Fine-Scale Ecological Dynamics of Closely Related Marine Microbes

by

Diana M. Chien

B.S., Princeton University (2010)

Submitted to the Microbiology Graduate Program in partial fulfillment
of the requirements for the degree of Doctor of Philosophy
at the Massachusetts Institute of Technology
January 31, 2016
[February 2016]

Author..... Signature redacted

Microbiology Graduate Program
January 2016

Certified by..... Signature redacted

Martin Polz
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by..... Signature redacted

Kristala Prather
Associate Professor of Chemical Engineering
Director of the Microbiology Graduate Program

Fine-Scale Ecological Dynamics
of Closely Related Marine Microbes

by

Diana M. Chien

Submitted to the Microbiology Graduate Program in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology,
January 31, 2016

Abstract

Microbial dynamics in the environment are major drivers of global biogeochemical cycles; hence there is great interest in characterizing their rates and causes. While dynamics are affected by processes across many spatiotemporal scales, and even closely related microbes are known to vary in their ecological distributions, most work has characterized dynamics at bulk scales and with low genetic resolution. Thus, little is known about the effects of finer structure. In this thesis, I show that characterizing the dynamics of environmental microbes with finer spatiotemporal and genetic resolution reveals otherwise concealed dynamics. I use the Vibrionaceae, an ecologically diverse family of marine heterotrophs, as a model system. First, I review past studies on environmental associations of the Vibrionaceae, showing that few abiotic parameters have consistent predictive value, and that observed patterns vary based on taxonomic resolution. Biotic associations, however, may represent more specific predictors for fine-scale Vibrionaceae taxa, reflecting their diverse lifestyles. I then characterize Vibrionaceae dynamics within a high-resolution environmental time-series, with three months of daily sampling across four habitat partitions, population-level resolution, and large datasets of potential biological correlates. These data reveal diverse and spatially structured population dynamics. Individual populations varied from consistently abundant generalists to rare populations that occasionally displayed brief but intense peaks of abundance. Free-living and particle-attached habitat partitions were distinct in terms of diversity, turnover, and biotic interactors. These results emphasize the ecological differentiation of the Vibrionaceae populations, and the extent to which spatial partitions can function as distinct ecological regimes. Finally, I use

sequence data from the Vibrionaceae populations to investigate a methodological question relating to phylogenetic resolution: how well does the standard taxonomic marker gene, 16S rRNA, resolve populations known to have distinct ecological distributions and dynamics? The analysis shows that even full-length 16S rRNA sequences collapse the majority of populations into only 2-3 taxa, concealing the breadth of ecological behavior within the family. Altogether, this thesis demonstrates that high-resolution sampling techniques reveal a wealth of otherwise unobserved ecological diversity even within one family of closely related microbes, and suggests that fine-scale turnover and structure may have an unappreciated impact on microbial dynamics.

Thesis Supervisor: Martin F. Polz, Professor of Civil and Environmental Engineering

Acknowledgements

I extend thanks...

To my advisor Martin Polz, whose lab and leadership have been the perfect fit for me here at MIT. I owe so much to Martin's clarity, decisiveness, broad-ranging biological insight, and generous feedback; in so many ways he has modeled an ideal scientist for me.

To the members of my committee – Eric Alm, Otto Cordero, and Ed DeLong – for their enthusiasm and support. I feel like the luckiest grad student to have received such warm guidance and collegiality from them, whether in committee meetings or coffee shops.

To the members of Polz Lab, past and present: Phil Arevalo, Nate Cermak, Chris Corzett, Michael Cutler, Manoshi Datta, Joseph Elsherbini, Jan-Hendrik Hehemann, Fatima Aysha Hussain, Kathryn Kauffman, Alison Takemura, Dave VanInsberghe, Hans Wildschutte, Hong Xue, and Joy Yang. For their scientific assistance and collaboration, I thank especially Michael, Joseph, Phil, Alison, Chris, and Dave. For their friendship and support, I thank especially Manoshi, Alison, and Fatima. All of you have filled me with admiration and inspiration, and offered assistance, encouragement, international snacks, and humor in the face of scientific travail. I count so many members of Alm Lab, too, as labmates: my thanks and affection to Sean Clarke, Brian Cleary, Sean Kearney, Scott Olesen, Ali Perrotta, and Chris Smillie.

To the many communities I have been lucky enough to be a part of here at MIT: the Microbiology Program, including "Auntie" BonnieLee Whang; Parsons Laboratory, where ideas will always trump facilities; the Biological Engineering Communication Lab, especially Jaime Goldstein; and the Sidney-Pacific Graduate Community, especially Steve Voinea.

To my constant friends beyond MIT, for their countless words and acts of love: Emily García, Karen Sun, Janice Chou, and Carly Inkpen.

To John Casey, for his love, strength, thoughtfulness, and all the ways in which he has inspired me to grow.

And finally, to my mother, father, and brother for their unwavering love and enthusiasm, and their lifelong nurturing of my pursuits and passions. I look back so

fondly and gratefully at all the years of nature summer camps, gardening, calculus tutoring, late-night physics talks, and endless library trips. Without my family, I would not be who I am or where I am.

This work has been funded by grants and a graduate fellowship from the National Science Foundation, and by a graduate fellowship from the MIT Energy Institute.

I gratefully acknowledge the assistance of Linda Amaral-Zettler and Victor Schmidt at Marine Biological Laboratory in sharing their PCR protocol for environmental *Vibrio* samples, which was invaluable for methods development for Chapter 2 of my thesis.

Table of Contents

List of Figures and Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 1: Introduction

Microbial dynamics – processes of growth, change, and turnover – are rich and intricate, woven from complex networks of interactions with other organisms and with environmental processes across many scales, from sinking marine particles to the turning seasons (Fuhrman *et al.*, 2015). These dynamics have numerous consequences of human interest, from their sizable contributions to global fluxes of carbon, nitrogen, and other vital nutrients (Azam and Malfatti, 2007; Pomeroy *et al.*, 2007), to their health impacts on both humans and wildlife. For example, as oceans warm in the wake of global climate change, there is great interest in predicting how this may alter the incidence of marine-borne diseases like cholera and coral reef infections (Ceccarelli and Colwell, 2013; Burge *et al.*, 2014). In order to make sense of such microbial dynamics, microbial ecologists seek to identify causative environmental forces, identify spatiotemporal patterns, and build predictive models (Larsen *et al.*, 2012; Fuhrman *et al.*, 2015; Faust *et al.*, 2015). But compared to traditional macroecological studies of dynamics, understanding the dynamics of environmental microbes presents additional challenges, including the relatively indirect means by which we must measure microbes' presence and/or activity, i.e., culture or molecular technologies; the vast genetic diversity of microbes, the phenotypic or ecological consequences of which is not always understood; their rapid lifespans and consequent rates of turnover; and the microscopic spatial scale on which their immediate interactions take place.

Here, I review factors that contribute to microbial dynamics and methods for interpreting them, with an emphasis on marine ecosystems. This review is organized according to consideration of genetic, spatial, and temporal resolution. First, I discuss the impact of phylogenetic resolution of microbial taxa when measuring or describing ecological behavior. Next, I review processes that contribute to marine microbial dynamics, at increasing spatial and temporal scales*. I conclude by summarizing the

* When sampling marine communities at a fixed location, as is often the case for time-series studies (Eulerian sampling), it is difficult to disentangle temporal dynamics from spatial variation as different "parcels" of water move past the study site. (This contrasts to the less scalable Lagrangian sampling, in which sampling attempts to follow a current or parcel of water.) For the purposes of this review, I discuss both spatial and temporal variation, since both give rise to marine microbial dynamics as we observe them.

aims of this thesis, which investigates ecological dynamics using marine Vibrionaceae as a model system.

1.1. Phylogenetic resolution matters when describing ecological dynamics

When investigating dynamics, or any other ecological phenomenon, it is crucial to consider the ecological significance of taxa at the level of phylogenetic resolution chosen for characterization. Does a taxon correspond to a fine-scale unit whose members can be expected to behave more or less identically in terms of habitat choice, resource usage, and so on, or a broader grouping encompassing some degree of diversity? Among macroorganisms, an ecologically cohesive unit is classically defined as a group of interbreeding organisms, i.e., the biological species concept. Among microorganisms, however, attempts to identify clear-cut units are muddied by the occurrence of horizontal gene transfer, which allows transmission of genetic material across phylogenetic lineages (Doolittle and Papke, 2006). Consequently, while closely related microbes might appear identical or near-identical in terms of 16S ribosomal rRNA, the highly conserved phylogenetic marker gene traditionally used to define microbial "species," genomics has confirmed that such close relatives can possess numerous divergent genes, plasmids, and other genetic elements, acquired from outside sources. For example, a comparison of three *Escherichia coli* genomes found that they shared only 39.2% protein content (Welch *et al.*, 2002).

If horizontally acquired genetic variation confers an adaptive benefit, it can in turn translate into ecological differentiation among close relative. Detecting such ecological distinctions and mapping them onto information about genetic relatedness is a powerful approach for identifying coherent microbial units. Microbes that spatiotemporally co-occur in the environment due to shared adaptive traits are expected to constitute both ecological and evolutionary units: i.e., an ecological population, with shared habitat, selective pressures, and gene pool. For example, while members of the *Prochlorococcus* genus can broadly be described as abundant marine cyanobacteria that are found in the euphotic zone of oligotrophic waters and are no more than 3% divergent in terms of the 16S rRNA gene (i.e., are a traditionally genetically defined microbial "species"), numerous ecological populations have been discovered within the genus, each with a distinct environmental distribution based on nutrient availability, light conditions, and other selective forces (Biller *et al.*, 2015). Heterotrophic microbes within

the family Vibrionaceae have similarly been shown in one coastal system to constitute ecological populations with distinct distributions across marine habitat partitions and resource types (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Szabo *et al.*, 2012). Hence, when asking where and when we expect to see a member of a given microbial taxon, *which* member can matter; investigators must employ both appropriate ecological measures and genetic resolution to tease apart such fine-scale ecological population structure.

When surveying environmental microbial distributions, next-generation sequencing of relatively short marker gene sequences (tag sequencing) is the most common and high-throughput approach for assessing phylogenetic composition. While highly conserved markers like 16S rRNA offer broad coverage, they are also lower resolution, as suggested above, and tend not to resolve below the family or genus level (Caporaso *et al.*, 2011). If seeking sensitivity to population structure within a narrower microbial community, investigators may then target more diverse marker sequences, e.g., in the *Prochlorococcus* system, the housekeeping genes *rpoC1, petB-D, ntcA* or *gyrB*, or the ribosomal internal transcribed spacer sequence (Biller *et al.*, 2015). Depending on the gene and taxonomic specificity desired, this may require the development of tailored methods, in contrast to "universal" PCR primers like those designed for capturing 16S rRNA diversity across the prokaryotes (Caporaso *et al.*, 2011).

Ultimately, when accompanied with measurements of relevant environmental variables, molecular surveys allow us to distinguish ecological patterns of microbial dynamics at varying levels of phylogenetic resolution. As microbial dynamics are often described in terms of compositional stability, resilience, or seasonality, these statements must be considered in light of the level of phylogenetic resolution measured. While higher taxa certainly possess some degree of ecological coherence (Philippot *et al.*, 2010), significant turnover could occur below that level of description, from populations to strains. As sequencing technologies advance, we are just beginning to investigate the extent to which such turnover of fine-scale taxa takes place in different systems. In addition, a host of other methods allows us to augment phylogenetic assessments of microbial diversity, with more direct insight into functional variation at various taxonomic levels: metagenomics, proteomics, transcriptomics, metabolomics and single-cell genomics.

2.1. Marine microbial dynamics over small spatial scales (microns to liters)

For practical reasons, sampling for marine microbial time-series is often performed with liters of water, yet microbial activity transpires on much smaller scales. From the perspective of a marine microbe, the water column is a patchy, rapidly shifting landscape, with diverse ecological opportunities, both solid (i.e. offering a surface for attachment) and dissolved: animals, particles of organic materials, and dissolved nutrients released by "sloppy grazing" by zooplankton, or exuded by phytoplankton or sinking particles (Alldredge and Silver, 1988; Stocker, 2012). Spatial architecture can exist on the scale of microns: e.g., the physical mixing of chemical gradients, which are sought by motile microbes, creates heterogeneity on this scale, resulting in heterogeneous microbial distribution and dynamics as well (Stocker, 2012).

Marine dynamics at micro spatial scales are technically difficult to resolve, and hence in the past they have infrequently been investigated. In the laboratory, microfluidics coupled with microscopy have recently emerged as a technology for visualizing and measuring chemotaxis, colonization, and other behaviors on sub-millimeter scales (Ahmed *et al.*, 2010; Son *et al.*, 2015). *In situ*, very few studies have attempted sampling on any kind of microscale, although two studies have actually compared microbial community structure among microliter-scale samples through simple, though nonscalable, technical means: sampling surface water with syringes or micropipettor (Bratbak *et al.*, 1996; Seymour *et al.*, 2005; discussed further in the next section). A scalable approach to taking account of fine-scale marine habitat structure *in situ* is through size fractionation of water samples, which allows separation of free-living and particle-attached cells; sequential size fractionation can provide further distinction of particles of defined size ranges (Hunt *et al.*, 2008; Sperling *et al.*, 2013; Lee *et al.*, 2015; Yang *et al.*, 2015). Thus, this approach affords insight into differential resource and habitat usage within a given water sample. Overall, however, the majority of time-series studies and other characterizations of dynamics do not take finer spatial structure into account, and so its contribution to larger-scale processes is not well understood.

By contrast, dynamic processes are easy to characterize at the mesoscale, e.g., using liter-scale mesocosms to recapitulate dynamics of algal blooms (Dam and Drapeau, 1995; Riemann *et al.*, 2000; Turner *et al.*, 2009), colonization of detrital particles (Abell and Bowman, 2005; Pedrotti *et al.*, 2009), grazing (Nejstgaard *et al.*, 1997; Tang *et al.*, 2006), or viral lysis (Brussaard *et al.*, 2005). And *in situ*, spatially separated mesoscale

samples can be compared to infer broader spatial heterogeneity in microbial composition, typically in relation to easily measurable environmental parameters like temperature, salinity, and dissolved nutrient content.

2.2. Marine microbial dynamics on short temporal scales (seconds to hours)

General biotic processes that could result in compositional turnover on the timescale of seconds to hours include growth; replication; cooperation, competition, and other intra-microbial processes; and predation by bacteriophage or eukaryotic grazers. Useful categories for characterizing relevant microbial behaviors (not necessarily limited to short timescales) include copiotrophy (or opportunitrophy) vs. oligotrophy, and generalization vs. specialization. The copiotroph-oligotroph spectrum describes nutrient-usage strategies: copiotrophs take advantage of short-lived nutrient pulses, and boast rapid uptake and growth rates under prime circumstances (Koch, 2001; Thompson and Polz, 2006; Dutta and Paul, 2012); the phyla Gammaproteobacteria, Flavobacteriia, and Alphaproteobacteria include many copiotrophs. By contrast, oligotrophs make a slower living by subsisting on a background of low but relatively constant nutrient availability (Koch, 2001; Dutta and Paul, 2012). Thus, these characterizations give an expectation of response time to environmental perturbations, and of consistency of abundance. The generalist-to-specialist continuum describes the relative breadth or specificity of a given taxon's ecological preferences, e.g., nutrient sources. This sets an expectation for in how many different habitats one might to expect to see such a taxon, with how many other organisms it might interact, the breadth of environmental perturbations it might respond to, and so on.

In marine systems specifically, other biotic dynamical forces on short timescales include diel cycling of phytoplankton productivity and vertical migration, which creates concomitant diel cycling of bacterial productivity, with bacteria increasing in abundance and activity late in the afternoon (Fuhrman *et al.*, 1985; Gasol *et al.*, 1998; Kuipers *et al.*, 2000). Physical marine forces include tidal movement and storm or wind action, which cause redistribution of nutrients, especially resuspension of detritus (and any associated microbes) from the sea floor.

While dynamics are, again, relatively easy to measure via sampling with meso-scale time resolution, microscale dynamics have rarely been sampled environmentally. Those few studies that have sampled on the scale of seconds to minutes, have

demonstrated high-frequency fluctuation and intermittent perturbations in viral abundance (Bratbak *et al.*, 1996), nutrients (Seuront *et al.*, 2002), and heterotrophic and cyanobacteria abundance and activity (Seymour *et al.*, 2005). The study of Seymour *et al.* (2005) moreover inferred different relationships between heterotrophs and cyanobacteria at different temporal sampling scales (10-second versus 30-minute intervals), based on the finding of significant correlation among the two communities during microscale sampling, but not small-scale sampling.

Findings like these emphasize the operation of different ecological processes at different spatiotemporal scales. While the kinds of microscale variation and intermittent perturbations described in the studies above could simply be considered short-term variation without larger consequences, they have both methodological and ecological consequences. First, conclusions drawn from measurements taken at longer intervals may misrepresent parameters or phenomena that significantly vary on more rapid timescales (Taylor and Howes, 1994). Second, sporadic perturbations like nutrient injections or wind events can in fact have significant impacts on broad-scale ecological dynamics like community production (Jenkins and Goldman, 1985; Dickey, 1990; Taylor and Howes, 1994; Morán and Estrada, 2001). Third, shorter term variation is presumably additive to some extent, since community self-similarity continually degrades over time (Fuhrman *et al.*, 2015). Finally, as reflected by the copiotrophic and oligotrophic strategies discussed above, microbes are adapted to different regimes in terms of variability of resources and other environmental parameters; varying ecological strategies result in different rates, variability, and magnitudes of input into community productivity, and hence overall nutrient fluxes.

3.1. Marine microbial ecology over kilometer spatial scales

Few studies have quantified horizontal spatial structure of microbial communities over the scale of kilometers. Two studies have both used community "fingerprinting" techniques (automated rRNA intergenic-spacer analysis and terminal restriction fragment length polymorphism analysis of the 18S rRNA gene) to assess broad composition of communities separated by different distances, and concluded that 2-20km patches represented parcels of "broadly homogeneous" microbial composition (Hewson *et al.*, 2006; Lie *et al.*, 2013). Although it is difficult to interpret how such community fingerprinting profiles may correspond to phylogenetic or ecological variation, this spatial structure presumably maps to similar environmental controls on

microbial diversity and dynamics – temperature, salinity, nutrient availability or composition – within each such patch. Except in the case of sharp fronts where characteristics like temperature alter more abruptly, any patch-like structure is unlikely to have clear-cut boundaries, as any neighboring patches will be continually mixed with one another, causing temperature mixing and import and export of organisms and chemicals.

4.2. Marine microbial dynamics on temporal scales of days to years

Abiotic marine dynamics on the scale of days to weeks include weather patterns and mesoscale physical processes, i.e. eddies and currents. Biotic factors include interactions with macrobiota like fish, trophic cascades resulting from zooplankton or macrobiotic activity (Jürgens *et al.*, 1994; Zöllner *et al.*, 2009), algal blooms and subsequent decay (Brussaard *et al.*, 1996; Riemann *et al.*, 2000; Castberg *et al.*, 2001), and longer-term patterns emerging from microorganismal interactions, e.g., predator-prey dynamics with phage and zooplankton grazers (Wikner and Hagström, 1988; Rassoulzadegan and Sheldon, 1986; Gavin *et al.*, 2006). Mathematical models are often used to predict outcomes of biotic interactions like robustness to perturbation, oscillating patterns, and existence of multiple stable states (Follows and Dutkiewicz, 2011; Damore and Gore, 2012; Faust and Raes, 2012; Larsen *et al.*, 2012; Stenuit and Agathos, 2015). Researchers also use network strategies to infer potential interactions and causality among community members (and environmental parameters), based on measured environmental dynamics (Faust *et al.*, 2015; Fuhrman *et al.*, 2015).

On the monthly to seasonal timescale, relevant abotic environmental factors include solar intensity; seasonal weather; temperature-driven vertical stratification of the water column; and upwelling events, which result in significant release of nutrients to surface waters (Fuhrman *et al.*, 2015). Biotic dynamics include seasonal macrobiotic turnover or migration, and resulting trophic cascades.

A number of long-term time-series (monthly samples for years) have given us insight into microbial community dynamics at longer time-scales, with sampling sites in a variety of systems: Bermuda, Hawaii, southern California, the western English Channel, and the northwestern Mediterranean, as reviewed by (Fuhrman *et al.*, 2015). Broadly, these studies find a weak seasonal signal of self-similarity, i.e., community composition is slightly more similar during matching seasons than during different

seasons. Any recurrent patterns are, however, limited by the overall temporal decline in community similarity: daily sampling shows that community self-similarity drops off steeply over the scale of one to two weeks, then plateaus thereafter, showing slower yet continual decrease in similarity (Needham *et al.*, 2013). One recent analysis found, in fact, that more community variation occurs within seasons than between, suggesting that short-term perturbations and turnover indeed have a major effect in structuring marine microbial communities (Hatosy *et al.*, 2013).

Aims of this Thesis

In this work, I investigate marine microbial dynamics within the family Vibrionaceae, which includes human and animal pathogens and hosts broad ecological diversity in terms of resource usage and lifestyles. Given that the family is also culturable and genetically tractable, we have previously developed a coastal Massachusetts Vibrionaceae community as a model system for microbial ecology. By identifying phylogenetic clusters with distinct ecological preferences, we have distinguished ecological populations, allowing us to interrogate fine-scale ecological behavior (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Preheim, Timberlake, *et al.*, 2011; Szabo *et al.*). Above, I identified gaps in terms of understanding dynamics at finer phylogenetic, spatial, and temporal scales; here, I characterize Vibrionaceae communities with a focus on illuminating dynamics at these scales.

Chapter 1 reviews past studies of distributions and dynamics of global Vibrionaceae, first quantifying the predictive power of bulk environmental variables and secondly evaluating evidence for more specific biological associations, which may more strongly inform dynamics at finer levels of taxonomic resolution.

Chapter 2 analyzes behavior of the previously defined Vibrionaceae populations throughout an environmental time-series with high temporal, spatial, and genetic resolution: 93 continuous days of daily sampling of different habitat partitions within the water column, with population-level resolution. This allows us to characterize the breadth and variability of different populations' dynamic behavior, as well as identify potential environmental causation and biological interactors.

Chapter 3 addresses the effect of phylogenetic resolution when characterizing microbial ecology and dynamics, by assessing the ability of the standard phylogenetic marker, the 16S rRNA gene, to resolve Vibrionaceae populations with distinct ecological distributions and dynamics, as compared to a more high-resolution marker gene.

Altogether, this thesis investigates the diversity of dynamic behavior among closely related wild microbes.

Chapter 2: Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level

Diana M. Chien*, Alison F. Takemura*, Martin F. Polz
Published in *Frontiers in Microbiology* 2014; 5:38.

2.1. Abstract

The Vibrionaceae, which encompasses several potential pathogens, including *V. cholerae*, the causative agent of cholera, and *V. vulnificus*, the deadliest seafood-borne pathogen, are a well-studied family of marine bacteria that thrive in a diverse habitats. To elucidate the environmental conditions under which vibrios proliferate, numerous studies have examined correlations with bulk environmental variables – e.g., temperature, salinity, nitrogen and phosphate – and association with potential host organisms. However, how meaningful these environmental associations are remains unclear because data are fragmented across studies with variable sampling and analysis methods. Here, we synthesize findings about *Vibrio* correlations and physical associations using a framework of increasingly fine environmental and taxonomic scales, to better understand their dynamics in the wild. We first conduct a meta-analysis to determine trends with respect to bulk water environmental variables, and find that while temperature and salinity are generally strongly predictive correlates, other parameters are inconsistent and overall patterns depend on taxonomic resolution. Based on the hypothesis that dynamics may better correlate with more narrowly defined niches, we review evidence for specific association with plants, algae, zooplankton, and animals. We find that *Vibrio* are attached to many organisms, though evidence for enrichment compared to the water column is often lacking. Additionally, contrary to the notion that they flourish predominantly while attached, *Vibrio* can have, at least temporarily, a free-living lifestyle and even engage in massive blooms. Fine-scale sampling from the water column has enabled identification of such lifestyle preferences for ecologically cohesive populations, and future efforts will benefit from similar

* Authors have contributed equally to this work.

analysis at fine genetic and environmental sampling scales to describe the conditions, habitats, and resources shaping *Vibrio* dynamics.

2.2. *Introduction*

The family Vibrionaceae (or vibrios for short) comprises a genetically and metabolically diverse group of heterotrophic bacteria that are routinely found in all ocean environments, ranging from coastal to open and surface to deep water (Thompson et al., 2004)(Thompson and Polz, 2006). Moreover, a few *Vibrio* species have extended their range beyond the marine environment, occurring predominantly in brackish and even freshwater environments (Thompson et al., 2004). The study of the environmental distribution and dynamics of vibrios has a long history, largely because many species contain potential human and animal pathogens (Thompson et al., 2004, Thompson et al., 2005). Hence there is considerable public health and economic interest in determining factors correlated to increased abundance of vibrios (Stewart et al., 2008). Moreover, vibrios are easily cultured on standard and selective media and thus were highly visible in the pre-molecular era of microbial ecology. In recent years, environmental dynamics have also been studied with culture-independent methods allowing for a more fine-scale assessment of environmental drivers of occurrence, and the vibrios have become a model for bacterial population biology and genomics. In fact, presently, the vibrios represent one of the best-studied models for the ecology and evolution of bacterial populations in the wild.

The early discovery that some fish species harbor high numbers of vibrios (e.g., Liston 1954, Liston 1957, Aiso *et al.*, 1968, Sera *et al.*, 1972) has led to the widespread notion that these bacteria are only transient members of microbial assemblages of the water column. Instead, vibrios were regarded as specifically associated with animals, and occurrence in water samples was thought to be primarily due to their excretion with fecal matter. This picture was enforced by the discovery that several luminescent *Vibrio* (*Allivibrio*) and related *Photobacterium* species form intimate symbioses with animals (e.g., fish, squid) (Ruby and Nealson 1976, Stabb 2006). More recent work has, however, revealed that the notion of vibrios being 'enterics of the sea' (Liston 1954) represents an oversimplification. Many *Vibrio* species grow actively in ocean water either in the free-living phase or associated with various types of organic particles, many of which are of non-animal origin (Lyons et al., 2007, Froelich et al., 2012). Thus although association with animals can be an important part of the life cycle of many *Vibrio* species, there are

others that only loosely associate with animals or not at all, an aspect we explore in detail in this review.

Another widely held belief about vibrios is that they play a relatively minor role in chemical transformations in the ocean, despite the wide range of metabolisms [e.g., chitin degradation (Hunt et al., 2008a; Grimes et al., 2009)] of which they are capable. This belief is largely based on low to medium average relative abundance of Vibrionaceae in ocean water. Yet three considerations suggest that the role of vibrios has been underestimated. First, it has been pointed out that although vibrios' abundances are generally only around 10^3 to 10^4 cells per ml seawater (i.e., on the order of few percent of total bacteria), they have very high biomass (Yooseph et al., 2010). For example, an actively growing *Vibrio* can have 100x the biomass of *Pelagibacter*, which, at ~10^5 cells per ml, is typically the most abundant heterotrophic member of bacterial assemblages in the ocean (Yooseph et al., 2010). Second, new time-series analysis shows that vibrios are capable of blooms in the water column during which they can even become the predominant members of the total bacterial assemblage (Gilbert et al., 2012). These blooms had been missed previously because they are of relatively short duration, yet they confirm that vibrios, which are capable of very rapid growth in laboratory media, can reach high doubling rates in the environment. Finally, vibrios might be disproportionately subject to predation by protozoa and viruses (Suttle 2007, Worden et al., 2006).], likely due to their comparatively large size. For example, cells were found in one study to measure more than three times the community average in volume, and, along with other similarly large genera, suffered especially high grazing mortality (Beardsley et al., 2003). Taken together, these considerations suggest that vibrios should be re-evaluated for their role in biogeochemical processes in the ocean since they have disproportionately high biomass that is subject to high turnover by rapid growth in concert with high predation.

The purpose of this review is to provide an overview of known environmental factors and ecological associations affecting *Vibrio* abundance and dynamics. We note that although we look at the dynamics of potentially pathogenic species, we purposefully exclude data on pathogenesis itself since this is outside the scope of this review. We first focus on total *Vibrio* (i.e., the assessment of occurrence of members of the genus or family), which have often been measured as a proxy for potential pathogen occurrence, asking whether they can be treated as an environmentally cohesive unit. To what extent do total vibrios correlate to specific environmental variables, and do these

measures have predictive power for individual species? To address this question, we present meta-analyses of the dynamics of *V. cholerae*, *V. parahaemolyticus*, and *V. vulnificus*, three species harboring genotypes potentially pathogenic to humans. The limitation to these three is necessary since public health interests have driven much of the research so that the literature is highly biased towards human pathogens. In this context, a further important question is to what extent easily measurable bulk parameters, such as temperature, salinity, nutrients, dissolved oxygen and/or chlorophyll a are good correlates for total vibrios or specific species, allowing easy and cost-effective risk assessment.

However, because our meta-analysis suggests poor or inconsistent performance of most bulk parameters, we researched alternative, frequently finer-scale environmental variables. These include associations with different animals, plants and algae, as well as organic polymers, which may occur as suspended particulate matter in the water column and provide resources for attached bacteria. Although such attached lifestyles are common for vibrios, recent research also suggests that many species can occur free-living at least part of the time and be engaged in relative short-lived blooms.

Finally, we summarize recent research aimed at defining habitat characteristics and phylogenetic bounds of ecologically cohesive populations among co-existing vibrios, using the water column and macroinvertebrates as examples of adaptive landscapes. This research demonstrates that such populations, which may or may not correspond to named (taxonomic) species, represent eco-evolutionary units that allow testing of hypotheses of how populations are structured by environmental selection and gene flow.

2.3. Environmental correlates of Vibrio presence and abundance

To better understand under what conditions vibrios occur and proliferate, most studies have investigated environmental variables that can be measured from bulk seawater such as temperature, salinity, dissolved oxygen, nitrogen, phosphorus, and chlorophyll a concentrations. These are attractive since they are easily measured and many are observable remotely by buoy or satellite [e.g., (Lobitz et al., 2000)] so that potential for presence of pathogenic vibrios might be easily assessed. In addition, several studies have extended measurements to more complex physicochemical and biotic variables, including dissolved organic carbon and zoo- and phyto-plankton taxa.

In the following, we first ask how informative these variables are by conducting a meta-analysis to compare correlations across studies, for both total *Vibrio* as well as the potential pathogens *V. cholerae, V. parahaemolyticus,* and *V. vulnificus,* and, second, determine if the genus and species levels exhibit similar patterns. To determine the potential impact of environmental variables, we looked at how strong their correlations are by comparing coefficient of determination values, R^2, reported in the literature. A goodness of fit parameter, R^2 varies from 0 (no explanation of variance in the dependent variable) to 1 (perfect explanation), giving us a means of assessing, for example, whether temperature better predicts abundance of total *Vibrio,* than salinity does. Studies included have regression analyses with associated R^2 values, or Spearman or Pearson correlations, whose rho values were squared to obtain R^2. Additionally, we compare how their abundances trend along gradients in two particularly well-studied variables, salinity and temperature.

2.3.1. Total Vibrio

When correlations across studies are compared, we see that the strongest environmental correlates to total *Vibrio* are temperature and salinity. These two variables most often explain the greatest amount of variance in total *Vibrio* abundance in the water column (Figure 1), whereas consideration of additional variables often makes only marginal improvements [e.g., in (Heidelberg et al., 2002; Oberbeckmann et al., 2012; Froelich et al., 2013)]. However, a minority of analyses has found temperature and salinity to be non-significant toward explaining *Vibrio* abundance. This inconsistency might be a result of the ranges considered; for instance, temperature may be found non-significant due to a narrow range observed, such that *Vibrio* abundance varies little. In fact, evidence supports this hypothesis; the correlation strength of temperature to vibrios varies by season (Oberbeckmann et al., 2012; Froelich et al., 2013), suggesting the magnitude of the correlation may depend on the temperature range examined. For instance, Oberbeckmann et al., (2012) and Froelich et al., (2013) both observed the highest correlation of temperature and *Vibrio* during the seasons with the broadest temperature ranges, spring and fall, respectively. Additionally, it is possible that at lower temperatures vibrios exhibit less variation in abundance; two studies assessing total vibrios in the cooler waters of the Baltic Sea and North Sea found non-significant correlations (Eiler et al., 2006; Oberbeckmann et al., 2012).

Compared to salinity and temperature, other environmental measures usually explain less variance in total *Vibrio*. Dissolved oxygen has had little explanatory power; for instance, in Figure 1, its largest R^2 was less than half that of temperature in the same analysis (Blackwell and Oliver, 2008). The same is true for nitrogen, whose highest R^2 was still less than temperature's (Blackwell and Oliver, 2008). In the environments examined, phosphate, pH, and turbidity explain little variance, and dissolved organic carbon (DOC) explains none at all, albeit the number of studies used for DOC in this meta-analysis is limited. Of interest, though not depicted, potential host organisms, copepods, decapods, and cyanobacteria, have been found to explain relatively little variance in total vibrios when considered in a model that already incorporates temperature (Turner et al., 2009; Vezzulli et al., 2009), and similarly for dinoflagellates when salinity is first considered (Eiler et al., 2006). Turner and colleagues (2009) did observe that diatoms explained more variance than temperature. While this might imply a physical association, the correlation was negative, suggesting that total *Vibrio*, at least as a whole, do not associate with diatoms.

Chlorophyll a, on the other hand, has had noted importance in two datasets: the spring and summer of the study by Oberbeckmann and colleagues (2012), with R^2 values of 60% and 26%, respectively. These were in fact higher than correlations to temperature or salinity in these seasons. Perhaps during this period, as temperature warms, growth conditions favor phytoplankton blooms that impact *Vibrio* abundance (Oberbeckmann et al., 2012). However, Froelich et al., (2013) did not make these same observations in their seasonal datasets. This inconsistency may be a product of the fact that different *Vibrio* species likely affiliate with or feed on exudates of specific algal taxa only, rather than algae in general, a subject further discussed in the section "The evidence for a planktonic, free-living lifestyle."

Given the frequent strength of temperature and salinity as correlates, we asked, how do total vibrios distribute with respect to these variables when their combined effect is considered? A few studies have modeled the bivariate relationship, finding that total *Vibrio* abundance increases as temperature and salinity increase (Hsieh et al., 2008; Turner et al., 2009; Froelich et al., 2013). The ranges investigated were also broad, lending confidence that these results are general; for example, Hsieh et al., (2008) modeled from 2.5 to 32.5°C and 0 to 27 ppt, respectively.

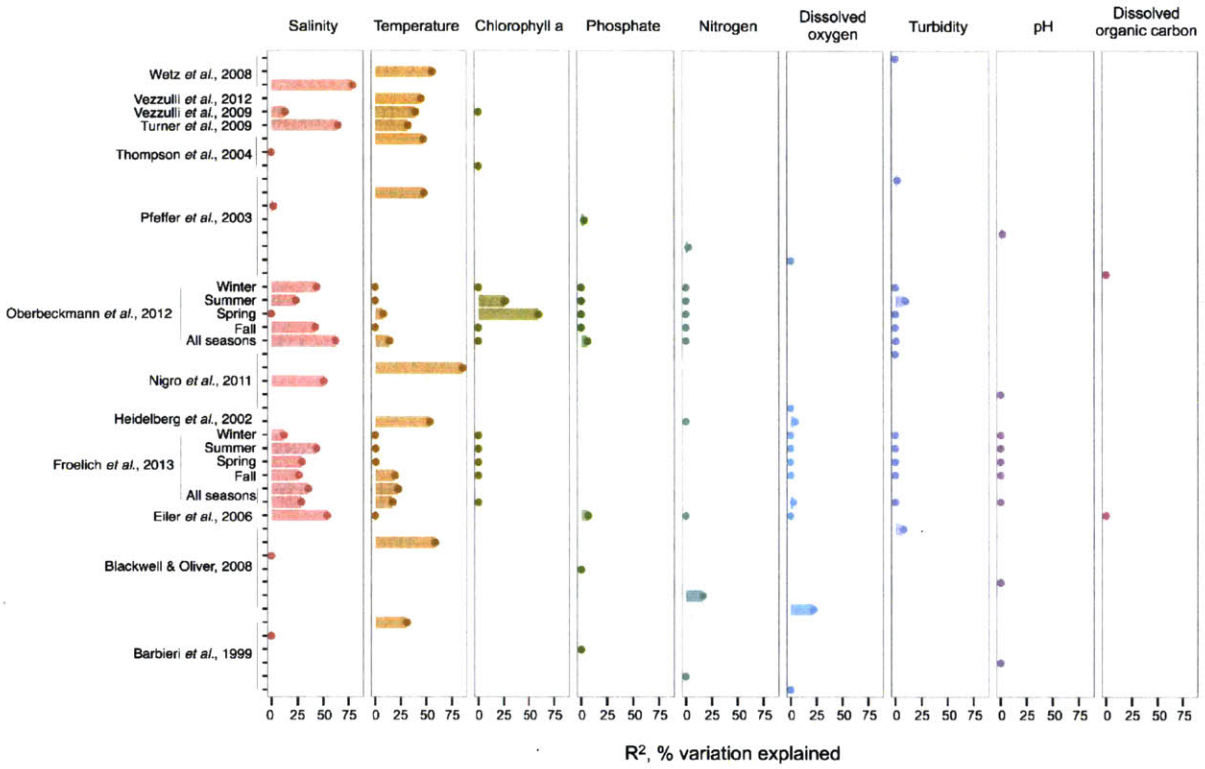


Figure 1. An overview of regression analyses indicates that temperature and salinity explain most variation in bulk-water total Vibrio abundance. The R^2, or pseudo-R^2, values associated with regression analyses are shown for selected environmental variables that are well-represented across studies. An individual study may perform multiple analyses because variables are considered for correlation independently (for ex. Wetz et al., 2008); because datasets are split (e.g. between seasons in Oberbeckmann et al., 2012); or because different sets of variables are considered sequentially (e.g. two variables versus six variables in the two All Seasons models from Froelich *et al.*, 2013). Dots indicate bar heights, and where a dot occurs without a bar, R^2 was non-significant (i.e. $R^2 = 0$). Variables may have been log or exponentially transformed in references.

2.3.2. V. cholerae, V. parahaemolyticus, and V. vulnificus

We compare environmental correlates and trends noted in total *Vibrio* to three species that have been well sampled across locales: *V. cholerae, V. parahaemolyticus,* and *V. vulnificus*. While it would also be interesting to consider species beyond potential pathogens, their environmental data is much more limited.

In *V. cholerae*, we see an interesting shift from total *Vibrio* in the strength of correlating environmental variables: some biotic variables are as strong or, in fact, stronger than temperature or salinity (Figure 2). Total *Vibrio*, congenerics *V. vulnificus*

and *V. parahaemolyticus*, as well as a dinoflagellate genus (*Prorocentrum*) and cladoceran species (*Diaphanosoma mongolianum*) have all significantly correlated to *V. cholerae* abundance (Eiler et al., 2006; Blackwell and Oliver, 2008; Kirschner et al., 2011; Prasanthan et al., 2011). Moreover, *V. parahaemolyticus* abundance has explained more *V. cholerae* abundance variance than nitrogen, temperature, or salinity in (Prasanthan et al., 2011), and dinoflagellate abundance has explained more variance than phosphorus, salinity, or temperature (Eiler et al., 2006). While correlations to plankton may represent direct associations, such high correlation of vibrios to each other is likely not indicative of causal interactions, but rather stems from overlap in environmental ranges and/or habitats (Blackwell and Oliver, 2008). *E. coli* and total coliforms have also correlated to *V. cholerae* abundance, though both groups may simply be responding to anthropogenic nutrient influxes favoring growth of heterotrophs (Blackwell and Oliver, 2008).

