nutrients, Chl *a*, T, S, P_{CO2}	Pipe in WCSH sampled directly from Netarts Bay (Fig. 4.1)	May - Sept 2015	13	62
Netarts Bay tidal flat	DNA (SW), DNA (sediment), FCM, nutrients, Chl *a*, T	Tidal channels, isolated seawater pools, seagrass beds, mudflats, and sand flats sampled at low tide (Fig. 4.1).	July - Aug 2015	3	56
WCSH larval tanks	DNA, FCM	Oyster larval rearing tanks at WCSH	May - Sept 2015	10	19
Yaquina Bay	DNA, FCM, nutrients, Chl *a*, T, S	OSU pumphouse dock; located near the mouth of Yaquina Bay (44.62°N, –124.04°W).	July - Sept 2015	6	22
Coastal OR NH10	DNA, FCM, nutrients, Chl *a*, T, S	Nearshore station; 80 m depth; Located at 44.65°N, – 124.29°W	Oct 2014	1	5
Coastal OR CE0405	DNA, FCM, nutrients, Chl *a*, T, S	Nearshore station; 588 m depth; Located at 44.37°N, –124.95°W	Oct 2014	1	8
Coastal OR NH5	DNA, FCM, nutrients, Chl *a*, T, S, P_{CO2}	Nearshore station; 59 m depth; Located at 44.65°N, –124.18 °W	Sept 2015	1	6
Coastal OR NH25	DNA, FCM, nutrients, Chl *a*, T, S, P_{CO2}	Nearshore station; 293 m depth; Located at 44.65°N, –124.65 °W	Sept 2015	1	6

Table 4.2: Table summarizing conditions of 19 larval tank samples collected from WCSH during summer 2015. * indicates replicate samples, with concentrations representing averages, and standard deviations presented in parentheses. "F" indicates failed PCR.

Tank ID	Date	Larvae size	Food	Time in tank	Bacteria cells mL^{-1}	*Vibrio* spp. cells mL^{-1}	V. *coralliilyticus* % *Vibrio*	V. *coralliilyticus* cells mL^{-1}
2	5 May	D-hinge	Yes	24 h	3.7E+06	2	F	F
4	12 Jul	None	No	< 1 d	1.1E+06	1012	22%	226
5	12 Jul	None	No	< 1 d	1.1E+06	775	17%	132
15	15 July	D-hinge	Yes	3 d	1.5E+06	406	F	F
16	15 July	D-hinge	Yes	3 d	9.9E+05	7	F	F
9 T=0	27 July	Eggs	No	2 h	6.2E+05	110	13%	15
9 T=24	28 July	Eggs	Yes	1 d	7.4E+05	1204	19%	228
13 T=0*	28 July	100-120	Yes	6 h	1.2E+06	11 (0.2)	F	F
9 T=24	29 July	Eggs	Yes	1 d	7.4E+05	1204	19%	228
13 T=40*	30 July	100-120	Yes	2 d	1.6E+06	7 (4)	0% / F	0 / F
10*	12 Aug	90-110	Yes	3 d	6.9E+05	77 (9)	F	F
7*	12 Aug	90-110	Yes	3 d	6.8E+05	29 (39)	F	F
9 T=0	8 Sept	Eggs	No	2 h	1.6E+06	294	37%	110
9 T=19	9 Sept	Eggs	No	1 d	1.4E+06	611	27%	167
3	9 Sept	D-hinge	Yes	3 d	1.6E+06	403	72%	290
9 T=4	10 Sept	Eggs	No	2 d	8.2E+05	322	42%	135

Table 4.3: Statistically significant explanatory variables [β: dissolved inorganic phosphorus (PO_4), daily wind stress (wind,), nitrate + nitrite (N+N), discrete tidal height (Tide, m) and temperature (Temp)], associated *p*-values, and relative contributions of each explanatory variable to the overall R^2 value from four separate linear regression models (rows). Note that all response variables (excluding the ratio of *V. coralliilyticus* to total heterotrophic bacteria) were log-transformed prior to regression analyses.

	Explanatory variables											
Response variable	**β1**	**β sign**	**β1 R^2**	**β1 *p*-value**	**β2**	**β sign**	**β2 R^2**	**β2 *p*-value**	**β3**	**β3 sign**	**β3 R^2**	**β3 *p*-value**
Vibrio spp. cells mL^{-1}	Temp	+	0.26	< 0.001	wind	−	0.14	0.001	PO_4	+	0.11	0.04
V. coralliilyticus cells mL^{-1}	PO_4	+	0.25	< 0.001	wind	−	0.10	0.04	Tide	−	0.10	0.04
V. coralliilyticus heterotrophic bacteria^{-1}	PO_4	+	0.23	< 0.001	wind	−	0.13	0.01	Tide	−	0.09	0.008
Heterotrophic bacteria cells mL^{-1}	PO_4	+	0.06	0.02	N+N	−	0.27	< 0.001	Temp	+	0.29	0.009

Table S4.1: Range of physical, chemical and biological conditions of seawater samples over the summer 2015 sampling period for Netarts Bay and Yaquina Bay, and over the depth profile for Coastal OR stations. ND indicates no data available.

See GradovilleMaryR2017_SupplementaryTables.xlsx

Table S4.2: Average coefficient of variation (CV, 100% × standard deviation / average) for concentrations of total *Vibrio* spp., *V. coralliilyticus*, and total heterotrophic bacterial observed at different temporal and spatial scales in this study. Note that these CV were calculated from absolute concentrations, whereas all regressions analyses used log-transformed concentrations.

See GradovilleMaryR2017_SupplementaryTables.xlsx

6. CONCLUSIONS AND FUTURE DIRECTIONS

Microbial communities exert a strong impact on the marine environment. On a large scale, microorganisms help to drive ocean carbon and nutrient cycles and exert a strong influence on Earth's climate. On smaller scales, certain marine microbes associate in symbioses with host organisms, where they can perform vital functions for the host (i.e. mutualism or commensalism) or invade the host as environmental pathogens. Recent advances in molecular technologies have enabled a detailed view into these microbial communities and the processes that they regulate. The three main body chapters of this dissertation all use genomic tools to examine the structure, function, and drivers of two groups of marine microorganisms, the diazotrophs and *Vibrio* spp. Several conclusions can be drawn from this work, bringing new insights into the ecology and environmental controls of these two groups.

In Chapter II, diazotrophic community structure and function were compared among three contrasting ocean regions. This chapter represents one of the first studies to use high-throughput *nifH* sequencing to examine diazotrophic diversity across ocean biomes; these analyses indicate strong biogeographical patterns for different diazotrophic groups. The primary motivation for this study was the recent controversy over the ecological importance of heterotrophic marine nitrogen fixation, as genes belonging to these organisms have been reported from a number of marine environments where nitrogen fixation would not be expected to occur (e.g. Hewson et al., 2007; Jayakumar et al., 2012). Results from Chapter II suggest that while heterotrophic diazotrophs may be nearly omnipresent in marine waters, their ecological significance is regionally restricted. Furthermore, the sensitivity analysis of nitrogen fixation rate measurements suggests that

some low nitrogen fixation rates previously reported in the mesopelagic may not be indicative of active diazotrophy. Heterotrophic marine diazotrophs appear to be ecologically significant in certain regions (e.g., the Eastern South Pacific) but not others (e.g., the Pacific Northwest coast of the United States). Future culture-based efforts are urgently needed in order to better understand the physiology and ecological controls of heterotrophic diazotrophs.

Chapter III examined the ecology of marine diazotrophs on a much smaller scale, using high-throughput sequencing to probe the diversity and functional potential of communities associated with *Trichodesmium* colonies. Results indicate that colonies included a diverse community of epibionts, which were distinct from microorganisms in the surrounding seawater. Furthermore, metagenomic evidence revealed the potential for epibionts to facilitate *Trichodesmium* nutrient acquisition, and for anaerobic processes (denitrification) within the colonies. Another interesting finding from this work was the presence of putative anaerobic, non-cyanobacterial diazotrophic genes within the colonies. As with heterotrophic communities from the Pacific Northwest in Chapter II, there was no evidence that these non-cyanobacterial diazotrophs were actively fixing nitrogen. However, there was evidence of active nitrogenase expression by heterocystous cyanobacteria associated with the colonies. The results from Chapter III indicate that *Trichodesmium* colonies are inhabited by complex microbial communities, and that community interactions have the potential to influence bulk colony carbon and nitrogen fixation rates.

Chapter IV investigated the diversity and environmental controls of oyster-pathogenic *Vibrio* spp. in two Oregon estuaries (Netarts Bay and Yaquina Bay) and in a

local shellfish hatchery. This was the first study to use non-culture based, high-throughput sequencing methods to examine the diversity of putative pathogenic *Vibrio* spp. in Pacific Northwest estuaries. Results indicated that *V. coralliilyticus* was the dominant putative-oyster pathogen, but samples also contained sequences closely related to other pathogenic sequences, including *V. cholerae*. Future work monitoring the abundances of *Vibrio* pathogens in Pacific Northwest waters may be advisable, as pathogenic *Vibrio* abundances and outbreaks are expected to increase in the coming century (Vezzulli et al., 2013). Furthermore, results from Chapter IV indicate that abundances of the *V. coralliilyticus* pathogen in Netarts Bay do not appear to be driven by the seeding of high concentrations from deep offshore waters, as proposed by Elston et al. (2008), but rather that favorable local conditions and sediment reservoirs likely drive the abundances of this organism in the bay. Finally, results suggest that a local shellfish hatchery is efficient at reducing concentrations of total *Vibrio* spp., but that the *V. coralliilyticus* pathogen appears particularly resistant to cleaning procedures.

Though the chapters of this dissertation explored different microbial communities and ecosystems, they also share several common themes. First, exploring microbial diversity beyond the functional type level was a crucial element of each of the studies, enabling the assessment of dominant diazotrophs across ecosystems (Chapter II), the finding that non-*Trichodesmium* diazotrophs associate with *Trichodesmium* colonies (Chapter III), and the tracking of an oyster-pathogenic *Vibrio* species through the summer upwelling season (Chapter IV). Second, all three chapters suggested that marine particles may create important microniches for marine bacteria. Marine particles may create favorable environments for the non-cyanobacterial, putative heterotrophic diazotrophs

observed in Chapter II (Bombar et al., 2016). Indeed, Chapter III described non-cyanobacterial diazotrophic genes associated with *Trichodesmium* colonies, which themselves constitute marine particles. Likewise, marine particles are common environments of *Vibrio* spp., and may help explain the high heterogeneity of *V. coralliilyticus* abundances in seawater observed in Chapter IV.

A final theme unifying the three chapters of this dissertation relates to microbial biogeography and the famous statement by Baas-Becking (1934): "Everything is everywhere but the environment selects." Chapter II presented an example of this through the heterotrophic diazotrophs, which were present in the Pacific Northwest coastal upwelling regime and the mesopelagic zone of the North Pacific Subtropical Gyre in spite of environmental conditions that would theoretically select against nitrogen fixation. Chapter III reported that *Trichodesmium* colonies were enriched in epibiotic taxa that were relatively rare in the surrounding seawater. Thus, the colonies may serve as an island of refugia for certain copiotrophs, and periodic dispersal from the colonies may help explain why these taxa are rare but present in the surrounding seawater, i.e. the "rare biosphere" (Pedrós-Alió, 2012). Finally, Chapter IV reported the presence of *V. coralliilyticus* in all environments surveyed, including deep shelf-break waters, though the growth of this organism is extremely temperature-dependent. The breadth of environments occupied by microorganisms such as the diazotrophs and pathogenic *Vibrio* spp. reinforces the need for microbial ecologists to better merge assessments of diversity and abundance with microbial rates and processes.

6.1 References

Baas-Becking, L.G.M. (1934). Geobiologie; of inleiding tot de milieukunde.

Bombar, D., Paerl, R.W., and Riemann, L. (2016). Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. *Trends in Microbiology* 24(11)**,** 916-927.

Elston, R.A., Hasegawa, H., Humphrey, K.L., Polyak, I.K., and Hase, C. (2008). Re-emergence of *Vibrio tubiashii* in bivalve shellfish aquaculture: severity, environmental drivers, geographic extent and management. *Diseases of Aquatic Organisms* 82(2)**,** 119.

Hewson, I., Moisander, P.H., Achilles, K.M., Carlson, C.A., Jenkins, B.D., Mondragon, E.A., et al. (2007). Characteristics of diazotrophs in surface to abyssopelagic waters of the Sargasso Sea. *Aquatic Microbial Ecology* 46(1)**,** 15.

Jayakumar, A., Al-Rshaidat, M., Ward, B.B., and Mulholland, M.R. (2012). Diversity, distribution, and expression of diazotroph nifH genes in oxygen-deficient waters of the Arabian Sea. *FEMS microbiology ecology* 82(3)**,** 597-606.

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Annual Review of Marine Science* 4**,** 449-466.

Vezzulli, L., Colwell, R.R., and Pruzzo, C. (2013). Ocean warming and spread of pathogenic vibrios in the aquatic environment. *Microbial Ecology* 65(4)**,** 817-825.

7. BIBLIOGRAPHY

Amin, A., Feng, G., Al-saari, N., Meirelles, P.M., Yamazaki, Y., Mino, S., et al. (2016). The first temporal and spatial assessment of *Vibrio* diversity of the surrounding seawater of coral reefs in Ishigaki, Japan. *Frontiers Microbiol.* 7, 1185.

Apprill, A., McNally, S., Parsons, R., Webe, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* 75, 129–137.

Armstrong, F., Stearns, C., and Strickland, J. (1967). The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. *Deep Sea Research and Oceanographic Abstracts* 14, 381-389.

Asplund, M.E., Baden, S.P., Russ, S., Ellis, R.P., Gong, N., and Hernroth, B.E. (2014). Ocean acidification and host–pathogen interactions: blue mussels, *Mytilus edulis*, encountering *Vibrio tubiashii*. *Env. Microbiol.* 16, 1029-1039.

Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983). The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* 10, 257-263.

Baas-Becking, L.G.M. (1934). Geobiologie; of inleiding tot de milieukunde.