Long thought to be a reservoir of toxigenic *V. cholerae*, zooplankton, and particularly copepods, are hypothesized to correlate to *V. cholerae* abundance. Surprisingly, however, when Magny and colleagues (2011) examined several zooplankton genera and species, including copepods *Cyclops* and *Diaptomus*, they did not find significant correlations to any zooplankter except the rotifer *Brachionus angularis* (not depicted in Figure 2, because Monte Carlo analysis did not yield R^2 values). While the association between *V. cholerae* O1/O139 and the copepod *Acartia tonsa* has also been studied (Huq et al., 2005; Lizárraga-Partida et al., 2009), quantitatively significant correlation in the environment has remained elusive. For instance, Lizárraga-Partida and colleagues (2009) demonstrated only a qualitative link between *V. cholerae* O1 presence coincident with an increase in *A. tonsa*, even though laboratory studies have shown ready attachment (e.g., Huq et al., 1984; Rawlings et al., 2007).

V. cholerae has also been hypothesized to correlate with chlorophyll a, a potential proxy of algal and zooplankton growth, and/or a eutrophic environment conducive to heterotroph growth, but chlorophyll a's general predictive value is unclear. While significant in Eiler et al., (2006), other studies have observed no correlation of chlorophyll a to *V. cholerae* abundance (Jiang and Fu 2001; Kirschner et al., 2008; Mishra *et al.*, 2012). Yet *V. cholerae* growth has been observed experimentally to depend on dissolved organic carbon, which could relate to phytoplankton abundance and thus chlorophyll a (Eiler et al., 2007). In microcosm experiments, Eiler and colleagues (2007) demonstrated that adding 2.1 mg carbon L^{-1} of cyanobacterial-derived dissolved organic matter influenced bacterial growth more than a 12 to 25°C change in temperature. The

inconsistency of chlorophyll a, and, incidentally, bulk DOC (which showed no significant correlation) (Eiler et al., 2006; Blackwell and Oliver, 2008; Kirschner et al., 2008; Neogi et al., 2012) as correlates might be due to the quality of exudates; its composition of refractory humic substances (Kirschner et al., 2008) or derivation from different algal species, differentially stimulating *V. cholerae* growth [(Worden *et al.*, 2006), see also section "The evidence for a planktonic, free-living lifestyle"]. Interestingly, the lack of clear support for chlorophyll a's influence on *V. cholerae* environmental abundance is in contrast to the fact that chlorophyll a can correlate with cholera *disease* incidence (Magny et al., 2008), and has been used in predictive models for cholera in Bangladesh (Bertuzzo et al., 2012; Jutla et al., 2013).

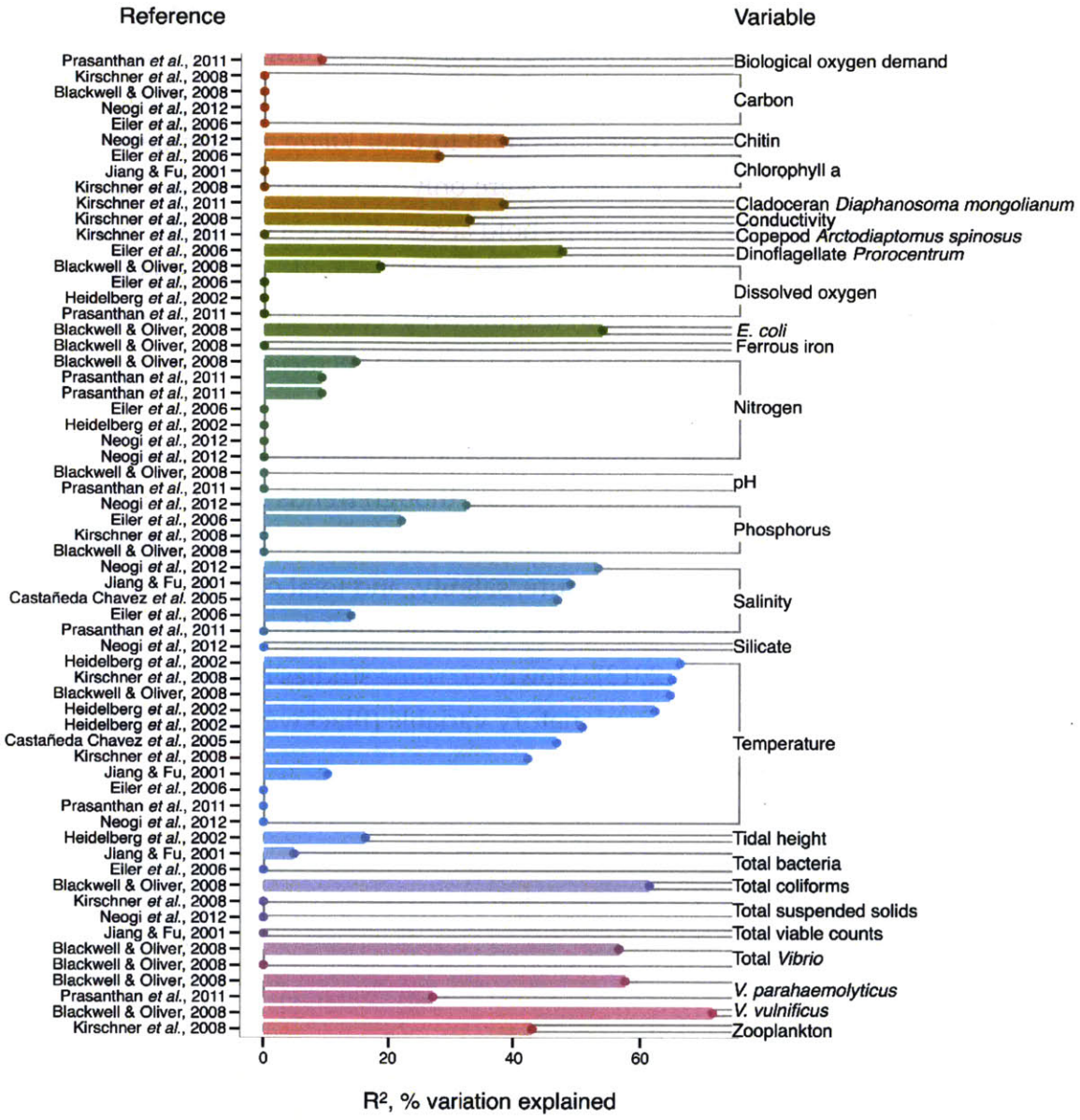


Figure 2. Variation in *V. cholerae* abundance or percent positive samples is best explained by temperature, other organisms, and salinity. R^2, or pseudo-R^2, values from analyses across studies are depicted grouped by variable, and then in rank order, with their associated reference. A reference may conduct multiple analyses for a given variable (e.g., on subsets of data or considering different variables combinations for data regression). Dots indicate bar heights, and where a dot occurs without a bar, R^2 was non-significant (i.e. $R^2 = 0$).

Like *V. cholerae*, *V. parahaemolyticus* abundance in water samples is also strongly correlated to temperature, and was found significant in all but one analysis reviewed

here (DePaola et al., 1990; Zimmerman et al., 2007; Blackwell and Oliver, 2008; Carbulotto et al., 2010; Deter et al., 2010; Johnson et al., 2010; Johnson et al., 2012; Böer et al., 2013), with maximal $R^2 = 50.6\%$ (Deter et al., 2010) (**Figure 3**). Blackwell and Oliver (2008) found that *V. parahaemolyticus* correlates both to total *Vibrio* and congenerics, as well as coliforms and *E. coli*. These variables were only considered in a single study, however, so it is not known if the relationships hold across different sampling locations. The significance of salinity is variable for *V. parahaemolyticus* with only three of seven studies having non-zero R^2 values (**Figure 3**) (Zimmerman et al., 2007; Carbulotto et al., 2010; Johnson *et al.*, 2010), but this may be due to *V. parahaemolyticus* colonizing a large salinity range, as detailed below (**Figure 6**).

Correlation to environmental variables has also frequently been studied for *V. parahaemolyticus* occurring in sediment and shellfish, though trends remain unclear. In sediment, considered a potential reservoir (Vezzulli et al., 2009), individual regressions of *V. parahaemolyticus* abundance to temperature, salinity, and total organic carbon have yielded moderate R^2 values, at times above 30% (Blackwell and Oliver, 2008; Deter et al., 2010; Johnson et al., 2012; Böer et al., 2013). However, some studies have found salinity or temperature to be a non-significant explanatory variable (Blackwell and Oliver, 2008; Deter et al., 2010; Johnson et al., 2010).

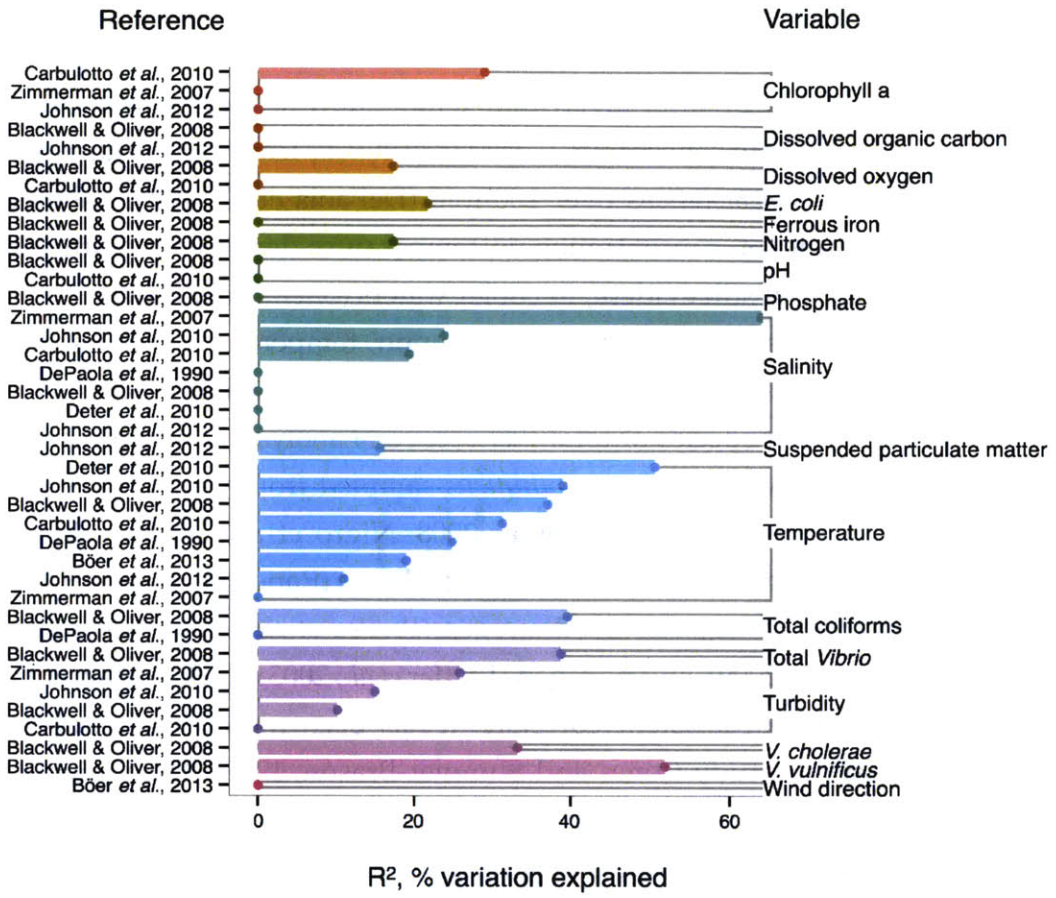


Figure 3. Variation in *V. parahaemolyticus* abundance or percent positive samples is best explained by temperature and other organisms. R^2, or pseudo-R^2, values from analyses across studies are depicted grouped by variable, and then in rank order, with their associated reference. A reference may conduct multiple analyses for a given variable (e.g., on subsets of data or considering different variables combinations for data regression). Dots indicate bar heights, and where a dot occurs without a bar, R^2 was non-significant (i.e. $R^2 = 0$).

In shellfish, a common vehicle of virulent vibrios to humans, the incidence of temperature and salinity as correlates to *V. parahaemolyticus* is also inconsistent. Salinity has been found explanatory in some studies, with R^2 as high as 42% (DePaola et al., 2003; Johnson et al., 2010; Johnson et al., 2012) and non-significant in others (Deepanjali et al., 2005; Deter et al., 2010; Sobrinho et al., 2010). Temperature can explain moderate amounts of variance in *V. parahaemolyticus* abundance (DePaola *et al.*, 1990; DePaola et al., 2003; Cook et al., 2002; Johnson et al., 2010; Sobrinho et al., 2010; Johnson et al., 2012),

with significant R^2 as high as 44% (Cook et al., 2002), though other studies have found little or no correlation (Deepanjali et al., 2005; Duan and Su, 2005; Deter et al., 2010). The absence of correlation is surprising, given that temperature's effect is amplified by influencing shellfish's ability to concentrate *V. parahaemolyticus* from surrounding water. Oysters can enrich *V. parahaemolyticus* over 100-fold (DePaola et al., 1990; Shen et al., 2009), and the magnitude of concentration is temperature-dependent, with effects greatest at 32°C and less, but still evident, in cooler waters (Shen et al., 2009).

For *V. vulnificus* isolated from the water column, temperature is the strongest correlate among measured environmental variables, and often explains more variance in *V. vulnificus* than for other species or total *Vibrio*; several analyses found temperature explained over 50% of the variance in *V. vulnificus* sampled from water (Motes et al., 1998; Randa et al., 2004; Blackwell and Oliver 2008; Nigro et al., 2011). Moreover, temperature has been a stronger correlate than chlorophyll a (Randa et al., 2004; Johnson et al., 2010; Johnson et al., 2012), dissolved oxygen (Pfeffer *et al.*, 2003; Blackwell and Oliver, 2008; Ramirez et al., 2009), and nitrogen (Pfeffer *et al.*, 2003; Blackwell and Oliver, 2008). While DOC is an inconsistent correlate, it has been more explanatory than temperature in at least one study (Jones and Summer-Brason, 1998). The variable pH, however, is not a significant correlate (Lipp et al., 2001; Pfeffer et al., 2003; Blackwell and Oliver, 2008; Ramirez et al., 2009; Franco et al., 2012), nor is phosphorus (Pfeffer et al., 2003; Blackwell and Oliver, 2008). Turbidity has been found non-significant in several studies (Lipp et al., 2001; Pfeffer et al., 2003; Wetz et al., 2008; Ramirez et al., 2009), or not as explanatory as temperature (Blackwell and Oliver, 2008). While salinity, when significant, has generally been less informative than temperature (Motes et al., 1998; Randa et al., 2004; Warner and Oliver, 2008; Johnson et al., 2010), it has, in one analysis, been more (Lipp et al., 2001).

Biotic correlates have also been identified for *V. vulnificus*. Total bacteria (Pfeffer et al., 2003; Randa et al., 2004; Blackwell and Oliver, 2008), enteroccous (Wetz et al., 2008; Ramirez et al., 2009), coliforms (Pfeffer et al., 2003; Blackwell and Oliver, 2008) and *E. coli* (Pfeffer 2003; Blackwell and Oliver, 2008; Wetz et al., 2008) have been studied only sporadically, but their correlation strength to *V. vulnificus* has usually been less than temperature's; one exception, however, is enterococcus in (Ramirez et al., 2009), potentially indicative of a surge in nutrients overtaking temperature's effect on growth. Interestingly, total *Vibrio* have explained substantial variance (R^2 = 43 to 54%) in *V. vulnificus* in more instances than for other *Vibrio* species (Pfeffer et al., 2003; Wetz et al.,

2008; Blackwell and Oliver, 2008), suggesting they are responding similarly to their environments under the conditions studied. However, instances do occur where total *Vibrio* and *V. vulnificus* do not correlate (Høi et al., 1998; Wetz et al., 2008), underscoring that a species is not a constant component of a genus, and may respond to environmental conditions independently.

Isolations of the three potentially pathogenic species across salinity and temperature gradients were also looked at, and found to exhibit different patterns. *V. cholerae* has a wide temperature range (~10 to 30°C) in brackish water (1 to 10 ppt), and generally decreases with increasing salinity over the entire range examined (0 to 40 ppt) (**Figure 5**). Observed *V. cholerae* abundance is greatest around 20°C and 0 to 10 ppt, on the order of 10^3 cells per mL. At less-favorable, higher salinities, *V. cholerae* has been found around this temperature, though in much lower abundances (on the order of 1 cell per mL). Interestingly, *V. cholerae*'s realized niche is much smaller than its fundamental one, as it has maximal temperature and salinity tolerances around 38°C and 75 ppt (Materna et al., 2012), suggesting other controls on its abundance in the environment.

V. parahaemolyticus contrasts *V. cholerae* by having a more constant abundance that is broadly spread out over salinities of 3 to 35 ppt in a narrow, much warmer temperature range, centered roughly around 29°C. (Figure 6). Consistent with this finding, it has been noted that this species prefers warmer waters (>20°C) (Martinez-Urtaza et al., 2012), and has been observed to grow best at 25°C *in vitro* (Nishina et al., 2004). However, isolations from shellfish can exhibit different trends from those observed in the water column; Martinez-Urtaza et al., (2008) detected *V. parahaemolyticus* in mussels gathered in much cooler, 15°C water, consistent with the potential for shellfish to concentrate *V. parahaemolyticus*.

A previous literature-based analysis showed *V. vulnificus* to have a more complicated relationship to temperature and salinity than either *V. cholerae* or *V. parahaemolyticus*. It has a narrow temperature range at higher salinities (>10 ppt) while at low salinities (between 5 and 10 ppt) its temperature range more than doubles – from 22-30°C to 10-32°C (Randa et al., 2004). This suggests that, in temperate climates, this species is found year-round in estuarine, low salinity environments but can expand into full strength seawater during warmer months. In the tropics, this species should be endemic to the ocean.

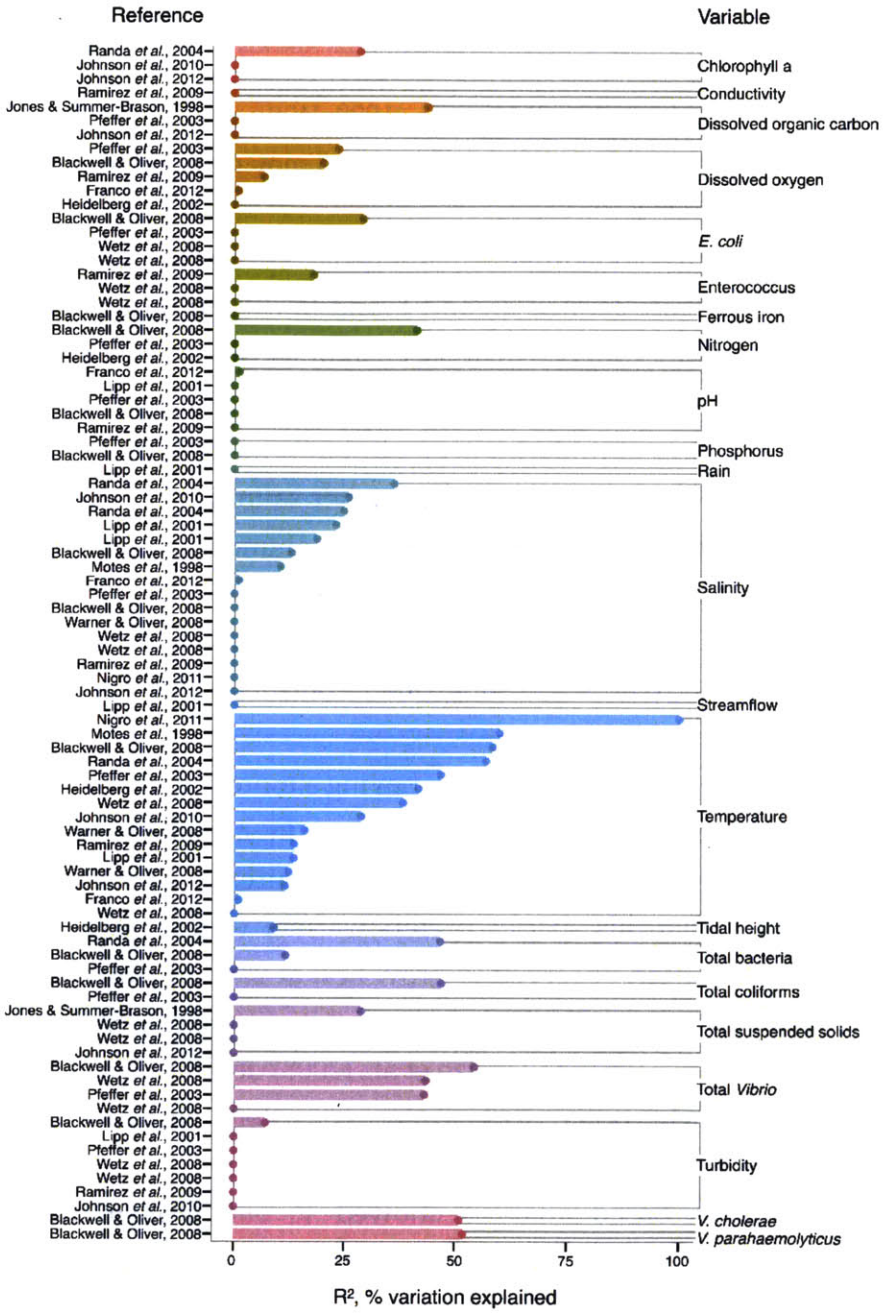


Figure 4. Variation in *V. vulnificus* abundance or percent positive samples is best explained by temperature, and other organisms, including Vibrio. R^2, or pseudo-R^2, values from analyses across studies are depicted grouped by variable, and then in rank order, with their associated reference. A reference may conduct multiple analyses for a given variable (e.g., on subsets of data or considering different variables combinations for data regression). Dots indicate bar heights, and where a dot occurs without a bar, R^2 was non-significant (i.e. $R^2 = 0$).

2.3.3. Conclusions from meta-analysis

From this meta-analysis, we find, first, that temperature and salinity often explain more variance than any other bulk water parameter, like phosphate, nitrogen, pH, or DOC. Yet some of the difficulty in making general statements regarding the relationship of vibrios to individual environmental variables likely stems from the fact that their strength can depend on the ranges examined, e.g. as for temperature, or in quality of the variable, such as DOC, which will encompass carbon derived from different sources that may impact *Vibrio* growth differentially. Second, we observe that trends that apply to the whole genus *Vibrio* do not necessarily reflect those of individual species. Total vibrios and the well-studied potential pathogens *V. cholerae*, *V. parahaemolyticus*, and *V. vulnificus* correlate with shared and distinct environmental variables. For *V. parahaemolyticus* and *V. vulnificus*, temperature often explains more variance than does salinity in the same analysis, and for *V. cholerae*, diverse biotic variables, including specific phyto- and zooplankton taxa, can be stronger correlates than abiotic variables. Unfortunately, biotic variables, particularly individual plankton taxa, have rarely been studied in more than one instance, making these observations difficult to generalize. But the correlations reviewed above hint that there may be ecological relationships between *Vibrio* and plankton that merit deeper investigation.

Across salinity and temperature gradients, the pattern also differs between total *Vibrio* and individual species, and species' patterns differ from each other. Indeed, differences may occur even within taxonomic species; *V. parahaemolyticus* pathogenic genotypes have been observed to be a variable fraction of total *V. parahaemolyticus* (Zimmerman et al., 2007). For example, at their Alabama site, total *V. parahaemolyticus* – detected via thermolabile hemolysin marker (tlh) – remained at a more constant concentration of between 1 and 10 cells per mL, while toxigenic genotypes – thermolabile hemolysin+ and thermostable direct hemolysin+ cells – fluctuated in a much wider range: between 0.0001 and 10 cells per mL. This result argues against using the total species to infer the potential pathogens. Taken together with the results from the meta-analysis, these findings suggest that finer-scale sampling – of both the environmental parameters and the *Vibrio* population of interest – is necessary to link ecological parameters to cellular abundances.

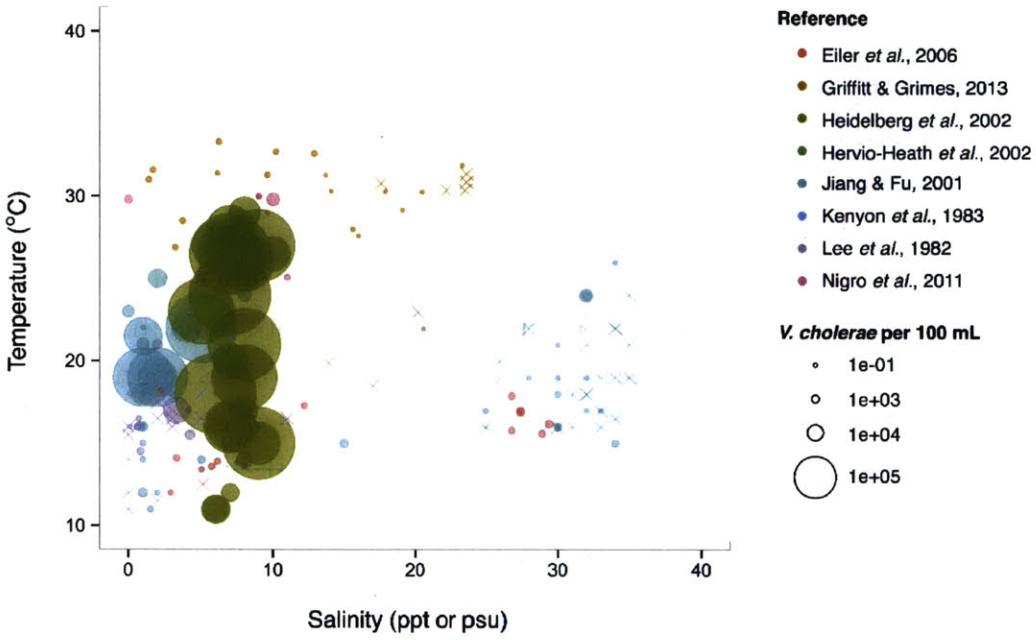


Figure 5. *V. cholerae* favors lower salinity and occupies a broad temperature range. *V. cholerae* concentrations, i.e. MPN-estimated CFU or molecular marker gene copies per 100 mL, reported in different studies are plotted against the temperature (°C) and salinity values (ppt or psu) at which they were found. All studies report *V. cholerae*, including O1/O139 and non-O1/non-O139, except for Heidelberg *et al.*, (2002), whose genetic marker detected *V. cholerae*/*V. mimicus*. Circle (◎) sizes correspond to concentrations, but note the breaks are scaled for clearer visualization, and not linearly. (X) indicates no *V. cholerae* found in that sample.

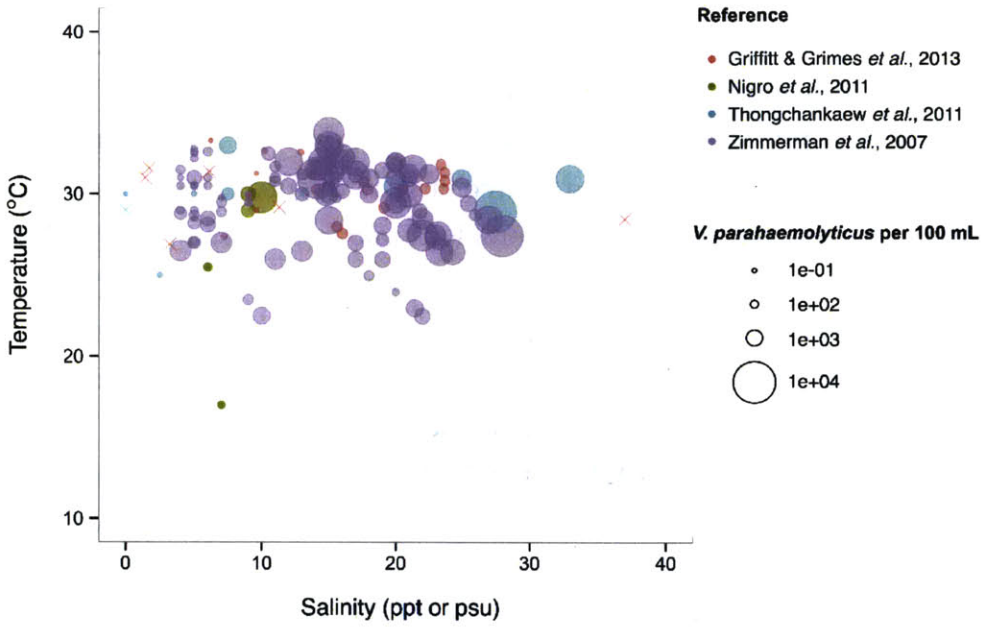


Figure 6. *V. parahaemolyticus* favors high temperatures but is relatively unconstrained by salinity. Concentrations, i.e. MPN-estimated CFU or molecular marker gene copies per 100 mL, reported in different studies are plotted against the temperature (°C) and salinity values (ppt or psu) at which they were found in bulk water samples. Circle (◎) correspond to concentrations, but note the breaks are scaled for clearer visualization, and not linearly. (X) indicates no *V. parahaemolyticus* found in that sample.

2.4. Associations with complex and particulate marine growth substrates

The previous sections demonstrate that, with the exception of temperature and salinity, parameters measured in bulk seawater have shown limited power in explaining the environmental dynamics of *Vibrio* species. This may, in part, be due to the narrow focus on only a few (potentially) pathogenic species, and frequently limited comparability of measured parameters across studies. It is also likely, however, that bulk measurements, such as dissolved oxygen, nitrogen and phosphate concentration in seawater, only poorly capture the ecological parameters that *Vibrio* populations are associated with or respond to. Vibrios are often presumed to primarily attach to biological surfaces, yet may also subsist on dissolved resources of biological origin while free-living. Taking these resource associations into account, their environmental dynamics may be somewhat decoupled from parameters measurable in bulk seawater, and may depend more on the concentration and properties of relevant solid or dissolved

resources. We review in the following sections the ample evidence for surface-associated niches, as well as more recent evidence for environmental dynamics including free-living states and formation of blooms.

From the perspective of bacteria attaching to surfaces, these are either metabolically inert or can be degraded as a source of growth substrates. Vibrios have the ability to attach to and degrade a considerable number of polymeric substrates (Johnson 2013), suggesting that specific association with surfaces is an important growth strategy. For example, nearly all vibrios can metabolize the abundant biopolymer chitin (present in both crustacean and diatom shells in the marine environment) (Hunt et al., 2008a; Grimes et al., 2009), and various representatives can metabolize an array of plant/algal polysaccharides: agar, alginate, fucoidan, mannan, cellulose, pectin, and laminarin (Goecke et al., 2010). In addition, vibrios may metabolize plastic wastes, as suggested by a recent study documenting that vibrios make up the majority of bacteria attached to plastic wastes floating in the ocean, with electron microscopy showing individual cells residing at the bottom of pits (Zettler et al., 2013). Although this suggests that these plastics, which had been thought to be largely biologically inert, could be degraded by vibrios, such activity remains to be confirmed.

Evidence is also accumulating that vibrios may play a role in oil spill degradation: *Vibrio* representatives can metabolize oil-derived compounds (West et al., 1984; Moxley and Schmidt, 2010), and have been found to comprise a sizable fraction of oil-associated microbial communities from the Deepwater Horizon spill, both from sea-surface samples (>31% in the molecular study of Hamdan and Fulmer, 2011) and salt-marsh plants contaminated with oil mousse (57% in the study of Liu and Liu, 2013). While a clear positive effect of crude oil on *Vibrio* growth has yet to be demonstrated in vitro, it appears that many vibrios can at least persist in the presence of oil (Stephens et al., 2013). *Vibrio* representatives furthermore show resistance to inhibition by the oil dispersant Corexit (Hamdan and Fulmer, 2011), which was widely used following the Deepwater Horizon spill; this resistance may additionally support an ability to persist after oil spills.

Most associations with specific surfaces have, however, been described for plants, algae, and animals, and the following section explores these organisms as potential biological niches for vibrios.

2.5. *Biological niches for Vibrio*

Vibrio have been detected on a plethora of aquatic biological surfaces, but which of these associations represent more than transient, incidental attachments? In the following sections we consider which aquatic plants (Table 1) and animals (Table 2) may represent sustained *Vibrio* niches, on the basis of (i) numerical enrichment compared to the surrounding medium, and (ii) knowledge of biological mechanisms, e.g., availability of nutrition and shelter, potentially supporting an association. In doing so, we also draw attention to the need for more quantitative and mechanistic approaches to understanding the ecological associations that allow vibrios to flourish – approaches that could underpin more powerful predictions of *Vibrio* dynamics arising from these diverse associations. We note also that many of the following observations are limited to *V. cholerae* because of its prominence as a pathogen, but the same niches may be available to other vibrios with similar biological activities.

2.5.1. Associations with Plants

Vibrio survival is enhanced in association with certain freshwater and estuarine plants (Table 1). Plant hosts can provide nutrition (Andrews and Harris, 2000) and the opportunity to form predation-resistant biofilms (Matz et al., 2005), and have been postulated to modulate unfavorably cold temperatures as well (Criminger et al., 2007). Two freshwater aquatic plants have been observed to support both in situ enrichment (in freshwater bodies of Bangladesh) and in vitro survival advantage for *V. cholerae*: duckweed, *Lemna minor* (Islam et al., 1990b), and water hyacinth, *Eichhornia crassipes* (Spira et al., 1981), with preference for roots of the latter. Concentration on *E. crassipes* roots may indicate that root exudate is a particularly rich nutritional source, but may also be an artifact of the fact that the roots represent the greatest area exposed to water, and hence to inoculation by planktonic *Vibrio*. By contrast, duckweed's minimal structure, lacking stem or developed leaves, means that almost the entire plant is in contact with the water and thus available for inoculation.

Among estuarine plants, nitrogen-fixing representatives of several *Vibrio* taxa – including *V. diazotrophicus, V. natriegens, V. cininnatiensis* (Urdaci et al., 1988), and *V. parahaemolyticus* (Criminger et al., 2007) – appear to be noteworthy members of the rhizosphere, given that they represent more than half of the culturable diazotrophs associated with the dominant marsh grasses *Spartina* sp. and *Juncus roemerianus* (Bagwell et al., 1998; Larocque et al., 2004), and the herb *Salicornia virginica* (Bergholz et al., 2001;

Criminger et al., 2007). While this numerical dominance may reflect culturing bias, later molecular studies of the *S. alterniflora* rhizosphere confirmed that vibrios (not taxonomically resolved below the level of the family) are stable constituents of the community (Lovell et al., 2008), with little seasonal fluctuation (Gamble et al., 2010). Nitrogen fixation thus appears to be an effective strategy supporting *Vibrio* survival in the anaerobic rhizosphere, demonstrating the ecological breadth granted by vibrios' facultatively anaerobic metabolism.

2.5.2. Associations with microalgae and filamentous cyanobacteria

While early culture-based studies have demonstrated numerical dominance of vibrios on phytoplankton surfaces compared to surrounding water, e.g., (Simidu et al., 1971), little is known about direct, physical associations with specific phytoplankton. Algal cells represent a nutritional opportunity in that they often excrete a high proportion of their photosynthetically fixed carbon, thereby creating a diffusive sphere (the phycosphere) around them, with elevated organic carbon concentration compared to the bulk (Paerl and Pinckney, 1996). However, in vitro survival advantage and persistence have been thus far been demonstrated only for *V. cholerae* in physical association with two microalgae: with the filamentous freshwater green alga *Rhizoclonium fontanum* (Islam et al., 1989), and inside the mucilaginous sheath of *Anabaena* sp. cyanobacteria under both freshwater (Islam et al., 1990a, 1999) and saline conditions (Ferdous, 2009) (Table 1).

Recent work has illuminated mechanistic details of the *V. cholerae* association with *Anabaena*, which may follow the canonical model of symbioses between heterotrophic bacteria and nitrogen-fixing freshwater cyanobacteria. In such associations, heterotrophs locate their hosts via chemotaxis and benefit from rich cyanobacterial exudate (Paerl and Gallucci, 1985). In return, their oxidative metabolism both relieves oxygen inhibition of nitrogen fixation (which would otherwise limit rapid algal growth), and generates carbon dioxide for photosynthetic assimilation (Paerl and Gallucci, 1985). For *V. cholerae*, chemotactic preference for components of the Anabaena mucilaginous sheath has been demonstrated (Mizanur et al., 2002). Furthermore, investigators have shown that both chemotaxis to and survival on *Anabaena* depend on *V. cholerae*'s expression of mucinase (Islam et al., 2002, 2006). The exact role of mucinase has yet to be defined, but activity of secreted mucinase might liberate from mucus the relevant chemotactic attractants, aid colonizing *Vibrio* in physical penetration of the mucilage,

and/or convert mucilage to nutritive compounds supplementary to the cyanobacterial exudate.

2.5.3. Associations with macroalgae

Numerous studies have shown that vibrios are one of the most abundant culturable constituents of macroalgal communities (Table 1): a recent meta-analysis of 161, predominantly culture-dependent macroalgal-bacterial studies determined that vibrios on average comprised 10% of these communities (Hollants et al., 2013), with 28%, 28% and 44% of them found on brown, green, and red macroalgae, respectively. While no molecular studies have yet quantified *Vibrio* within macroalgal communities, numerical enrichment of culturable vibrios has been demonstrated for the brown algae *Ascophyllum nodosum* (Chan and McManus, 1969), and *Laminaria longicruris* (Laycock 1974); the red algae *Hypnea* sp. (Lakshmanaperumalsamy and Purushothaman, 1982), *Polysiphonia lanosa* (Chan and McManus, 1969), and *Porphyra yezoensis* (Duan et al., 1995); and the green algae *Chaetomorpha* sp. (Lakshmanaperumalsamy and Purushothaman, 1982), *Enteromorpha* sp. (Lakshmanaperumalsamy and Purushothaman, 1982), and *Ulva pertusa* (Duan et al., 1995). For *V. cholerae,* in vitro survival advantage has been shown on the green algae *Ulva lactuca* and *Enteromorpha intestinalis* and the red alga *Polysiphonia lanosa* (Islam et al., 1988).

As mentioned above, vibrios can metabolize many algal polysaccharides; they have furthermore been implicated in several other biological activities facilitating symbiosis with macroalgal hosts. These include antagonism directed towards potential bacterial or algal competitors for host surface area (Dobretsov et al., 2002; Kanagasabhapathy et al., 2008), developmental morphogenic effects on *Ulva pertusa* (Nakanishi et al., 1996), and stimulation of spore germination for *Ulva* sp. (Patel et al., 2003; Tait et al., 2005). Hence multiple lines of evidence point to significant *Vibrio* association with *Ulva* sp. (enrichment, survival, morphogenesis and spore modulation) and *Polysiphonia* sp. (enrichment, survival) in particular.

Table 1. Plant and algae hosts for vibrio, as demonstrated by numerical enrichment and biological mechanisms supporting association.