Barton, A., Hales, B., Waldbusser, G.G., Langdon, C., and Feely, R.A. (2012). The Pacific oyster, *Crassostrea gigas*, shows negative correlation to naturally elevated carbon dioxide levels: Implications for near-term ocean acidification effects. *Limnol. and Oceanogr.* 57, 698-710.

Barton, A., Waldbusser, G.G., Feely, R.A., Hales, B., and Langdon, C.J. (2015). Impacts of coastal acidification on the Pacific Northwest shellfish industry and adaptation strategies implemented in response. *Oceanography* 28, 146-159.

Ben-Haim, Y., and Rosenberg, E. (2002). A novel *Vibrio* sp. pathogen of the coral *Pocillopora damicornis*. *Marine Biol.* 141, 47-55.

Ben-Haim, Y., Thompson, F., Thompson, C., Cnockaert, M., Hoste, B., Swings, J., et al. (2003a). *Vibrio coralliilyticus* sp. nov., a temperature-dependent pathogen of the coral *Pocillopora damicornis*. *International Journal of Systematic and Evolutionary Microbiology* 53(1), 309-315.

Ben-Haim, Y., Zicherman-Keren, M., and Rosenberg, E. (2003b). Temperature-regulated bleaching and lysis of the coral *Pocillopora damicornis* by the novel pathogen *Vibrio coralliilyticus*. *Applied and Environmental Microbiology* 69, 4236-4242.

Benavides, M., et al. (2016). Basin. Basin-wide N_2 fixation in the deep waters of the Mediterranean Sea. *Global Biogeochem. Cycles* 30, 952–961.

Benavides, M., Moisander, P.H., Berthelot, H., Dittmar, T., Grosso, O., Bonnet, S. (2015). Mesopelagic N_2 fixation related to organic matter composition in the Solomon and Bismarck Seas (Southwest Pacific). *PloS one* 10: e0143775.

Bentzon-Tilia, M., I. Severin, L. H. Hansen, Riemann, L. (2015). Genomics and ecophysiology of heterotrophic nitrogen-fixing bacteria isolated from estuarine surface water. *MBio* 6: e00929–00915, doi: 10.1128/mBio.00929-15

Bombar, D., Paerl, R.W., Riemann, L. (2016). Marine non-cyanobacterial diazotrophs: moving beyond molecular detection. *Trends Microbiol.* 24, 916-927.

Bonnet, S. and others. 2013. Aphotic N_2 Fixation in the Eastern Tropical South Pacific Ocean. *PLoS One* 8, e81265, doi:10.1371/journal.pone.0081265

Borstad, G., Borstad, L. (1977). "The *Oscillatoria erythraea* (Cyanophyta) community of associates," in *Cooperative Investigations of the Caribbean and Adjacent Regions-II*, *FAO Fish. Rep., 200*, ed. HB Stewart, 51–57.

Böttjer, D., D. M. Karl, R. M. Letelier, D. A. Viviani, Church, M. J. (2014). Experimental assessment of diazotroph responses to elevated seawater pCO_2 in the North Pacific Subtropical Gyre. *Global Biogeochem. Cycles* 28, 601–616.

Böttjer, D., Dore, J.E., Karl, D.M., Letelier, R.M., Mahaffey, C., Wilson, S.T., et al. (2016). Temporal variability of nitrogen fixation and particulate nitrogen export at Station ALOHA. *Limnol. Oceanogr.* doi: 10.1002/lno.10386

Bryant, J.A., Clemente, T.M., Viviani, D.A., Fong, A.A., Thomas, K.A., Kemp, P., et al. (2016). Diversity and activity of communities inhabiting plastic debris in the North Pacific Gyre. *mSystems* 1: e00024–00016.

Bybee, S.M., Bracken-Grissom, H., Haynes, B.D., Hermansen, R.A., Byers, R.L., Clement, M.J., et al. (2011). Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. *Genome Biology and Evolution* 3, 1312-1323.

Campbell, L., Nolla, H., Vaulot, D. (1994). The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean. *Limnol. Oceanogr.* 39, 954–961.

Cangelosi, G.A., Freitag, N.E., and Riley-Buckley, M.S. (2004). *From outside to inside: environmental microorganisms as human pathogens.* Washington, D.C. : American Society for Microbiology.

Capone, D.G., Burns, J.A., Montoya, J.P., Subramaniam, A., Mahaffey, C., Gunderson, T., et al. (2005). Nitrogen fixation by *Trichodesmium* spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Glob. Biogeochem. Cycles* 19, GB2024.

Capone, D.G., Ferrier, M.D., Carpenter, E.J. (1994). Amino acid cycling in colonies of the planktonic marine cyanobacterium *Trichodesmium thiebautii*. *Appl. Env. Microbiol.* 60, 3989–3995.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7, 335–336.

Chappell, P.D., Webb, E.A. (2010). A molecular assessment of the iron stress response in the two phylogenetic clades of *Trichodesmium*. *Env. Microbiol.* 12, 13–27.

Chase, E., Young, S., and Harwood, V.J. (2015). Sediment and vegetation as reservoirs of *Vibrio vulnificus* in the Tampa Bay Estuary and Gulf of Mexico. *Applied Env. Microbiol.* 81, 2489-2494.

Church, M.J., B.D. Jenkins, D.M. Karl, Zehr, J.P. (2005a). Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. *Aquat. Microb. Ecol.* 38, 3–14.

Church, M. J., C. M. Short, B. D. Jenkins, D. M. Karl, Zehr, J.P. (2005b). Temporal patterns of nitrogenase gene (nifH) expression in the oligotrophic North Pacific Ocean. *Appl. Environ. Microbiol.* 71, 5362–5370.

Church, M. J., K. M. Björkman, D. M. Karl, M. A. Saito, Zehr, J.P. (2008). Regional distributions of nitrogen-fixing bacteria in the Pacific Ocean. *Limnol. Oceanogr.* 53, 63-77.

Codispoti, L. An oceanic fixed nitrogen sink exceeding 400 Tg N a^{-1} vs the concept of homeostasis in the fixed-nitrogen inventory. *Biogeosciences* 4, 233–253.

Coles, V.J., Hood, R.R., Pascual, M., Capone, D.G. (2004). Modeling the impact of *Trichodesmium* and nitrogen fixation in the Atlantic Ocean. *J. Geophys. Res.* 109, C06007.

Colwell, R.R. (1993). Nonculturable but still viable and potentially pathogenic. *Zentralblatt für Bakteriologie* 279, 154-156.

Colwell, R.R., Huq, A., Islam, M.S., Aziz, K., Yunus, M., Khan, N.H., et al. (2003). Reduction of cholera in Bangladeshi villages by simple filtration. *PNAS* 100, 1051-1055.

Crump, B.C., Armbrust, E.V., Baross, J.A. (1999). Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl. Env. Microbiol.* 65, 3192–3204.

Dabundo, R. et al. (2014). The Contamination of Commercial $^{15}N_2$ Gas Stocks with ^{15}N-Labeled Nitrate and Ammonium and Consequences for Nitrogen Fixation Measurements. *PLoS One* 9: e110335, doi: 10.1371/journal.pone.0110335

Davis, C.S., McGillicuddy, D.J. (2006). Transatlantic abundance of the N_2-fixing colonial cyanobacterium *Trichodesmium*. *Science* 312, 1517–1520.

de O Santos, E., Alves, N., Dias, G.M., Mazotto, A.M., Vermelho, A., Vora, G.J., et al. (2011). Genomic and proteomic analyses of the coral pathogen *Vibrio coralliilyticus* reveal a diverse virulence repertoire. *ISME* 5, 1471-1483.

Dekaezemacker, J., S. Bonnet, O. Grosso, T. Moutin, M. Bressac, Capone, D. (2013). Evidence of active dinitrogen fixation in surface waters of the eastern tropical South Pacific during El Niño and La Niña events and evaluation of its potential nutrient controls. *Global Biogeochem. Cycles* 27, 768–779.

DeLong, E.F., Franks, D.G., Alldredge, A.L. (1993). Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* 38, 924–934.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.U., et al. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311, 496–503.

Deutsch, C., J. L. Sarmiento, D. M. Sigman, N. Gruber, Dunne, J. P. (2007). Spatial coupling of nitrogen inputs and losses in the ocean. *Nature* 445, 163-167.

Ducklow, H.W., Steinberg, D.K., Buesseler, K.O. (2001). Upper ocean carbon export and the biological pump. *Oceanography* 14, 50-58.

Dugdale, R. (1961). Nitrogen fixation in the Sargasso Sea. *Deep Sea Res. Pt. II:* 7, 297–300.

Dyhrman, S., Chappell, P., Haley, S., Moffett, J., Orchard, E., Waterbury, J., et al. (2006). Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* 439, 68–71.

Dyhrman, S.T., Benitez-Nelson, C.R., Orchard, E.D., Haley, S.T., Pellechia, P.J. (2009). A microbial source of phosphonates in oligotrophic marine systems. *Nat. Geosci.* 2, 696–699.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460-2461.

Eilers, H., Pernthaler, J., and Amann, R. (2000). Succession of pelagic marine bacteria during enrichment: a close look at cultivation-induced shifts. *Applied Env. Microbiol.* 66, 4634-4640.

Elston, R.A., Hasegawa, H., Humphrey, K.L., Polyak, I.K., Hase, C. (2008). Re-emergence of *Vibrio tubiashii* in bivalve shellfish aquaculture: severity, environmental drivers, geographic extent and management. *Diseases of Aquatic Organisms* 82, 119.

Estes, R.M., Friedman, C.S., Elston, R.A., Herwig, R.P. (2004). Pathogenicity testing of shellfish hatchery bacterial isolates on Pacific oyster *Crassostrea gigas* larvae. *Diseases of Aquatic Organisms* 58, 223-230.

Falkowski, P.G. (1997). Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* 387, 272-275.

Falkowski, P.G., and Raven, J.A. (2013). *Aquatic photosynthesis.* Princeton University Press.

Farnelid, H. et al. (2011). Nitrogenase gene amplicons from global marine surface waters are dominated by genes of non-cyanobacteria. *PloS one* 6: e19223, doi: 10.1371/journal.pone.0019223

Feely, R.A., Sabine, C.L., Hernandez-Ayon, J.M., Ianson, D., Hales, B. (2008). Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science* 320, 1490-1492.

Feely, R.A., S.R. Alin, B. Hales, G.C. Johnson, R.H. Byrne, W.T. Peterson, Greeley, D. Chemical and hydrographic profile measurements during the 2013 West Coast Ocean Acidification Cruise (WCOA2013, August 3-29, 2013). http://cdiac.ornl.gov/ftp/oceans/WCOA2013/. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee. doi: 10.3334/CDIAC/OTG.COAST_WCOA2013

Fernandez, C. L., et al. (2013). Community N_2 fixation and *Trichodesmium* spp. abundance along longitudinal gradients in the eastern subtropical North Atlantic. *Ices J. Mar. Sci.* fss142, doi: 10.1093/icesjms/fss142

Fernandez, C., L. Farías, Ulloa, O. (2011). Nitrogen fixation in denitrified marine waters. *PloS one* 6, e20539, doi: 10.1371/journal.pone.0020539

Fernandez, C., M. L. González, C. Muñoz, V. Molina, Farias, L. (2015). Temporal and spatial variability of biological nitrogen fixation off the upwelling system of central Chile (35–38.5° S). *J. Geophys. Res.: Oceans* 120, 3330-3349.

Ferrón S, del Valle DA, Björkman KM, Quay PD, Church MJ, Karl DM (2016). Application of membrane inlet mass spectrometry to measure aquatic gross primary production by the ^{18}O in vitro method. *Limnology and Oceanography: Methods*; e-pub ahead of print 15 July 2016, doi: 10.1002/lom3.10116

Follows, M.J., Dutkiewicz, S., Grant, S., and Chisholm, S.W. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* 315, 1843-1846.

Fontanez, K.M., Eppley, J.M., Samo, T.J., Karl, D.M., DeLong, E.F. (2015). Microbial community structure and function on sinking particles in the North Pacific Subtropical Gyre. *Front. Microbiol.* 6, 469.

Foster, R.A., Goebel, N.L., Zehr, J.P. (2010). Isolation of *Calothrix rhizosoleniae* (cyanobacteria) strain SC01 from *Chaetoceros* (bacillariophyta) spp. diatoms of the Subtropical North Pacific ocean. *J. Phycol.* 46, 1028–1037.

Foster, R.A., Zehr, J.P. (2006). Characterization of diatom–cyanobacteria symbioses on the basis of *nifH*, *hetR* and 16S rRNA sequences. *Env. Microbiol.* 8, 1913–1925.

Galloway, J. N. et al. (2004). Nitrogen cycles: past, present, and future. *Biogeochemistry* 70, 153-226.

Gay, M., Berthe, F., and Leroux, F. (2004). Screening of *Vibrio* isolates to develop an experimental infection model in the Pacific oyster *Crassostrea gigas*. *Diseases of Aquatic Organisms* 59, 49-56.

Gharaibeh, D.N., Hasegawa, H., Häse, C.C. (2009). Development of a quantitative real-time PCR assay for detection of *Vibrio tubiashii* targeting the metalloprotease gene. *Journal of Microbiological Methods* 76, 262-268.

Giovannoni, S.J., Cameron Thrash, J., Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565.

Gomez-Gil, B., Thompson, C.C., Matsumura, Y., Sawabe, T., Iida, T., Christen, R., et al. (2014). "The Family Vibrionaceae," in *The Prokaryotes – Gammaproteobacteria,* ed. E. Rosenberg. (Berlin: Springer-Verlag), 659-747.

Gradoville, M.R., White, A.E., Böttjer, D., Church, M.J., and Letelier, R.M. (2014). Diversity trumps acidification: Lack of evidence for carbon dioxide enhancement of *Trichodesmium* community nitrogen or carbon fixation at Station ALOHA. *Limnol. Oceanogr.* 59, 645-659.

Green, S.J., Venkatramanan, R., and Naqib, A. (2015). Deconstructing the polymerase chain reaction: understanding and correcting bias associated with primer degeneracies and primer-template mismatches. *PLoS One* 10, e0128122.

Grime, J. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *J. Ecol.* 86, 902-910.

Grömping, U. (2006). Relative importance for linear regression in R: the package Relaimpo. *Journal of Statistical Software* 17, 1-27.