Host	Reference, study site	Associated vibrios	Enumeration method	Enrichment, survival advantage	Host site, mechanism of association
Plants, freshwater					
***Eichhornia crassipes* (water hyacinth)**	Spira et al., 1981: Bangladesh, freshwater bodies	*V. cholerae* O1 El Tor	Culture	In situ enrichment: 84% incidence on plants, 16% in water only. In vitro survival advantage: enriched by 10^2-10^3 compared to surrounding water.	Possible preference for root exudate
***Lemna minor* (duckweed)**	Islam et al., 1990b: in vitro	*V. cholerae* O1: one clinical strain, one environmental (from Australian river water)	Culture	In vitro survival advantage: >27 days survival of attached cells, versus 15-21 days for cells in surrounding water.	Whole plant; mechanism untested
Plants, estuarine					
***Spartina alterniflora, Spartina patens* (marsh grass)**	Bagwell et al., 1998; Lovell et al., 2008 ; Gamble et al., 2010: South Carolina estuary, USA	Spp. including *V. alginolyticus, anguillarum, diazotrophicus, parahaemolyticus*	Culture; molecular	In situ enrichment: >50% of culturable diazotrophs; molecular evidence (Gamble et al., 2010) demonstrates stable abundance across seasons	Root association; anaerobic diazotrophy
***Juncus roemarianus* (marsh grass)**	Larocque et al., 2004: South Carolina estuary, USA	Vibrionaceae	Culture	In situ enrichment: >50% of culturable diazotrophs	Root association; anaerobic diazotrophy
***Salicornia viginica* (marsh herb)**	Bergholz et al., 2001; Criminger et al., 2007: South Carolina estuary, USA	Vibrionaceae	Culture	In situ enrichment: >50% of culturable diazotrophs	Root association; anaerobic diazotrophy
Microalgae and filamentous cyanobacteria, freshwater					
***Rhizoclonium fontanum* (filamentous green alga)**	Islam et al., 1989: in vitro	*V. cholerae* O1 strains from Australian and Bangladeshi surface water; O1 Bangladeshi clinical isolates	Culture	In vitro survival advantage: 21 days survival of attached cells, compared to 3 days in surrounding water and in no-algae control.	Mechanism untested
Anabaena variabilis	Islam et al., 1990; Mizanur et al., 2002; Islam et al, 2002, 2006: in vitro	*V. cholerae* O1 Bangladeshi environmental isolates	Culture	In vitro survival advantage: up to 5 days survival of attached cells; >6 survival in associated water. Persist as VBNC inside algal sheath up to 15 months.	Mucilaginous sheath, with possible preference for heterocysts. Possible mechanism: benefiting from algal exudate while relieving oxygen inhibition of N_2 fixation and contributing CO_2. Demonstrated mechanisms: chemotaxis to host mucus components; mucinase dependence of both chemotaxis and survival with host.
Macroalgae, marine					
Brown algae					

Ascophyllum nodosum	Chan and McManus 1969: Canada	*Vibrio* spp.	Culture	In situ enrichment: Dominant culturable bacteria; enriched by 10^2-10^4 compared to water column	Algal polysaccharide metabolism
***Laminaria* spp.**	Laycock 1974: Nova Scotia, Canada; Wang et al., 2008	Spp. incl. *V. tasmaniensis*	Culture	In situ enrichment: Dominant culturable bacteria	Algal polysaccharide metabolism; laminaranolytic activity in particular demonstrated
Red algae					
***Hypnea* spp.**	Lakshmanaperumalsamy and Purushothaman, 1982: tropical estuary, Africa	*Vibrio* spp.	Culture	In situ enrichment: Dominant culturable bacteria	Algal polysaccharide metabolism
Polysiphonia lanosa	Chan and McManus 1969: Canada. Islam et al., 1988: in vitro.; Wang et al., 2008	*Vibrio* spp., incl. *V. tasmaniensis, splendidus*; in vitro experiments with *V. cholerae* O1	Culture	In situ enrichment: Dominant culturable bacteria; enriched by 10^2-10^4 compared to water column. In vitro survival advantage demonstrated.	Algal polysaccharide metabolism
Porphyra yezoensis	Duan et al., 1995: China	*Vibrio* spp.	Culture, scanning electron microscopy	In situ enrichment: Dominant microscopically identifiable and culturable bacteria	Algal polysaccharide metabolism
Green algae					
***Chaetomorpha* spp.**	Laksmanaperumalsamy and Purushothaman, 1982: tropical estuary, Africa	*Vibrio* spp.	Culture	In situ enrichment: Dominant culturable bacteria	Algal polysaccharide metabolism
Enteromorpha intestinalis, linza	Laksmanaperumalsamy and Purushothaman, 1982: tropical estuary, Africa Islam et al., 1988: in vitro	*Vibrio* spp.; in vitro experiments with *V. cholerae* O1	Culture	In situ enrichment: Dominant culturable bacteria. In vitro survival advantage demonstrated.	Algal polysaccharide metabolism
Ulva lactuca, pertusa	Islam et al., 1988: in vitro; Duan et al., 1995: China; Nakanishi et al., 1996; Patel et al., 2003; Tait et al., 2005	*Vibrio* spp.; in vitro experiments with *V. cholerae* O1	Culture, scanning electron microscopy	In situ enrichment: Dominant microscopically identifiable and culturable bacteria. In vitro survival advantage demonstrated.	Algal polysaccharide metabolism; modulation of host processes: developmental morphogenic effects, spore germination stimulation

2.5.4. Associations with animals

Vibrio interactions with animals include both specific, stable symbioses, and less well-defined associations (Table 2). Stable symbioses have been described for luminescent *V. fischeri* (*Aliivibrio*) with sepiolid squids (*Euprymna scolopes*) and loligonoid squids (Ruby and Lee, 1998), and for various luminescent *Vibrio* with flashlight fishes (Anamalopidae) and anglerfishes (Ceratioidei) (Haygood and Distel, 1993). The dynamics of the *V. fischeri-Euprymna symbiosis* have been particularly well

explicated: *V. fischeri* from surrounding waters colonize the developing squid light organ, successfully outcompeting nonsymbionts in this process, which triggers a developmental program in the host. Once established, the symbionts undergo daily cycles of expulsion and regrowth (Ruby and Lee, 1998; Stabb, 2006). Thus the symbiosis regularly seeds the water column, such that luminous *V. fischeri* are enriched in the water surrounding *E. scolopes* (Ruby and Lee, 1998). This expedites continual recolonization of immature squid, which is likely further facilitated by *V. fischeri* chemotaxis toward squid mucus (DeLoney-Marino et al., 2003).

Some *Vibrio* have also been deemed facultative intracellular symbionts of *Acanthamoeba* protozoa: *Vibrio cholerae* O1 and O139, and *Vibrio mimicus* (Abd et al., 2005; Abd et al., 2007; Sandström et al., 2010; Abd et al., 2010). These vibrios can replicate intracellularly for at least 14 days without affecting host health, at least in nutrient-replete artificial medium, and have been observed in both cytoplasm and cysts of the protozoa. Like several other microbial taxa, then, most famously the pathogen *Legionella* (Rowbotham, 1980), vibrios appear capable of evading *Acanthamoeba* endocytosis to shelter intracellularly. Thus they gain protection from antibiotics (Abd et al., 2005; Abd et al., 2007, 2010), predation, and perhaps other adverse conditions, e.g. cold temperatures. Still to be investigated are the questions of why some *Acanthamoeba* cells encyst their *Vibrio* inhabitants while others do not; why the *Vibrio* do not appear to be detrimental to host survival; and how often *Vibrio* might be released following host lysis, or even actively ejected, thus returning to the water column. Moreover, all studies of the *Vibrio-Acanthamoeba* relationship have been experimental: in situ surveys are necessary to establish the environmental relevance of this potential symbiosis, and assess any effects on *Vibrio* population dynamics.

Vibrios may be neutral or benign inhabitants of coral hosts: they have been shown to comprise a significant portion of the mucus-dwelling bacterial community of healthy corals [e.g., (Koren and Rosenberg, 2006; Kvennefors et al., 2010)], being able to subsist on coral mucus as their sole carbon and nitrogen source (Sharon and Rosenberg, 2008). *V. splendidus*, for example, constituted 50-68% of clone libraries derived from *Oculina patagonica* coral mucus, but was scarce in the coral tissue itself (Koren and Rosenberg, 2006). Moreover, nitrogen-fixing *Vibrio* representatives, primarily *V. harveyi* and *V. alginolyticus*, have been found to dominate the culturable diazotrophs of the coral *Mussimilia hispida* (Chimetto et al., 2008), and likely share fixed nitrogen with either or both coral and zooxanthellae. Evidence also suggests immune interaction between *Vibrio*

and coral hosts: adaptation of *Vibrio* commensals to coral antimicrobials has been suggested by significant antibiotic-resistance gene cassette content of their integrons (Koenig et al., 2011), while one *V. harveyi* coral isolate has been found to help defend its host by inhibiting colonization by a pathogen (Krediet et al., 2013).

In freshwater habitats, *V. cholerae* have been found to proliferate on egg masses of the abundant, widely distributed chironomid midges (Broza and Halpern, 2001; Halpern et al., 2008). These egg masses are embedded in thick, gelatinous material, which *V. cholerae* can use as a sole carbon source (Broza and Halpern, 2001); their degradation of the gelatinous matrix via secreted hemagglutinin/protease appears to be the primary cause of egg mass disintegration (Halpern et al., 2003). Accordingly, Halpern et al., (2006) were able to show correlations of chironomid egg mass with the abundance of attached *V. cholerae*, although they have not yet investigated any correlation of *V. cholerae* dynamics in the surrounding aquatic environment.

Zooplankton, primarily estuarine copepods such as *Acartia* and *Eurytemora*, have been investigated as a major reservoir of *V. cholerae* in particular, but while attachment has been demonstrated, it remains unclear whether the association is specific, and whether attached vibrios are consistently enriched compared to surrounding waters. Individual copepods have been shown to be able to host up to 10^5 *V. cholerae* cells (Colwell, 1996; Mueller et al., 2007), with preference often shown for attachment to the oral region and egg sac (next to the anal pore) – that is, regions offering close access to host exudates (Huq et al., 1983, 1984). Culture-based studies have detected enriched *Vibrio* occurrence on copepods compared to the surrounding water column (e.g., Simidu et al., 1971; Sochard et al., 1979), and one culture-based study showed *Vibrio* dominance of wild copepods' surface- and gut-attached bacterial communities (Sochard et al., 1979). However, other studies, both in vitro and in situ, have observed *V. cholerae* remaining predominantly free-living in the presence of copepods (Worden et al., 2006; Neogi et al., 2012) or attaching with greater preference to phytoplankton (Tamplin et al., 1990). Additionally, one culture-independent environmental study detected greater concentrations of Vibrio, including *V. cholerae*, in water compared to zooplankton (Heidelberg et al., 2002ab). Perhaps such variability of association with copepods helps explain the difficulty in detecting correlated *Vibrio*-copepod dynamics, as mentioned above in the section "Environmental correlates of *Vibrio* presence and abundance."

Other uncertainties regarding *Vibrio* association with copepods exist. There is a lack of quantitative evidence demonstrating long-term proliferation of copepod-

attached *Vibrio*: existing studies assessing survival advantage of *Vibrio* cultured with copepods have only demonstrated increased abundance of *Vibrio* in surrounding water, without monitoring attached abundance (Huq et al., 1983, 1984). Finally, it is not clear whether vibrios prefer colonizing live or dead copepods. While several in vitro studies have noted *V. cholerae* attachment preference for dead or detrital copepods (Huq et al., 1990; Tamplin et al., 1990; Mueller et al., 2007), one study instead observed survival advantage only upon association with live copepods, and found little attachment to dead copepods (Huq et al., 1983). Perhaps this question could be resolved by investigating from which part(s) exactly of the copepod vibrios derive nutrition: from oral/anal exudates or gut contents of actively feeding copepods, from degradation of the chitinaceous exoskeleton (which for live copepods is protected by a waxy epicuticle that resists attachment (Tarsi and Pruzzo, 1999), or from degradation of other copepod detritus. In addition, variable host traits such as immune defenses, age, and time since molting or death (which likely affect epicuticle condition) should be taken into account. As of yet, evidence of association with live copepods as an ecological specialization has been demonstrated for only one *Vibrio* sp. nov. (F10) (Preheim et al., 2011a).

In addition, zooplankton other than copepods may represent potential *Vibrio* hosts as well. Kirschner and colleagues (2011) found cladoceran *Diaphanosoma mongolianum* to enhance growth more than the copepod *Arctodiaptomus spinosus* in microcosm experiments; when cladocerans were added, they enhanced the growth of *V. cholerae* strains in the surrounding medium relative to controls where cladocerans were excluded, while copepods did not. In addition, the number of cells attached to cladocerans per individual was on average 100 times higher than on copepods. When a back-of-the-envelope calculation is done to consider whether *V. cholerae* is enriched on zooplankton, however, we find that they are not, even on cladocerans; from six microcosms, 10^5 - 10^7 cells were estimated attached and 10^6 - 10^7 cells not attached, a result suggesting that cladocerans might enhance overall growth with frequent dispersal, rather than supporting exclusively attached growth.

For other animals in which *Vibrio* have been found to be abundant –fish, and shellfish – it has not yet been determined whether vibrios form specific, lasting associations as gut microbiota, or are merely transient occupants, temporarily proliferating on favorable nutrients until excreted or otherwise detached. In marine fish, numerous studies, both culture-dependent and -independent, have demonstrated that *Vibrio* are major gut inhabitants, often dominating the community, and hence are

substantially enriched compared to surrounding seawater. Surveyed fish include flatfish (Liston, 1957; Xing et al., 2013), jackmackerel (Aiso et al., 1968), bluefish (Newman et al., 1972), salmonids (Yoshimizu and Kimura, 1976), sea bream (Muroga et al., 1987), and various coral reef fishes (Sutton and Clements, 1988; Smriga et al., 2010). Notably, *Vibrio* abundances often appear comparable between culture-based and -independent studies: e.g., 35-74% and 83.4%, respectively, of flatfish inhabitants (Liston, 1957; Xing et al., 2013). The ability of *Vibrio* representatives to resist low pH and bile supports their survival within the fish gut (Yoshimizu and Kimura, 1976). Whether food or water intake is the greater source of inoculation is an open question: some studies have found a strong effect of food source on gut *Vibrio* composition (e.g., Grisez et al., 1997), whereas others found a stronger influence of *Vibrio* representation in the water column (e.g., Blanch et al., 2009). Conversely, *Vibrio* content of the fish gut has also been shown to be responsible for increasing *Vibrio* abundance in surrounding water when fish were introduced into a tank that did not otherwise support *Vibrio* growth, demonstrating significant excretion of viable cells from the fish gut (Sugita et al., 1985). Hence, regardless of length of association, the fish gut appears to represent a favorable refuge where *Vibrio* can rapidly proliferate, prior to being released again to the water column. Indeed, the bioluminescence of marine microbes, including many vibrios, has been suggested to be an adaptation encouraging fish ingestion: fish preferentially predate zooplankton that are glowing after having grazed bioluminescent *Photobacterium* (Zarubin et al., 2012).

Among shellfish, high *Vibrio* abundance has been reported on surfaces and in tissues of hosts including oysters (e.g., Murphree and Tamplin, 1995; Froelich and Oliver, 2013), abalone (Sawabe, 2006), and blue crabs (Davis and Sizemore, 1982), with uptake and population dynamics particularly well documented for *V. vulnificus* in association with oysters (Froelich and Oliver, 2013). *V. haliotis* has been suggested to stably associate with gut of the herbivorous *Haliotis* abalone on the basis of reproducibly specific occurrence: it has never been isolated from other seaweed-consuming invertebrates (reviewed in Sawabe, 2006). Being alginolytic, *V. haliotis* has also been suggested to aid its host's digestion of algal polysaccharides (Sawabe, 2006). Otherwise, it is not clear whether copious *Vibrio* representation might solely be the result of nonspecific uptake from food or water, particularly for filter-feeding shellfish, whose highly efficient filtration has been reported to increase *Vibrio* concentrations by up to 4 orders of magnitude in oysters compared to surrounding waters (Froelich and Oliver,

2013). Furthermore, filter feeders can produce copious amounts of mucus, which rapidly and efficiently removes associated microbes, so that their turnover may be high. Consequently, it is challenging to prove specific association on the basis of abundance. In the next section, we will review a metapopulation study that more explicitly addresses the problem of assessing *Vibrio* host specificity by analyzing population structure across and within macroinvertebrate hosts. Future application of the approach described could help to resolve the question of whether *Vibrio* colonization of animal hosts like fish and crabs is specific, or driven more by indiscriminate uptake from the water column.

Table 2. Animal hosts for vibrio, as demonstrated by numerical enrichment and biological mechanisms supporting association.

Host	Reference, study site	Associated vibrios	Enumeration method	Enrichment, survival advantage	Host site, mechanism of association
Invertebrates					
Freshwater					
***Acanthamoeba* protozoa**	Abd et al., 2005; Abd et al., 2007; Sandström et al., 2010; Abd et al., 2010: in vitro	*V. cholerae* O1, O139; *V. mimicus*	Culture, microscopy	In vitro survival advantage: replicate intracellularly >14 days	Cytoplasm, cysts; protected from antibiotics and predation.
Chironomid midge egg masses	Broza and Halpern, 2001; Halpern et al., 2003; Halpern et al., 2008: in vitro	*V. cholerae* isolates from Israeli rivers and waste-stabilization ponds	Culture	In vitro survival advantage: 10^3 greater cell counts compared to growth in medium alone.	Gelatinous egg matrix; can use gelatinous material as sole carbon source, degrading via secreted hemagglutinin/protease
Zooplankton: cladoceran *Diaphanosoma mongolianum*, from alkaline lake, Germany	Kirschner et al, 2011: in vitro	*V. cholerae* non-O1/non-O139 isolate from alkaline lake, Germany	Fluorescence in situ hybridization	In vitro survival advantage, but not enrichment: up to 6-fold increase in growth rate of cells in surrounding medium; 10^5-10^7 cells attached compared to 10^6-10^7 cells in surrounding medium.	Probable use of host exudates
Estuarine and marine					
Zooplankton: Estuarine copepods, espp. *Acartia* and *Eurytemora*	Simidu et al., 1971: Japan; Sochard et al., 1979: Gulf of Mexico; Huq et al., 1983, 1984: in vitro; Colwell, 1996: in vitro; Mueller et al., 2007: in vitro; Preheim et al., 2011a: Massachusetts estuary, USA	*Vibrio* spp., espp. *V. cholerae*	Culture	In situ and in vitro enrichment shown in some cases, with up to 10^5 cells per host. Can dominate culturable surface- and gut-attached communities.	Possible preference for oral region and egg sac, due to proximity to host exudates; preference for live versus dead hosts unclear.
Corals, incl. *Acropora hyacinthus*, *Oculina patagonica*, *Mussimilia hispida*,	Koren and Rosenberg, 2006: Israel; Kvennefors et al., 2010: Great Barrier Reef; Sharon and Rosenberg, 2008; Chimetto et al., 2008;	Spp. incl. *V. alginolyticus, harveyi, splendidus,*	Culture, molecular	In situ enrichment: Can dominate mucus community, according to both culturing and molecular methods; can dominate culturable diazotrophs (found for *Mussimilia hispida*).	Mucus. Metabolize mucus; diazotrophs likely contribute nitrogen to hosts; may adapt to host antimicrobials via antibiotic-resistance gene

Stylophora pistillata	Koenig et al., 2011; Krediet et al., 2013				acquisition; can inhibit pathogen colonization.
Shellfish: blue crabs, *Callinectes sapidus*	Davis and Sizemore, 1982: Texas, USA	Spp. incl. *V. cholerae, vulnificus, parahaemolyticus*	Culture	In situ enrichment: Dominant culturable bacteria in hemolymph	Hemolymph; mechanism untested
Shellfish: oysters	Murphree and Tamplin, 1995; Froelich and Oliver, 2013	Spp. incl. *V. cholerae, parahaemolyticus, vulnificus*	Culture	In situ enrichment, via host filtration: can be concentrated by up to 10^4 compared to surrounding water	Gut; unclear whether true gut microbionts, or transient occupants concentrated from food and water
Shellfish: abalone, *Haliotis*	Reviewed in Sawabe et al., 2006	*V. haliotis*	Culture	In situ enrichment: ~70% of culturable gut bacteria; reproducibly specific association	Gut; may contribute to host seaweed digestion via alginolytic activity
Squids: Sepiolid (*Euprymna scolopes*) and loligonoid	Reviewed in Ruby and Lee, 1998; Stabb, 2006	*V. fischeri*	Culture, molecular	Exclusive light organ symbiotes	Bioluminescent symbiotes of nutrient-rich light organ. Colonize immature squid; in mature fish, are expelled and recolonize daily, outcompeting nonsymbiotes.
Vertebrates					
Bluefish	Newman et al., 1972: New York, USA	*Vibrio* spp.	Culture	In situ enrichment: Can dominate gut bacteria	
Coral reef fishes, incl. surgeonfish *Acanthurus nigricans*, parrotfish *C. sordidus*, snapper *Lutjanus bohar*	Sutton and Clements, 1988; Smriga et al., 2010: Palmyra Atoll, northern Pacific	Spp. including *V. agarivorans, coralliilyticus, fortis, furnissii, ponticus, qinhuangdaoora, nigripulchritudo; Photobacterium* spp.	Culture, molecular	In situ enrichment: Can dominate gut bacteria, according to both culturing and molecular methods. Molecular quantification: 10% of *A. nigricans* gut community, 71% of *C. sordidus*, 76% of *L. bohar*	Gut; unclear whether true gut microbionts, or transient occupants ingested from food (i.e. coral, for parrotfish) and water
Flashlight fishes (Anamalopidae) and anglerfishes (Ceratioidei)	Haygood and Distel, 1993	Novel *Vibrio* spp.	Molecular	Exclusive light organ symbiotes	Bioluminescent symbiotes of nutrient-rich light organ
Flatfishes incl. Rajidae skate, lemon sole *Microstomus kitt*, turbot *Scopthalmus maximus*	Liston, 1957: Scotland, UK; Xing et al., 2013: fish farm, China	Spp. incl. *V. cholerae, parahaemolyticus, cholerae; Photobacterium* spp.	Culture, molecular	In situ enrichment: Can dominate gut bacteria, according to both culturing (35-74%, *M. kitt*) and molecular (~80%, *S. maximus*) methods.	Gut; unclear whether true gut microbionts, or transient occupants ingested from food and water
Jackmackerel *Trachurus japonicus*	Aiso et al., 1968: Japan	*Vibrio* spp.	Culture	In situ enrichment: 27% of stomach culturable bacteria, 100% of intestine.	Gut; unclear whether true gut microbionts, or transient occupants ingested from food and water
Salmonidae, incl. pink salmon *Onchorhynchus gorbuscha*, chum salmon	Yoshimizu and Kimura, 1976: Japanese coast, East Bering Sea	*Vibrio* spp.	Culture	In situ enrichment: dominate gut bacteria of saltwater-dwelling (but not freshwater) salmonids; on average represent 69% of saltwater gut community	Gut; unclear whether true gut microbionts, or transient occupants ingested from food and water.

O. keta*, sockeye salmon *O. nerka*, Chinook salmon *O. tshawytscha					
Sea bream *Pagrus major*, *Acanthopagrus schlegeli*	Muroga et al., 1987: Japan	*Vibrio* spp.	Culture	In situ enrichment: ~45% of culturable gut bacteria	Gut; unclear whether true gut microbionts, or transient occupants ingested from food and water

2.5.5. Population dynamics associated with macroinvertebrate hosts

In a metapopulation study by Preheim and colleagues (2011a), relative abundances of *Vibrio* groups were compared across different shellfish and parts of shellfish. The study found that macroinvertebrates do not appear to be a strongly selective habitat for vibrios, when contrasted to preceding metapopulation studies of the water column, where differential associations of genotype clusters revealed ecologically distinct populations (described in detail in the section "Using ecology to define cohesive populations"). When different body parts of mussels and crabs were sampled by Preheim and colleagues (2011a), little host preference was evident, and the diversity and frequency of populations (identified by multi-locus sequence analysis) resembled that in water samples. For example, *V. splendidus* represented the dominant population in the water and on both animals. For mussels, which can retain particles when filter feeding (Vahl 1972), the similarity between water column and animal-associated populations was particularly high, and there appeared to be relatively little difference when gills, stomach and gut walls and contents were compared. This was interpreted as population assembly being largely driven by filter-feeding activity, as was posited in the section above. In contrast to mussels' highly uniform population structure across individual hosts, crabs showed high variance in associated *Vibrio* populations, although composition across individuals' body parts was still similar to that in the water column. What causes the high variance among individual crabs is not known, although there was some evidence suggesting that they may be inoculated by food items, which could be of variable composition given their scavenging lifestyle.

The apparent lack of specificity for the animals was surprising considering that ecological theory predicts that habitats that are long-lived and stable compared to the colonizing species should be dominated by specialists (Kassen, 2002). Yet with regard to mussels and crabs as habitats, vibrios appear to be generalists whose population

dynamics may be determined by direct inoculation from the water or via food items (Preheim et al., 2011a). A similar dynamic has recently been suggested to drive *V. vulnificus* accumulation in oysters (Froelich et al., 2010). These can only retain larger particles when filter feeding, and hence enrich pathogenic ecotypes of *V. vulnificus* that are particle-associated as compared to ecotypes that are predominantly free-living.

Overall, these studies demonstrate that colonization may be a complex process strongly influenced by dispersal. In contrast to water column populations, which showed varying degrees of specificity towards microhabitats (e.g., organic particles, zooplankton), *Vibrio* populations on larger invertebrates (mussels and crabs) showed little specificity either for host or host body parts. Whether similar patterns exist for other animals remains unknown; it will be valuable to test fish to determine whether their *Vibrio* inhabitants are true gut microflora. The above studies stress the importance of taking into account potential *Vibrio* sources, i.e. water and food, when assessing host association. For example, *V. splendidus* was the dominant population on both crabs and mussels, and on particles in the water column; had only mussels been sampled, *V. splendidus* may have appeared to have been a mussel specialist. Such erroneous conclusions can be avoided by "mass balancing" populations in a particular location by determining their frequency across different microhabitats or patches that are potentially connected by migration.

2.6. Vibrio proliferation in the water column

Ocean water is a heterogeneous landscape of varying ecological opportunities on small scales, with a highly patchy distribution of resources that may represent microhabitats for vibrios. Some of these are hotspots of soluble organic material, which originates from exudates or excretions of larger organisms, while others are particulates of various origins. For example, as mentioned above, algal cells exude a zone of enriched organic material (Bell and Mitchell, 1972; Paerl and Pinckney, 1996). Several other processes can also generate ephemeral patches of dissolved nutrients, and it is likely that many bacteria, including vibrios, can chemotax towards these and take advantage of the elevated nutrient concentrations (e.g., for vibrios, Sjoblad and Mitchell, 1979; Mizanur et al., 2001, 2002). In addition, diverse processes are responsible for the formation of suspended particulate organic matter that can be colonized and degraded by bacteria. This includes dead biomass of small planktonic organisms, fecal pellets, and aggregates (marine snow) formed from polymers and other, smaller particles.

This section will address two main subjects, both seeking to situate *Vibrio* within the marine water column. Here, we will first review both experimental and environmental evidence that blooms of *Vibrio* can and do occur, despite their typically low representation in marine assemblages. Second, we will review the evidence for proliferation of *Vibrio* in the planktonic, free-living phase, expanding the view of their niche range beyond the longstanding proposition that their lifestyle is predominantly attached.

2.6.1. Vibrio blooms

Thompson and Polz (2006) summed up three key *Vibrio* traits supporting the ability to bloom on sporadic nutrient pulses: *Vibrio* can (i) survive long-term under resource-limited conditions, as indicated by continued respiratory activity in mesocosms (Ramaiah et al., 2002; Armada et al., 2003); (ii) recover from starvation and grow rapidly in response to substrate pulses, enabled by maintenance of high ribosome content (Hood et al., 1986; Flärdh et al., 1992; Kramer and Singleton, 1992; Eilers et al., 2000; Pernthaler et al., 2001); and (iii) actively seek out nutrient patches via chemotaxis (Bassler et al., 1991; Yu et al., 1993), including under starvation conditions (Gosink et al., 2002; Larsen et al., 2004).

Vibrio proliferation on natural dissolved resources alone has been experimentally demonstrated by rapid growth of inocula in mesocosms or microcosms of filtered water from algal blooms. *V. cholerae* strain N19691 grew at a rate of up to 2.6 d-1 in dinoflagellate (*Lingulodinium polyedrum*) bloom water (Mouriño-Pérez et al., 2003), and up to 1.73 d-1 in water from a dense picophytoeukaryote and dinoflagellate bloom, surpassing the 0.76 d-1 average growth rate of the separately incubated native bacterial assemblage (Worden et al., 2006).

Experiments have furthermore demonstrated conditions where algal resources were sufficient for *Vibrio* to overcome competition and/or grazing pressure. Taking competition into account, but in the absence of predation, strains of both *V. cholerae* and *V. vulnificus* have been shown capable of increasing in relative abundance when in direct competition with the total bacterial community for filtered homogenate of a cyanobacteria bloom (dominated by *Nodularia spumigena*) (Eiler et al., 2007). Meanwhile, *V. cholerae* N19691 has been shown to overcome substantial protozoan grazing when proliferating on filtrate of a particularly dense algal bloom (Worden et al., 2006). Ample algal dissolved organic material may have permitted this *V. cholerae* growth by relieving

resource competition, as the *V. cholerae* inocula grew at the same rate with or without the whole bacterial community filtered out from their bloom-water amendments. Similarly, an analysis of *Vibrio* dynamics sampled from the Arabian Sea suggested that algal resource supply can be a more significant control on *Vibrio* abundance than predation, enabling rapid turnover (Asplund et al., 2011).

Reinforcing these experimental findings, (Gilbert et al., 2011) observed an explosive *Vibrio* bloom in the environment, demonstrating that their potential for rapid growth is indeed relevant in the context of a full marine community. In one month, a single *Vibrio* sp., otherwise comprising only 0-2% of total rRNA genes, grew to constitute 54% of the community – the largest bloom of any bacterial group observed over the course of a six-year time series. Furthermore, there was a correlated bloom of the diatom *Chaetoceros compressus*, itself typically rare within the phytoplankton community. Hence, nutrients exuded by the unusually proliferating diatom taxon may have sparked the *Vibrio* bloom, whether by specifically appealing to the species' metabolic palate, relieving resource competition, diluting protozoan grazing pressure by stimulating rapid growth of the surrounding bacterial community, or some combination of the three. Luminescent *Vibrio* blooming in association with algae have even been suggested to be responsible for the phenomenon dubbed "milky seas," where significant stretches of surface water are rendered white with bioluminescence (Lapota et al., 1988; Nealson and Hastings, 2006); one recent case was expansive enough (>17,700 km^2) to be detectable by satellite. Whether such bloom events are rare remains unknown due to currently infrequent sampling and lack of time series; however, the observations cited above provide evidence that *Vibrio* are capable of rapid growth in the environment.

2.6.2. The evidence for a planktonic, free-living lifestyle

The two mesocosm/microcosm studies discussed above (Mouriño-Pérez et al., 2003; Worden et al., 2006), both furnish evidence that vibrios can thrive while free-living. Mourino-Perez et al., (2003) demonstrate the ability of a *V. cholerae* strain to flourish purely on dissolved compounds derived from an algal bloom. Even more strikingly, Worden et al., (2006) observed *V. cholerae* N19691 remaining free-living in four out of their five seawater mesocosm experiments: one initiated from non-bloom seawater, and the other three initiated from seawater collected during distinctly different phytoplankton blooms. Notably, in two of these four experiments, *V. cholerae* attachment to cohabiting copepods was assessed and found to be insignificant (e.g., <1 *V. cholerae*

cell found per copepod, averaged over a sampling of ten copepods, in one of the experiments). This stands in contrast to the theory that *V. cholerae* preferentially attach to copepods, as discussed above in the section on animal associations. In the remaining experiment of (Worden et al., 2006), by contrast to the mesocosms in which *Vibrio* remained free-living, the *V. cholerae* inoculum was initially almost entirely free-living, but, as bloom decay progressed and algal detrital particles increased in size, the population became almost entirely particle-attached, presumably in response to nutrient limitation.

The factors determining whether *Vibrio* remain free-living versus particle-attached are still unknown, but both environmental and genetic determinants could come into play. Past studies have demonstrated effects of temperature, pH, ion concentration, and starvation state (Hood and Winter, 1997); salinity (Kumazawa et al., 1991; Hsieh et al., 2007); and growth-stage-dependent chitin content of diatom cell walls (Frischkorn et al., 2013) on *Vibrio* attachment. Perhaps encounters with relevant biological compounds, e.g., a specific algal cell wall component or polysaccharide, might also trigger lifestyle changes. Even less is known about the genetic mechanisms, diversity, and dynamics underlying vibrio lifestyle association; this remains a rich field of inquiry. For example, Shapiro et al., (2012) recently discovered genomic patterns underlying the ongoing ecological differentiation of two *V. cyclitrophicus* populations: the population with preference for association with larger particles possessed genes for attachment and biofilm formation that were absent from the preferentially free-living population. Such evidence of genetic bases for habitat specificity will provide invaluable insights into selective pressures exerted by different marine microhabitats.

The findings described above suggest great flexibility in *Vibrio* lifestyle, permitting many lines of attack on marine substrates, with different ecological implications for vibrios' dynamics in the water column. For example, biofilm attachment on particulate resources can decrease susceptibility to protozoan predation (Matz et al., 2005), while association with larger particles might increase probability of ingestion by macrofaunal predators, which could in turn facilitate rapid proliferation and dispersal, as discussed above in the section on fish associations. Given vibrios' possibilities for rapid growth and association with diverse marine niches and resources, their impacts on marine nutrient cycling and trophic structure might be much greater than previously believed. Understanding their dynamics will help to elucidate these fundamental

marine processes, as well as *Vibrio*-specific models of pathogen persistence and transmission.

2.7. Using ecology to define cohesive populations

The studies summarized above suggest potential for association of vibrios with plants, algae and animals as well as growth response to specific classes of particulate and dissolved organic matter; however, they have targeted primarily a single, taxonomically defined species, leaving several important questions unanswered. First, do such taxonomic species correspond to ecologically cohesive units, i.e., do they comprise several ecologically distinct populations or should they be merged with others to form one ecologically cohesive population? Second, if we can define such populations, do these partition resources or compete with each other? Finally, are vibrios primarily ecological generalists or specialists?

A series of studies explored to what extent ecologically coherent groups of vibrios could be distinguished by determining the distribution patterns of genotypes among different potential microhabitats in the coastal ocean (Hunt et al., 2008a; Preheim et al., 2011a, 2011b; Szabo et al., 2012). Initially, this was done by isolation of vibrios from four consecutive size fractionations of ocean water, collected in the spring and fall, to distinguish free-living and attached genotypes (Hunt et al., 2008a). The rationale of this sampling scheme was that different types of microhabitats (e.g., organic particles of various origin, zoo- and phytoplankton) have characteristic size spectra and hence will be enriched in a specific size fraction. Consequently, bacteria specifically associated with a microhabitat should be enriched in the same specific size fraction. Further, because ecological associations may evolve on faster time scales than rRNA genes, isolates were also characterized at higher genotypic resolution using several protein coding genes in a multilocus sequence analysis (MLSA) scheme, to better capture the eco-evolutionary dynamics of environmental populations. Because of the complexity of the data, a statistical clustering algorithm (AdaptML) was developed that allows identification of groups of related genotypes with distinct and characteristic distributions among the sampled parameters (size fractions and seasons) (Hunt et al., 2008a).

The analysis of >1,000 isolates identified a large number of genotypic clusters with clear microenvironmental preferences, consistent with the notion of an ecological population (Hunt et al., 2008a). Seasonal differentiation was particularly strong, with

little overlap between spring and fall samples, supporting the observed significant correlation of some species to temperature discussed in above sections. The study also revealed that several populations appear free-living or predominantly free-living, again supporting the notion that vibrios can pursue, at least temporarily (e.g., during a bloom), unattached lifestyles. Most populations, however, displayed various preferences for size fractions enriched in different types of organic particles or zoo- and phytoplankton. For example, *V. calviensis* appeared almost entirely free-living, while *V. alginolyticus* had significant representation in both the free-living and large-particle fractions, and *V. fischeri* occurred on small and large particle size fractions. Most strikingly, *V. splendidus* was broken up into several, very closely related populations with distinct distributions. Overall, 25 distinct populations could be identified in the two seasonal samplings. (Hunt et al., 2008a), demonstrating the fine-scale resource partitioning co-existing vibrios are engaged in.

To what extent does the commonly used rRNA marker gene resolve these populations? The *V. splendidus* example and several others demonstrate that at least some ecologically distinct genotypic clusters may not be resolved by rRNA analysis and do require high resolution protein-coding genes to identify genotypic clusters whose environmental distributions can be assessed (Preheim et al., 2011b; Shapiro et al., 2012). Most populations, however, were manifest as deeply divergent protein-coding gene clusters (Hunt et al., 2008a) that correspond to microdiverse rRNA gene clusters previously postulated to represent ecological populations (Acinas et al., 2004). Although overall reassuring for rRNA gene-based environmental surveys, variable performance of marker genes is expected since they are slowly evolving and may not capture populations at early stages of divergence (Shapiro et al., 2012).

Additional studies carried out at the same coastal site refined the habitat resolution for several populations, allowed identification of ecological generalists and specialists, and also demonstrated reproducible associations (Preheim et al., 2011a, 2011b; Szabo et al., 2012). The actual microhabitat of several attached populations was identified by hand-picking under the microscope visually identifiable types of particles and zooplankton (Preheim et al., 2011a, 2011b). This revealed high habitat specificity for several populations while others occurred more broadly, indicating different levels of ecological specialization. For example, *V. breoganii* occurred on algal derived detritus while a not yet formally described species (*Vibrio* F10) was highly specific for living zooplankton. On the other hand, *V. crassostreae* was associated with both zooplankton

and algal detritus. Metabolic potential in these species, measured by growth assays and comparative genomics, reflects these associations. Both *V. breoganii* and *V. crassostreae* are able to exploit alginate, a brown algal cell wall component, as the sole carbon source, yet the algae-associated *V. breoganii* has acquired the ability to grow on the algal storage polysaccharide laminarin but has lost the ability to grow on chitin, a trait ancestral to vibrios (Hunt et al., 2008b). Moreover, such high specificity for algal derived material was unexpected for vibrios, which are reputed to be animal associated, and supports the evidence provided above that vibrios encompass algal specialists.

A recent study that attempted to reproduce the original size fractionation of ocean water collected at a similar time point, but three years after the initial sampling, showed that population structure was preserved for many of the originally detected populations, but also revealed populations as dynamic and environmentally responsive entities (Szabo et al., 2013). For example, *V. breoganii, V. crassostreae* and *V. splendidus,* which range in ecological specialization from specialist to generalist, had highly reproducible distributions indicative of similar habitat associations. The study, however, also showed that several populations were nearly absent in either of the samplings, possibly due to the lower frequency of their habitat in the water samples. Moreover, some populations had shifted distributions among the size fractions. This was the case for a recently diverged population of *V. cyclitrophicus* that was associated with larger particles or organisms in the first study, but was highly represented in the free-living fraction in the second sampling. It was hypothesized that this shift represented a population expansion following a diatom bloom because the relative frequency of *V. cyclitrophicus* increased coincident with a shift from a copepod- to a diatom-dominated eukaryotic plankton community. Similarly, bloom dynamics, as have previously been observed for total vibrios in the water column, may cause the variable representation of several additional populations. Overall, the comparison of the two studies supports highly predictable population-habitat linkage but also provides additional support for the notion that vibrios may be subject to rapid population expansions or blooms in response to often overlooked or unknown environmental factors.

2.7.1. Populations as ecological, genetic and social units

Populations as defined here are genotypic clusters (evident by MLSA) that act as ecologically cohesive units, i.e., their ecology is more similar within the cluster than

between. Defining populations in this way has afforded the opportunity to test the hypothesis that, akin to sexual eukaryotes, gene flow boundaries across such clusters are strong enough for adaptive genes or alleles to spread in a population-specific manner. A population genomic analysis of two very recently diverged populations of *V. cyclitrophicus*, which are ecologically distinct but remain >99% similar in average nucleotide composition across their genomes, showed that specific genome regions have swept each of the two populations (recently reviewed by Polz et al., 2013). Moreover, annotation of these genome regions as well as behavioral and growth analysis suggest that these genome regions are adaptive for differential lifestyles (Shapiro et al., 2012, Yawata et al., submitted).

A second study showed that ecologically defined populations may also act as social units. This was evident in a test for potential of antagonistic interactions mediated by antibiotics between individuals from different ecological populations of vibrios (Cordero et al., 2012). Because of higher niche overlap among close relatives, it was expected that antagonism be more advantageous if directed against members of the same population. In stark contrast, however, antagonism was primarily directed against members of other populations while members of the same population were resistant to antibiotics produced within their own populations. This suggests synergism on the population level, especially since multiple antibiotics were produced within each population but each only by relatively few members.

Overall, this research shows agreement between ecological, genetic and social population structure and suggests that, in many ways, populations can be regarded as species-like units in the wild. Importantly, these units are non-clonal, and their genetic exchange and social structure suggest that populations frequently coexist and re-assemble on small-scale habitats.