Großkopf, T., et al. (2012). Doubling of marine dinitrogen-fixation rates based on direct measurements. *Nature* 488, 361-364.

Gruber, N. (2004). The dynamics of the marine nitrogen cycle and its influence on atmospheric CO_2 variations, p. 97-148. In Follows, M. and T. Oguz [eds.], The ocean carbon cycle and climate. Springer.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59, 307-321.

Halm, H., et al. (2011). Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. *ISME J* 6: 1238-1249.

Hasegawa, H., Lind, E.J., Boin, M.A., and Häse, C.C. (2008). The extracellular metalloprotease of *Vibrio tubiashii* is a major virulence factor for Pacific oyster (*Crassostrea gigas*) larvae. *Applied Env. Microbiol.* 74, 4101-4110.

Heidelberg, J., Heidelberg, K., and Colwell, R. (2002). Bacteria of the γ-subclass Proteobacteria associated with zooplankton in Chesapeake Bay. *Applied Env. Microbiol.* 68, 5498-5507.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477-483.

Hewson, I., et al. (2007). Characteristics of diazotrophs in surface to abyssopelagic waters of the Sargasso Sea. *Aquat. Microb. Ecol.* 46, 15-30.

Hewson, I., Poretsky, R.S., Dyhrman, S.T., Zielinski, B., White, A.E., Tripp, H.J., et al. (2009). Microbial community gene expression within colonies of the diazotroph, *Trichodesmium*, from the Southwest Pacific Ocean. *ISME J.* 3, 1286–1300.

Hill, J., Wheeler, P. (2002). Organic carbon and nitrogen in the northern California current system: comparison of offshore, river plume, and coastally upwelled waters. *Progress Oceanogr.* 53, 369-387.

Hmelo, L.R., Van Mooy, B.A.S., Mincer, T.J. (2012). Characterization of bacterial epibionts on the cyanobacterium *Trichodesmium*. *Aquat. Microb. Ecol.* 67, 1–14.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., et al. (2016). A new view of the tree of life. *Nature Microbiol.* 1, 16048.

Hunt, D.E., Gevers, D., Vahora, N.M., and Polz, M.F. (2008). Conservation of the chitin utilization pathway in the Vibrionaceae. *Applied Env. Microbiol.* 74(1), 44-51.

Huntemann, M., Ivanova, N.N., Mavromatis, K., Tripp, H.J., Paez-Espino, D., Palaniappan, K., et al. (2015). The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v. 4). *Stand. Genomic Sci.* 10, 86.

Hutchins, D, Fu, FX, Zhang, Y, Warner, M, Feng, Y, Portune, K, et al. (2007). CO_2 control of *Trichodesmium* N_2 fixation, photosynthesis, growth rates, and elemental ratios: Implications for past, present, and future ocean biogeochemistry. *Limnol. Oceanogr.* 52, 1293–1304.

Hutchins, D.A., Fu, F.-X., Webb, E.A., Walworth, N., Tagliabue, A. (2013). Taxon-specific response of marine nitrogen fixers to elevated carbon dioxide concentrations. *Nature Geoscience* 6, 790-795.

Hutchinson, G.E. (1961). The paradox of the plankton. *The American Naturalist* 95, 137-145.

Hynes, A.M., Webb, E.A., Doney, S.C., Waterbury, J.B. (2012). Comparison of cultured *Trichodesmium* (Cyanophyceae) with species characterized from the field. *J. Phycol.* 48, 196–210.

Islam, M., Drasar, B., and Bradley, D. (1990). Long-term persistence of toxigenic *Vibrio cholerae* 01 in the mucilaginous sheath of a blue-green alga, *Anabaena variabilis*. *The Journal of Tropical Medicine and Hygiene* 93, 133-139.

Jannasch, H.W., and Jones, G.E. (1959). Bacterial populations in sea water as determined by different methods of enumeration. *Limnol. Oceanogr.* 4, 128-139.

Janson, S., Bergman, B., Carpenter, E.J., Giovannoni, S.J., Vergin, K. (1999). Genetic analysis of natural populations of the marine diazotrophic cyanobacterium *Trichodesmium*. *FEMS Microbiol. Ecol.* 30, 57–65.

Jayakumar, A., M. Al-Rshaidat, B. B. Ward, Mulholland, M. R.. (2012). Diversity, distribution, and expression of diazotroph *nifH* genes in oxygen-deficient waters of the Arabian Sea. *FEMS Microbiol. Ecol.* 82, 597-606.

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

Karl, D., et al. (2002). Dinitrogen fixation in the world's oceans. Biogeochemistry 57, 47-98.

Karl, D., Letelier, R., Hebel, D., Tupas, L., Dore, J., Christian, J., et al. (1995). Ecosystem changes in the North Pacific subtropical gyre attributed to the 1991–92 El Nino. *Nature* 373, 230–234.

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J., Hebel, D. (1997). The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* 388, 533–538.

Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., et al. (2002). Dinitrogen fixation in the world's oceans. *Biogeochemistry* 57, 47–98.

Kimes, N.E., Grim, C.J., Johnson, W.R., Hasan, N.A., Tall, B.D., Kothary, M.H., et al. (2012). Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *ISME J* 6, 835-846.

Kirchman, D.L. (2010). *Microbial ecology of the oceans.* John Wiley & Sons.

Knapp, A. N. (2012). The sensitivity of marine N_2 fixation to dissolved inorganic nitrogen. *Front. Microbiol.* 3, 374, doi:10.3389/fmicb.2012.00374.

Knapp, A. N., J. Dekaezemacker, S. Bonnet, J. A. Sohm, and D. G. Capone. (2012). Sensitivity of *Trichodesmium erythraeum* and *Crocosphaera watsonii* abundance and N_2 fixation rates to varying NO_3^- and PO_4^{3-} concentrations in batch cultures. *Aquat. Microb. Ecol.* 66, 223-236.

Knapp, A.N., Casciotti, K.L., Berelson, W.M., Prokopenko, M.G., Capone, D.G. (2016). Low rates of nitrogen fixation in eastern tropical South Pacific surface waters. *PNAS* 113, 4398–4403.

Koch, R. (1884). An address on cholera and its bacillus. *British Medical Journal* 2, 453.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Env. Microbiol.* 79, 5112–5120.

Langlois, R., T. Großkopf, M. Mills, S. Takeda, Laroche, J. (2015). Widespread distribution and expression of Gamma A (UMB), an uncultured, diazotrophic, γ-proteobacterial *nifH* phylotype. *PloS one* 10: e0128912, doi: 10.1371/journal.pone.0128912

Langmead, B., Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359.

Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., et al. (2009). The genomic basis of trophic strategy in marine bacteria. *PNAS* 106, 15527–15533.

Lefever S., F. Pattyn, J. Hellemans, Vandesompele, J. (2013). Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clin. Chem.* 59: 1470–1480.

Legendre, L., Gosselin, M. (1997). Estimation of N or C uptake rates by phytoplankton using ^{15}N or ^{13}C: revisiting the usual computation formulae. *J. Plankton. Res.* 19, 263–271.