2.8. Conclusion

In this review, we examine what is known about *Vibrio* ecology at increasingly fine environmental and taxonomic scales, to reveal factors with potential for greater predictive and explanatory power for *Vibrio* dynamics.

We find that while bulk environmental variables are often inconsistent in their ability to explain variance in *Vibrio* abundances, at both the genus and species levels, temperature and salinity are usually the strongest abiotic correlates. Yet total *Vibrio* trends do not necessarily capture species-level trends, and thus it is necessary to monitor

populations of interest directly to capture their dynamics. Correlations of species to specific plankton – like those of *V. cholerae* to dinoflagellate (Eiler et al., 2006), cladoceran (Kirschner et al., 2011) and rotifer (Magny et al., 2011) taxa – can provide the bases for hypotheses of biological associations - as was demonstrated by Kirschner and colleagues (2011) for cladoceran the *D. mongolianum*. Further investigation is necessary to confirm reproducibility and biological significance of such correlations.

Indeed, the breadth of vibrios' metabolic and attachment abilities mean that they can appear quite generalist in their ecological associations, making it difficult to discern which relationships with other organisms are specific and stable, rather than simply the product of promiscuous attachment followed by proliferation. Among the diverse biological associations that we review, some may be true mutualisms, on the basis of vibrios exchanging benefits with their hosts. The symbioses of luminescent vibrios with certain squid and fish are well attested, while possible symbioses with other organisms are suggested by potentially mutual metabolic exchange (salt marsh plants, cyanobacteria, corals), or *Vibrio* modulation of host processes like development and reproduction (macroalgae), and response to infection (corals). Notably, diazotrophy may facilitate relationships with both marsh plants and corals. In numerous other cases, vibrios may simply be taking advantage of hosts as nutrient sources, and perhaps only temporarily and opportunistically be associated with microalgae, zooplankton, fish, shellfish, and chironomid egg masses, or as intracellular occupants of protozoa. Of these, we argue that evidence points towards a particularly significant ecological impact of *Vibrio* interactions with algae, given the abundant laboratory and environmental observations of vibrios' ability to live on algal exudates - including blooms as free-living cells, a historically underappreciated *Vibrio* lifestyle. Nonetheless, much work remains to be done in resolving more specific *Vibrio*-algae associations.

In light of these studies, we have several recommendations. Previous surveys of *Vibrio* abundance are predominantly culture-dependent; going forward, molecular methods, such as fluorescent in-situ hybridization or quantitative PCR, can be used to gain less biased quantitative data. Such techniques also enable targeting of specific genotypic groupings, allowing better discrimination of pathogenic variants or ecologically meaningful populations than traditional taxonomic assays of species identity. Furthermore, to distinguish specialized association from incidental attachment, a "mass-balanced" approach is necessary: are vibrio enriched on a given microhabitat (e.g., a specific organic particle type or zooplankton) compared to the surrounding

water? Or, is the habitat enriched in *Vibrio* compared to other habitats? This approach has provided support for many of the potential symbioses noted above, and enabled identification of specialist *Vibrio* populations, e.g., *V. breoganii* for macroalgae-derived material and *V.* F10 for zooplankton (Hunt et al., 2008a; Preheim et al., 2011a, 2011b; Szabo et al., 2012). It provides a strong basis from which to proceed to more detailed and, ideally, mechanistic elucidation of *Vibrio* associations: for example, identifying chemotactic preferences for or proliferation on host or host exudates, or taking advantage of vibrios' genetic tractability to demonstrate dependence of an association on particular metabolic pathways.

When considering the question of to what extent environmental affiliations may be shared among or within *Vibrio* taxa, we also explore the shifting perspective on the nature of microbial groupings: recent work has moved towards discerning ecologically cohesive *Vibrio* populations, rather than relying on named species as the unit of inquiry. Pursuing this approach, whereby habitat associations are mapped onto genotypic clusters, has been successful in identifying ecological, genetic and social units among vibrios in the wild. We stress, however, that the initial identification of environment-genotypic cluster associations by the "mass-balanced" approach outlined above must be treated as a hypothesis of population structure to be further explored by more mechanistic investigation of, for example, dynamic habitat associations, biological interactions and gene flow boundaries. As demonstrated above, this approach has already helped to resolve apparently generalist *Vibrio* taxa into specialized populations and to identify mechanisms of how adaptive genes spread amongst nascent, ecologically differentiated populations. By sampling the environment at fine scales and molecularly characterizing associated *Vibrio*, we will gain a deeper understanding of the ways in which vibrios live in the environment. Such a population-based framework serves as a means of understanding the ecology of microorganisms in general.

Chapter 2 References

Abd, H., Saeed, A., Weintraub, A., Nair, G. B., and Sandström, G. (2007). *Vibrio cholerae* O1 strains are facultative intracellular bacteria, able to survive and multiply symbiotically inside the aquatic free-living amoeba *Acanthamoeba castellanii*. *FEMS Microbiol. Ecol.*60, 33–39.

Abd, H., Valeru, S. P., Sami, S. M., Saeed, A., Raychaudhuri, S., and Sandström, G. (2010). Interaction between *Vibrio mimicus* and *Acanthamoeba castellanii*. *Environ. Microbiol. Reports* 2, 166–171.

Abd, H., Weintraub, A., and Sandström, G. (2005). Intracellular survival and replication of *Vibrio cholerae* O139 in aquatic free-living amoebae. *Environ. Microbiol.*7, 1003–1008.

Acinas, S. G., Klepac-Ceraj, V., Hunt, D. E., Pharino, C., Ceraj, I., Distel, D. L., and Polz, M. F. (2004). Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430, 551–554.

Aiso, K., Simidu, U., and Hasuo, K. (1968). Microflora in the digestive tract of inshore fish in Japan. *J. Gen. Microbiol.* 52, 361–364.

Andrews, J. H., and Harris, R. F. (2000). The Ecology and Biogeography of Microorganisms on Plant Surfaces. *Annual Review of Phytopathology* 38, 145–180.

Armada, S. P., Farto, R., Pérez, M. J., and Nieto, T. P. (2003). Effect of temperature, salinity and nutrient content on the survival responses of *Vibrio splendidus* biotype I. *Microbiol.* 149, 369–375.

Asplund, M. E., Rehnstam-Holm, A.-S., Atnur, V., Raghunath, P., Saravanan, V., Härnström, K., Collin, B., Karunasagar, I., and Godhe, A. (2011). Water column dynamics of *Vibrio* in relation to phytoplankton community composition and environmental conditions in a tropical coastal area. *Environ. Microbiol.*13, 2738–2751.

Barbieri, E., Falzano, L., Fiorentini, C., Pianetti, A., Baffone, W., Fabbri, A., Matarrese, P., Casiere, A., Katouli, M., Kuhn, I., et al., (1999). Occurrence, diversity, and pathogenicity of halophilic *Vibrio* spp. and non-O1 *Vibrio cholerae* from estuarine waters along the Italian Adriatic coast. *Appl. Environ. Microbiol.* 65, 2748–2753.

Bassler, B. L., Yu, C., Lee, Y. C., and Roseman, S. (1991). Chitin utilization by marine bacteria. Degradation and catabolism of chitin oligosaccharides by *Vibrio furnissii*. *J. Biol. Chem.* 266, 24276–24286.

Bell, W., and Mitchell, R. (1972). Chemotactic and growth responses of marine bacteria to algal extracellular products. *Biological Bulletin*, 265–277.

Beardsley, C., Pernthaler, J., Wosniok, W., and Amann, R. (2003). Are readily culturable bacteria in coastal North Sea waters suppressed by selective grazing mortality? *Appl. Environ. Microbiol.* 69, 2624–2630.

Bergholz, P. W., Bagwell, C. E., and Lovell, C. R. (2001). Physiological Diversity of Rhizoplane Diazotrophs of the Saltmeadow Cordgrass, *Spartina patens*: Implications for Host Specific Ecotypes. *Microb Ecol* 42, 466–473.

Bertuzzo, E., Mari, L., Righetto, L., Gatto, M., Casagrandi, R., Rodriguez-Iturbe, I., and Rinaldo, A. (2012). Hydroclimatology of dual-peak annual cholera incidence: Insights from a spatially explicit model. *Geophys. Res. Lett.* 39.

Blackwell, K. D., and Oliver, J. D. (2008). The ecology of *Vibrio vulnificus*, *Vibrio cholerae*, and *Vibrio parahaemolyticus* in North Carolina Estuaries. *J. Microbiol.* 46, 146–153.

Blanch, A. r., Hispano, C., Bultó, P., Ballesté, E., González-López, J. j., and Vilanova, X. (2009). Comparison of *Vibrio* spp. populations found in seawater, in exhibition aquaria, in fish intestine and in fish feed. *J. Appl. Microbiol.* 106, 57–65.

Böer, S. I., Heinemeyer, E.-A., Luden, K., Erler, R., Gerdts, G., Janssen, F., and Brennholt, N. (2013). Temporal and spatial distribution patterns of potentially pathogenic *Vibrio* spp. at recreational beaches of the German north sea. *Microb. Ecol.* 65, 1052–1067.

Broza, M., and Halpern, M. (2001). Pathogen reservoirs: Chironomid egg masses and *Vibrio cholerae*. *Nature* 412, 40–40.

Caburlotto, G., Haley, B. J., Lleò, M. M., Huq, A., and Colwell, R. R. (2010). Serodiversity and ecological distribution of *Vibrio parahaemolyticus* in the Venetian Lagoon, Northeast Italy. *Environ. Microbiol.Reports* 2, 151–157.

Chan, E. C. S., and McManus, E. A. (1969). Distribution, characterization, and nutrition of marine microorganisms from the algae *Polysiphonia lanosa* and *Ascophyllum nodosum*. *Can. J. Microbiol.* 15, 409–420.

Chimetto, L. A., Brocchi, M., Thompson, C. C., Martins, R. C. R., Ramos, H. R., and Thompson, F. L. (2008). Vibrios dominate as culturable nitrogen-fixing bacteria of the Brazilian coral *Mussismilia hispida*. *Sys. Appl. Microbiol.* 31, 312–319.

Colwell, R. R. (1996). Global Climate and Infectious Disease: The Cholera Paradigm*. *Science* 274, 2025–2031.

Cook D.W., Bowers J.C., DePaola A. (2002). Density of total and pathogenic (tdh+) *Vibrio parahaemolyticus* in Atlantic and Gulf coast molluscan shellfish at harvest. *J. Food. Prot.* 65:1873–80.

Cordero, O. X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., Le Roux, F., Mincer, T., and Polz, M. F. (2012). Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. *Science* 337, 1228–1231.

Criminger, J. D., Hazen, T. H., Sobecky, P. A., and Lovell, C. R. (2007). Nitrogen fixation by *Vibrio parahaemolyticus* and its implications for a new ecological niche. *Appl. Environ. Microbiol.*. 73, 5959–5961.

Davis, J. W., and Sizemore, R. K. (1982). Incidence of Vibrio species associated with blue crabs (*Callinectes sapidus*) collected from Galveston Bay, Texas. *Appl. Environ. Microbiol.*. 43, 1092–1097.

Deepanjali, A., Kumar, H. S., Karunasagar, I., and Karunasagar, I. (2005). Seasonal variation in abundance of total and pathogenic *Vibrio parahaemolyticus* bacteria in oysters along the southwest coast of India. *Appl. Environ. Microbiol.*. 71, 3575–3580.

Del Refugio Castañeda Chávez, M., Sedas, V. P., Orrantia Borunda, E., and Reynoso, F. L. (2005). Influence of water temperature and salinity on seasonal occurrences of *Vibrio cholerae* and enteric bacteria in oyster-producing areas of Veracruz, México. *Marine Pollution Bulletin* 50, 1641–1648.

DeLoney-Marino, C. R., Wolfe, A. J., and Visick, K. L. (2003). Chemoattraction of *Vibrio fischeri* to Serine, Nucleosides, and N-Acetylneuraminic Acid, a Component of Squid Light-Organ Mucus. *Appl. Environ. Microbiol.* 69, 7527–7530.

DePaola, A., Hopkins, L. H., Peeler, J. T., Wentz, B., and McPhearson, R. M. (1990). Incidence of *Vibrio parahaemolyticus* in US coastal waters and oysters. *Appl. Environ. Microbiol.*. 56, 2299–2302.

DePaola, A., Nordstrom, J. L., Bowers, J. C., Wells, J. G., and Cook, D. W. (2003). Seasonal abundance of total and pathogenic *Vibrio parahaemolyticus* in Alabama oysters. *Appl. Environ. Microbiol.* 69, 1521–1526.

Deter, J., Lozach, S., Véron, A., Chollet, J., Derrien, A., and Hervio-Heath, D. (2010). Ecology of pathogenic and non-pathogenic *Vibrio parahaemolyticus* on the French Atlantic coast. Effects of temperature, salinity, turbidity and chlorophyll a. *Environ. Microbiol.*12, 929–937.

Dobretsov, S. V., and Qian, P.-Y. (2002). Effect of Bacteria Associated with the Green Alga *Ulva reticulata* on Marine Micro- and Macrofouling. *Biofouling* 18, 217–228.

Duan, D., Xu, L., Fei, X., and Xu, H. (1995). Marine organisms attached to seaweed surfaces in Jiaozhou Bay, China. *World Journal of Microbiology & Biotechnology* 11, 351–352.

Duan, J., and Su, Y.-C. (2005). Occurrence of *Vibrio parahaemolyticus* in two Oregon oyster-growing bays. *J. Food Sci.* 70, M58–M63.

Eiler, A., Gonzalez-Rey, C., Allen, S., and Bertilsson, S. (2007). Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water. *FEMS Microbiol. Ecol.*60, 411–418.

Eiler, A., Johansson, M., and Bertilsson, S. (2006). Environmental influences on *Vibrio* populations in northern temperate and boreal coastal waters (Baltic and Skagerrak Seas). *Appl. Environ. Microbiol.* 72, 6004–6011.

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

Ferdous, Z. (2009). Survival of Vibrio cholerae O139 in Association with Anabaena variabilis in Four Different Microcosoms. *American J. Appl. Sci.* 6, 439–444.

Flärdh, K., Cohen, P. S., and Kjelleberg, S. (1992). Ribosomes exist in large excess over the apparent demand for protein synthesis during carbon starvation in marine Vibrio sp. strain CCUG 15956. *J. Bacteriol.* 174, 6780–6788.

Franco, S. L. M., Swenson, G. J., and Long, R. A. (2012). Year round patchiness of *Vibrio vulnificus* within a temperate Texas bay. *J. Appl. Microbiol.* 112, 593–604.

Frischkorn, K. R., Stojanovski, A., and Paranjpye, R. (2013). Vibrio parahaemolyticus type IV pili mediate interactions with diatom-derived chitin and point to an unexplored mechanism of environmental persistence. *Environ. Microbiol.*15, 1416–1427.

Froelich, B., Ayrapetyan, M., and Oliver, J. D. (2012). Integration of Vibrio vulnificus into Marine Aggregates and Its Subsequent Uptake by Crassostrea virginica Oysters. *Appl. Environ. Microbiol.*. 79, 1454–1458.

Froelich, B., and Oliver, J. D. (2013a). The Interactions of Vibrio vulnificus and the Oyster Crassostrea virginica. *Microbial ecology*, 1–10.

Froelich, B., Bowen, J., Gonzalez, R., Snedeker, A., and Noble, R. (2013b). Mechanistic and statistical models of total *Vibrio* abundance in the Neuse River Estuary. *Water Research* 47, 5783–5793.

Froelich, B., Ringwood, A., Sokolova, I., and Oliver, J. (2010). Uptake and depuration of the C-and E-genotypes of *Vibrio vulnificus* by the Eastern Oyster (Crassostrea virginica). *Environ. Microbiol.Reports* 2, 112–115.

Gamble, M. D., Bagwell, C. E., LaRocque, J., Bergholz, P. W., and Lovell, C. R. (2010). Seasonal Variability of Diazotroph Assemblages Associated with the Rhizosphere of the Salt Marsh Cordgrass, *Spartina alterniflora*. *Microb. Ecol.* 59, 253–265.

Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., Huse, S., McHardy, A. C., Knight, R., and Joint, I. (2011). Defining seasonal marine microbial community dynamics. *The ISME Journal* 6, 298–308.

Goecke, F., Labes, A., Wiese, J., and Imhoff, J. F. (2010). Chemical interactions between marine macroalgae and bacteria. *Mar Ecol Prog Ser* 409, 267–299.

Gosink, K. K., Kobayashi, R., Kawagishi, I., and Häse, C. C. (2002). Analyses of the roles of the three cheA homologs in chemotaxis of *Vibrio cholerae*. *Journal of Bacteriology* 184, 1767–1771.

Griffitt, K. J., and Grimes, D. J. (2013). Abundance and distribution of *Vibrio cholerae, V. parahaemolyticus,* and *V. vulnificus* following a major freshwater intrusion into the Mississippi Sound. *Microb. Ecol.* 65, 578–583.

Grimes, D. J., Johnson, C. N., Dillon, K. S., Flowers, A. R., Noriea, N. F., and Berutti, T. (2009). What Genomic Sequence Information Has Revealed About *Vibrio* Ecology in the Ocean—A Review. *Microb. Ecol.* 58, 447–460.

Grisez, L., Reyniers, J., Verdonck, L., Swings, J., and Ollevier, F. (1997). Dominant intestinal microflora of sea bream and sea bass larvae, from two hatcheries, during larval development. *Aquac.* 155, 387–399.

Halpern, M., Gancz, H., Broza, M., and Kashi, Y. (2003). *Vibrio cholerae* Hemagglutinin/Protease Degrades Chironomid Egg Masses. *Appl. Environ. Microbiol.* 69, 4200–4204.

Halpern, M., Raats, D., Lavion, R., and Mittler, S. (2006). Dependent population dynamics between chironomids (nonbiting midges) and *Vibrio cholerae*. *FEMS Microbiol. Ecol.*55, 98–104.

Hamdan, L. J., and Fulmer, P. A. (2011). Effects of COREXIT® EC9500A on bacteria from a beach oiled by the Deepwater Horizon spill. *Aquat. Microb. Ecol.* 63, 101.

Haygood, M. G., and Distel, D. L. (1993). Bioluminescent symbionts of flashlight fishes and deep-sea anglerfishes form unique lineages related to the genus *Vibrio*. *Nature* 363, 154–156.

Heidelberg, J. F., Heidelberg, K. B., and Colwell, R. R. (2002a). Seasonality of Chesapeake Bay bacterioplankton species. *Appl. Environ. Microbiol.* 68, 5488–5497.

Heidelberg, J. F., Heidelberg, K. B., and Colwell, R. R. (2002b). Bacteria of the γ-Subclass Proteobacteria Associated with Zooplankton in Chesapeake Bay. *Appl. Environ. Microbiol.* 68, 5498–5507.

Hervio-Heath, D., Colwell, R. R., Derrien, A., Robert-Pillot, A., Fournier, J. M., and Pommepuy, M. (2002). Occurrence of pathogenic vibrios in coastal areas of France. *J. Appl. Microbiol.* 92, 1123–1135.

Høi, L., Larsen, J. L., Dalsgaard, I., and Dalsgaard, A. (1998). Occurrence of *Vibrio vulnificus* biotypes in Danish marine environments. *Appl. Environ. Microbiol.* 64, 7–13.

Hollants, J., Leliaert, F., De Clerck, O., and Willems, A. (2013). What we can learn from sushi: a review on seaweed–bacterial associations. *FEMS Microbiol. Ecol.*83, 1–16.

Hood, M. A., and Winter, P. A. (1997). Attachment of *Vibrio cholerae* under various environmental conditions and to selected substrates. *FEMS Microbiol. Ecol.*22, 215–223.

Hood, M. A., Guckert, J. B., White, D. C., and Deck, F. (1986). Effect of nutrient deprivation on lipid, carbohydrate, DNA, RNA, and protein levels in Vibrio cholerae. *Appl. Environ. Microbiol.*. 52, 788–793.

Hsieh, J. L., Fries, J. S., and Noble, R. T. (2007). Vibrio and phytoplankton dynamics during the summer of 2004 in a eutrophying estuary. *Ecological Applications* 17, S102–S109.

Hsieh, J. L., Fries, J. S., and Noble, R. T. (2008). Dynamics and predictive modelling of *Vibrio* spp. in the Neuse River Estuary, North Carolina, USA. *Environ. Microbiol.*10, 57–64.

Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J., and Polz, M. F. (2008a). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320, 1081–1085.

Hunt, D. E., Gevers, D., Vahora, N. M., and Polz, M. F. (2008b). Conservation of the chitin utilization pathway in the Vibrionaceae. *Appl. Environ. Microbiol.*. 74, 44–51.

Huq, A., Small, E. B., West, P. A., Huq, M. I., Rahman, R., and Colwell, R. R. (1983). Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. *Appl. Environ. Microbiol.* 45, 275–283.

Huq, A., West, P. A., Small, E. B., Huq, M. I., and Colwell, R. R. (1984). Influence of water temperature, salinity, and pH on survival and growth of toxigenic *Vibrio cholerae* serovar 01 associated with live copepods in laboratory microcosms. *Appl. Environ. Microbiol.*. 48, 420–424.

Huq, A., Colwell, R. R., Rahman, R., Ali, A., Chowdhury, M. A., Parveen, S., Sack, D. A., and Russek-Cohen, E. (1990). Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and culture methods. *Appl. Environ. Microbiol.* 56, 2370–2373.

Huq, A., Sack, R. B., Nizam, A., Longini, I. M., Nair, G. B., Ali, A., Morris, J. G., Khan, M. N. H., Siddique, A. K., Yunus, M., et al., (2005). Critical factors influencing the occurrence of *Vibrio cholerae* in the environment of Bangladesh. *Appl. Environ. Microbiol.* 71, 4645–4654.

Islam, M. S., Drasar, B. S., and Bradley, D. J. (1988). Survival and attachment of toxigenic *Vibrio cholerae* O1 in association with four marine algae. *Bangladesh J Microbiol* 5, 41Á.

Islam, M. S., Drasar, B. S., and Bradley, D. J. (1989). Attachment of toxigenic *Vibrio cholerae* 01 to various freshwater plants and survival with a filamentous green alga, *Rhizoclonium fontanum*. *Journal of Tropical Medicine and Hygiene* 92, 396–401.

Islam, M. S., Drasar, B. S., and Bradley, D. J. (1990a). Survival of toxigenic *Vibrio cholerae* O1 with a common duckweed, Lemna minor, in artificial aquatic ecosystems. *Transactions of the Royal Society of tropical medicine and hygiene* 84, 422–424.

Islam, M. S., Drasar, B. S., and Bradley, D. J. (1990b). Long-term persistence of toxigenic *Vibrio cholerae* O1 in the mucilaginous sheath of a blue-green alga, Anabaena variabilis. *Journal of Tropical Medicine and Hygiene* 93, 133–139.

Islam, M. S., Rahim, Z., Alam, M. J., Begum, S., Moniruzzaman, S. M., Umeda, A., Amako, K., Albert, M. J., Sack, R. B., Huq, A., et al., (1999). Association of *Vibrio cholerae* O1 with the cyanobacterium, *Anabaena* sp., elucidated by polymerase chain reaction and transmission electron microscopy. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 93, 36–40.

Islam, M. S., Goldar, M. M., Morshed, M. G., Khan, M. N., Islam, M. R., and Sack, R. B. (2002). Involvement of the hap gene (mucinase) in the survival of *Vibrio cholerae* O1 in association with the blue-green alga, *Anabaena* sp. *Canadian journal of microbiology* 48, 793–800.

Islam, M. S., Goldar, M. M., Morshed, M. G., Bakht, H. B. M., and Sack, D. A. (2006). Chemotaxis between *Vibrio cholerae* O1 and a blue-green alga, *Anabaena* sp. *Epidemiology and Infection* 134, 645.

Jiang, S. C., and Fu, W. (2001). Seasonal abundance and distribution of *Vibrio cholerae* in coastal waters quantified by a 16S-23S intergenic spacer probe. *Microb. Ecol.* 42, 540–548.

Johnson, C. N., Flowers, A. R., Noriea, N. F., Zimmerman, A. M., Bowers, J. C., DePaola, A., and Grimes, D. J. (2010). Relationships between environmental factors and pathogenic vibrios in the northern Gulf of Mexico. *Appl. Environ. Microbiol.* 76, 7076–7084.

Johnson, C. N., Bowers, J. C., Griffitt, K. J., Molina, V., Clostio, R. W., Pei, S., Laws, E., Paranjpye, R. N., Strom, M. S., Chen, A., et al., (2012). Ecology of *Vibrio parahaemolyticus* and *Vibrio vulnificus* in the coastal and estuarine waters of Louisiana, Maryland, Mississippi, and Washington (United States). *Appl. Environ. Microbiol.* 78, 7249–7257.

Johnson, C. N. (2013). Fitness Factors in Vibrios: a Mini-review. *Microb Ecol* 65, 826–851.

Jones, S. H., and Summer-Brason, B. (1998). Incidence and detection of pathogenic *Vibrio* sp. in a northern new England Estuary, USA. *Journal of Shellfish Research* 17, 1665–1669.

Jutla, A., Akanda, A. S., Huq, A., Faruque, A. S. G., Colwell, R., and Islam, S. (2013). A water marker monitored by satellites to predict seasonal endemic cholera. *Remote Sensing Letters* 4, 822–831.

Kanagasabhapathy, M., Sasaki, H., and Nagata, S. (2008). Phylogenetic identification of epibiotic bacteria possessing antimicrobial activities isolated from red algal species of Japan. *World J Microbiol Biotechnol* 24, 2315–2321.

Kassen, R. (2002). The experimental evolution of specialists, generalists, and the maintenance of diversity. *Journal of Evolutionary Biology* 15, 173–190.

Kenyon, J. E., Gillies, D. C., Piexoto, D. R., and Austin, B. (1983). *Vibrio cholerae* (non-O1) isolated from California coastal waters. *Appl. Environ. Microbiol.* 46, 1232–1233.

Kirschner, A. K., Schlesinger, J., Farnleitner, A. H., Hornek, R., Süß, B., Golda, B., Herzig, A., and Reitner, B. (2008). Rapid growth of planktonic *Vibrio cholerae* non-O1/non-O139 strains in a large alkaline lake in Austria: dependence on temperature and dissolved organic carbon quality. *Appl. Environ. Microbiol.*. 74, 2004–2015.

Kirschner, A. K. T., Schauer, S., Steinberger, B., Wilhartitz, I., Grim, C. J., Huq, A., Colwell, R. R., Herzig, A., and Sommer, R. (2011). Interaction of *Vibrio cholerae* non-O1/non-O139 with copepods, cladocerans and competing bacteria in the large alkaline lake Neusiedler See, Austria. *Microb Ecol* 61, 496–506.

Koenig, J. E., Bourne, D. G., Curtis, B., Dlutek, M., Stokes, H. W., Doolittle, W. F., and Boucher, Y. (2011). Coral-mucus-associated *Vibrio* integrons in the Great Barrier Reef: genomic hotspots for environmental adaptation. *ISME J* 5, 962–972.

Koren, O., and Rosenberg, E. (2006). Bacteria Associated with Mucus and Tissues of the Coral *Oculina patagonica* in Summer and Winter. *Appl. Environ. Microbiol.* 72, 5254–5259.

Kramer, J. G., and Singleton, F. L. (1992). Variations in rRNA content of marine *Vibrio* spp. during starvation-survival and recovery. *Appl. Environ. Microbiol.*. 58, 201–207.

Krediet, C. J., Ritchie, K. B., Alagely, A., and Teplitski, M. (2013). Members of native coral microbiota inhibit glycosidases and thwart colonization of coral mucus by an opportunistic pathogen. *ISME J* 7, 980–990.

Kumazawa, N. H., Fukuma, N., and Komoda, Y. (1991). Attachment of *Vibrio parahaemolyticus* strains to estuarine algae. *The Journal of veterinary medical science / the Japanese Society of Veterinary Science* 53, 201–205.

Kvennefors, E. C. E., Sampayo, E., Ridgway, T., Barnes, A. C., and Hoegh-Guldberg, O. (2010). Bacterial Communities of Two Ubiquitous Great Barrier Reef Corals Reveals Both Site- and Species-Specificity of Common Bacterial Associates. *PLoS ONE* 5, e10401.

Lakshmanaperumalsamy, P., and Purushothaman, A. (1982). Heterotrophic bacteria associated with seaweed. *Proc. Indian Acad. Sci. (Plant Sci.)* 91, 487–493.

Lapota, D., Galt, C., Losee, J. R., Huddell, H. D., Orzech, J. K., and Nealson, K. H. (1988). Observations and measurements of planktonic bioluminescence in and around a milky sea. *Journal of Experimental Marine Biology and Ecology* 119, 55–81.

Larocque, J. R., Bergholz, P. W., Bagwell, C. E., and Lovell, C. R. (2004). Influence of host plant-derived and abiotic environmental parameters on the composition of the diazotroph assemblage associated with roots of *Juncus roemerianus*. *Antonie Van Leeuwenhoek* 86, 249–261.

Larsen, M. H., Blackburn, N., Larsen, J. L., and Olsen, J. E. (2004). Influences of temperature, salinity and starvation on the motility and chemotactic response of *Vibrio anguillarum*. *Microbiol.*150, 1283–1290.

Laycock, R. A. (1974). Bacteria associated with the surface of *Laminaria* fronds. *Mar. Biol.* 25, 223–231.

Lee, J. V., Bashford, D. J., Donovan, T. J., Furniss, A. L., and West, P. A. (1982). The incidence of *Vibrio cholerae* in water, animals and birds in Kent, England. *J. Appl. Microbiol.* 52, 281–291.

Liston, J. (1954). A group of luminous and nonluminous bacteria from the intestine of flatfish. *J. Gen. Microbiol.* 12.

Liston, J. (1957). The occurrence and distribution of bacterial types on flatfish. *J. Gen. Microbiol.* 16, 205–216.

Liu, Z., and Liu, J. (2013). Evaluating bacterial community structures in oil collected from the sea surface and sediment in the northern Gulf of Mexico after the Deepwater Horizon oil spill. *Microbiol. Open* 2, 492–504.

Lizárraga-Partida, M. L., Mendez-Gómez, E., Rivas-Montaño, A. M., Vargas-Hernández, E., Portillo-López, A., González-Ramírez, A. R., Huq, A., and Colwell, R. R. (2009). Association of *Vibrio cholerae* with plankton in coastal areas of Mexico. *Environ. Microbiol.*11, 201–208.

Lobitz, B., Beck, L., Huq, A., Wood, B., Fuchs, G., Faruque, A. S. G., and Colwell, R. (2000). Climate and infectious disease: use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. *Proc. Natl. Acad. Sci. U.S.A.*97, 1438–1443.

Lovell, C. R., Decker, P. V., Bagwell, C. E., Thompson, S., and Matsui, G. Y. (2008). Analysis of a diverse assemblage of diazotrophic bacteria from *Spartina alterniflora* using DGGE and clone library screening. *J. Microbiol. Methods* 73, 160–171.

Lyons, M. M., Lau, Y.-T., Carden, W. E., Ward, J. E., Roberts, S. B., Smolowitz, R., Vallino, J., and Allam, B. (2007). Characteristics of Marine Aggregates in Shallow-water Ecosystems: Implications for Disease Ecology. *EcoHealth* 4, 406–420.

Magny, G. C. de, Murtugudde, R., Sapiano, M. R. P., Nizam, A., Brown, C. W., Busalacchi, A. J., Yunus, M., Nair, G. B., Gil, A. I., Lanata, C. F., et al., (2008). Environmental signatures associated with cholera epidemics. *Proc. Natl. Acad. Sci. U.S.A.*105, 17676–17681.

Magny, G. C. de, Mozumder, P. K., Grim, C. J., Hasan, N. A., Naser, M. N., Alam, M., Sack, R. B., Huq, A., and Colwell, R. R. (2011). Role of zooplankton diversity in *Vibrio cholerae* population dynamics and in the incidence of cholera in the Bangladesh Sundarbans. *Appl. Environ. Microbiol.* 77, 6125–6132.

Martinez-Urtaza, J., Lozano-Leon, A., Varela-Pet, J., Trinanes, J., Pazos, Y., and Garcia-Martin, O. (2008). Environmental Determinants of the occurrence and distribution of *Vibrio parahaemolyticus* in the Rias of Galicia, Spain. *Appl. Environ. Microbiol.* 74, 265–274.

Martinez-Urtaza, J., Blanco-Abad, V., Rodriguez-Castro, A., Ansede-Bermejo, J., Miranda, A., and Rodriguez-Alvarez, M. X. (2012). Ecological determinants of the occurrence and dynamics of *Vibrio parahaemolyticus* in offshore areas. *ISME J* 6, 994–1006.

Materna, A. C., Friedman, J., Bauer, C., David, C., Chen, S., Huang, I. B., Gillens, A., Clarke, S. A., Polz, M. F., and Alm, E. J. (2012). Shape and evolution of the fundamental niche in marine *Vibrio*. *ISME J* 6, 2168–2177.

Matz, C., McDougald, D., Moreno, A. M., Yung, P. Y., Yildiz, F. H., and Kjelleberg, S. (2005). Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16819–16824.

Mishra, A., Taneja, N., and Sharma, M. (2012). Environmental and epidemiological surveillance of *Vibrio cholerae* in a cholera-endemic region in India with freshwater environs. *J. Appl. Microbiol.* 112, 225–237.

Mizanur, R. M., Khan, S. I., Islam, M. S., Sultana, I., Rahim, Z., and Albert, M. J. (2001). Role of Chemotaxis in the attachment of *Vibrio cholerae* 0139 with different aquatic flora and fauna. *Pak. J. Biol. Sci* 4, 1395–1399.

Mizanur, R. M., Islam, M. S., Khan, S. I., and Rahim, Z. (2002). The Chemotactic Response of *Vibrio cholerae* 0139 to The Mucilaginous Sheath of Blue-Green Algae Is Mediated by a Combination of Sheath Components. *Microbes and Environments* 17, 18–25.

Motes, M. L., DePaola, A., Cook, D. W., Veazey, J. E., Hunsucker, J. C., Garthright, W. E., Blodgett, R. J., and Chirtel, S. J. (1998). Influence of water temperature and salinity on *Vibrio vulnificus* in Northern Gulf and Atlantic coast oysters (*Crassostrea virginica*). *Appl. Environ. Microbiol.* 64, 1459–1465.

Mouriño-Pérez, R. R., Worden, A. Z., and Azam, F. (2003). Growth of *Vibrio cholerae* O1 in Red Tide Waters off California. *Appl. Environ. Microbiol.* 69, 6923–6931.

Moxley, K., and Schmidt, S. (2010). Preliminary characterization of an estuarine, benzoate-utilizing *Vibrio* sp. isolated from Durban Harbour, South Africa. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, 1249–1254.

Mueller, R. S., McDougald, D., Cusumano, D., Sodhi, N., Kjelleberg, S., Azam, F., and Bartlett, D. H. (2007). *Vibrio cholerae* Strains Possess Multiple Strategies for Abiotic and Biotic Surface Colonization. *J. Bacteriol.* 189, 5348–5360.

Muroga, K., Higashi, M., and Keitoku, H. (1987). The isolation of intestinal microflora of farmed red seabream (*Pagrus major*) and black seabream (*Acanthopagrus schlegeli*) at larval and juvenile stages. *Aquac.* 65, 79–88.

Murphree, R. L., and Tamplin, M. L. (1995). Uptake and retention of *Vibrio cholerae* O1 in the Eastern oyster, *Crassostrea virginica*. *Appl. Environ. Microbiol.*. 61, 3656–3660.

Nakanishi, K., Nishijima, M., Nishimura, M., Kuwano, K., and Saga, N. (1996). Bacteria that Induce Morphogenesis in *Ulva pertusa* (Chlorophyta) Grown Under Axenic Conditions. *J. Phycol.* 32, 479–482.

Nealson, K. H., and Hastings, J. W. (2006). Quorum Sensing on a Global Scale: Massive Numbers of Bioluminescent Bacteria Make Milky Seas. *Appl. Environ. Microbiol.* 72, 2295–2297.

Neogi, S. B., Islam, M. S., Nair, G. B., Yamasaki, S., and Lara, R. J. (2012). Occurrence and distribution of plankton-associated and free-living toxigenic *Vibrio cholerae* in a tropical estuary of a cholera endemic zone. *Wetlands Ecology and Management* 20, 271–285.

Newman Jr., J. T., Cosenza, B. J., and Buck, J. D. (1972). Aerobic microflora of the bluefish (*Pomatomus saltatrix*) intestine. *Journal of the Fisheries Board of Canada* 29, 333–336.

Nigro, O. D., Hou, A., Vithanage, G., Fujioka, R. S., and Steward, G. F. (2011). Temporal and spatial variability in culturable pathogenic *Vibrio* spp. in Lake Pontchartrain, Louisiana, following hurricanes Katrina and Rita. *Appl. Environ. Microbiol.*. 77, 5384–5393.

Nishina, T., Wada, M., Ozawa, H., Hara-Kudo, Y., Konuma, H., Hasegawa, J., and Kumagai, S. (2004). Growth kinetics of *Vibrio parahaemolyticus* O3: K6 under varying conditions of pH, NaCl concentration and temperature. *Shokuhin eiseigaku zasshi. J. Food Hyg. Soc. Jpn.* 45, 35.

Oberbeckmann, S., Fuchs, B. M., Meiners, M., Wichels, A., Wiltshire, K. H., and Gerdts, G. (2012). Seasonal dynamics and modeling of a *Vibrio* community in coastal waters of the North Sea. *Microb Ecol* 63, 543–551.

Oliver, J. D., Warner, R. A., and Cleland, D. R. (1983). Distribution of *Vibrio vulnificus* and other lactose-fermenting vibrios in the marine environment. *Appl. Environ. Microbiol.* 45, 985–998.

Paerl, H. W., and Gallucci, K. K. (1985). Role of Chemotaxis in Establishing a Specific Nitrogen-Fixing Cyanobacterial-Bacterial Association. *Science* 227, 647–649.

Paerl, H. W., and Pinckney, J. L. (1996). A mini-review of microbial consortia: Their roles in aquatic production and biogeochemical cycling. *Microb. Ecol.* 31, 225–247.

Patel, P., Callow, M. E., Joint, I., and Callow, J. A. (2003). Specificity in the settlement–modifying response of bacterial biofilms towards zoospores of the marine alga *Enteromorpha*. *Environ. Microbiol.*5, 338–349.

Pernthaler, A., Pernthaler, J., Eilers, H., and Amann, R. (2001). Growth patterns of two marine isolates: adaptations to substrate patchiness? *Appl. Environ. Microbiol.*. 67, 4077–4083.

Pfeffer, C. S., Hite, M. F., and Oliver, J. D. (2003). Ecology of *Vibrio vulnificus* in estuarine waters of eastern North Carolina. *Appl. Environ. Microbiol.* 69, 3526–3531.

Polz, M. F., Alm, E. J., and Hanage, W. P. (2013). Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* 29, 170–175.

Polz, M. F., and Cordero, O. X. (Submitted.) Ecological interactions, population structure and the genome dynamics of bacteria and archaea.

Prasanthan, V., Udayakumar, P., Sarathkumar, and Ouseph, P. P. (2011). Influence of abiotic environmental factors on the abundance and distribution of *Vibrio* species in coastal waters of Kerala, India. *Indian J. Mar. Sci.* 40, 587–592.

Preheim, S. P., Boucher, Y., Wildschutte, H., David, L. A., Veneziano, D., Alm, E. J., and Polz, M. F. (2011a). Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ. Microbiol.*13, 265–275.

Preheim, S. P., Timberlake, S., and Polz, M. F. (2011b). Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Appl. Environ. Microbiol.*. 77, 7195–7206.

Ramaiah, N., Ravel, J., Straube, W. L., Hill, R. T., and Colwell, R. R. (2002). Entry of *Vibrio harveyi* and *Vibrio fischeri* into the viable but nonculturable state. *J. Appl. Microbiol.* 93, 108–116.

Ramirez, G. D., Buck, G. W., Smith, A. K., Gordon, K. V., and Mott, J. B. (2009). Incidence of *Vibrio vulnificus* in estuarine waters of the south Texas Coastal Bend region. *J. Appl. Microbiol.* 107, 2047–2053.

Randa, M. A., Polz, M. F., and Lim, E. (2004). Effects of temperature and salinity on *Vibrio vulnificus* population dynamics as assessed by quantitative PCR. *Appl. Environ. Microbiol.* 70, 5469–5476.

Rawlings, T. K., Ruiz, G. M., and Colwell, R. R. (2007). Association of *Vibrio cholerae* O1 El Tor and O139 Bengal with the copepods *Acartia tonsa* and *Eurytemora affinis*. *Appl. Environ. Microbiol.* 73, 7926–7933.

Rowbotham, T. J. (1980). Preliminary report on the pathogenicity of *Legionella pneumophila* for freshwater and soil amoebae. *J. Clin. Pathol.* 33, 1179–1183.

Ruby, E. G., and Nealson, K. H. (1976). Symbiotic association of *Photobacterium fischeri* with the marine luminous fish *Monocentris japonica*: a model of symbiosis based on bacterial studies. *The Biological Bulletin* 151, 574–586.

Ruby, E. G., and Lee, K. H. (1998). The *Vibrio fischeri-Euprymna scolopes* light organ association: current ecological paradigms. *Appl. Environ. Microbiol.*. 64, 805–812.

Sandström, G., Saeed, A., and Abd, H. (2010). *Acanthamoeba polyphaga* is a possible host for *Vibrio cholerae* in aquatic environments. *Exp. Parasitol.* 126, 65–68.

Sawabe, T. (2006). "The Mutual Partnership between Vibrio halioticoli and Abalones," in *Biology of Vibrios*, eds. F. L. Thompson, B. Austin, and J. Swings (Washington: Amer Soc Microbiology), 219–230.

Shapiro, B. J., Friedman, J., Cordero, O. X., Preheim, S. P., Timberlake, S. C., Szabó, G., Polz, M. F., and Alm, E. J. (2012). Population genomics of early events in the ecological differentiation of bacteria. *Science* 336, 48–51.

Sharon, G., and Rosenberg, E. (2008). Bacterial growth on coral mucus. *Curr. Microbiol.* 56, 481–488.

Shen, X., Cai, Y., Liu, C., Liu, W., Hui, Y., and Su, Y.-C. (2009). Effect of temperature on uptake and survival of *Vibrio parahaemolyticus* in oysters (*Crassostrea plicatula*). *Int. J. Food Microbiol.* 136, 129–132.

Shiba, T., and Taga, N. (1980). Heterotrophic bacteria attached to seaweeds. *J. Exp. Mar. Biol. Ecol.* 47, 251–258.

Simidu, U., Ashino, K., and Kaneko, E. (1971). Bacterial flora of phyto-and zoo-plankton in the inshore water of Japan. *Can. J. Microbiol.* 17, 1157–1160.

Sjoblad, R. D., and Mitchell, R. (1979). Chemotactic responses of Vibrio alginolyticus to algal extracellular products. *Can. J. Microbiol.* 25, 964–967.

Smriga, S., Sandin, S. A., and Azam, F. (2010). Abundance, diversity, and activity of microbial assemblages associated with coral reef fish guts and feces. *FEMS Microbiol. Ecol.*73, 31–42.

Sobrinho, P. de S. C., Destro, M. T., Franco, B. D. G. M., and Landgraf, M. (2010). Correlation between environmental factors and prevalence of *Vibrio parahaemolyticus* in oysters harvested in the southern coastal area of Sao Paulo State, Brazil. *Appl. Environ. Microbiol.* 76, 1290–1293.

Sochard, M. R., Wilson, D. F., Austin, B., and Colwell, R. R. (1979). Bacteria associated with the surface and gut of marine copepods. *Appl. Environ. Microbiol.*. 37, 750–759.

Spira, W. M., Huq, A., Ahmed, Q. S., and Saeed, Y. A. (1981). Uptake of *Vibrio cholerae* Biotype eltor from Contaminated Water by Water Hyacinth (*Eichornia crassipes*). *Appl. Environ. Microbiol.* 42, 550–553.

Stabb, E. V. (2006). *The Vibrio fischeri-Euprymna scolopes* light organ symbiosis in *Biology of Vibrios*, eds. F. L. Thompson, B. Austin, and J. Swings (Washington: Amer. Soc. Microbiol.), 204–218.

Stephens, E. L., Molina, V., Cole, K. M., Laws, E., and Johnson, C. N. (2013). In situ and in vitro impacts of the Deepwater Horizon oil spill on *Vibrio parahaemolyticus*. *Mar. Poll. Bull.* 75, 90–97.

Sugita, H., Ushioka, S., Kihara, D., and Deguchi, Y. (1985). Changes in the bacterial composition of water in a carp rearing tank. *Aquac.* 44, 243–247.

Suttle CA. 2007. Marine viruses - major players in the global ecosystem. *Nat. Rev. Microbiol.* 5:801-812.

Sutton, D. C., and Clements, K. D. (1988). Aerobic, heterotrophic gastrointestinal microflora of tropical marine fishes. in *Proceedings of the 6th International Coral Reef Symposium*, 185–190.

Szabo, G., Preheim, S. P., Kauffman, K. M., David, L. A., Shapiro, J., Alm, E. J., and Polz, M. F. (2013). Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J.* 7, 509–519.

Tait, K., Joint, I., Daykin, M., Milton, D. L., Williams, P., and Camara, M. (2005). Disruption of quorum sensing in seawater abolishes attraction of zoospores of the green alga *Ulva* to bacterial biofilms. *Environ. Microbiol.*7, 229–240.

Tamplin, M. L., Gauzens, A. L., Huq, A., Sack, D. A., and Colwell, R. R. (1990). Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl. Environ. Microbiol.*. 56, 1977–1980.

Tarsi, R., and Pruzzo, C. (1999). Role of Surface Proteins in *Vibrio cholerae* Attachment to Chitin. *Appl. Environ. Microbiol.* 65, 1348–1351.

Thompson F. L., Iida T., and Swings J. 2004. Biodiversity of vibrios. *Microbiol. Mol. Biol. Rev.* 68:403-431.

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. *Appl. Environ. Microbiol.* 70, 4103–4110.

Thompson, J. R., Marcelino, L. A., and Polz, M. F. (2005). "Diversity, Sources, and Detection of Human Bacterial Pathogens in the Marine Environment," in Oceans and Health: Pathogens in the Marine Environment, eds. S. Belkin and R. R. Colwell (Springer US), 29–68.

Thompson, J. R., and Polz, M. F. (2006). "Dynamics of *Vibrio* populations and their role in environmental nutrient cycling," in *The Biology of Vibrios*, ed. F. L. Thompson, B. Austin, and J. Swings (Washington, DC: ASM Press), 190–203.

Thongchankaew, U., Mittraparp-Arthorn, P., Sukhumungoon, P., Tansila, N., Nuidate, T., Nishibuchi, M., and Vuddhakul, V. (2011). Occurrence of potentially pathogenic vibrios and related environmental factors in Songkhla Lake, Thailand. *Canadian Journal of Microbiology* 57, 867–873.

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.

Urdaci, M. C., Stal, L. J., and Marchand, M. (1988). Occurrence of nitrogen fixation among *Vibrio* spp. *Arch. Microbiol.* 150, 224–229.

Vahl, O. (1972). Efficiency of particle retention in *Mytilus edulis L. Ophelia* 10, 17–25.

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

Vezzulli, L., Brettar, I., Pezzati, E., Reid, P. C., Colwell, R. R., Höfle, M. G., and Pruzzo, C. (2012). Long-term effects of ocean warming on the prokaryotic community: evidence from the vibrios. *ISME J.* 6, 21–30.

Wang, Z., Xiao, T., Pang, S., Liu, M., and Yue, H. (2009). Isolation and identification of bacteria associated with the surfaces of several algal species. *Chin. J. Ocean. Limnol.* 27, 487–492.

Warner, E., and Oliver, J. D. (2008). Population structures of two Genotypes of *Vibrio vulnificus* in oysters (*Crassostrea virginica*) and seawater. *Appl. Environ. Microbiol.*. 74, 80–85.

West, P. A., Okpokwasili, G. C., Brayton, P. R., Grimes, D. J., and Colwell, R. R. (1984). Numerical taxonomy of phenanthrene-degrading bacteria isolated from the Chesapeake Bay. *Appl. Environ. Microbiol.*. 48, 988–993.

Wetz, J. J., Blackwood, A. D., Fries, J. S., Williams, Z. F., and Noble, R. T. (2008). Trends in total *Vibrio* spp. and *Vibrio vulnificus* concentrations in the eutrophic Neuse River Estuary, North Carolina, during storm events. *Aquat. Microb. Ecol.* 53, 141.

Worden, A. Z., Seidel, M., Smriga, S., Wick, A., Malfatti, F., Bartlett, D., and Azam, F. (2006). Trophic regulation of Vibrio cholerae in coastal marine waters. *Environ. Microbiol.*8, 21–29.

Xing, M., Hou, Z., Yuan, J., Liu, Y., Qu, Y., and Liu, B. (2013). Taxonomic and functional metagenomic profiling of gastrointestinal tract microbiome of the farmed adult turbot (*Scophthalmus maximus*). *FEMS Microbiol. Ecol.*

Yawata, Y., Cordero, O. X., Menolascina, F., Hehemann, J.-H., Polz, M. F., Stocker, R. (Submitted.) A competition-dispersal trade-off ecologically differentiates recently speciated marine bacterioplankton populations.

Yooseph S, Nealson KH, Rusch DB, McCrow JP, Dupont CL, Kim M, Johnson J, Montgomery R, Ferriera S, Beeson K, et al., 2010. Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468, 60-66.

Yoshimizu, M., and Kimura, T. (1976). Study on the intestinal microflora of salmonids. *Fish Pathology* 10, 243–259.

Yu, C., Bassler, B. L., and Roseman, S. (1993). Chemotaxis of the marine bacterium *Vibrio furnissii* to sugars. A potential mechanism for initiating the chitin catabolic cascade. *Journal of Biological Chemistry* 268, 9405–9409.

Zarubin, M., Belkin, S., Ionescu, M., and Genin, A. (2012). Bacterial bioluminescence as a lure for marine zooplankton and fish. *Proc. Natl. Acad. Sci. U.S.A.* 109, 853–857.

Zettler, E. R., Mincer, T. J., and Amaral-Zettler, L. A. (2013). Life in the "plastisphere": microbial communities on plastic marine debris. *Environ. Sci. Technol.* 47, 7137–7146.

Zimmerman, A. M., DePaola, A., Bowers, J. C., Krantz, J. A., Nordstrom, J. L., Johnson, C. N., and Grimes, D. J. (2007). Variability of total and pathogenic *Vibrio parahaemolyticus* densities in Northern Gulf of Mexico water and oysters. *Appl. Environ. Microbiol.*. 73, 7589–7596.

Chapter 3: Community Dynamics of Marine Microbes at High Genetic, Temporal, and Spatial Resolution

Abstract

In the effort to characterize the dynamics of environmental microbes, most time-series studies have employed monthly sampling of whole-water samples, and typically resolve taxa only to the family or genus level at best, as constrained by the low phylogenetic resolution of tag sequencing. Hence, little is known about the extent to which dynamics may differ at finer genetic and ecological scales. Here, we present a three-month daily marine environmental time-series wherein we resolved the Vibrionaceae, a family of diverse marine heterotrophs, to the population level and tracked their dynamics among four different habitat partitions, using targeted sequencing of a high-resolution marker gene. We show that populations display a broad range of dynamic behavior that does not assort with phylogenetic distance, reflecting sympatric differentiation of close relatives. Community dynamics are moreover structured by habitat: small size fractions, representing microbes that are free-living or attached to small particles, demonstrate remarkably high daily turnover, including a dramatic yet brief population expansion wherein the Vibrionaceae cumulatively rose from ~2% of the microbial community to nearly 30% over a few days, induced by warm weather and abundant macroalgal material. By contrast, the community associated with large particles was far more stable and far lower in diversity, and its dynamics did not reflect the bloom in the smaller size fractions. We were furthermore able to identify likely abiotic drivers and eukaryotic and bacterial interactors of the Vibrionaceae populations using complementary datasets. We found that while most populations shared abiotic drivers across the fractions, biotic interactors were nearly unique to each population, and distinct in composition between the smaller size fractions (dominated by interactions with eukaryotic heterotrophs) and the large particle fraction (dominated by bacteria-bacteria interactions), again reflecting differing ecological regimes between the habitats. Thus, our results demonstrate that even close relatives sharing the same environment manifest substantial and ecologically structured diversity in their dynamic behavior, suggesting that typical time-series sampling resolution significantly underrepresents the dynamism and diversity of marine microbes.

Chapter 3 Introduction

Heterotrophic bacteria account for nearly a quarter of the surface ocean's biomass and dominate its metabolic activity (Azam and Malfatti, 2007; Pomeroy *et al.*, 2007). Consequently, understanding microbial dynamics is crucial to developing a full picture of marine energy and nutrient fluxes. However, elucidating all of the factors underlying dynamics is challenging in the mutable environment of the ocean, where diverse microbes compete for diverse nutrient sources while undergoing bacteriophage and eukaryotic predation, all in a physically shifting environment.

When trying to characterize such complex ecosystems, time-series studies allow us to capture a systems-level view and extract ecological patterns: different microbes' environmental preferences, their probable positive and negative interactions with other organisms, and properties like stability and seasonal variation (Fuhrman *et al.*, 2015; Faust *et al.*, 2015). This stands in contrast to extrapolating from the snapshot of a single timepoint, or of inferring microbial capabilities based on genomic content or other molecular surveys, which represent only potential ecological function.

Across time-series studies, the pattern evident is that communities tend to be broadly predictable when characterized at longer time intervals (seasons to years) and higher taxonomic levels (genus or above, as is typical for surveys conducted with 16S rRNA and other highly conserved marker genes) (Shade *et al.*, 2013; Fuhrman *et al.*, 2015), although observed patterns of seasonality and stability nonetheless break down over time. Still, finer-scale dynamics remain more difficult to interpret and predict. On timescales shorter than weeks, communities turn over rapidly due to factors that are difficult to measure at bulk scales: biological interactions, and short-term perturbations like phytoplankton blooms and storms (e.g., Brussaard *et al.*, 1996; Castberg *et al.*, 2001; Teeling *et al.*, 2012). The complexity of dynamics at this scale is multiplied when considering the ecological diversity that can be concealed below genus level, such that close relatives may have distinct environmental distributions and react differently to the same perturbation (Wu and Hahn, 2006; Carlson *et al.*, 2008; Hunt *et al.*, 2008; Malmstrom *et al.*, 2010; Preheim *et al.*, 2011; Szabo *et al.*, 2012). To gain insight into finer-scale dynamics, then, communities must be characterized at appropriate scales.

Here, we present a three-month time-series dataset distinguished by its fine-scale resolution along genetic, temporal, and spatial axes: population-level resolution within a single microbial family, the Vibrionaceae; daily sampling; and size-fractionated water column samples representing distinct potential habitats. The Vibrionaceae are ecologically diverse copiotrophs (Takemura *et al.*, 2014), able to

exploit many nutrient types and mount rapid physiological responses to changes in nutrient availability. The coastal Vibrionaceae community sampled for this time-series has previously been developed as a model for microbial ecology, revealing the coexistence of ~20 populations that are closely related, yet have distinct ecological preferences: free-living versus attached; generalist versus specialized; use of animal and/or algal resources (Hunt *et al.*, 2008; Preheim *et al.*, 2011; Szabo *et al.*). This study is the first to situate these finely resolved populations in a dynamic context. A previous meta-analysis of Vibrionaceae community dynamics revealed that while broad environmental variables like temperature and salinity have strong predictive power for Vibrionaceae as a family, biological variables like chlorophyll a concentration add little or inconsistent information, again suggesting that low taxonomic resolution obscures population-specific biological affiliations (Takemura *et al.*, 2014).

Here, in order to achieve population-level resolution of the Vibrionaceae community within our time-series samples, we developed a pipeline for Vibrionaceae-specific sequencing of a high-resolution marker gene and taxonomic assignment via phylogenetic placement. We were able to tease apart diverse ecological dynamics within the Vibrionaceae, and correlate them with environmental parameters and dynamics of the co-occurring bacterial and eukaryotic communities in order to discover potential drivers and interactors. Our genetic, temporal, and spatial resolution ultimately show that Vibrionaceae population dynamics are diverse, rapid, and spatially structured: individual Vibrionaceae populations display a broad range of dynamic behavior, from abundant generalists to rare "one-hit wonders" that manifest brief yet intense peaks, while the size fractions reflect free-living versus particle-attached lifestyles, with distinct signatures of dynamics, diversity, and predicted abiotic and biotic interactors, and near-daily turnover occurring in many of them. These results shed light on previously undersampled ecological regimes, and suggest that typical lower-resolution time-series sampling strongly underrepresent the dynamism and diversity of marine community dynamics.

Chapter 3 Results and Discussion

1. Validation of Vibrionaceae hsp60 amplicon libraries and taxonomy assignment

In order to characterize Vibrionaceae community composition within environmental DNA samples from our time-series dataset, we developed a protocol for targeted PCR amplification and Illumina sequencing of Vibrionaceae *hsp*60; *hsp*60 is a housekeeping gene, which has previously been developed as a marker for Vibrionaceae population structure because its phylogenetic resolution is greater than 16S rRNA's (Hunt *et al.*, 2008). In addition, we designed a Vibrionaceae *hsp*60 reference library for population-level taxonomy assignment by phylogenetic placement. This reference library was designed to encompass 18 ecological populations previously identified by (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Preheim, Timberlake, *et al.*, 2011; Szabo *et al.*), 42 additional species represented by type-strain gene sequences, and 16 potentially novel taxa (environmental strain isolates) whose *hsp*60 sequences did not fall into clusters with named representatives; these were named based on the nearest named representative in the reference library tree, e.g., *Vibrio sp. nov. pacinii*-like. An additional three taxonomic labels represented intermixtures of 2-3 closely related populations that could not always be distinguished using our 215bp *hsp*60 amplicon region. Only one of these aggregate labels was significantly abundant in our system, "*Vibrio splendidus tasmaniensis*." We later added sequences under this label to *V. tasmaniensis*, as their dynamics were more similar than those of *V. splendidus* (average Spearman correlation of 0.52 with *V. tasmaniensis*, vs. 0.34 with *V. splendidus*; see Methods for more detail). Hence, despite relatively short amplicon length, we were ultimately able to resolve all previously ecologically characterized populations that reached significant abundance in our dataset. In addition, after analyzing our environmental amplicon data we discovered additional taxa that did not correspond to previously named representatives, two of which were occasionally abundant and were also named as *Vibrio sp. nov.* based on location in the *hsp*60 reference tree.

To quantify any biases in community composition resulting from amplification, sequencing, and/or computational sequence processing, and to validate the performance of the reference library, we created a mock community comprising genomic DNA from environmental Vibrionaceae isolates, representing 25 different populations. We subjected the community to the same *hsp*60 library construction and *in silico* pipeline as the environmental samples. Analysis of the mock communities'

taxonomic compositions showed only minor deviations from the expected composition (Supp. Fig. 1), supporting the accuracy of the amplicon libraries' representation of input sequence diversity. First, increasing the number of PCR amplification cycles used across the range of 21-25 did not create any significant variation in community composition (Mann-Whitney tests, each observed composition versus the expected: p-values of 0.85-0.98). Second, observed abundance for each Vibrionaceae population represented in the community varied from expected abundance by a median factor of only 1.6, i.e., populations were rarely under- or overrepresented by a factor of more than 2, which we deemed a reasonable amount of variation to tolerate. Third, when testing *in silico* sequence-filtering parameters, we found that excluding singleton OTUs often resulted in substantial underrepresentation of the *V. splendidus* and *V. tasmaniensis* groups (which were abundant in the mock community, to echo previously observed community compositions) (Supp. Fig. 1); as a result, we decided to retain singletons. Examination of phylogenetic distances among mock-community output sequences suggests that this effect occurred because the closely related *V. splendidus* and *tasmaniensis* are prone to producing chimeras when PCR-amplified, resulting in large numbers of OTUs that are low-abundance, yet still accurate reflections of input *V. splendidus* and *V. tasmaniensis* quantities (data not shown). This unexpectedly strong bias introduced by potential chimaera formation among closely related sequences should be investigated in more detail in the future. Overall, the analysis of the mock communities supported the accuracy of both our experimental and computational procedures in reading out Vibrionaceae community composition based on the 215bp *hsp*60 amplicon used here.

One shortcoming of our *hsp*60-sequencing method is that since the primers are not perfectly specific, the sequencing libraries contained a significant fraction of non-Vibrionaceae diversity, requiring several filtering steps (see Methods). Approximately 90% of sequences per daily library were discarded, with most being lost when we discarded sequences with no BLAST hit to a Vibrionaceae reference sequence. Other discarded sequences were close Gammaproteobacteria relatives, which were identified based on phylogenetic placement with outgroup sequences. Despite these filtering losses, we were able to recover a median of thousands of Vibrionaceae sequences per day.

2. Dynamics within the Vibrionaceae community, compared among water-column size fractions

Using Vibrionaceae-specific *hsp*60 amplicon libraries, we analyzed Vibrionaceae community composition across a 93-day marine coastal time-series (July 23 – October 23, 2010). One unfractionated and four sequentially filtered sample sets allowed us to assess specificity of communities to marine habitat partitions of defined size range: unfractionated samples represented the whole water column, 0.2μm-filtered samples represented free-living cells, 1μm-filtered samples represented larger free-living cells and small organic particles, 5μm-filtered samples represented intermediate particles, and 63μm represented large particles. Particles could be of animal or algal origin, whether detrital material or live microorganisms, i.e., zooplankton and phytoplankton in the largest size fraction. In this section, we discuss dynamics within the Vibrionaceae community only; below, we situate these dynamics in the context of the full bacterial community, which was previously quantified with 16S rRNA sequencing by (Platero *et al.*).

In the unfractionated samples, i.e., lacking spatial resolution, the whole Vibrionaceae community appeared relatively stable, and was dominated by a few high-abundance populations (Fig. 1, Fig. 2, Fig. 3, Supp. Fig. 2). The major populations were *V. cyclitrophicus, V. splendidus,* and *Vibrio* sp. F10 (average abundances of 25, 22, and 18%), with *Aliivibrio fischeri* and *V. tasmaniensis* making up much of the remainder (Fig. 1, Fig. 2). Stability was evidenced by a median between-day Bray-Curtis similarity of 0.83, where 1 would denote identical communities (Fig. 3). However, three populations demonstrated more sporadic dynamics, appearing midway through the time-series: *V. kanaloae,* which remained present at a low abundance thereafter, and *V. pacinii* and *V. cincinnatiensis,* which both appeared over the same approximately 2-week period (ordinal days 244-260), each reaching up to 20% of the community. Neither *V. pacinii* nor *V. cincinnatiensis* were discovered at high frequency in past isolate collections from this system, which could be explained by the sporadic dynamics shown here.

Of the typically dominant populations in the unfractionated samples, *V. splendidus* is known to be an abundant ecological generalist based on past collections of isolates from this system, so its abundance was unsurprising; *V. cyclitrophicus* and *V. tasmaniensis* have also been discovered in significant amounts (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Szabo *et al.*). More surprising was the abundance of *Vibrio* sp. F10, which was previously inferred to be a zooplankton specialist, and *A. fischeri.* These

tended to be low-abundance in past strain collections; *Vibrio* sp. F10 also appeared at low abundances in the size fractions (Fig. 2).

Before discussing the fractionated samples in more detail, we note that the unfractionated sample set cannot be considered a linear sum or average of all of the size fractions, since 1) there may be differential contributions of the fractions to the bulk water column, and 2) the 63μm samples integrate over much more water volume than the other fractions (they were derived from a 100L sample, vs. 1L for each of the other size fractions).

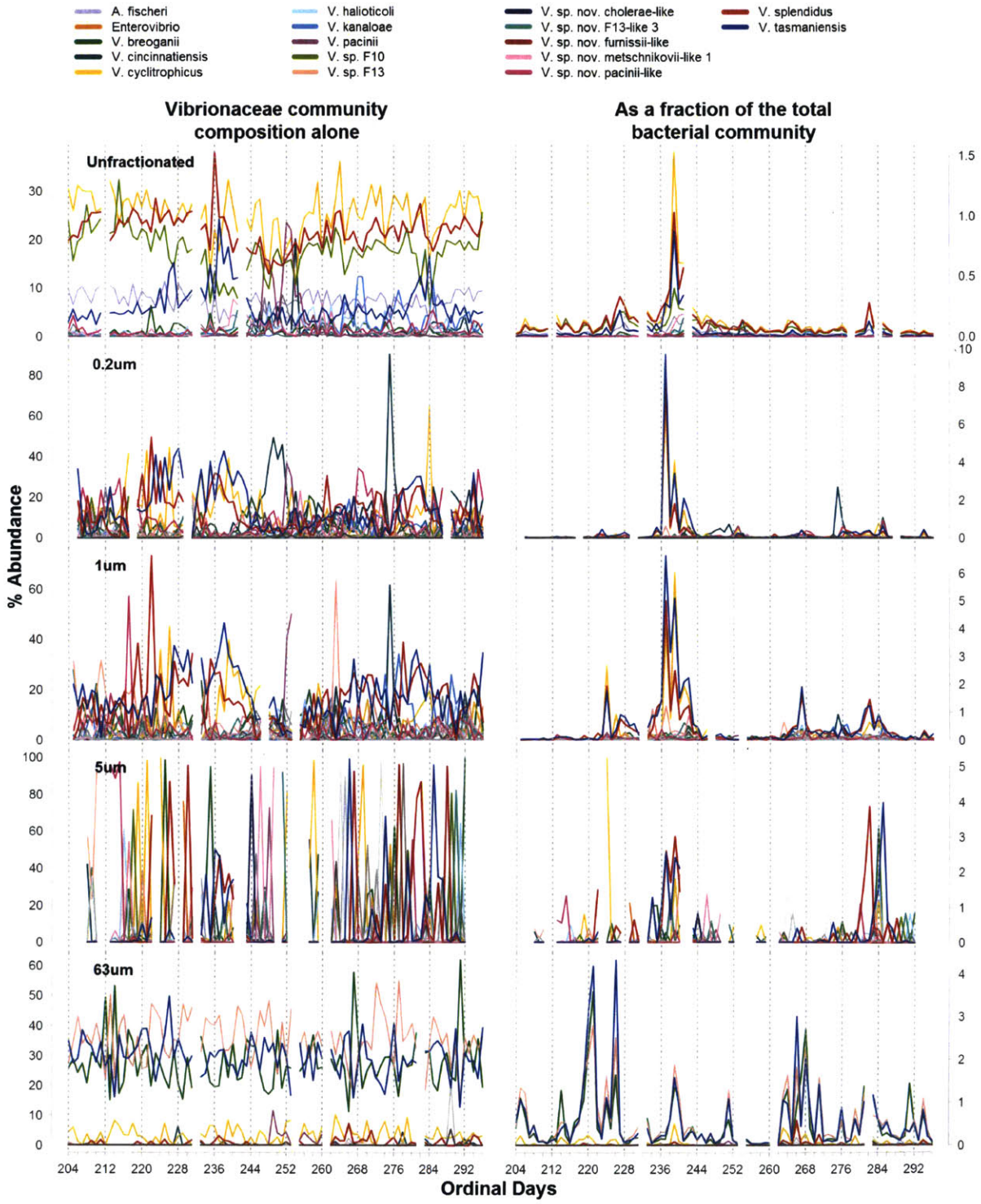


Figure 1: Vibrionaceae communities within different marine habitat partitions demonstrate differential compositions and dynamics across a 93-day time-series. Community

composition was quantified through Vibrionaceae-specific amplification of the hsp60 marker gene from daily DNA samples, and population-level taxonomy was assigned via phylogenetic placement. Left column shows dynamics within the Vibrionaceae community alone; right column shows these dynamics normalized to the fraction of Vibrionaceae within the total bacterial community (as quantified by 16S rRNA gene sequencing). Unfractionated samples represent the whole water column, 0.2μm-filtered represents free-living cells, 1μm represents larger cells and small organic particles, 5μm represents organic particles, and 63μm represents large particles. Blanks indicate samples that were missing, or filtered out due to low recovery of Vibrionaceae sequences. For clarity, only shown are populations that reached a minimum abundance of 5% in the unfractionated and 63μm samples, or 10% in the other samples.

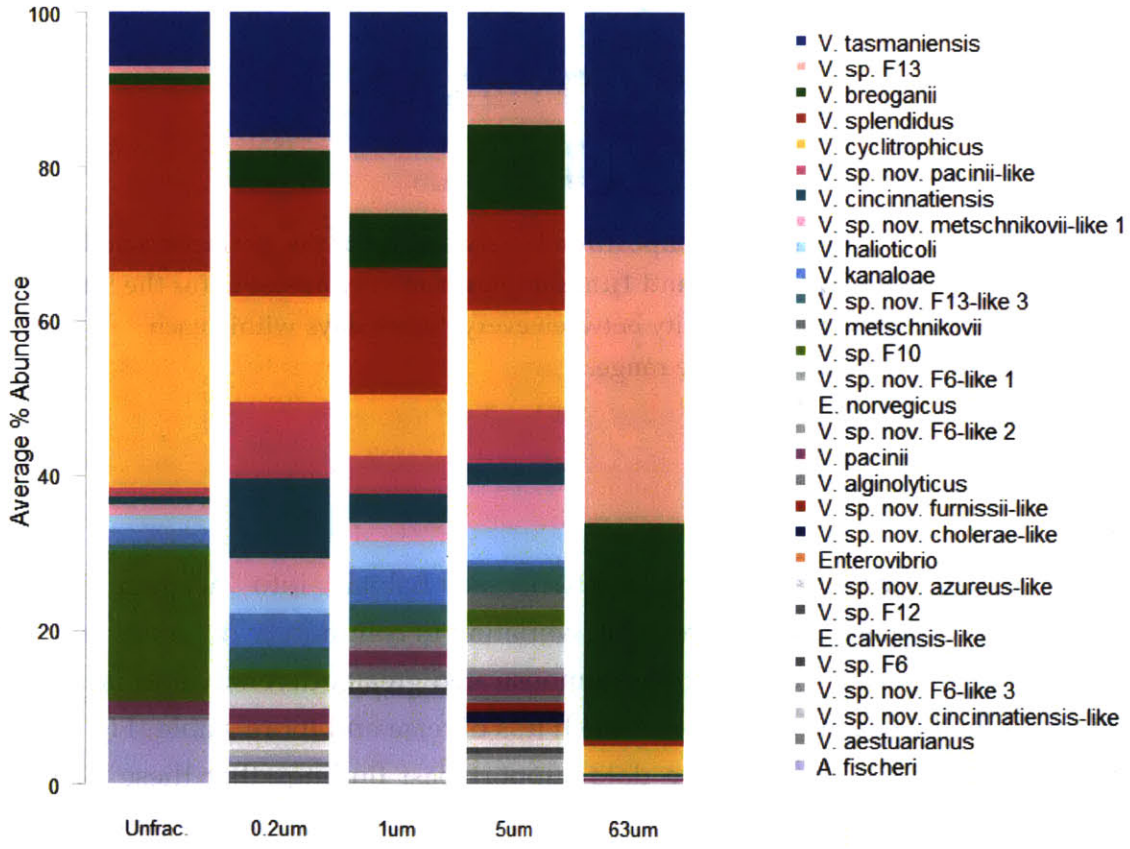


Figure 2: Average Vibrionaceae community composition is similar for the 0.2μm, 1μm, and 5μm size fraction, while the 63μm size fraction has a distinctive, low-diversity composition. The unfractionated sample set represents the whole water column. Community compositions were averaged across all days. Population colors are the same as in Fig. 1.

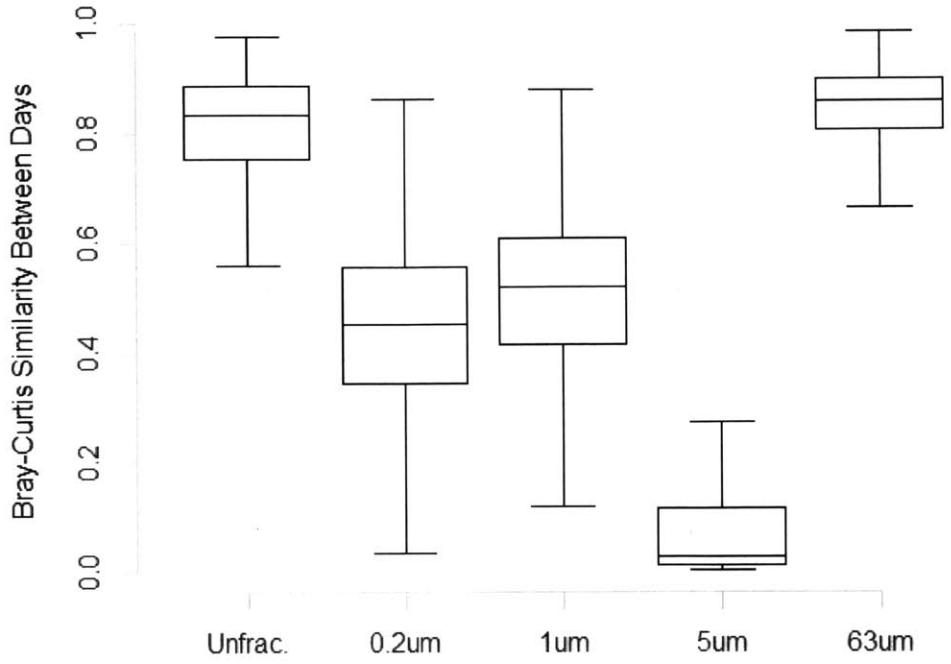


Figure 3: Daily Vibrionaceae community composition is very stable for the unfractionated and 63μm samples, less stable for the 0.2m and 1μm samples, and very unstable for the 5μm samples. Distributions of Bray-Curtis similarity between every pair of days within each sample set; whiskers indicate 1.5 interquartile range.

Taking spatial structure of the Vibrionaceae habitat into account by characterizing the fractionated samples revealed variation in community diversity and stability. The 0.2μm, 1μm, and 5 μm fractions demonstrated greater diversity and rapid daily turnover, while the 63μm fraction was much less diverse and more stable (Fig. 1, Fig. 2, Fig. 3, Supp. Fig. 2). Hence, these data support past findings that these size fractions represent distinct ecological landscapes for Vibrionaceae, which vary in distribution from specialist to generalist (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Szabo *et al.*).

Among the smaller size fractions, the 0.2μm and 1μm dynamics appeared very similar, suggesting that these fractions represent equivalent habitat partitions, i.e., the free-living planktonic community. Daily turnover was high, with median between-day Bray-Curtis similarities of 0.42 and 0.52 (Fig. 3) and numerous short-lived abundance

peaks visible. *V. tasmaniensis, V. splendidus,* and *V. cyclitrophicus* were again major constituents on all days (with respective average abundances across the two fractions of 16%, 15%, and 14%). Lesser but still consistent constituents included *Vibrio sp. nov. pacinii*-like, *V. breoganii* and *Vibrio* sp. F13 (7%, 5%, and 5%). The later appearance of *V. kanaloae* and the *V. pacinii* and *V. cincinnatiensis* blooms noted in the Unfractionated samples were evident here as well, with an additional one-day *V. cincinnatiensis* peak of 76% abundance (averaged across the two fractions) occurring a few weeks later. Overall, the dynamism of the smaller size fractions suggests a landscape of patchy and rapidly shifting ecological opportunities, as different dissolved nutrient sources became available to free-living microbes within the water column, and as the water mass at the sample site itself turned over physically, bringing with it different resource opportunities.

The 5μm fraction, representing Vibrionaceae associated with particles of moderate size, demonstrates even more striking turnover than the two smallest size fractions: although its average composition was very similar to theirs (Fig. 2), the community turned over almost entirely from day to day (median between-day Bray-Curtis similarity of 0.03; Fig. 3) and nearly a third of the days were almost entirely dominated by one taxon. Nonetheless, a few dynamics noted in the smaller size fractions were still reflected here: overlapping *V. splendidus, V. tasmaniensis,* and *V. cyclitrophicus* peaks around days 234-244, and again from day 266 onward; and a *V. pacinii* peak around day 248. In order to test whether the extremely uneven community composition noted in this fraction might be an artifact of uneven sequence coverage, we compared Shannon diversity to sequencing depth for each day, but found no consistent trend: days with low sequence coverage (<500 Vibrionaceae sequences) did not have notably lower Shannon diversity, and days with very high sequencing coverage (>10,000 Vibrionaceae sequences) could have near-zero Shannon diversity (Supp. Fig. 3). These community dynamics could suggest that on many days, the 5μm particle fraction is characterized by distinct patches that are each nearly exclusive to one Vibrionaceae population. Such patchiness could be generated by rapid temporal turnover (material in the 5μm particle fraction being generated and degraded on a rapid basis), or by spatial heterogeneity in 5μm particle composition within water masses. A further question is to what extent the incidence of near-exclusive dominance by single Vibrionaceae populations is stochastic or deterministic. Past isolate collections from this size fraction, which were each collected over one to two days, did show more even diversity (Hunt *et al.*, 2008; Preheim,

Boucher, *et al.*, 2011; Szabo *et al.*). This may reflect the fraction of days (approximately one-third) wherein diversity was more similar to that in the other size fractions (Supp. Fig. 2).

Finally, the 63μm fraction was heavily dominated by three populations: *Vibrio* sp. F13, *V. breoganii*, and *V. tasmaniensis* (35%, 30%, and 28% on average), with minor representation by *V. cyclitrophicus* and *V. splendidus*. Although less is known about the ecology of *V. tasmaniensis*, previously described ecology helps justify the presence of the rest of these populations: *Vibrio* sp. F13 is a generalist that can exploit both animal and algal resources; *V. breoganii* is a specialist for attachment to algal detritus; and *V. cyclitrophicus* and *V. splendidus* are known to encompass both particle-attached subpopulations and free-living subpopulations (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Preheim, Timberlake, *et al.*, 2011; Szabo *et al.*, 2012; Yawata *et al.*, 2014). Altogether, these populations appear to demonstrate a competitive advantage over other Vibrionaceae for exploitation of particulate animal and/or algal matter. In this fraction, we note again the presence of *V. pacinii* peaks around day 248.

Overall, Vibrionaceae populations showed a wide range of patterns of abundance and temporal variability across habitat partitions in this system; Fig. 4 compares these two characteristics for each population and sample set, with variability quantified as dynamic range, i.e., maximum abundance over minimum abundance. Temporally, populations could be consistently abundant (*V. splendidus, V. tasmaniensis, V. cyclitrophicus, V. sp.* F13), consistently rare, or typically rare yet capable of peaking dramatically within the Vibrionaceae (*V. pacinii, V. cincinnatiensis*). Spatially, while some populations were strongly represented across all size fractions, reflecting more generalized ecological associations, as with *V. tasmaniensis*, others were far more fraction-specific, like the algal specialist *V. breoganii* in the large size fraction.

Finally, taking phylogeny of the populations into account, patterns of dynamic behavior and spatial association did not assort by genetic relatedness, revealing the ecological diversity that can arise among close relatives sharing a habitat (Fig. 4). For example, among the very closely related *V. splendidus, V. tasmaniensis,* and *V. kanaloae,* while *V. splendidus* and *V. tasmaniensis* are both abundant and spatially generalized, *V. tasmaniensis* appeared more competitive for large particle attachment during this time-series, and *V. kanaloae* was far rarer than both. Such differentiation is expected to arise as a product of speciation processes, when differential environmental association evolves and leads to a separation of gene pools (Shapiro *et al.*, 2012). Such processes can

lead to accumulation of population-specific adaptations, and likely require trade-offs in adaptive traits, as has been observed for recently speciated populations that displayed different propensity to attach firmly to particles or rapidly disperse (Yawata *et al.*, 2014). Over time, as subpopulations become increasingly adapted and specific to different habitats, gene flow is depressed between them and speciation ensues (Shapiro *et al.*, 2012; Shapiro and Polz, 2014).