Lemire, A., Goudenège, D., Versigny, T., Petton, B., Calteau, A., Labreuche, Y., et al. (2015). Populations, not clones, are the unit of *Vibrio* pathogenesis in naturally infected oysters. *ISME J* 9, 1523-1531.

Letelier, R.M., Karl, D.M. (1996). Role of *Trichodesmium* spp. in the productivity of the subtropical North Pacific Ocean. *Mar. Ecol. Prog. Ser.* 133, 263–273.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Loescher, C. R. et al. (2014). Facets of diazotrophy in the oxygen minimum zone waters off Peru. *ISME J* 8, 2180–2192.

Lomas, M., Hopkinson, B., Losh, J., Ryan, D., Shi, D., Xu, Y., et al. (2012). Effect of ocean acidification on cyanobacteria in the subtropical North Atlantic. *Aquat. Microb. Ecol.* 66, 211–222.

Lundgren, P., Janson, S., Jonasson, S., Singer, A., Bergman, B. (2005). Unveiling of novel radiations within *Trichodesmium* cluster by *hetR* gene sequence analysis. *Appl. Env. Microbiol.* 71, 190–196.

Luo, Y-W., Lima, I., Karl, D., Deutsch, C., Doney, S. (2014). Data-based assessment of environmental controls on global marine nitrogen fixation. *Biogeosciences* 11, 691–708.

Luo, Y. -W. et al. (2012). Database of diazotrophs in global ocean: abundances, biomass and nitrogen fixation rates. *Earth System Science Data* 4, 47-73.

Letunik, I., Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245, doi: 10.1093/nar/gkw290

Lyons, M.M., Lau, Y.-T., Carden, W.E., Ward, J.E., Roberts, S.B., Smolowitz, R., et al. (2007). Characteristics of marine aggregates in shallow-water ecosystems: implications for disease ecology. *EcoHealth* 4, 406-420.

Lyons, T.W., Reinhard, C.T., Planavsky, N.J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506, 307-315.

Mahaffey, C., Michaels, A.F., Capone, D.G. (2005). The conundrum of marine N_2 fixation. *Am. J. Sci.* 305, 546–595.

Marie, D., Partensky, F., Jacquet, S., and Vaulot, D. (1997). Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Appl. Env. Microbiol.* 63, 186-193.

McKibben, S.M., Peterson, W., Wood, A.M., Trainer, V.L., Hunter, M., White, A.E. (2017). Climatic regulation of the neurotoxin domoic acid. *PNAS* 114, 239-244.

McKibben, S.M., Watkins-Brandt, K.S., Wood, A.M., Hunter, M., Forster, Z., Hopkins, A., et al. (2015). Monitoring Oregon Coastal Harmful Algae: Observations and implications of a harmful algal bloom-monitoring project. *Harmful Algae* 50, 32-44.

Mersni-Achour, R., Cheikh, Y.B., Pichereau, V., Doghri, I., Etien, C., Dégremont, L., et al. (2015). Factors other than metalloprotease are required for full virulence of French *Vibrio tubiashii* isolates in oyster larvae. *Microbiology* 161, 997-1007.

Mohr, W., Großkopf, T., Wallace, D.W.R., LaRoche, J. (2010). Methodological underestimation of oceanic nitrogen fixation rates. *PloS one* 5, e12583.

Moisander, P. H., R. A. Beinart, M. Voss, Zehr, J.P. (2008). Diversity and abundance of diazotrophic microorganisms in the South China Sea during intermonsoon. *ISME J* 2, 954–967.

Mollet, C., Drancourt, M., Raoult, D. (1997). *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol. Microbiol.* 26, 1005–1011.

Momper, L.M., Reese, B.K., Carvalho, G., Lee, P., Webb, E.A. (2015). A novel cohabitation between two diazotrophic cyanobacteria in the oligotrophic ocean. *ISME J.* 9, 882–893.

Montoya, J.P., Voss, M., Kahler, P., Capone, D.G. (1996). A simple, high-precision, high-sensitivity tracer assay for N_2 fixation. *Appl. Env. Microbiol.* 62, 986–993.

Moonsamy, P., Williams, T., Bonella, P., Holcomb, C., Höglund, B., Hillman, G., et al. (2013). High throughput HLA genotyping using 454 sequencing and the Fluidigm Access Array™ system for simplified amplicon library preparation. *Tissue Antigens* 81, 141-149.

Mulholland, M.R., Bronk, D.A., Capone, D.G. (2004). Dinitrogen fixation and release of ammonium and dissolved organic nitrogen by *Trichodesmium* IMS101. *Aquat. Microb. Ecol.* 37, 85–94.

Nalven, S.G. (2016). Decoding DOM Degradation with metatranscriptomics: How do sunlight and microbial communities interact to degrade dissolved organic matter in Arctic freshwaters? [dissertation/master's thesis]. [Corvallis (OR)]: Oregon State University.

O'Neil, J.M. (1998). The colonial cyanobacterium *Trichodesmium* as a physical and nutritional substrate for the harpacticoid copepod *Macrosetella gracilis*. *J. Plankton. Res.* 20, 43–59.

Orcutt, K., Rasmussen, U., Webb, E.A., Waterbury, J.B., Gundersen, K., Bergman, B. (2002). Characterization of *Trichodesmium* spp. by genetic techniques. *Appl. Env. Microbiol.* 68, 2236–2245.

Paerl, H. W., K. M. Crocker, Prufert, L.E. (1987). Limitation of N_2 fixation in coastal marine waters: Relative importance of molybdenum, iron, phosphorus, and organic matter availability. *Limnol. Oceanogr.* 32, 525-536.

Paerl, H.W., Pinckney, J. (1996). A mini-review of microbial consortia: their roles in aquatic production and biogeochemical cycling. *Microbial Ecology* 31, 225-247.

Paerl, H.W., Bebout, B.M. (1988). Direct measurement of O_2-depleted microzones in marine *Oscillatoria*: relation to N_2 fixation. *Science* 241, 442–445.

Paerl, H.W., Bebout, B.M., Prufert, L.E. (1989a). Bacterial associations with marine *Oscillatoria* sp. (*Trichodesmium* sp.) populations: Ecological implications. *J. Phycol.* 25, 773–784.

Paerl, H.W., Priscu, J.C., Brawner, D.L. (1989b). Immunochemical localization of nitrogenase in marine *Trichodesmium* aggregates: Relationship to N_2 fixation potential. *Appl. Env. Microbiol.* 55, 2965–2975.

Paillard, C., Le Roux, F., and Borrego, J.J. (2004). Bacterial disease in marine bivalves, a review of recent studies: trends and evolution. *Aquatic Living Resources* 17, 477-498.

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Ann. Rev. Mar. Sci.* 4, 449-466.

Peterson, W., Bond, N., and Robert, M. (2016). The Blob (Part Three): Going, going, gone? *PICES Press* 24, 46.

Prouty, M.G., and Klose, K.E. (2006). *Vibrio cholerae*: the genetics of pathogenesis and environmental persistence. *The Biology of Vibrios*, 311-339.

Rahav, E. et al. (2013). Dinitrogen fixation in aphotic oxygenated marine environments. *Front. Microbiol.* 4, 227.

Raimbault, P., Garcia. N. (2008). Evidence for efficient regenerated production and dinitrogen fixation in nitrogen-deficient waters of the South Pacific Ocean: impact on new and export production estimates. *Biogeosciences* 5, 323–338.

Richards, G.P., Watson, M.A., Needleman, D.S., Church, K.M., Häse, C.C. (2015). Mortalities of Eastern and Pacific oyster larvae caused by the pathogens *Vibrio coralliilyticus* and *Vibrio tubiashii*. *Appl. Env. Microbiol.* 81, 292-297.

Riemann, L., H. Farnelid, and G. F. Steward. (2010). Nitrogenase genes in non-cyanobacterial plankton: prevalence, diversity and regulation in marine waters. *Aquat. Microb. Ecol.* 61, 235–247.

Rii, Y. M., S. Duhamel, R. R. Bidigare, D. M. Karl, D. J. Repeta, Church, M. J. (2016). Diversity and productivity of photosynthetic picoeukaryotes in biogeochemically distinct regions of the South East Pacific Ocean. *Limnol. Oceanogr.* 61, 806-824.

Roe, K.L., Barbeau, K., Mann, E.L., Haygood, M.G. (2012). Acquisition of iron by *Trichodesmium* and associated bacteria in culture. *Env. Microbiol.* 14, 1681–1695.

Rosenzweig, M. L., Abramsky, Z. (1993). How are diversity and productivity related?, p 52–65. In Ricklefs, R. E. and Schluter D. [eds.], Species diversity in ecological communities. University of Chicago Press.

Rouco, M., Haley, S.T., Dyhrman, S.T. (2016). Microbial diversity within the *Trichodesmium* holobiont. *Env. Microbiol.* 18, 5151–5160.

Rouco, M., Warren, H.J., McGillicuddy, D.J., Waterbury, J.B., Dyhrman, S.T. (2014). *Trichodesmium* sp. clade distributions in the western North Atlantic Ocean. *Limnol. Oceanogr.* 59, 1899–1909.

Sainz-Hernández, J.C., and Maeda-Martínez, A.N. (2005). Sources of *Vibrio* bacteria in mollusc hatcheries and control methods: a case study. *Aquaculture Research* 36, 1611-1618.

Sargent, E. C., A. Hitchcock, S. A. Johansson, R. Langlois, C. M. Moore, J. LaRoche, A. J. Poulton, Bibby, T. S. (2016). Evidence for polyploidy in the globally important diazotroph *Trichodesmium*. *FEMS Microbiol. Lett.* 362. doi: 10.1093/femsle/fnw244

Sawabe, T., Ogura, Y., Matsumura, Y., Feng, G., Amin, A.R., Mino, S., et al. (2013). Updating the *Vibrio* clades defined by multilocus sequence phylogeny: proposal of eight new clades, and the description of *Vibrio tritonius* sp. nov. *Front. Microbiol.* 4, 414.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Env. Microbiol.* 75, 7537-7541.

Severin I, Bentzon-Tilia M, Moisander PH, Riemann L. (2015). Nitrogenase expression in estuarine bacterioplankton influenced by organic carbon and availability of oxygen. *FEMS Microbiol. Lett.* 362: fnv105, doi: 10.1093/femsle/fnv105

Sheridan, C., Steinberg, D., Kling, G. (2002). The microbial and metazoan community associated with colonies of *Trichodesmium* spp.: a quantitative survey. *J. Plankton. Res.* 24, 913–922.

Shiozaki, T., M. Ijichi, T. Kodama, S. Takeda, Furuya, K. (2014). Heterotrophic bacteria are major nitrogen fixers in the euphotic zone of the Indian Ocean. *Global Biogeochem. Cycles* 28, 1096–1110.

Siboni, N., Balaraju, V., Carney, R., Labbate, M., and Seymour, J.R. (2016). Spatiotemporal dynamics of *Vibrio* spp. within the Sydney Harbour Estuary. *Front. Microbiol.* 7, 460.

Siddiqui, P., Bergman, B., Carpenter, E. (1992). Filamentous cyanobacterial associates of the marine planktonic cyanobacterium *Trichodesmium*. *Phycologia* 31, 326–337.

Sohm, J.A., Hilton, J.A., Noble, A.E., Zehr, J.P., Saito, M.A., and Webb, E.A. (2011). Nitrogen fixation in the South Atlantic Gyre and the Benguela upwelling system. *Geophysical Research Letters* 38, L16608.

Staley, J.T., and Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Ann. Rev. Microbiol.* 39, 321-346.

Stihl, A., Sommer, U., Post, A.F. (2001). Alkaline phosphatase activities among populations of the colony-forming diazotrophic cyanobacterium *Trichodesmium* spp. (cyanobacteria) in the Red Sea. *J. Phycol.* 37, 310–317.

Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., et al. (2013). *IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.* Cambridge University Press.

Subramaniam, A., Brown, C.W., Hood, R.R., Carpenter, E.J., Capone, D.G. (2001). Detecting *Trichodesmium* blooms in SeaWiFS imagery. *Deep Sea Res. Pt. II* 49, 107–121.

Subramaniam, A., C. Mahaffey, W. Johns, and N. Mahowald. (2013). Equatorial upwelling enhances nitrogen fixation in the Atlantic Ocean. *Geophys. Res. Lett.* 40, 1766-1771.

Sussman, M., Mieog, J.C., Doyle, J., Victor, S., Willis, B.L., Bourne, D.G. (2009). *Vibrio* zinc-metalloprotease causes photoinactivation of coral endosymbionts and coral tissue lesions. *PLoS One* 4, e4511.

Sutton, D., Garrick, R. (1993). Bacterial disease of cultured giant clam *Tridacna gigas* larvae. *Diseases of Aquatic Organisms* 16, 47-53.

Takemura, A.F., Chien, D.M., and Polz, M.F. (2014). Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Front. Microbiol.* 5, doi: 10.3389/fmicb.2014.00038.

Thompson, F.L., Austin, B., and Swings, J. (2006). *The Biology of Vibrios.* ASM press.

Thompson, J.R., Randa, M.A., Marcelino, L.A., Tomita-Mitchell, A., Lim, E., Polz, M.F. (2004). Diversity and dynamics of a North Atlantic coastal *Vibrio* community. *Applied and Environmental Microbiology* 70, 4103-4110.

Tout, J., Siboni, N., Messer, L.F., Garren, M., Stocker, R., Webster, N.S., et al. (2015). Increased seawater temperature increases the abundance and alters the structure of natural *Vibrio* populations associated with the coral *Pocillopora damicornis*. *Front. Microbiol.* 6, 432.

Turk K.A., et al. (2011). Nitrogen fixation and nitrogenase (*nifH*) expression in tropical waters of the eastern North Atlantic. *ISME J* 5, 1201–1212.

Turk-Kubo K. A., K. M. Achilles, T. R. Serros, M. Ochiai, J. P. Montoya, Zehr, J. P. (2012). Nitrogenase (*nifH*) gene expression in diazotrophic cyanobacteria in the Tropical North Atlantic in response to nutrient amendments. *Front. Microbiol.* 3, 386, doi: 10.3389/fmicb.2012.00386

Turk-Kubo, K.A., Frank, I.E., Hogan, M.E., Desnues, A., Bonnet, S., Zehr, J.P. (2015). Diazotroph community succession during the VAHINE mesocosm experiment (New Caledonia lagoon). *Biogeosciences* 12, 7435–7452.