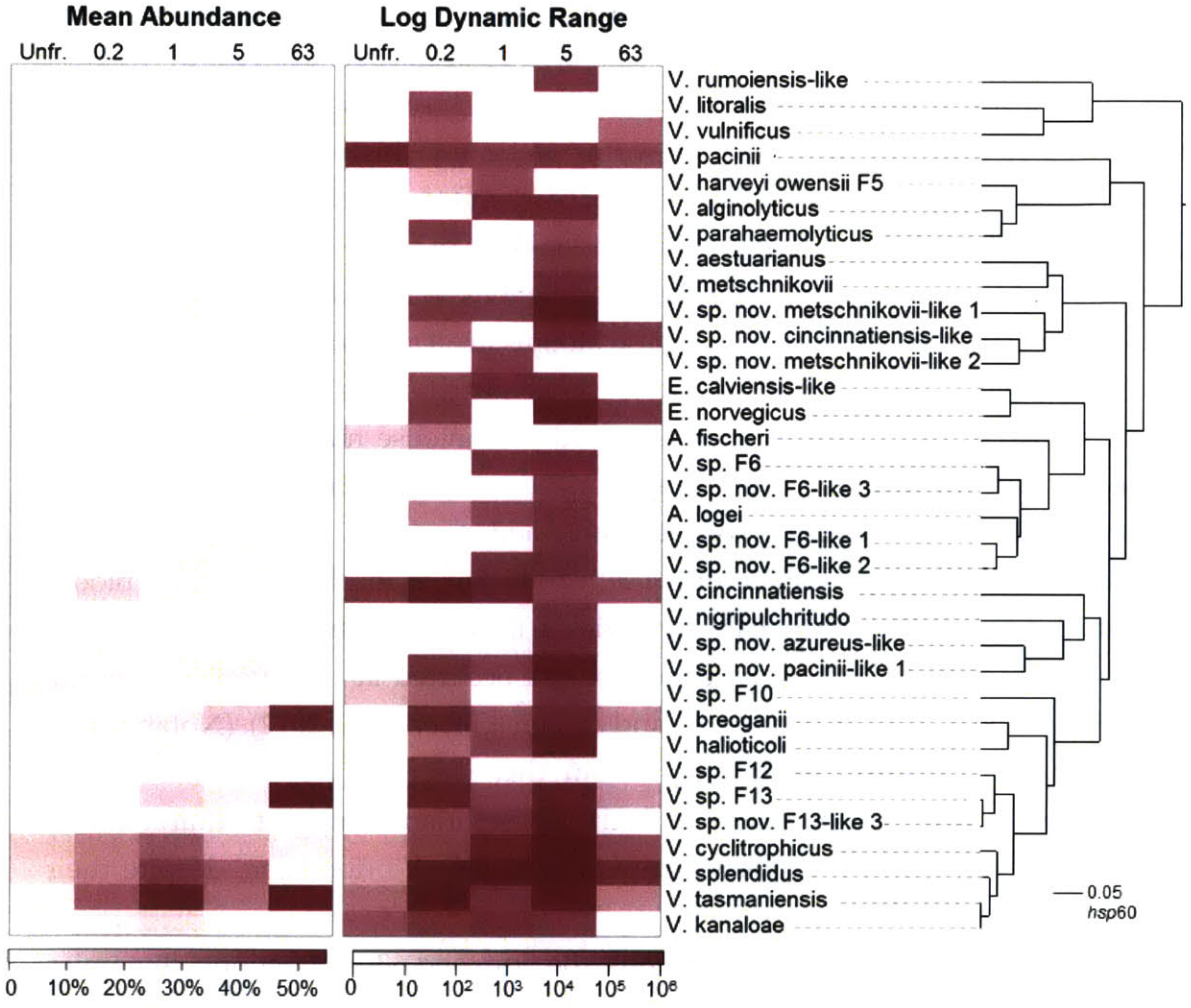


Figure 4: Vibrionaceae populations' mean abundances and dynamic ranges do not assort with phylogenetic distance. For each population, mean abundance across all days per sample (Unfractionated, 0.2μm, 1μm, 5μm, 63μm), and log dynamic range (maximum abundance / minimum abundance) across all days per sample. Phylogenetic tree represents distance among populations' most recent common ancestors in the *hsp*60 amplicon reference tree. (Because the reference tree was not ultrametric, not all "*Vibrio* sp. nov. …-like" populations cluster here with their namesakes.)

3. Dynamics of the Vibrionaceae within the total bacterial community

Using previously obtained 16S rRNA diversity data (Platero *et al.*), we were able to normalize our Vibrionaceae population dynamics data to the fractional abundance occupied by total Vibrionaceae within the larger bacterial community (Fig. 1). Within this time-series, the most striking feature of Vibrionaceae activity was a dramatic bloom around days 235-245, which was particularly pronounced in the free-living fraction: total Vibrionaceae abundance peaked at 28% in the 0.2μm fraction and 19% in the 1μm fraction, in contrast to their typical representation of ~1% within each fraction. This bloom was associated with a period of warm water and high macroalgal detritus availability, conditions which favor the Vibrionaceae, and terminated by the passage of Hurricane Earl on day 247, which caused a sudden drop in water temperature (Platero *et al.*). We found that the already abundant populations *V. splendidus, V. tasmaniensis,* and *V. cyclitrophicus* strongly dominated these blooms in all size fractions but the 63μm (Fig. 1). This suggests that in addition to their generalized spatial distributions, these populations possess the ability to respond rapidly to intense nutrient pulses. While Vibrionaceae have historically been characterized as predominantly carrying out attached lifestyles, the free-living bloom seen here adds to several past observations of intense Vibrionaceae blooms (Takemura *et al.*, 2014). For example, the largest bloom observed during the course of a 6-year time-series in the English Channel was of a single *Vibrio* taxon (based on 16S rRNA sequencing), and occurred at the same time as the highest observed peak in organic nitrogen and carbon (Gilbert *et al.*, 2012). (Notably, this bloom occurred in August, as did the bloom that we observed, reaffirming the importance of warm temperatures.) Such observations suggest that certain environmental regimes allow explosive growth of certain Vibrionaceae, despite their typically low abundance.

By contrast to the smaller size fractions, the 5μm and 63μm fractions' dynamics were not solely dominated by the day 235-245 bloom; each of these two fractions had several other population expansions of similar or greater intensity, emphasizing that the different fractions encompass communities that respond to different ecological stimuli. While the 5μm fraction's blooms were again dominated by *V. splendidus, V. tasmaniensis,* and/or *V. cyclitrophicus* (not always in concert), the algal specialist *V. breoganii* also contributed one peak (day 284), perhaps due to a proliferation of small algal detrital

particles on that day. Meanwhile, even during the day 235-245 bloom, the 63μm fraction remained stably dominated by its usual constituents, *V. tasmaniensis, V. breoganii,* and *V. cyclitrophicus,* suggesting that the composition of the large-particle-associated community is robust to environmental perturbations experienced by the smaller fractions.

Finally, we note that of the Vibrionaceae taxa described above as typically rare but occasionally abundant, *V. cincinnatiensis* also attained some abundance within the whole bacterial community, e.g., reaching nearly 3% of the whole community on day 275 in the 0.2μm fraction despite negligible abundance on nearly every other day. Hence, such peaks of rare Vibrionaceae can have significance within the larger community as well, even if minor compared to the dynamics of the major, bloom-prone Vibrionaceae populations.

4. Associations of Vibrionaceae populations with abiotic environmental parameters

We next investigated environmental parameters that could be driving the observed Vibrionaceae population dynamics. We used Granger causality to identify abiotic parameters with significant predictive power for Vibrionaceae dynamics, testing applied time lags of 1-7 days. Populations for which we found variables with significant Granger causality ($p < 0.01$) are plotted along with their dynamics in Fig. 5A; connected parameters and populations are represented as networks in Fig. 5B. Generally, we found that chlorophyll, macroalgae, and wave height were the most frequent environmental parameters with significant Granger causality, especially for the major populations, *V. splendidus, V. tasmaniensis, V. cyclitrophicus,* and *Vibrio* sp. F13. This supports the proposition that these populations are associated with availability of algal resources; wave height is likely correlated with the amount of resuspension of benthic algal detritus.

As an exception to this trend, however, we found no significant Granger-causing parameters for the major populations in the 63μm fraction. That they should not be associated with macroalgal presence suggests that the free-living and particle-attached Vibrionaceae use algal resources in distinct ways: the free-living populations thrive on exudates from large, live macroalgae (macroalgal presence was qualitatively scored based on the appearance of large amounts of material at the sampling site), while the attached populations may prefer smaller, detrital material. This resource distinction corresponds with the observation that 63μm fraction's dynamics were not dominated

by the day 235-245 bloom (Fig. 1). More generally, our failure to find Granger-causing variables for the major 63μm populations might indicate that for this particular fraction, the timespan of this dataset was insufficient to sample environmentally predictable trend. Additionally, other controls might be more significant to their dynamics: unmeasured abiotic variables which might control zooplankton or animal-derived particles; and/or biological interactions. Their dynamics could also depend on a combination of factors such that, for example, one peak of abundance might follow an algal-related event, while another might follow a zooplankton-related event.

We then assessed the extent to which different populations responded to the same environmental parameters by examining network representations of the Granger causality results (Fig. 5B). Overall, few populations had unique environmental parameters, showing that the Vibrionaceae broadly respond to the same abiotic factors. As notable exceptions, we found that the sparse bloomers *V. pacinii* and *V. cinnatiensis* are Granger-caused by silicate, suggesting a potential relationship with diatom presence; several diatom taxa were indeed later predicted to be potential interactors (see below). Overall, the Vibrionaceae-environmental networks are much more interconnected than the biotic interactions investigated below, supporting the notion that bulk environmental parameters are generally predictive of broad community dynamics, while biotic interactions are much more taxon-specific (Takemura *et al.*, 2014).

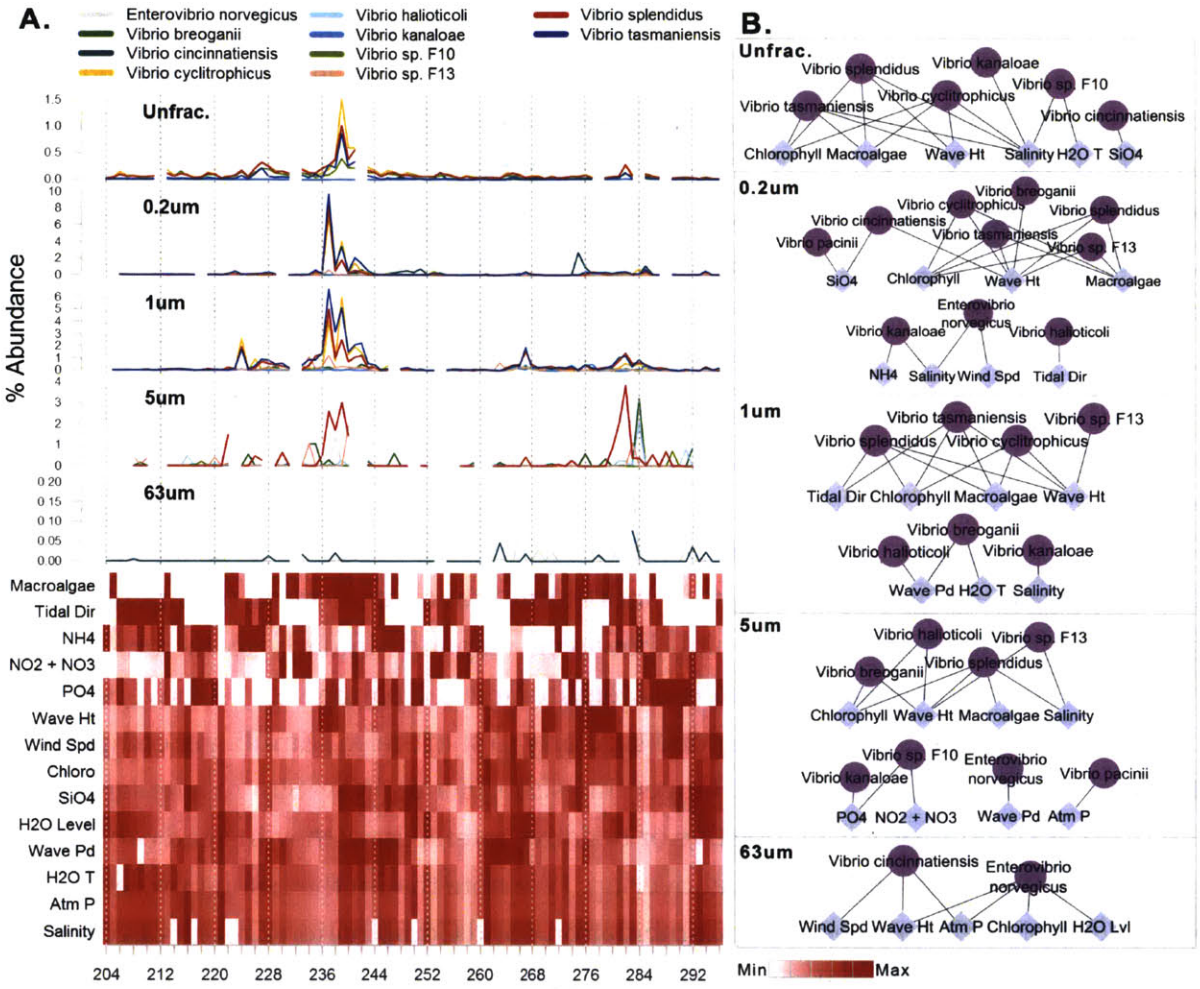


Figure 5: Identification of abiotic environmental parameters with significant Granger causality for Vibrionaceae populations reveals that many populations respond to the same parameters, especially chlorophyll, macroalgae, and wave height. A. Dynamics are shown for those populations for which we found significant Granger-causing parameters ($p < 0.01$). Heatmap represents each environmental parameter normalized to its maximum value, and hierarchically clustered by similarity of dynamics. Parameters are macroalgae, tidal direction, ammonium (NH_4), nitrites and nitrates ($NO_2 + NO_3$), phosphates (PO_4), wave height, wind speed, chlorophyll concentration, silicate (SiO_4), water level, wave period, water temperature, atmospheric pressure, and salinity. White bars represent missing data for wave period, water temperature, atmospheric pressure, and salinity; all other white bars represent minimal data values. **B.** Populations and Granger-causing parameters represented as networks: a line indicates that a given parameter (grey diamond node) Granger-causes a given population (purple circular node).

5. Associations of Vibrionaceae populations with bacterial and eukaryotic taxa

We next investigated potential interactions of Vibrionaceae populations with other bacteria and with eukaryotes, based on previously sequenced 16S and 18S rRNA data (Platero *et al.*). We used two approaches to identify potential biotic interactors of the Vibrionaceae among the bacterial and eukaryotic taxa: Granger causality, as applied above to the environmental metadata; and WaveSim, a new approach based on wavelet decomposition, which unlike Granger causality is robust to non-stationary data (Platero *et al.*). WaveSim identifies pairs of potential interactors by decomposing dynamic data into several spectra at different frequencies (low to high), and then searching for all taxon pairs with 1) anticorrelated dynamics at some characteristic low frequency and correlated dynamics at some characteristic high frequency, or 2) correlated low-frequency dynamics and anticorrelated high-frequency dynamics. Combination 1 is expected to represent taxa with negative short-term biotic interactions (competition, predation) whose longer-term dynamics, however, respond similarly to factors that change on a slower timescale (mutualisms, environmental parameters). Conversely, Combination 2 represents taxa that respond similarly to short-term stimuli, yet compete or otherwise negatively interact over a longer timescale. Hence, WaveSim results cannot be described as straightforwardly positive or negative interactions; they instead capture taxa likely to have significant ecological overlap. (We note that in the results presented below, we do not distinguish between Combination 1 and Combination 2 results.) When previously applied to bacterial and eukaryotic data from this time-series, WaveSim, followed by Markov clustering (WaveClust), was able to discern distinct communities of densely interacting organisms that responded cohesively to physicochemical drivers, showing the method's utility for identifying organisms with coordinated behavior (Platero *et al.*). We note further that WaveSim and Granger causality are complementary in that pairwise WaveSim will not capture organisms with dynamics that are perfectly in phase, whereas Granger causality will.

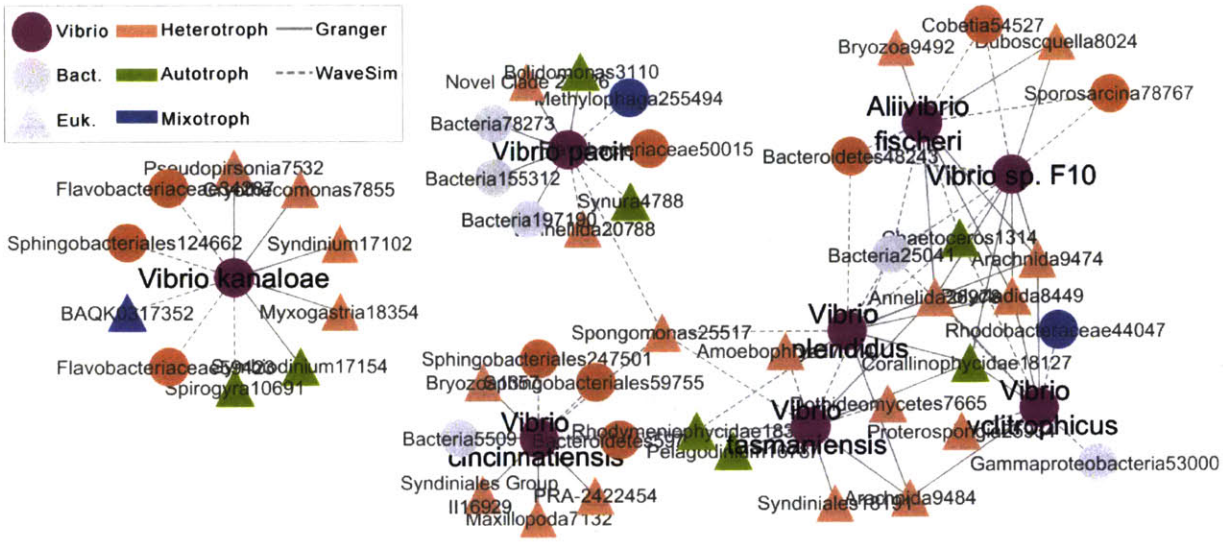


Figure 6: Unfractionated Vibrionaceae populations, top 5 predicted bacterial and eukaryotic connections for Granger and WaveSim. Bacterial and eukaryotic taxa are named with their highest-resolution possible taxonomic label, followed by a unique numerical OTU identifier.

After completing the Granger and WaveSim analyses, we investigated the extent to which different Vibrionaceae populations within each sample set shared predicted interactors by representing Granger and WaveSim results together as network structures (top 5 results per method, per population, showing only the 15 Vibrionaceae populations that were most abundant on average across all fractions; Figs. 6-10). We found that while the dominant populations per fraction consistently shared a subset of their interactors, reflecting likely habitat overlap, the remaining populations tended to have completely unique sets of interactors. Hence, these findings were consistent with observations of both the dynamics and relationships with environmental parameters: the distinct dynamics of the rarer populations corresponded to distinct sets of potential interactors.

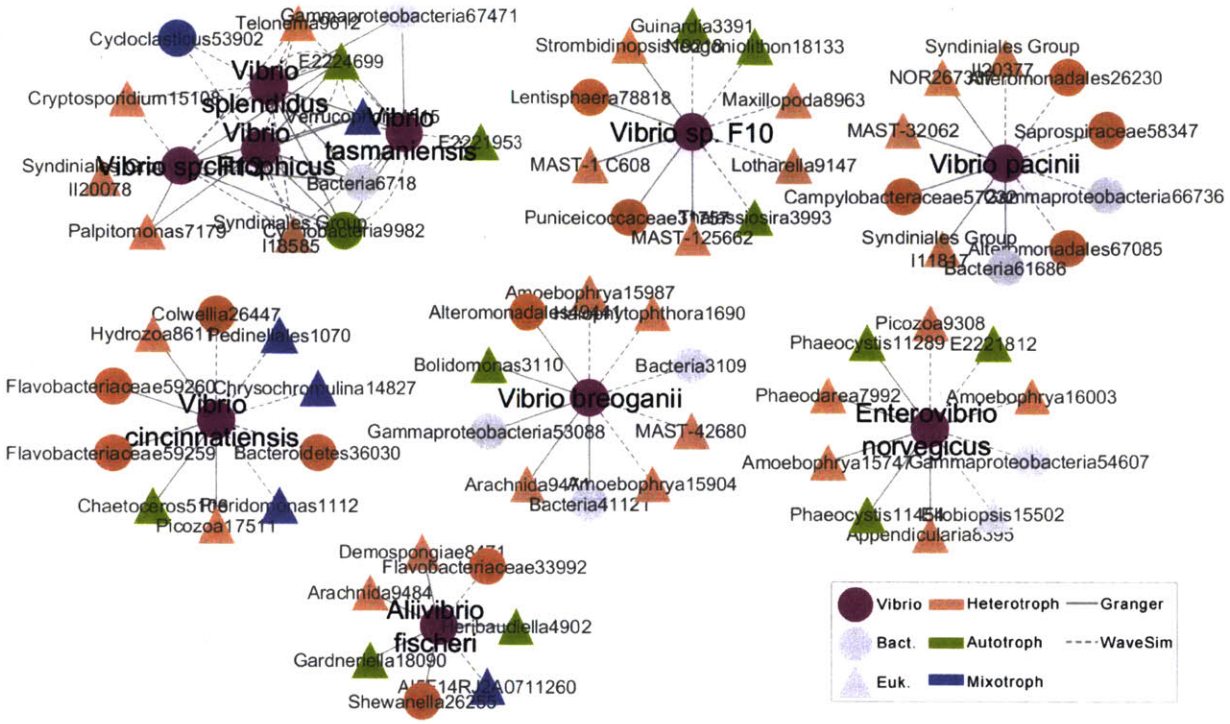


Figure 7: 0.2µm fraction Vibrionaceae populations, top 5 predicted bacterial and eukaryotic connections for Granger and WaveSim. Bacterial and eukaryotic taxa are named with their highest-resolution possible taxonomic label, followed by a unique numerical OTU identifier.

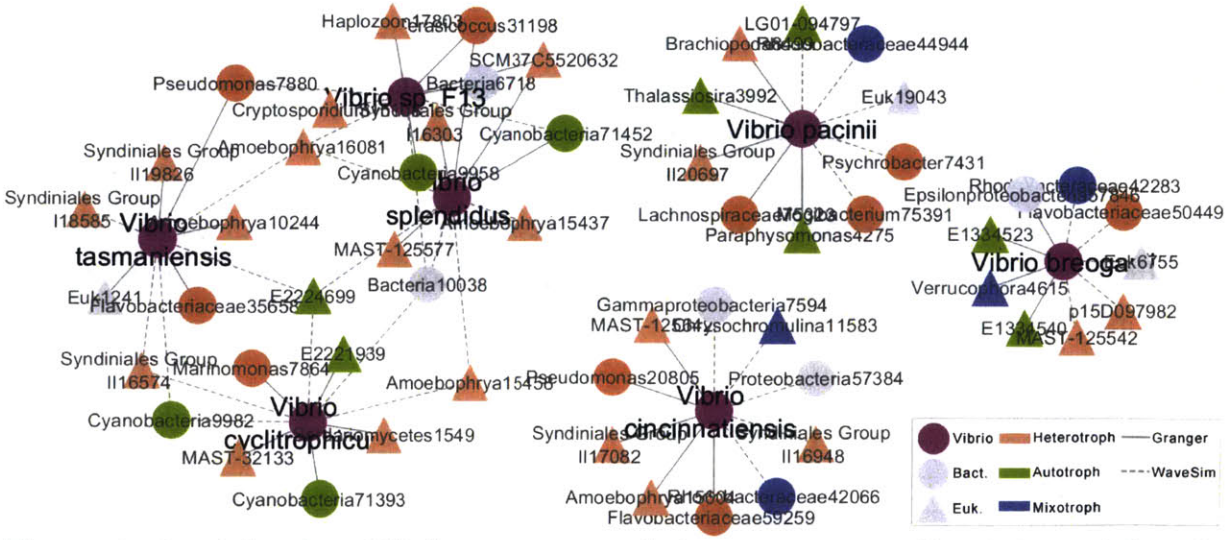


Figure 8: 1µm fraction Vibrionaceae populations, top 5 predicted bacterial and eukaryotic connections for Granger and WaveSim. Bacterial and eukaryotic taxa are named with their highest-resolution possible taxonomic label, followed by a unique numerical OTU identifier.

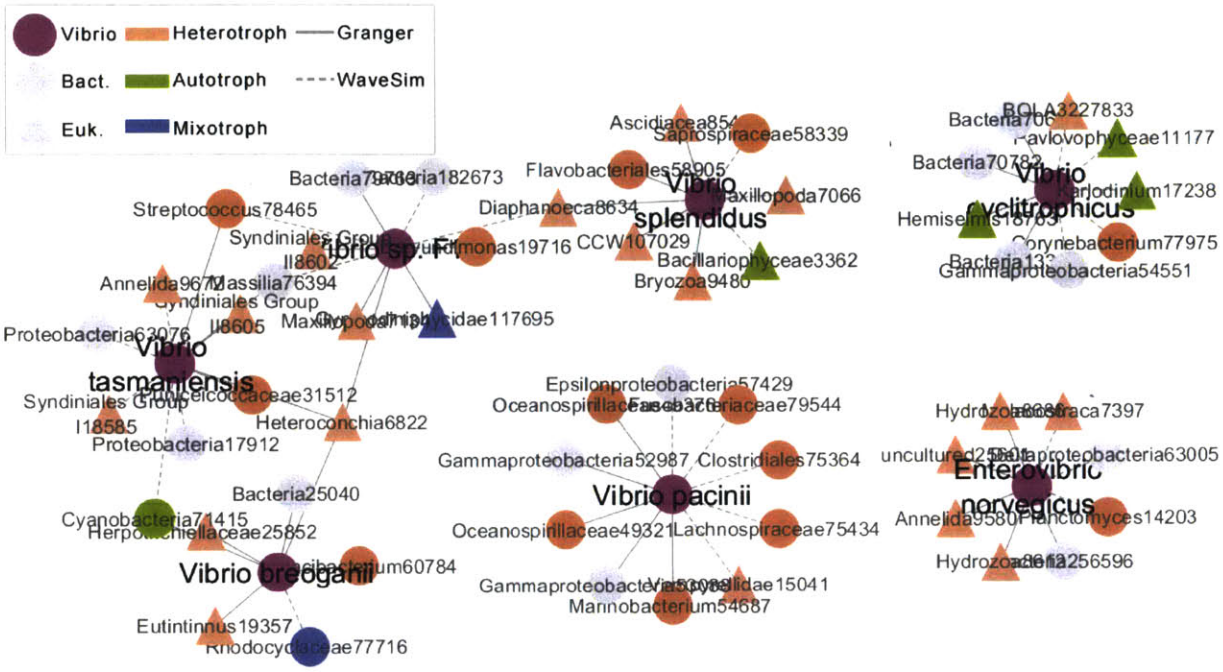


Figure 9: 5µm fraction Vibrionaceae populations, top 5 predicted bacterial and eukaryotic connections for Granger and WaveSim. Bacterial and eukaryotic taxa are named with their highest-resolution possible taxonomic label, followed by a unique numerical OTU identifier.

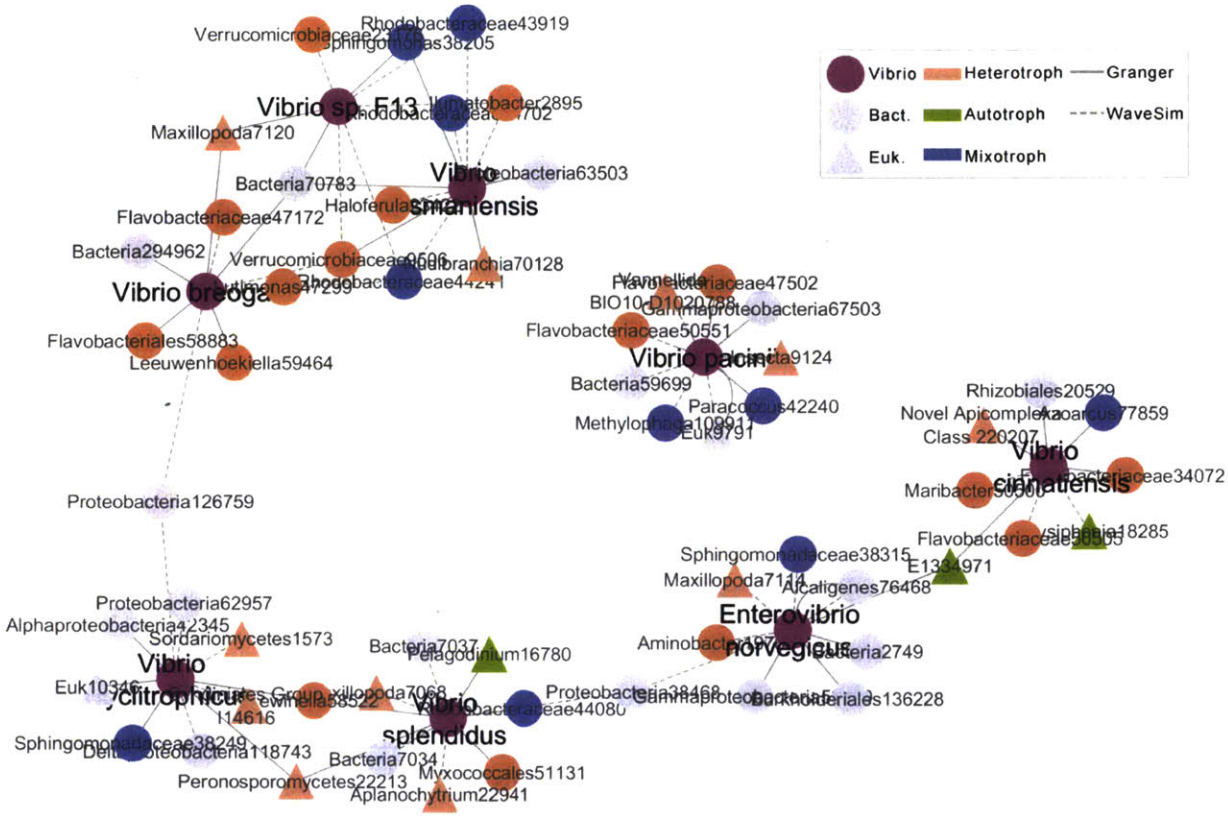


Figure 10: 63µm fraction Vibrionaceae populations, top 5 predicted bacterial and eukaryotic connections for Granger and WaveSim. Bacterial and eukaryotic taxa are named with their highest-resolution possible taxonomic label, followed by a unique numerical OTU identifier.

5.1. Functional characterizations of predicted eukaryotic and bacterial interactors

To investigate potential functional roles of the predicted interactors, we assigned metabolic annotations based on taxonomy, here examining the top 10 results per method per population. We found the most to least common metabolic categories were eukaryotic heterotrophs (34% averaged across all populations and samples), bacterial heterotrophs (22%), eukaryotic autotrophs (13%), bacterial mixotrophs (6.4%), eukaryotic mixotrophs (2.7%), and bacterial autotrophs (2.7%); bacteria and eukaryotes of unknown metabolism contributed the remaining 18% and 1%, respectively (Fig. 11A). (Findings were similar between methods at the phylum level; see Supp. Fig. 4A.) Our observation that interactions with heterotrophic eukaryotes predominate contrasts with findings in other systems of predominant bacteria-bacteria interactions (Gilbert *et al.*, 2012), but is consistent with past analyses within this system (Platero *et al.*).

Heterotrophic eukaryotes may represent mutualistic or predatory relationships, and hence predicted heterotroph-Vibrionaceae relationships will require further characterization to infer potential mechanism. By contrast, autotrophic interactors are easier to interpret as potential sources of algal exudates as nutrients.

We note a striking exception to the above ranking of metabolic categories: in the 63μm fraction, bacterial heterotrophs dominated over eukaryotic heterotrophs; bacteria of unknown metabolism were also more abundant than in the other fractions (Fig. 11A). This contributes another line of evidence that the large-particle-associated Vibrionaceae occupy a distinct ecological niche compared to the smaller size fractions. Here, the predominance of bacterial interactors suggests that relationships among microbes colonizing the same particles are a stronger structuring force for these Vibrionaceae than, e.g., eukaryotic grazing pressure. Several mechanisms could contribute to this trend: first, close proximity and high concentrations of microbes attached to the same particle would heighten the effect of local biotic interactions; second, biofilm formation can offer resistance to protozoan grazing, thus lowering its impact on community dynamics (Matz and Kjelleberg, 2005).

We further investigated the composition of the predicted interactors by quantifying their taxonomic composition at the phylum level, and found all samples except the 63μm to be broadly similar (Fig. 11B; Supp. Fig. 4B. for breakdown by method). Major phyla on average across all samples included eukaryotic Alveolata (i.e., dinoflagellates, which are largely mixotrophic; 25%), eukaryotic Stramenopiles (i.e., diatoms; 17%), Proteobacteria (12%), Bacteroidetes (11%), metazoa (8.8%), eukaryotic Rhizaria (heterotrophic amoeboids; 5.2%), Verrucomicrobia (2.7%), and Cyanobacteria (2.5%). Like the Vibrionaceae, many of the Proteobacteria (i.e., Alphaproteobacteria including Rhodobacterales, and other Gammaproteobacteria) and Bacteroidetes (i.e., Flavobacteriia) are also generalist heterotrophs often characterized as copiotrophs, so their inferred association is logical (Fuhrman *et al.*, 2015). Similarly, members of the Verrucomicrobia have been characterized as significantly contributing to aquatic degradation of polysaccharides (Martinez-Garcia *et al.*, 2012; Cardman *et al.*, 2014). By examining phylum composition here, we were also able to discern that the 63μm fraction's predominance of heterotrophic bacterial interactors was accounted for by greater relative abundance of Proteobacteria, Bacteroidetes, Verrucomicrobia, and Planctomycetes, and correspondingly decreased Alveolata and Stramenopiles (Fig. 11B).

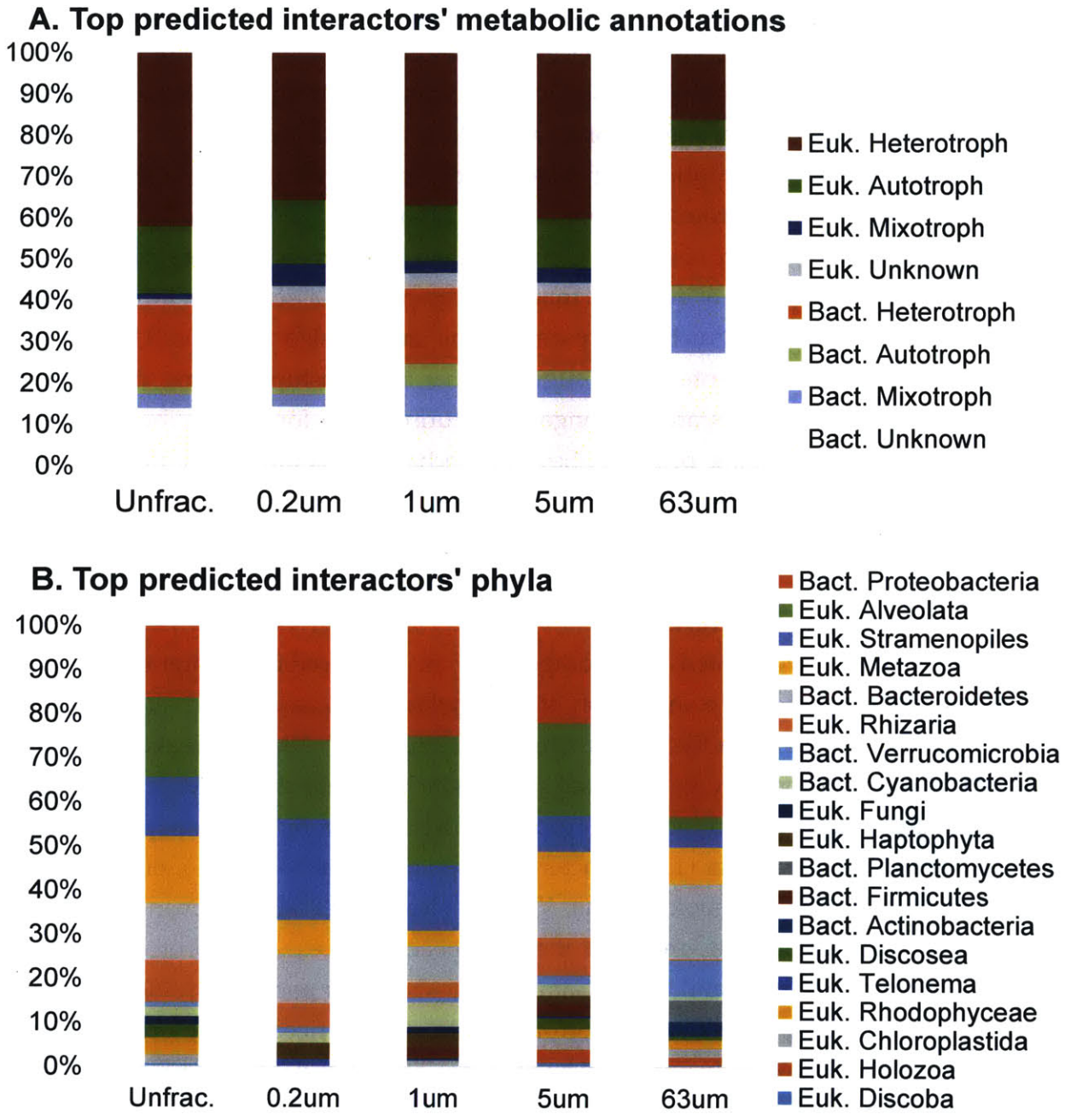


Figure 11: Predicted interacting taxa for Vibrionaceae populations are dominated by heterotrophic eukaryotes in all sample sets but the 63µm, which is dominated by Proteobacteria. For each sample set, we summed the A. metabolic annotations or B. phyla for the top 10 Granger and top 10 WaveSim results for every major population, then averaged across the populations.

Despite broad commonalities in composition of interactors at the phylum level, examination of the Vibrionaceae-interactor network structure ultimately emphasizes that taxon-level affiliations within these phyla tend to be extremely specific to each Vibrionaceae population (Figs. 6-10). We also examined to what extent a given population's connections might be specific to each size fraction by comparing *V. tasmaniensis'* top 10 Granger-predicted interactors across all five of the sample sets. We found that all 50 predicted interactors were unique at the OTU level, and only two taxa defined below the phylum level occurred across multiple fractions: Flavobacteriaceae in both the 1μm and 5μm fractions, and parasitic dinoflagellates of the order Syndiniales in all samples but the 63μm, which are believed to have generalized distribution across many marine hosts, including fish, protozoans, algae, and other dinoflagellates (Guillou *et al.*, 2008). Various Syndiniales representatives were noted as interactors for many of the other populations as well, and may be indicative of generalized association with animals or algae, or detritus derived from them.

5.2. Examples of predicted interactors' dynamics

As examples of a fairly simply interpretable dynamic relationship, we briefly discuss here predicted interactors whose dynamics were highly in-phase with those of the Vibrionaceae. We were able to discover such interactors by examining Granger results; such taxa were often Gammaproteobacteria relatives like *Arcobacter, Colwellia, Shewanella,* and *Pseudoalteromonas.* (We note that since Granger causality does not include a metric for effect size, a more straightforward way to identify such closely overlapping taxa would be to use a correlation metric, e.g., Spearman.) One *Pseudoalteromonas* OTU was particularly striking in its similarity, which suggests that its members occupy nearly identical habitat space throughout the fractions to the Vibrionaceae, especially the major populations, e.g., *V. splendidus* and *V. tasmaniensis* as shown in Fig. 12. Distinction of other potential ecological relationships among the Vibrionaceae and their predicted interactors will require both visual examination and further method developments, e.g., determination of directionality of time-lagged dynamics to determine which taxa precede and which succeed Vibrionaceae peaks, and distinction among the low-frequency/high-frequency correlated/anti-correlated combinations for the wavelet results in order to infer the timescales of potential negative and positive interactions.

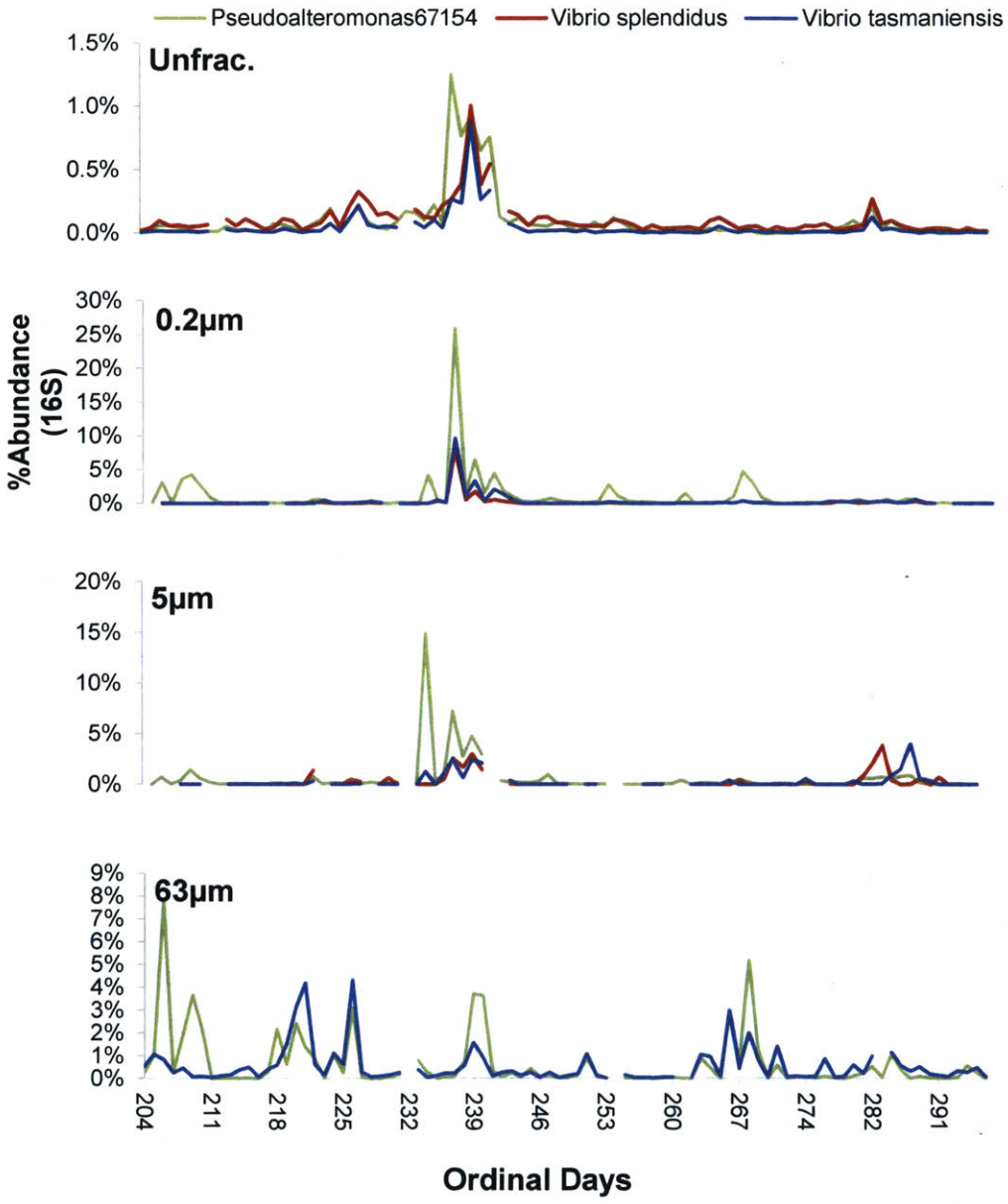


Figure 12: One Pseudoalteromonas taxon identified via Granger causality has strikingly similar dynamics to Vibrionaceae populations throughout all of the fractions, suggesting significant ecological overlap.

Altogether, quantitative predictions of interacting taxa reflected ecological overlap of the Vibrionaceae with numerous other heterotrophic taxa, both eukaryotic

and bacterial, with many of those expected to also have generalized ecological capabilities or distributions. Results also emphasized the greater population-level specificity of biological correlates compared to abiotic factors. The above findings furthermore reinforce the view that the 63μm fraction is ecologically distinct from the other fractions: here, local interactions among microbes co-colonizing large particulate resources are expected to have a much stronger structuring effect on community dynamics than other biotic interactions.

Chapter 3 Conclusions

By achieving population-level resolution of cohabiting and closely related yet ecologically distinct Vibrionaceae via high-resolution tag sequencing and phylogenetic placement, we were able to mine a multifaceted time-series dataset for insight into both ecological dynamics and spatial partitioning among fine-scale genetic units in a dynamic context. We showed that population-level turnover is remarkably high: samples could look almost completely different when separated by only one or two days. Hence, sampling with lower genetic or temporal resolution can drastically underrepresent community turnover. While Vibrionaceae abundance is often assumed to be consistently low in marine communities, with our high temporal resolution, we were also able to show that several of their populations are capable of dramatic yet transient blooms in the free-living fraction, stimulated by warm weather and sudden input of macroalgal material. By exploring temporally associated taxa, we were also able to identify a *Pseudoalteromonas* with nearly identical behavior, showing that multiple taxa are capable of such dramatic growth; that is, low-abundance taxa cannot always be assumed to be low-abundance, or negligible contributors to metabolic activity and nutrient cycling. Indeed, one recent study across multiple ecosystems suggested that such "conditionally rare taxa" contribute disproportionately to community turnover (Shade *et al.*, 2014).

Multiple lines of evidence also pointed to the extent to which the water column, far from being a homogeneous medium, offers spatially and ecologically distinct habitats, among which the Vibrionaceae populations manifested diverse patterns of abundance over time. We especially observed clear ecological distinctions in dynamics and inferred ecology between the smaller size fractions and the large particle fraction: while the 5μm registered some "spillover" of events in the free-living fractions, i.e. the major bloom, and was compositionally more similar to them, no such signals were

observed in the large size fraction, which appears to be governed by a different regime of abiotic (scantily predictive, for this timespan) and biotic (far more predictive power) structuring forces or selective pressures. Hence, environmental microstructure gives rise to a heterogeneous adaptive landscape, which gives rise to opportunities for differentiation.

An interesting future direction of investigation that this dataset offers would be to strive for even greater genetic resolution, and track strain-level distributions and behaviors using same *hsp*60 amplicon data. For example, is there further genetic substructure among representatives of the same population that are present among multiple actions – can we detect temporally consistent assortment of different genotypes by fraction? Conversely, can we detect "transitions" among the size fractions, wherein a given genotype relocates from one fraction to another, suggesting a spatial and temporal shift in, for example, the distribution of the most desirable nutrient sources? A further, classic question of microbial ecology to explore would be, are observed blooms accounted for by clonal expansion, or by coincident growth of multiple lineages? Such investigations will require the overcoming of technical impediments such as the high abundance of likely chimeric sequences within our sequencing data, yet should offer an exceptionally high-resolution look into the fine-scale dynamics of environmental microbes.

Chapter 3 Methods

1. *Environmental sampling and data*
2. *16S and 18S rRNA amplicon libraries*
3. *Vibrionaceae hsp60 amplicon libraries*
4. *Detecting associations with environmental parameters, eukaryotes, and bacteria*

1. Environmental sampling and data

1.1. Environmental sampling

Sampling procedures are described in detail in (Kauffman, 2014). Briefly, to obtain a time-series of marine microbial communities, we collected triplicate whole-seawater samples (4L) from Canoe Cove in Nahant, MA for 93 consecutive days (July 23 – October 23, 2010; ordinal days 204-296). Prior to Day 259, triplicates were collected at an undefined location. From Day 259 forward, triplicates were collected at 3 defined stations within Canoe Cove. The first station was distinguished by being close to a rocky outcropping covered in macroalgae. Samples were collected by wading into the littoral zone to a depth of several feet; absolute sampling location on each day varied depending on tidal height.

1.2. Unfractionated seawater DNA samples

Triplicate 4L samples of unfractionated seawater were collected in screw-top polypropylene bottles and filtered by peristaltic pump through 0.2μm Sterivex filters. Genomic DNA for PCR analysis was recovered from the filters as detailed in (Platero *et al.*), and summarized as follows. Cells were lysed with 750 μL of Cell Lysis Solution (Qiagen, USA), and bead-beating was performed at 5000 rpm for 60 seconds in a Mini Beadbeater-1 (BioSpec Products, Inc., Bartlesville, OK). Samples were incubated at 80°C for 5 minutes, then cooled to room temperature. RNA was digested by adding 4μL of RNAse A (4mg/mL), mixing by inversion, and incubation at 37°C for 30 minutes. DNA extraction was performed by adding 250 μl of Protein Precipitation Solution (Qiagen, USA), vortexing for 20 seconds, incubating on ice for 5 minutes, and centrifugation at 13,000 rpm for 5 minutes. Supernatant was transferred to a new tube and centrifuged

again to remove any remaining sediment. To precipitate DNA, 750µl of isopropanol was added to 750µl of supernatant, and mixed by inversion. To improve DNA extraction yield, the DNA was left precipitating overnight at -20°C. Finally, DNA was recovered by centrifugation at 13,000 rpm for 5 minutes, washed with 700µl of 70% ethanol, dried, and resuspended in 100 µl of DNA Hydration Solution (Qiagen, USA).

1.3. Fractionated seawater DNA samples: 63µm, 5µm, 1µm, 0.2µm

63µm samples were collected independently of the other fractions by pouring approximately 100L of water through a plankton net (63µm Nitex Turtox Tow Net, Wild Co., 426-A38), resulting in a 30mL concentrate of the retained particles, collected into an Oakridge tube. Concentrated samples were homogenized with tissue grinders (47732-450, VWR). For DNA, 10mL of homogenized concentrate was filtered through a 0.2µm filter (polycarbonate or Sterivex), and DNA was extracted as described for the unfractionated samples.

The 5µm, 1µm, and 0.2µm samples were obtained by filtration through successive cup and frit filtration towers. 1L of water was first collected by pre-filtration through the 63µm plankton net. The water was then transferred and gravity-filtered through a 5µm filter, the filtrate was transferred and gravity-filtered through a 1µm, and that filtrate was transferred and pumped through a 0.2µm Sterivex filter. Finally, DNA was extracted from the different filters as described for the unfractionated samples.

1.4: Environmental data

Environmental parameters were measured or collected from external sources as detailed in Table 1.

Table 1: Abiotic environmental parameters

Parameter(s)	Measure-ment site	Methods
Air temperature Water temperature	On-site	Averaged measurements from digital and analog thermometers
Salinity	On-site	Stored water samples, measured with a Reichert digital refractometer calibrated to a standard salt solution
Ammonium Phosphate	On-site	Water samples were collected in 250mL polypropylene bottles, filtered through ashed GF/F filters, collected into 20mL scintillation vials and

Nitrite+Nitrate Silicate		stored frozen. Nutrients were analyzed by the Nutrient Facility at Woods Hole Oceanographic Institution. Correction factors were applied for all cases where volume loss occurred during freezing; correction factors for silicate were assumed to be linear with change in salinity as previously shown for phosphate and nitrate (Karam, 2012).
Macroalgal abundance	On-site	Ranked based on qualitative daily field notes regarding macroalgae presence in the water. Rankings were assigned with these rules: (0) Macroalgae not mentioned in qualitative notes or explicit statement that there is no algae, "very little" algae, or "little to no" algae; (1) Mention of a "little" algae or wrack; (2) Mention of algae or algal filaments or wrack without quantifier, or described as "some" algae; (3) Mention of "lots" of algae or multiple rafts of algae or algal filaments; (4) Mention of superlatively much algae. For cases where highly abundant macroalgae is noted as being washed up by wave action but not explicitly mentioned in the notes as being present in the water column, photographs were consulted to assess presence in the water column. A downgrade from "4" to "3" was applied for a case of a superlatively abundant small planktonic algae because the contribution to biomass was not comparable with other days ranked as "4".
Atmospheric pressure Chlorophyll concentration Dominant wave period Wave height Wind speed	NOAA Station 44013, 16km east of Boston, MA	Downloaded from "Ocean Data" (chlorophyll) and "Standard Meteorological Data" (all others), via http://www.ndbc.noaa.gov/station_history.php?station=44013. Times were converted from UTC to EST; all no-data values were removed. 24-hr averages were calculated for each time point by averaging values for parameter at a given time point and all time points in the previous 24 hours. Time points for days with >10 missing hourly values in the preceding 24hr window were discarded. The time point used for each day was 10:50 AM EST for all datatypes except chlorophyll concentration, for which 11:00 AM EST was used.
Tidal direction Water level	NOAA Boston Harbor station 8443970	Verified six-minute water level data for each month of the time series were downloaded from: http://opendap.co-ops.nos.noaa.gov/axis/webservices/waterlevelverifiedsixmin/index.jsp. A new vector for tidal direction was generated assigning each time point as either "Incoming" (1) or "Outgoing" (0), depending on whether the water level was higher or lower than the previous 6-minute time point, respectively. The time point for each day differed and was dependent on arrival and departure time from the sampling site each day. Briefly, if both an arrival and departure time were recorded then the time point used was the half-way point between these two times, if no departure was recorded then the time point used was calculated as the arrival time plus 30 minutes, half of the average sampling time.

2. 16S and 18S rRNA amplicon libraries

2.1: 16S and 18S rRNA amplicon libraries: construction and sequencing

To quantify bacterial and eukaryotic diversity, respectively, 16S and 18S rRNA amplicon paired-end Illumina libraries were created from each daily genomic DNA sample, as detailed in (Platero *et al.*) and summarized as follows. We used a two-step PCR approach according to (Preheim *et al.*, 2013). First, we amplified with primer pairs targeting the ribosomal gene and incorporating partial Illumina adapters: PE16S_V4_U515_F / PE16S_V4_E786_R for 16S rRNA, V4 variable region; and Il_18S_V9_1391F / Il_18S_V9_EukB for 18S rRNA, V9 variable region (Table 2). Second, we used overlapping primers to incorporate the full Illumina adapter and a sample-specific 9bp barcode for library identification.

To normalize template concentrations and avoid over-cycling, we used real-time PCR to quantify the threshold cycle of each template using the Step 1 PCR primers. This real-time PCR was carried out in a final volume of 25 µl containing 5 µl of 5x HF buffer, 250 µM of dNTPs, 0.3 µM of each primer (PE16S_V4_U515_F / PE16S_V4_E786_R or Il_18S_V9_1391F / Il_18S_V9_EukB), 0.5x SYBR Green I nucleic acid stain (Invitrogen™), 2.5 U of Phusion® High-Fidelity DNA Polymerase (New England BioLabs Inc.), and 20 ng of template DNA. The amplification program consisted of an initial denaturing step at 98°C for 3 min; an amplification step of 45 cycles of 30 s at 98°C, 30 s at 52°C, and 30 s at 72°C; and a final extension of 5 min at 72°C. According to the threshold cycle, we normalized all template concentrations to the most dilute sample for subsequent library construction. Step 1 PCR was done with normalized templates in quadruplicates in the same conditions as the previous real-time PCR, excluding the SYBR Green and using the number of cycles determined in the real-time PCR.

To clean up Step 1 PCR products, the quadruplicates were pooled and purified by magnetic beads (AgenCourt® AMPure® XP, Beckman Coulter). 100 µl of PCR product was mixed with 85.5 µl of beads, incubated 13 min for DNA binding, incubated 15 min on a magnet (SPRIplate® 96-Ring), and then washed with 100 µl of ethanol 70%. After drying the ethanol, the DNA was eluted with 40 µl of EB buffer (Qiagen, USA), with 7 min incubation for elution and 15 min on magnet to separate DNA from the beads.

The sample-specific barcodes were included in the Step 2 PCR. This PCR was carried out in a final volume of 25 ul containing 5 µl of 5x HF buffer, 250 µM of dNTPs,

0.4 μM of the primers PE-III-PCR-F and PE-III-PCR-01-096, 2.5 U of Phusion® High-Fidelity DNA Polymerase (New England BioLabs Inc.), and 4 ul of the previous purified PCR as template. The amplification program consisted of an initial denaturing step at 98°C for 2 min; an amplification step of 9 cycles of 30 s at 98°C, 9 s at 70°C, and 30 s at 72°C; and a final extension of 2 min at 72°C. This PCR was also done in quadruplicate, pooled after amplification, and purified by AgenCourt® AMPure® XP magnetic beads as described above.

Ribosomal gene libraries were multiplexed in groups of 96 samples. To this end, multiplexing ratios were estimated by real-time PCR with Illumina sequencing primers. This PCR was carried out in a final volume of 25 ul containing 12.5 μl of 2x QuantiTec® SYBR® Green PCR kit mastermix (QIAGEN), 0.2 μM of the primers PE-seq-F / PE-seq-R (Table 2), and 5 μl of each libraries as template. The amplification program consisted of an initial denaturing step at 95 ºC for 15 min; an amplification step of 45 cycles of 10 s at 95°C, 20 s at 60°C, and 30 s at 72°C, and a final extension of 5 min at 72°C. Multiplexed libraries were submitted for Illumina sequencing at the Biomicro Center (MIT, Cambridge, MA).

Ribosomal gene libraries were sequenced with paired-end reads of either 100-bp on the Illumina HiSeq or Illumina GAIIx, or 150-bp on the Illumina MiSeq, by the BioMicro Center (MIT, Cambridge, MA). 150-bp reads were subsequently trimmed to 100 bp to make analyzed sequences uniform in length.

Table 2: Primers used for 16S rRNA, 18S rRNA, and hsp60 gene amplification

Primer purpose	Name	Sequence (5'-3')	Source / Notes
Step 1, 16S rRNA amplification: V4 variable region	PE16S_V4_U515_F	ACACG ACGCT CTTCC GATCT YRYRG TGCCA GCMGC CGCGG TAA	U515F (Caporaso *et al.*, 2011) *
	PE16S_V4_E786_R	CGGCA TTCCT GCTGA ACCGC TCTTC CGATC TGGAC TACHV GGGTW TCTAA T	E786R (Caporaso *et al.*, 2011)
Step 1, 18S rRNA amplification: V9 variable region	Il_18S_V9_1391F	ACACG ACGCT CTTCC GATCT YRYRG TACAC ACCGC CCGTC	1391F (Stoeck *et al.*, 2010) *
	Il_18S_V9_EukB	CGGCA TTCCT GCTGA ACCGC TCTTC CGATC TTGAT CCTTC TGCAG GTTCA CCTAC	EukB (Stoeck *et al.*, 2010)
Step 1, Vibrionaceae *hsp60* amplification	PE-V-hsp60-105-F	ACACGA CGCTC TTCCG ATCTY RYRAC CCDAT GGATC TKAAG CGYGG	*
	PE-V-hsp60-364-R	CGGCA TTCCT GCTGA ACCGC TCTTC CGATC TCCRC GRTCG AACTG CATVC CTT	
Step 2, Illumina adapter and barcode incorporation	PE-III-PCR-F	AATGA TACGG CGACC ACCGA GATCT ACACT CTTTC CCTAC ACGAC GCTCT TCCGA TCT	
	PEIII-PCR-001-096	CAAGC AGAAG ACGGC ATACG AGATN NNNNN NNNCG GTCTC GGCAT TCCTG CTGAA CCGCT CTTCC GATCT	N's stand for 96 sample-specific barcodes.
Illumina sequencing primers	PE-seq-F	ACACT CTTTC CCTAC ACGAC GCTCT TCCGA TCT	
	PE-seq-R	CGGTC TCGGC ATTCC TGCTG AACCG CTCTT CCGAT CT	

*Plus partial Illumina adapter at 5', and YRYR as a complexity region to aid cluster distinction during Illumina sequencing. YRYR was also swapped with TGGTC to allow multiplexing of 2x96 samples.

2.2: 16S and 18S rRNA amplicon libraries: sequence processing and OTU calling

Raw 16S Illumina FASTQ files were quality filtered using QIIME 1.3 (Caporaso *et al.*, 2010), filtering with a PHRED score threshold of 23 and subsequently retaining only 99-bp sequences. Primers were then trimmed using mothur 1.31 (Schloss *et al.*, 2009), yielding 76bp forward and reverse 16S reads. For the remainder of the analysis, for 16S only the forward read was used, due to the lack of overlap between the paired-end reads. For 18S, paired-end reads were overlapped and trimmed to 120bp-long sequences.

OTUs were called using the Distribution-Based Clustering (DBC) method of (Preheim *et al.*, 2013), which takes into account both genetic distance and the distribution of sequences across samples. To create distribution-based clusters, sequences were first progressively clustered into 90% identity clusters with USEARCH (Edgar, 2010).

Bacterial taxonomy was assigned to the resulting OTUs using the RDP classifier algorithm (Wang *et al.*, 2007), through the QIIME 1.3 toolkit (Caporaso *et al.*, 2010), and the 12_10 greengenes 97% reference OTU collection (http://greengenes.secondgenome.com/downloads/database/12_10/gg_12_10_otus.tar.gz). A 0.80 confidence threshold was used to assign each possible taxonomic level to a given OTUs representative sequence. Triplicate 16S and 18S rRNA data were averaged together for subsequent following analyses.

3. *Vibrionaceae amplicon libraries*

3.1: Vibrionaceae hsp60 sequencing pipeline development: design of primers and reference library for phylogenetic placement

In order to target Vibrionaceae *hsp*60, a single-copy marker gene, an *hsp*60 primer pair was designed as follows. We downloaded all >100bp sequences matching the query "hsp60" from the European Nucleotide Archive database (ENA, http://www.ebi.ac.uk/; retrieved October 10, 2010), and aligned them using MUSCLE V.3.8.31 (Edgar, 2004), with default parameters. Next, only sequences with identifiers from Vibrionaceae genera were extracted from the alignment. Degenerate primers for Illumina sequencing were designed based on a consensus sequence generated from these aligned Vibrionaceae sequences (PE-V-hsp60-105-F / PE-V-hsp60-364R, Table 2). These primers yield a 259bp *hsp*60 amplicon (215bp with primers trimmed).

In order to be able to assign taxonomy to environmental Vibrionaceae *hsp*60 sequences, we developed a pipeline for taxonomic identification via phylogenetic placement using the pplacer v.1.1 package (Matsen *et al.*, 2010), as BLAST does not perform well with very closely related sequences. Phylogenetic placement works by inserting a query sequence into a fixed reference tree; the query receives the taxonomic label of the clade into which it is placed. First, we assembled a reference library of Vibrionaceae *hsp*60 sequences from the following sources: 1) sequences from the coastal Vibrionaceae strain collections of (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Preheim, Timberlake, *et al.*, 2011; Szabo *et al.*, 2012); 2) all matches from Vibrionaceae

genera for the query "hsp60," from the ENA (http://www.ebi.ac.uk/; retrieved May 2014); and 3) all Vibrionaceae type-strains were identified from Microbial Earth Project's listings (http://microbial-earth.org; retrieved December 2014), available genomes were downloaded using NCBI's Genome Browser (http://www.ncbi.nlm.nih.gov/genome/browse/; retrieved December 2014), and BLASTed to yield *hsp*60 hits. Furthermore, sequences from the closely related Gammaproteobacteria genera *Alteromonas, Aeromonas, Escherichia, Psychromonas,* and *Shewanella* were added as outgroups. After trimming to the 215bp primer-trimmed amplicon region and dereplicating, this yielded a reference library of 674 unique sequences.

Taxonomic labels for every sequence in the reference library were curated as follows, to determine whether any populations were intermixed based on this 215bp amplicon, to assess the taxonomic labeling of ENA sequences since not all derived from well-characterized sources, and because some sequences came from unnamed environmental populations. First, we aligned the sequences via MUSCLE (Edgar, 2004) and made a phylogenetic tree with FastTree (Price *et al.*, 2010), using *Alteromonas* as the root. Taxonomic labels for each reference sequence were then manually curated based on examination of taxonomic composition within sequence clusters in the resulting tree, taking into account the taxonomic labels of sequences subsumed by each representative. Monophyletic clusters with >95% *hsp*60 sequence identity were treated as the relevant unit for naming, based on the observation that 95% is the average identity within the ecological Vibrionaceae populations defined by (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Preheim, Timberlake, *et al.*, 2011; Szabo *et al.*)). Such clusters were named with the following rules: 1) A label was trusted and propagated to unnamed or low-abundance sequences with conflicting names in the same cluster, if the label a) had been confirmed by Multi-Locus Sequencing Analysis as described in (Preheim, Boucher, *et al.*, 2011), or b) originated from a type-strain genome. Low-abundance conflicting sequences, e.g., 1 *Vibrio* sp. F10 within a cluster of 61 *V.* sp. F13, could be accounted for by strain mislabeling or horizontal gene transfer. 2) >95% identity clusters that combined multiple very closely related taxa in significant amounts were given combined taxonomic designations, i.e., *V. harveyi owensii* F5 and *V. anguillarum ordalii*; the case of *V. splendidus* and *V. tasmaniensis* is described in greater detail below. 3) >95% identity clusters containing only unnamed sequences were given names of the format "*Vibrio* sp. nov. ...-like," with "..." being the closest named taxon in the reference tree. Multiple

>95% identity clusters near the same named taxon received numbers, e.g., *Vibrio* sp. nov. *pacinii*-like 1 and 2.

Altogether, this resulted in a reference library of 674 sequences with 69 different taxonomic labels: 53 named and 16 "*Vibrio* sp. nov. ...-like." We note that overall, few naming ambiguities or conflicts occurred. The only populations that are abundantly represented in this environment that were difficult to distinguish with this 215bp amplicon were the very closely related *V. splendidus* and *V. tasmaniensis*, whose short branch lengths resulted in significant admixture within the tree. Given their significant abundance within this Vibrionaceae community, we attempted to differentiate their reference sequences as best as possible. After closely examining the reproducibility of branch topology for these two populations, we labeled those sequences by attempting to create the most reproducible separation possible of subclades that were predominantly *V. splendidus* or *V. tasmaniensis*.

Finally, the reference tree and taxonomic labels were converted into a reference package for use with the pplacer suite, using the taxtastic package (https://github.com/fhcrc/taxtastic). Some manual edits were made to the minimal subtaxonomy outputted by taxtastic in order to obtain desired behavior, e.g., having *V. splendidus* and *V. tasmaniensis* behave as "subspecies" of a "*V. splendidus tasmaniensis*" taxon in order to capture sequences that could not be differentiated.

3.2. Vibrionaceae hsp60 amplicon libraries: construction and sequencing

To quantify population-level Vibrionaceae diversity within the environmental time-series, Vibrionaceae-targeted *hsp*60 amplicon libraries were created from each daily genomic DNA sample. Vibrionaceae *hsp*60 amplicon libraries were constructed with a protocol similar to that described for the 16S and 18S rRNA amplicon libraries in section 2, but with modifications to improve amplification, because Vibrionaceae *hsp*60 was a low-abundance target in these environmental DNA samples.

First, we used Platinum™ Taq DNA Polymerase (Invitrogen™) for normalization and amplification, as the high-fidelity Phusion® Polymerase (New England BioLabs) used above failed to amplify, likely due to less robust target binding. Reactions for the real-time PCR normalization step contained the following, in a final reaction volume of 25 µl: 2.5 µl of 10x Platinum Taq buffer (Invitrogen™), 200 µM of dNTPs, 0.4 µM of each primer (PE-V-hsp60-105-F / PE-V-hsp60-364-R, Table 2), 0.5x SYBR Green I nucleic acid

stain (Invitrogen™), 4 U of Platinum™ Taq DNA Polymerase (Invitrogen™), and 13.4 ng of template DNA. The amplification program consisted of an initial denaturing step at 94°C for 1 min 30 s, an amplification step of 45 cycles of 30 s at 94°C, 30 s at 57°C, and 25 s at 72°C, and a final extension of 2 min at 72°C. Step 1 PCR was done with normalized templates in quadruplicate in the same conditions as the previous qPCR, excluding the SYBR Green I nucleic acid stain reagent and using the number of cycles determined in the qPCR. Additionally, because the larger size fractions contained higher quantities of PCR inhibitors, 10 mg/mL BSA (New England BioLabs Inc.) was used as an enhancer for both the normalization and Step 1 reactions. Step 1 products were pooled and purified with AgenCourt® AMPure® XP magnetic beads as described above.

Step 2 PCR, incorporating full Illumina adapters and sample-specific barcodes, used the following reaction mixture, in a final volume of 25 µl: 5 µl of 5x HF buffer, 250 µM of dNTPs, 0.4 µM of the primers PE-III-PCR-F and PE-III-PCR-01-096, 2.5 U of Phusion® High-Fidelity DNA Polymerase (New England BioLabs Inc.), and 4 ul of the previous purified PCR as template. The amplification program consisted of an initial denaturing step at 98 ºC for 2 min followed by an amplification step of 9 cycles of 30 s at 98 ºC, 9 s at 70 ºC, and 30 s at 72 ºC, and a final extension of 2 min at 72 ºC. This PCR was also done in quadruplicates, pooled after amplification, and purified by AgenCourt® AMPure® XP magnetic beads as described above.

Finally, libraries were multiplexed in batches of 96 or 192 according to ratios estimated by qPCR with Illumina sequencing primers. This PCR was carried out in a final volume of 25 ul containing 12.5 µl of 2x QuantiTec® SYBR® Green PCR kit mastermix (QIAGEN), 0.2 µM of the primers PE-seq-F / PE-seq-R, and 5 µl of each library as template. The amplification program was identical to that described for the 16S and 18S rRNA libraries. Multiplexed libraries were sequenced with Illumina HiSeq2000 using 150bp paired-end reads (achieving overlapping coverage of the amplicon), by the Yale Center for Genomic Analysis (New Haven, CT).

3.3. Vibrionaceae hsp60 amplicon libraries: sequence processing

Raw Illumina paired-end *hsp*60 FASTQ reads were processed with the Caravan pipeline (https://github.com/swo/caravan). We trimmed primers with default settings (2 mismatches allowed per primer); merged forward and reverse reads, allowing for any mismatches but no size variation from 215bp; then quality-filtered with default settings

(discard reads with >2 expected errors). Finally, sequences were pooled across all days for all of the sample sets (unfractionated and fractions) and then dereplicated, yielding 100% OTUs. We then undertook a pipeline for assigning taxonomy, refining low-resolution taxonomy assignments, and several filtering steps for excluding irrelevant sequences.

First, because the *hsp*60 primers captured a large fraction of non-Vibrionaceae diversity, OTUs were filtered by BLASTing against a custom database containing all ENA *hsp*60 sequences retrieved as described in section 3.1. We discarded any OTU whose top BLAST hit was not a Vibrionaceae sequence.

To assign taxonomy to the retained Vibrionaceae *hsp*60 100% OTU representatives, we used pplacer v.1.1 (Matsen *et al.*, 2010) with the reference package described in section 3.1 and a likelihood cutoff for placements of 0.8. The pplacer algorithm attempts to find a placement location for a query sequence that maximizes the likelihood of the resulting tree; queries are placed at the lowest possible taxon level that fulfills the likelihood cutoff. Following placement, we discarded OTUs that were placed within the outgroup taxa *Alteromonas*, *Aeromonas*, *Escherichia*, *Psychromonas*, or *Shewanella*, as well as sequences placed only as the high-level taxa "Vibrionales" or "Vibrionaceae," since manual phylogenetic placement of these sequences (i.e., examination of a tree created with the sequences and the reference library) found that they were in fact outgroups to the Vibrionaceae.

Next, because we found that pplacer placed only as "*Vibrio*" OTUs that actually occurred within the Vibrionaceae yet were outgroups to named sequence clusters, we used the following protocol to group and assign names to these sequences. All "*Vibrio*" OTUs were made into a tree with the reference library sequences, and the topology of the resulting tree was compared to that of the reference library alone. "*Vibrio*" OTUs were then categorized according to their Most Recent Common Ancestor on the reference tree. For examples, OTUs that occurred as an outgroup to both *V. cholerae* and *V. mimicus*, yet still within their clade, were categorized as "*V. cholerae V. mimicus*." In this way, all OTUs could be categorized, even those without a named representative extant in the reference library. These OTUs then received names of the form "Vhsp60-", short for *Vibrio hsp*60, followed by a numeral. We chose this naming convention rather than the MRCA names because this reference tree is not an exhaustive phylogenetic reference, due to the short amplicon length used.

We also examined three other groups of less-informative sequence placements: sequences placed as the *V. splendidus tasmaniensis* group (as described in section 3.1), genus *Aliivibrio*, and genus *Enterovibrio*. We resolved *V. splendidus tasmaniensis* by adding it to *V. tasmaniensis*, on the basis that their dynamics across the time-series datasets looked more similar than *V. splendidus*' (Spearman correlations for *V. splendidus tasmaniensis* + *V. tasmaniensis* across the Unfractionated, 0.2μm, and 1 μm time-series: 0.43, 0.56, 0.58; for *V. splendidus tasmaniensis* + *V. splendidus*: 0.29, 0.24, 0.50.). However, we maintained *Aliivibrio* and *Enterovibrio* as separate calls since their dynamics did not appear similar to any of their respective member species; sequences called as these genera may represent other species from the genera.

Finally, in order to eliminate days that might be skewed by low sequencing coverage, we filtered out any days with fewer than 100 Vibrionaceae sequences recovered.

3.4. Vibrionaceae hsp60 sequencing pipeline development: validation with mock communities

In order to asses any biases in community composition created by amplification with the Vibrionaceae *hsp*60 primers, we created and analyzed a set of mock communities, in parallel with development of the sequencing pipeline for the actual time-series samples. The mock community contained genomic DNA from 87 Vibrionaceae strains collected from this environment by (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Szabo *et al.*, 2012), representing 23 different populations. To ensure that strains were equally represented in terms of *hsp*60 abundance, input DNA amounts were normalized across the strains based on amount of *hsp*60 template, as determined with real-time PCR using the conditions described in section 3.2. After creating the DNA cocktail, we diluted it to yield a template concentration similar to those of the environmental samples.

To test for biases created by increasing numbers of amplification cycles, we created five different *hsp*60 amplicon libraries from this mock community, each amplified to 21, 22, 23, 24, or 25 cycles, otherwise mimicking the entire amplification protocol described for *hsp*60 library construction above. This range of cycles had been determined to be within the linear amplification range for samples from the unfractionated time-series. The resulting libraries were sequenced with Illumina HiSeq2000 as described above, and output community composition analyzed with a

protocol similar to that described in section 3.3., except that 95% OTUs were called rather than 100%.

4. Detecting associations among Vibrionaceae populations, environmental parameters, bacteria, and eukaryotes.

We used two methods to identify potential abiotic environmental drivers and biological interactors of the Vibrionaceae: Granger causality, a standard technique in analysis of biological time-series; and WaveSim, an implementation of the wavelet decomposition method, which has also begun to be applied to biological time-series. (Platero *et al.*, in preparation). We applied Granger causality to both the abiotic and biotic datasets, but WaveSim only to the biotic datasets. Before applying these methods, we filtered out negligible taxa by including only the top 15 Vibrionaceae populations by average abundance across all of the fractions, and only bacterial and eukaryotic taxa that with a minimum abundance of 0.1% within their respective datasets. 16S rRNA-defined Vibrionaceae taxa were removed, and replaced instead with the *hsp60*-defined Vibrionaceae population data, normalized to the fraction occupied by total Vibrionaceae within the 16S rRNA data.

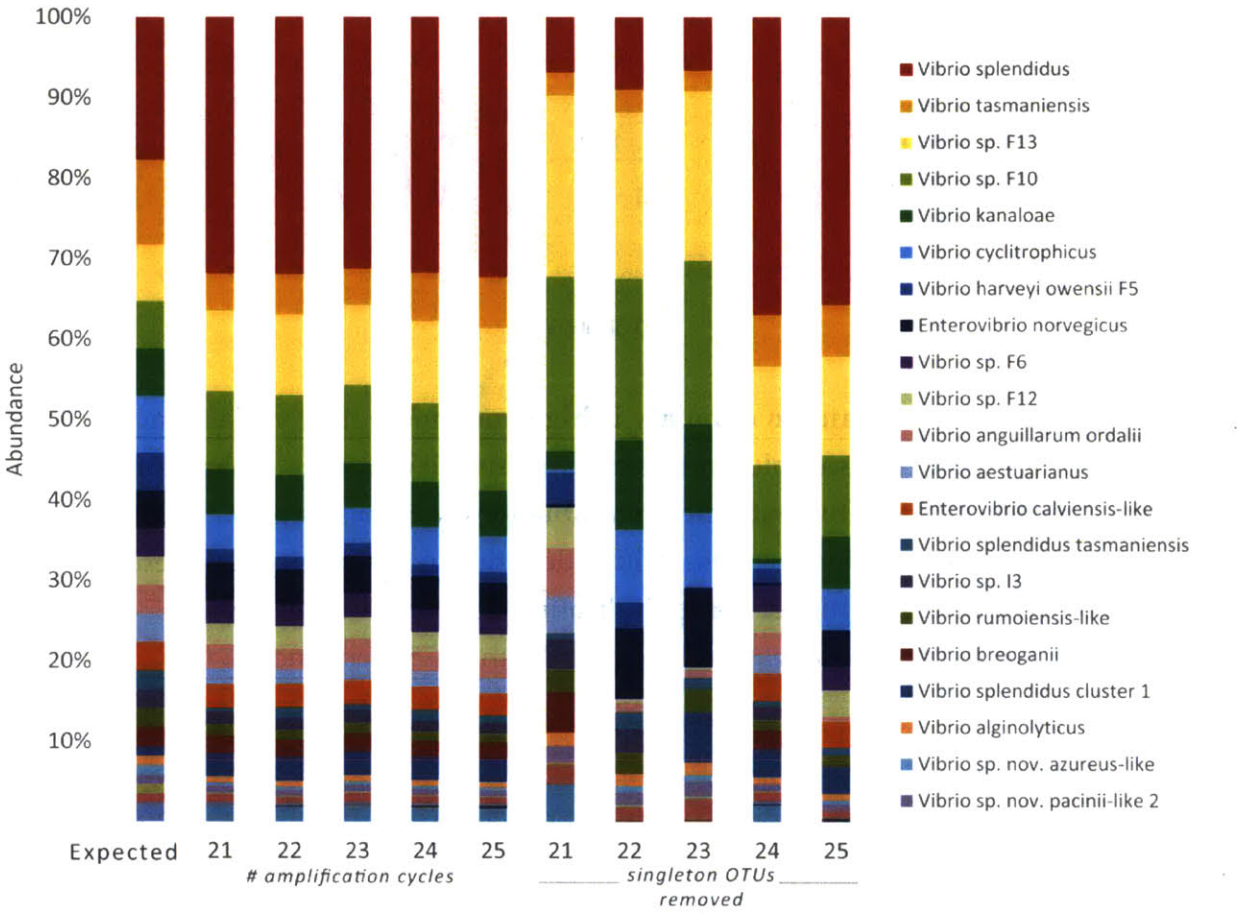
Granger causality determines whether one time-series provides statistically significant predictive information about future values of another time-series, when time lags are applied between the two; if so, the first time-series is considered to "Granger-cause" the second. Granger causality does not determine an effect size, only a p-value. Here, we tested whether environmental parameters Granger-caused the Vibrionaceae, with time lags of 1-7 days.