Turk-Kubo, K. A., M. Karamchandani, D. G. Capone, and J. P. Zehr. (2014). The paradox of marine heterotrophic nitrogen fixation: Abundances of heterotrophic diazotrophs do not account for nitrogen fixation rates in the Eastern Tropical South Pacific. *Environ. Microbiol.* 16: 3095–3114.

Turner, J.W., Good, B., Cole, D., and Lipp, E.K. (2009). Plankton composition and environmental factors contribute to *Vibrio* seasonality. *ISME J* 3, 1082-1092.

Urakawa, H., and Rivera, I.N.G. (2006). Aquatic Environment. *The Biology of Vibrios*, 175-189.

Van Mooy, B.A., Hmelo, L.R., Sofen, L.E., Campagna, S.R., May, A.L., Dyhrman, S.T., et al. (2011). Quorum sensing control of phosphorus acquisition in *Trichodesmium* consortia. *ISME J.* 6, 422–429.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304, 66-74.

Vezzulli, L., Colwell, R.R., and Pruzzo, C. (2013). Ocean warming and spread of pathogenic vibrios in the aquatic environment. *Microb. Ecol.* 65, 817-825.

Vezzulli, L., Pezzati, E., Moreno, M., Fabiano, M., Pane, L., Pruzzo, C., et al. (2009). Benthic ecology of Vibrio spp. and pathogenic Vibrio species in a coastal Mediterranean environment (La Spezia Gulf, Italy). *Microb. Ecol.* 58, 808-818.

Wagner, G.P., Kin, K., Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* 131, 281–285.

Volk, T., and Hoffert, M.I. (1985). Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO2 changes. *The Carbon Cycle and Atmospheric CO: Natural Variations Archean to Present*: 99-110.

Waldbusser, G.G., and Salisbury, J.E. (2014). Ocean acidification in the coastal zone from an organism's perspective: multiple system parameters, frequency domains, and habitats. *Annual Review of Marine Science* 6, 221-247.

Waldbusser, G.G., Hales, B., Langdon, C.J., Haley, B.A., Schrader, P., Brunner, E.L., et al. (2014). Saturation-state sensitivity of marine bivalve larvae to ocean acidification. *Nature Climate Change* 5, 273–280.

Walsby, A. (1978). The properties and buoyancy-providing role of gas vacuoles in *Trichodesmium* Ehrenberg. *Brit. Phycol. J.* 13, 103–116.

Walsby, A. (1992). "The gas vesicles and buoyancy of *Trichodesmium*," in *Marine pelagic cyanobacteria: Trichodesmium and other diazotrophs*, eds. E.J. Carpenter, D.G. Capone, J.G. Rueter. (Springer Netherlands), 141–161.

Walworth, N., Pfreundt, U., Nelson, W.C., Mincer, T., Heidelberg, J.F., Fu, F., et al. (2015). *Trichodesmium* genome maintains abundant, widespread noncoding DNA in situ, despite oligotrophic lifestyle. *PNAS* 112, 4251–4256.

Weber, T., Deutsch, C. (2014). Local versus basin-scale limitation of marine nitrogen fixation. *Proc. Natl. Acad. Sci.* U.S.A. 111, 8741–8746.

Weber, T., and Deutsch, C. (2012). Oceanic nitrogen reservoir regulated by plankton diversity and ocean circulation. *Nature* 489, 419-422.

Welschmeyer, N.A. (1994). Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. *Limnol. Oceanogr.* 39, 1985-1992.

Wendling, C.C., Batista, F.M., Wegner, K.M. (2014). Persistence, seasonal dynamics and pathogenic potential of *Vibrio* communities from Pacific oyster hemolymph. *PLoS One* 9, e94256.

Wietz, M., Gram, L., Jørgensen, B., Schramm, A. (2010). Latitudinal patterns in the abundance of major marine bacterioplankton groups. *Aquat. Microb. Ecol.* 61, 179-189.

Wilson, B., Muirhead, A., Bazanella, M., Huete-Stauffer, C., Vezzulli, L., Bourne, D.G. (2013). An improved detection and quantification method for the coral pathogen *Vibrio coralliilyticus*. *PloS one* 8, e81800. doi: 10.1371/journal.pone.0081800.

Wilson, S.T., Böttjer, D., Church, M.J., Karl, D.M. (2012). Comparative assessment of nitrogen fixation methodologies, conducted in the oligotrophic North Pacific Ocean. *Appl. Env. Microbiol.* 78, 6516–6523.

Woese, C.R. (1987). Bacterial evolution. *Microbiological Reviews* 51, 221.

Wright, S. (1991). Improved HPLC method for the analysis of chlorophylls and carotenoids from marine phytoplankton. *Mar. Ecol. Prog. Ser.* 77, 183-196.

Wyman, M., Hodgson, S., Bird, C. (2013). Denitrifying Alphaproteobacteria from the Arabian Sea that express *nosZ*, the gene encoding nitrous oxide reductase, in oxic and suboxic waters. *Appl. Env. Microbiol.* 79, 2670–2681.

Yong, L., Guanpin, Y., Hualei, W., Jixiang, C., Xianming, S., Guiwei, Z., et al. (2006). Design of *Vibrio* 16S rRNA gene specific primers and their application in the analysis of seawater *Vibrio* community. *Journal of Ocean University of China* 5, 157-164.

Zani, S., Mellon, M.T., Collier, J.L., Zehr, J.P. (2000). Expression of *nifH* genes in natural microbial assemblages in Lake George, New York, detected by reverse transcriptase PCR. *Appl. Env. Microbiol.* 66, 3119–3124.

Zehr, J. P., and L. A. Mcreynolds. (1989). Use of degenerate oligonucleotides for amplification of the nifH gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl. Environ. Microbiol.* 55, 2522–2526.

Zehr, J. P., L. L. Crumbliss, M. J. Church, E. O. Omoregie, Jenkins, B. D.. (2003a). Nitrogenase genes in PCR and RT-PCR reagents: implications for studies of diversity of functional genes. *Biotechniques* 35, 996–1013.

Zehr, J. P., M. T. Mellon, and S. Zani. (1998). New nitrogen-fixing microorganisms detected in oligotrophic oceans by amplification of nitrogenase (*nifH*) genes. *Appl. Environ. Microbiol.* 64, 3444–3450.

Zehr, J.P., Jenkins, B.D., Short, S.M., Steward, G.F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Env. Microbiol.* 5, 539-554.

Zehr J. and P. Turner. (2001). Nitrogen fixation: Nitrogenase genes and gene expression. *Methods Microbiol.* 30, 271-286.

Zhang, Y., Z. Zhao, J. Sun, Jiao, N. (2011). Diversity and distribution of diazotrophic communities in the South China Sea deep basin with mesoscale cyclonic eddy perturbations. *FEMS Microbiol. Ecol.* 78, 417–427.

Zhu, J., Miller, M.B., Vance, R.E., Dziejman, M., Bassler, B.L., and Mekalanos, J.J. (2002). Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*. *PNAS* 99, 3129-3134.