WaveSim identifies pairs of potential interactors by decomposing dynamic data into several spectra at different frequencies (here, intervals of 11, 7, 4, and 2 days, roughly), and then searching for all taxon pairs with 1) anticorrelated dynamics at some characteristic low frequency and correlated dynamics at some characteristic high frequency, or 2) correlated low-frequency dynamics and anticorrelated high-frequency dynamics. Combination 1 is expected to represent taxa with negative short-term biotic interactions (competition, predation) whose longer-term dynamics, however, respond similarly to factors that change on a slower timescale (mutualisms, environmental parameters). Conversely, Combination 2 represents taxa that respond similarly to short-term stimuli, yet compete or otherwise negatively interact over a longer timescale. Interaction strength is quantified with a metric that corresponds to the average of the

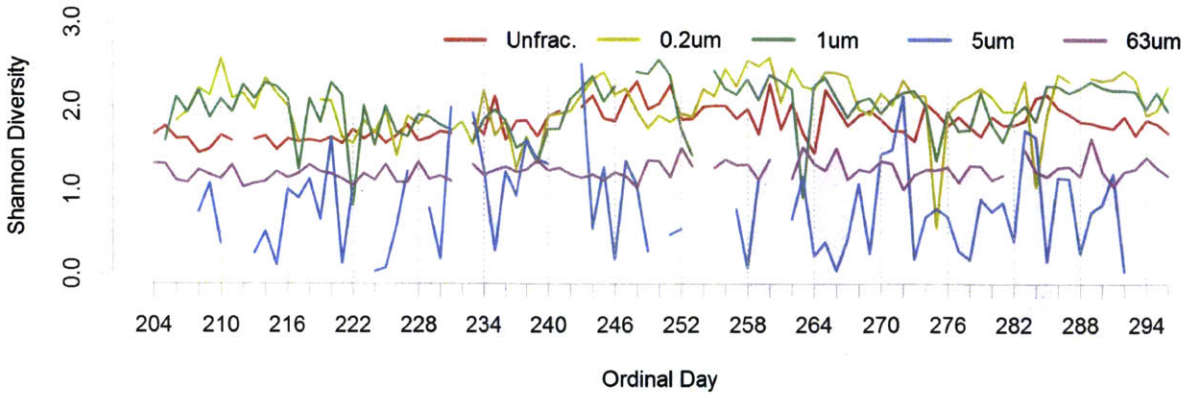
correlation and anticorrelation coefficients at the characteristic frequencies discovered to be significant, with weighting to favor relationships where the in- and out-of-phase frequencies are very close (e.g. in phase at 4 days, out of phase at 2 days).

After obtaining Granger and WaveSim data, we filtered results to the top 5 (for visualization via network representations) and top 10 (for taxonomic and metabolic characterizations) per method. We considered only Granger results with $p < 0.01$ and WaveSim results with an interaction strength metric > 0.6.

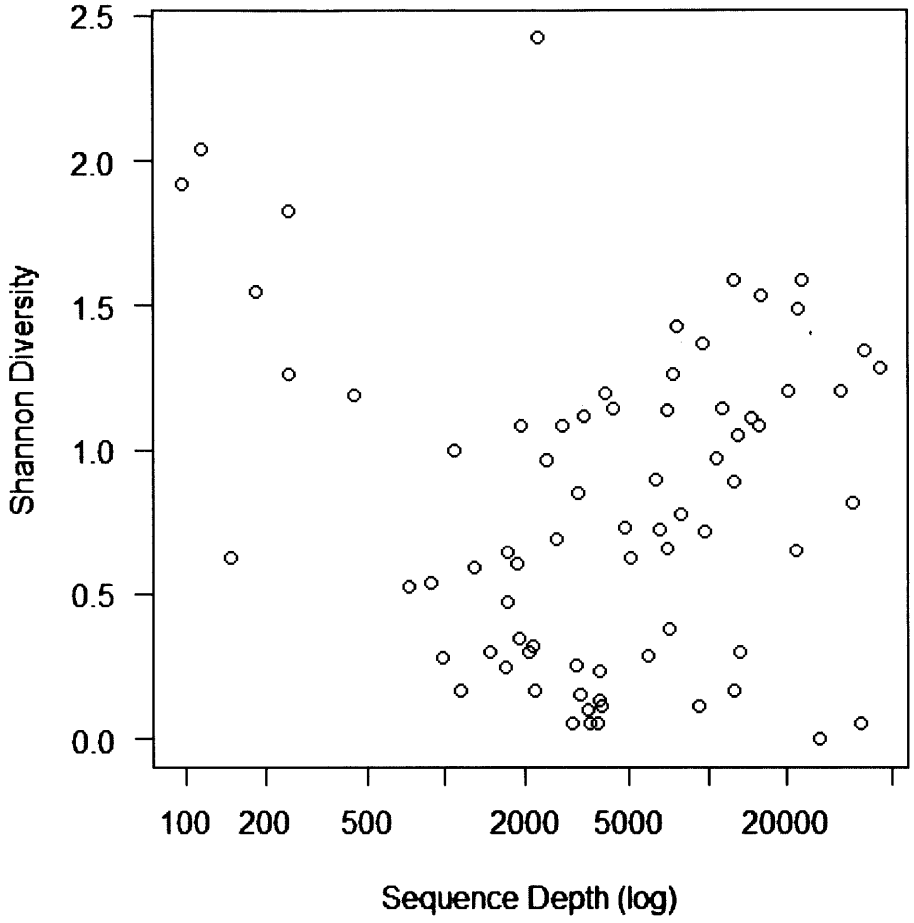
In order to infer functional significance of predicted biotic interactions, we assessed both taxonomic composition and metabolic annotations of the bacterial and eukaryotic taxa. Metabolic annotations were compiled mainly based on *Bergey's Manual of Systematic Bacteriology* (Brenner *et al.*, 2005) and on *The Revised Classification of Eukaryotes* (Adl *et al.*, 2012).



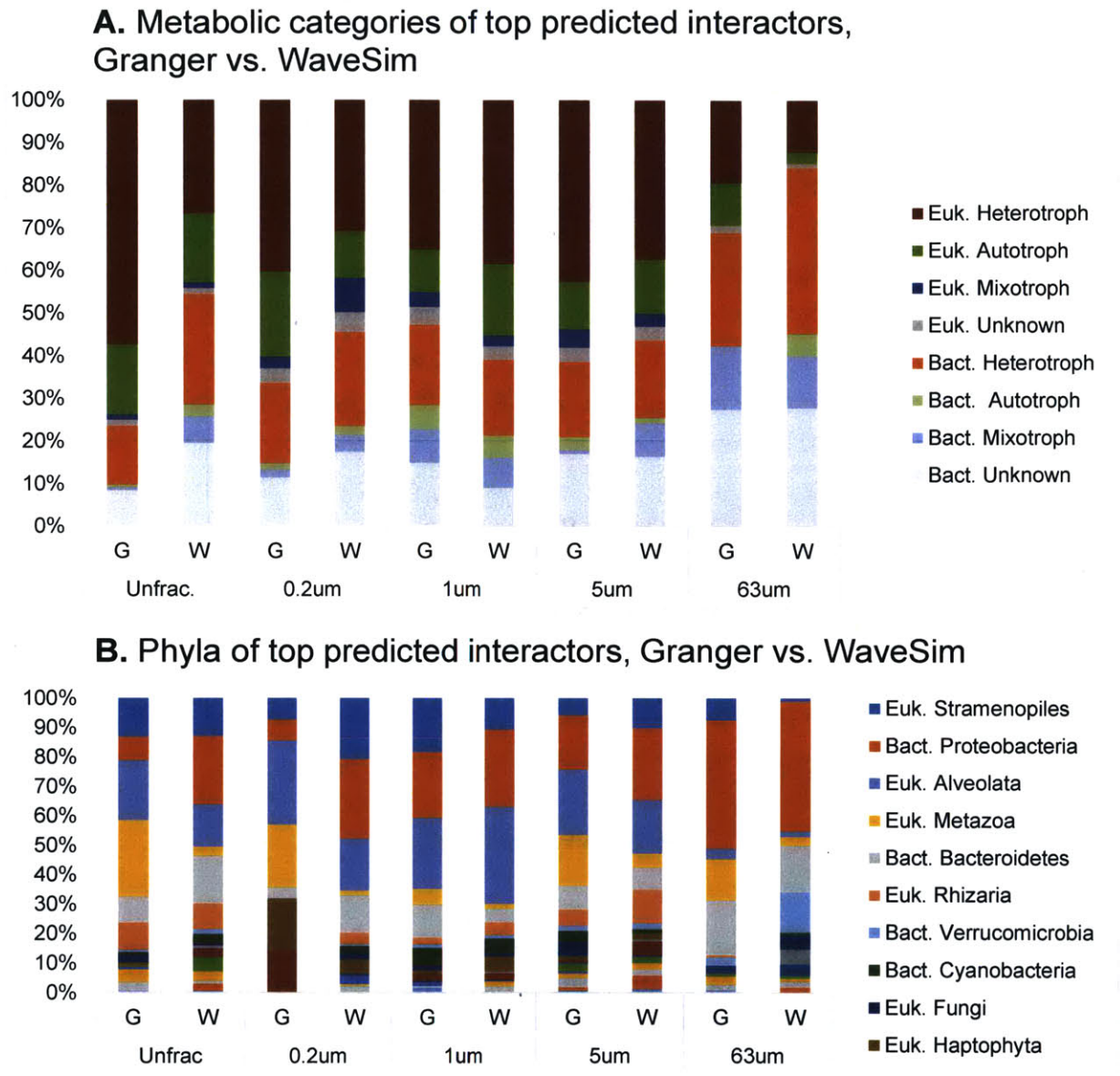
Supplementary Figure 1: Mock communities with known Vibrionaceae composition, amplified to 21-25 cycles, show little deviation from the expected composition following PCR, sequencing, and taxonomic assignment via phylogenetic placement. However, exclusion of singleton OTUs when quality-filtering sequences distorts community composition.



Supplementary Figure 2: Shannon diversity is high and moderately variable across time for the 0.2µm and 1µm samples, moderate and invariant for the unfractionated and 63µm samples, and low but extremely variable for the 5µm samples. Shannon diversity quantifies community evenness, with 0 denoting a community completely dominated by one taxon and higher values indicating richer and more even communities. Each daily sample was rarefied to a depth of 750 sequences.



Supplementary Figure 3: For the 5μm fraction, per-day Shannon diversity of Vibrionaceae population composition does not scale with sequencing depth. Each point represents one day of the time-series; sequencing depth here refers to the number of Vibrionaceae sequences found.



Supplementary Figure 4: Among top interactors predicted by Granger causality (G) and WaveSim (W), compositions of metabolic categories tend to be similar between the two methods; phylum composition is more variable. Sum of the top 10 results for each method, per population per fraction.

Chapter 3 References

Adl SM, Simpson AG, Lane CE, Lukeš J, Bass D, Bowser SS, *et al.* (2012). The revised classification of eukaryotes. *J Eukaryot Microbiol* **59**: 429–514.

Brenner DJ, Krieg NR, Garrity GM, Staley JT. (2005). Bergey's manual of systematic bacteriology. New York : Springer, c2005.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* **108**: 4516–4522.

Cardman Z, Arnosti C, Durbin A, Ziervogel K, Cox C, Steen AD, *et al.* (2014). Verrucomicrobia: candidates for polysaccharide-degrading bacterioplankton in an Arctic fjord of Svalbard. *Appl Environ Microbiol* AEM.00899–14.

Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Fuhrman JA, Cram JA, Needham DM. (2015). Marine microbial community dynamics and their ecological interpretation. *Nat Rev Microbiol* **13**: 133–146.

Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, *et al.* (2012). Defining seasonal marine microbial community dynamics. *ISME J* **6**: 298–308.

Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, Massana R, *et al.* (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ Microbiol* **10**: 3349–3365.

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081–1085.

Karam HN. (2012). Seawater circulation in coastal aquifers : processes and impacts. *Mass Inst Technol* 130–143.

Kauffman AKM. (2014). Demographics of lytic viral infection of coastal ocean Vibrio. Massachusetts Institute of Technology http://dspace.mit.edu/handle/1721.1/90046 (Accessed October 19, 2015).

Martinez-Garcia M, Brazel DM, Swan BK, Arnosti C, Chain PSG, Reitenga KG, *et al.* (2012). Capturing Single Cell Genomes of Active Polysaccharide Degraders: An Unexpected Contribution of Verrucomicrobia. *PLoS ONE* **7**: e35314.

Matsen FA, Kodner RB, Armbrust EV. (2010). pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **11**: 538.

Matz C, Kjelleberg S. (2005). Off the hook – how bacteria survive protozoan grazing. *Trends Microbiol* **13**: 302–307.

Platero AMM, Cleary B, Kauffman K, Alm EJ, Polz MF. High resolution time series reveals community-level organization and rapid turnover in coastal plankton.

Preheim SP, Boucher Y, Wildschutte H, David LA, Veneziano D, Alm EJ, *et al.* (2011). Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ Microbiol* **13**: 265–275.

Preheim SP, Perrotta AR, Martin-Platero AM, Gupta A, Alm EJ. (2013). Distribution-Based Clustering: Using Ecology To Refine the Operational Taxonomic Unit. *Appl Environ Microbiol* **79**: 6593–6603.

Preheim SP, Timberlake S, Polz MF. (2011). Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Appl Environ Microbiol* **77**: 7195–7206.

Price MN, Dehal PS, Arkin AP. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **5**: e9490.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N, *et al.* (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *MBio* **5**: e01371–14.

Shapiro BJ, Friedman J, Cordero OX, Preheim SP, Timberlake SC, Szabó G, *et al.* (2012). Population genomics of early events in the ecological differentiation of bacteria. *science* **336**: 48–51.

Shapiro BJ, Polz MF. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol* **22**: 235–247.

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner H-W, *et al.* (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **19 Suppl** 1: 21–31.

Szabo G, Preheim, SP, Kauffman, AKM, David, LA, Shapiro, BJ, Alm, EJ, *et al.* Reproducibility of Vibrionaceae population structure in coastal bacterioplankton (Submitted).

Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ, *et al.* (2012). Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J* **7**: 509–519.

Takemura AF, Chien DM, Polz MF. (2014). Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Aquat Microbiol* **5**: 38.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Yawata Y, Cordero OX, Menolascina F, Hehemann J-H, Polz MF, Stocker R. (2014). Competition–dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proc Natl Acad Sci* **111**: 5622–5627.

Chapter 4: How well do 16S rRNA sequences resolve ecologically distinct populations?

4.1. Abstract

The 16S ribosomal RNA gene is the standard marker gene for classifying and quantifying microbial communities, from the marine environment to the human microbiome. However, it is unclear how well the gene is able to resolve closely related, ecologically distinct populations, given its high conservation and the frequent usage of sequence identity cluster and/or short sequences. Yet recent studies have shown that closely related bacterial species or populations can harbor significant ecological differentiation. Here, we use a system of 13 closely related wild microbial populations of Vibrionaceae with distinct habitat preferences to quantify 16S rRNA sequences' ability to resolve ecological populations. We find that even maximally informative 16S rRNA sequence clusters (100% identity clusters of full-length sequences) can resolve only a minority of the populations, and that with any decrease in sequence length or clustering identity cutoff, almost no populations can be resolved. By contrast, a higher-resolution protein-coding gene, *hsp*60, can resolve all or nearly all populations. We use amplicon data from an environmental time-series to show that 16S rRNA sequences conceal dynamic behavior as well, collapsing all sequences from this community into only 2-4 taxa that fail to reflect the wide range of dynamic behavior exhibited by the *hsp*60-defined taxa. While 16S rRNA remains a vital tool for broad surveys of microbial diversity, our work illustrates its limits in resolving ecology, and supports the importance of targeting high-resolution marker genes when characterizing closely related populations.

4.2. Introduction

The 16S ribosomal RNA gene has been used as a phylogenetic marker gene since the 1980's to classify and quantify microbial diversity because of its sequence conservation and universal distribution (Tringe and Hugenholtz, 2008). The recent adoption of next-generation sequencing techniques has cemented 16S rRNA gene sequencing as the standard for surveying microbial communities (Caporaso *et al.*, 2011). However, the high conservation of 16S rRNA creates a tradeoff in terms of phylogenetic resolution, especially when coupled with the typical practice of calling sequence identity clusters as operational taxonomic units (e.g., 97% OTU's). Even full-length 16S rRNA gene sequences can fail to resolve species that are closely related, and with clustered partial reads, as in tag sequencing, genus-level assignments can vary depending on which region of the gene is used (Caporaso *et al.*, 2011). This lack of resolution may obscure finer-scale patterns within microbial communities, at the level of species or population.

Recent research has demonstrated that OTUs with >97% 16S rRNA identity can subsume populations with finely differentiated behavior or capabilities (e.g., Hunt *et al.*, 2008; Denef *et al.*, 2010; Biller *et al.*, 2015). For example, the marine cyanobacterium *Prochlorococcus*, the world's most abundant photosynthesizer, retains ~97% 16S rRNA identity across the genus, yet has over a dozen clades with distinct environmental distributions according to light, nutrients, and other parameters (Biller *et al.*, 2015). Identification of such clades allows inference of ecological populations: closely related, co-occurring organisms, which have also been shown to act as gene-flow units through which adaptive alleles and genes can spread (Shapiro and Polz, 2014). Accordingly, these function as basic units of ecology and evolution. However, the extent to which 16S rRNA gene profiling may conflate or obscure such ecological units is unclear, which limits understanding of the ecological significance of 16S rRNA OTUs.

Here, we quantify the ability of 16S rRNA gene sequences to resolve ecological populations, using a system of previously identified wild Vibrionaceae populations (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Szabo *et al.*, 2012), as compared to a higher-resolution marker gene, *hsp*60, for which we have developed a tag-sequencing method (see Chapter 3 of this thesis). To identify the most informative usage of 16S rRNA sequences, we include comparisons of full-length 16S rRNA versus variable regions, and different OTU cutoffs. Furthermore, we demonstrate resolving ability of

16S rRNA versus *hsp*60 with real amplicon data from a time-series, showing the extent to which 16S rRNA sequence clusters may conflate taxa with distinct ecological dynamics.

4.3. Results and Discussion

In order to compare taxonomic resolution of 16S rRNA and *hsp*60, we created in silico reference libraries of sequences for both genes from a collection of coastal marine Vibrionaceae isolates. The Vibrionaceae are heterotrophs with diverse ecological lifestyles (Takemura *et al.*, 2014); previous multilocus sequencing analysis in this system has identified approximately 20 populations with distinct distributions among microhabitats within the same water samples, i.e., free-living or associated with organic particles, algal or animals (Hunt *et al.*, 2008; Preheim, Boucher, *et al.*, 2011; Szabo *et al.*, 2012). *hsp*60 was established as a marker gene for these Vibrionaceae as it is single-copy within them, but more variable than 16S rRNA.

To obtain 16S rRNA gene sequences from the Vibrionaceae collection, nearly full-length 16S rRNA sequences were retrieved from isolates' genomes via BLAST, aligned with an *Escherichia coli* sequence, and curated to exclude sequences with terminal gaps. This yielded a 1551-basepair alignment with 482 sequences (henceforth referred to as "full-length 16S rRNA sequences"), representing 13 populations (Supp. Table 1). (Sequenced genomes were not available for members of the other populations from this system.) To serve as an outgroup, we added 29 sequences BLASTed from *Shewanella* type-strain genomes. Finally, to characterize frequently studied variable regions, we trimmed the alignment to yield the V1-V3 region (here, 8F-534R, with 534R chosen as the 3' cutoff according to the Human Microbiome Project Consortium's protocol version 4.2.2 [http://www.hmpdacc.org/]), and the V4 region (515F-926R, according to the Earth Microbiome Project's protocol version 4_13 [http://www.earthmicrobiome.org]).

For the *hsp*60 reference library, we used a preexisting collection of 2641 422-basepair Sanger sequences from the same 14 populations (including *Shewanella*; Supp. Table 1). We additionally analyzed a 215-basepair region within these *hsp*60 sequences, to correspond to the amplicon region analyzed with the tag-sequencing method mentioned above.

To assess the taxonomic composition of 16S rRNA and *hsp*60 sequence clusters, we clustered each sequence collection with the following identity cutoffs via USEARCH (Edgar, 2010): 16s rRNA full-length, 100, 99, 97%; 16S rRNA V1-V3 and V4, 100%; *hsp*60

422-bp, 100%; and *hsp*60 215-bp, 100, 95%. Taxonomic names for each isolate analyzed had been assigned based on Multi-Locus Sequencing Analysis (MLSA) or *hsp*60 sequence (Preheim, Timberlake, *et al.*, 2011). Based on the clustering analysis, we identified which populations possessed distinctly identifying gene sequence clusters (clusters not intermixed with sequences from any other population). Finally, to illustrate the extent to which phylogenetically and/or ecologically divergent populations were distinguished by these clusters, we arrayed the clustering information against a phylogeny of the 14 populations based on 52 concatenated ribosomal proteins, and information about the populations' habitat associations (Fig. 1).

We found that at best, with 100% identity clusters of full-length sequences, 16S rRNA sequences are able to resolve six of the more deeply branching populations among the 13 Vibrionaceae (*Enterovibrio norvegicus, Aliivibrio fischeri,* and *Vibrio* sp. *rumoiensis*-like, sp. *alginolyticus,* sp. F6, and sp. F10) (Fig. 1). Each of the other seven populations is merged with at least one other population in at least one gene cluster (Fig. 1, Supp. Fig. 1). Among these seven unresolved populations, though six are recently diverged and very closely related, four of those six are known to nonetheless possess distinct ecological distributions, i.e., *V. splendidus* is a broad generalist abundant in the spring, whereas *Vibrio* sp. F13, *V. tasmaniensis,* and *V.* sp. F12 are abundant in the fall and have narrower habitat distributions. Hence, even 100%-clustered full-length 16S rRNA sequences obscure this information.

16S rRNA sequences substantially lose resolving power when either cluster sequence identity cutoff is relaxed or sequence length decreases. Full-length gene sequences clustered at 99% and 97% identity resolve only two of the more deeply branching populations (*A. fischeri, V. rumoiensis*-like), while 100% clusters of the V1-V3 region resolve only the outgroup, and 100% clusters of the V4 region resolve only one Vibrionaceae population (*V. rumoiensis*-like). In addition, 16S rRNA clusters overall do not always reflect phylogeny: the deeply branching *E. norvegicus* and *V. breoganii* are confounded with much more recently diverged taxa across nearly all or all clustering cutoffs, respectively.

By contrast, 422-bp *hsp*60 sequences uniquely distinguish all 13 Vibrionaceae populations when clustered at 100% identity, while decreasing sequence length to 215 bp loses resolution of only two of the most closely related populations (*V. splendidus* and *V. tasmaniensis*). Relaxing clustering stringency to 95% clusters of 215-bp sequences does not further decrease resolution, which accords with the observation that 95% is the

average within-population *hsp*60 identity in this system, and hence should reflect population structure (data not shown).

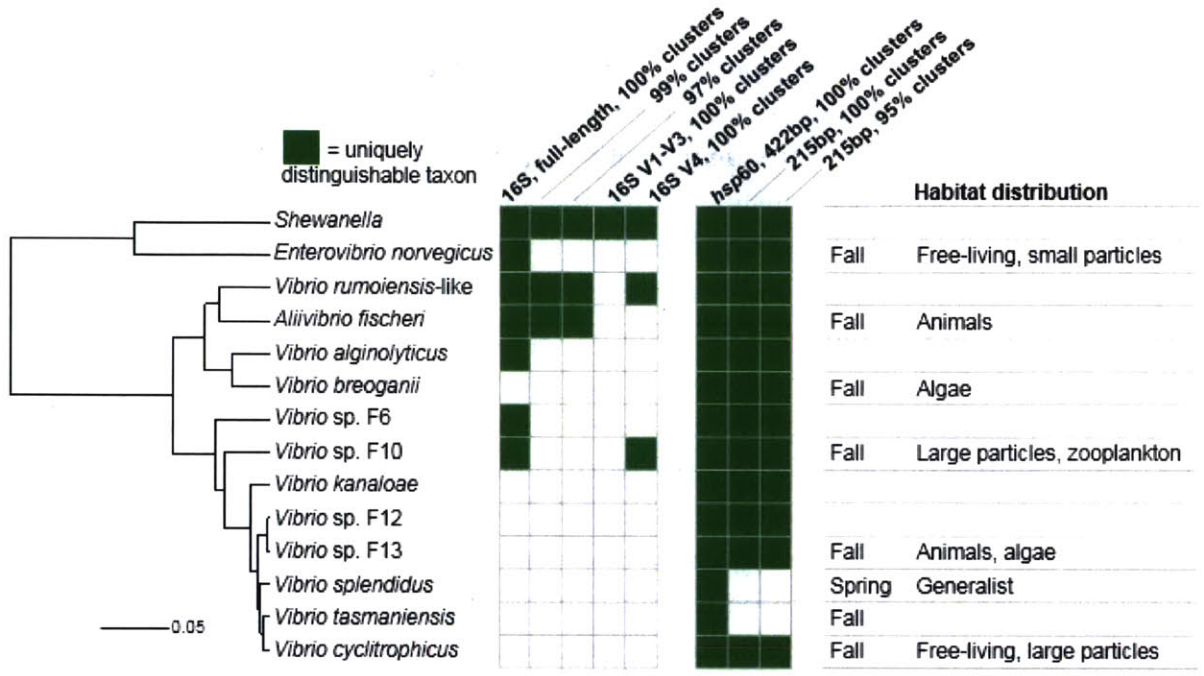


Figure 1: 16S rRNA gene sequence clusters can distinguish 0-7 of 14 ecologically distinct Vibrionaceae populations depending on sequence length and clustering cutoff, while hsp60 clusters can distinguish all or nearly all. Phylogeny is based on from 52 concatenated ribosomal proteins. A shaded box indicates that a taxon can be uniquely distinguished with the given gene length and clustering method, while a white box indicates that a taxon is merged with at least one other taxon, in at least one gene cluster. (See Supp. Fig. 1 for full representation of which taxa are merged per cluster.) Habitat distribution descriptions are derived from a quantitative analysis of populations' distributions across three different sample sets by (Preheim, 2010). Taxa without habitat descriptions were excluded from that analysis because of limited sampling. Spatial distribution of *V. tasmaniensis* is excluded as it was less clearly characterizable.

We then compared resolution by 16S rRNA versus *hsp*60 sequences in a dynamic context with real data, using amplicon data from a 93-day daily environmental time-series sampled from the same coastal environment as our Vibrionaceae collection (see Chapter 3 of this thesis). For 16S rRNA, we analyzed 76-bp sequences (V4 region, 515-591) that had been BLAST-assigned as Vibrionaceae after clustering with 100%, 98%, and 97% sequence identity cutoffs, and with the Distribution-Based Clustering (DBC) method of (Preheim *et al.*, 2013), which identifies sequence clusters that covary across

samples rather than using an identity cutoff, creating more ecologically informed units. DBC is expected to perform similarly to 100% identity cutoffs, but can account for operon-level variation within the same genome. For *hsp*60, we analyzed 215-bp sequences amplified with Vibrionaceae-specific primers, confirmed to be Vibrionaceae based on phylogenetic placement against a reference library, and accordingly categorized by population (see Chapter 3 of this thesis). Finally, to quantify the range of information captured by each set of sequence clusters, we calculated the dynamic range (ratio of maximum to minimum frequency within the Vibrionaceae) versus the average frequency of every sequence cluster across the time-series (Fig. 2). To filter out sequencing noise, we included only 16S rRNA clusters that reached a minimum abundance of 0.01% within the 16S dataset, and *hsp*60 clusters that occurred at least at 0.1% on at least 20 days within their Vibrionaceae-specific dataset.

The 16S rRNA sequences, clustered with all methods, produced less informative clusters than *hsp*60: they produced fewer clusters with higher average frequencies, and with a narrower span of dynamic ranges. Hence, they resulted in OTUs that captured less variable behavior across the time series. DBC produced the most (6) and most dynamic clusters, with a wider range of both dynamic ranges (20.2-337) and average frequencies (2.2-54%). 100% cutoffs performed similarly, except for producing one fewer low-frequency cluster. 98% and 97% cutoffs produced progressively fewer clusters (3 and 2), with correspondingly narrower span of dynamic ranges, and higher average frequencies. At worst, with 97% clustering, two clusters comprised nearly 85 and 15% of the Vibrionaceae sequences, effectively obscuring almost all variation within the Vibrionaceae community. Overall, the 16S sequence clusters tend to demonstrate a slight positive correlation between dynamic range and average frequency, i.e., larger OTUs are more variable. Additionally, consistent with the analysis of population resolution above, the 16S rRNA sequences could only be BLAST-identified taxonomically to the family or genus level.

By contrast, the 14 populations identified by *hsp*60 capture a wider range of behavior: their dynamic ranges span nearly four orders of magnitude, and their frequency distribution is less skewed towards a few extremely large clusters; the maximum average frequency is only 26%, compared to 54%, 58%, 62%, and 85% frequency for DBC, 100%, 98%, and 97% clusters respectively. Furthermore, the populations do not demonstrate 16S rRNA clusters' positive correlation between dynamic range and average frequency. Instead, there are five more abundant clusters (6.3-26%) with moderate dynamic ranges (3-41), while the remaining clusters' dynamic ranges span over 3 orders of magnitude. This suggests that this particular community is

dominated by a few rather stable populations, while the remainder are rare and fluctuate widely.

Hence, our study shows that when attempting to capture ecological diversity below the family level, 16S rRNA gene sequences should be used with greater sequence length and more stringent or ecologically informed clustering methods, while the traditional metric of 97% sequence identity can conceal nearly all intra-family variation. While 100% cutoffs for tag sequencing offer good resolution, they can result in sequencing artifacts and operons being counted as separate OTUs. As a possible solution, ecologically informed approaches should be able to decrease such artifacts by grouping sequence variants (including operons) with similar behavior. In addition to the Distribution-Based Clustering method employed above, the oligotyping method discriminates ecologically distinct taxa based on a subset of highly ecologically informative nucleotide positions, discarding the rest (Eren *et al.*, 2013). For example, one recent study used oligotyping to detect Vibrionaceae taxa with different habitat preferences across multiple 16S rRNA datasets collected from different environments (Schmidt *et al.*, 2014).

Because 16S rRNA sequences still fall short of resolving below the genus level, regardless of clustering methods, our analysis also points to the importance of identifying 1) suitably high-resolution marker genes and validating sequencing approaches for them, and 2) incorporating fine-scale ecology into analysis, in order to resolve gene clusters acting as ecological units. Here, *hsp*60 provided a much richer view of Vibrionaceae ecological diversity and dynamics than 16S rRNA, and has previously been shown to perform nearly as well as MLSA for the same system (Preheim, Timberlake, *et al.*, 2011). In other systems, for example *Prochlorococcus*, intergenic spacer sequences are often employed to distinguish ecological populations (Biller *et al.*, 2015). Thus, microbial ecologists seeking insight into fine-scale ecology can consider two strategies: using appropriate clustering methods to mine ecological insight from broad-coverage 16S rRNA datasets, or narrowing focus to higher-specificity sequencing targets for a particular system. In either case, sequence data should be accompanied by appropriate environmental metadata, to illuminate the identity of microbial taxa not just as sequence clusters, but as diverse, dynamic populations interacting with a rich environment.

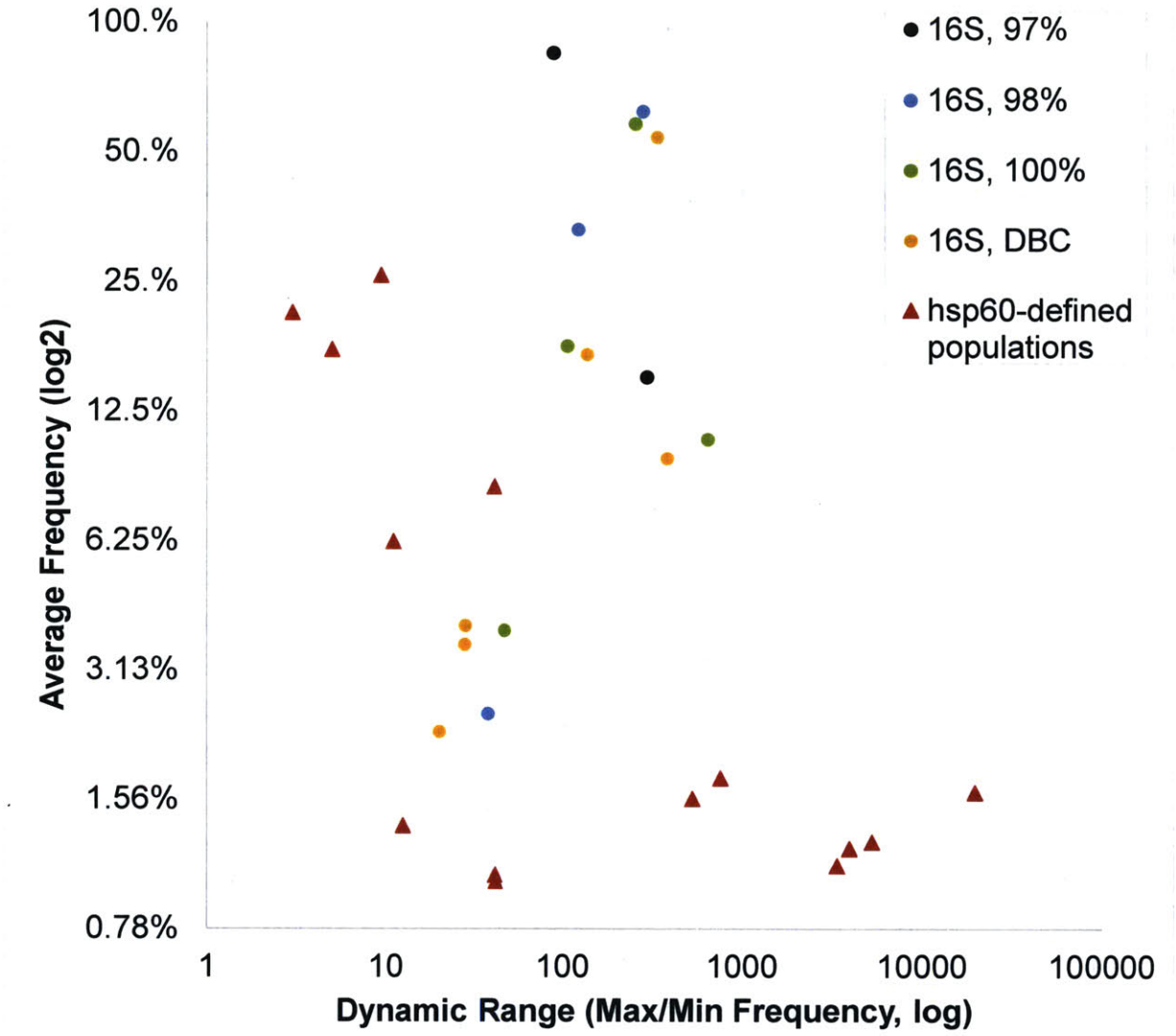


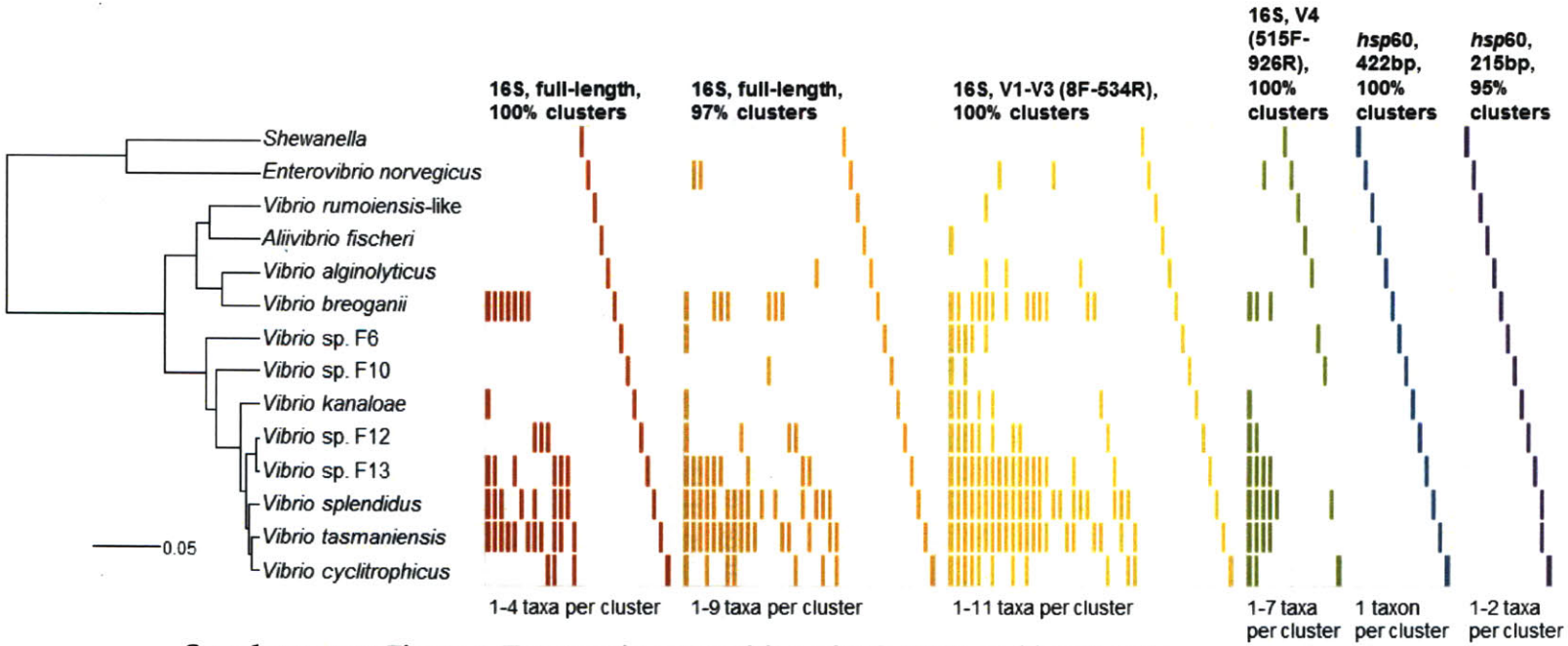
Figure 2: In an environmental time-series dataset, 16S rRNA Vibrionaceae gene sequence clusters reflect a narrower range of dynamic behavior than hsp60 sequence clusters, and are heavily dominated by a few highly abundant clusters. 76-bp 16S rRNA amplicons identified as Vibrionaceae were clustered by sequence identity cutoffs and by Distribution-Based Clustering (DBC, Preheim *et al.*, 2013), then filtered to only those with a minimum abundance of 0.01% within the 16S dataset. 215-bp Vibrionaceae-specific *hsp*60 amplicons were clustered at 100% and filtered to only those with a frequency of at least 0.1% on at least 20 days within the Vibrionaceae *hsp*60 dataset.

Chapter 4 Supplementary Figures

	16S rRNA sequences, 1500bp	***hsp*60 sequences, 422bp**
Shewanella	29	9
Enterovibrio norvegicus	11	157
Vibrio rumoiensis-like	1	4
Aliivibrio fischeri	1	148
Vibrio alginolyticus	1	24
Vibrio breoganii	85	14
Vibrio sp. F6	2	17
Vibrio sp. F10	4	51
Vibrio kanaloae	9	133
Vibrio sp. F12	12	88
Vibrio sp. F13	62	63
Vibrio splendidus	93	922
Vibrio tasmaniensis	123	831
Vibrio cyclitrophicus	78	180
Total	511	2641

Supplementary Table 1: Number of marker gene sequences obtained via BLAST from genomes of isolated strain representatives of 13 environmental Vibrionaceae populations. Populations were previously defined on the basis of multilocus sequencing analysis and distinct ecological distributions.

16S rRNA sequences from the outgroup *Shewanella* were obtained from genomes of the following type strains (with accession code in parentheses): *Shewanella pealeana* ATCC 700345 (CP000851.1), *Shewanella piezotolerans* WP3 (CP000472.1), *Shewanella woodyi* ATCC 51908 (CP000961.1); all retrieved from GenBank on December 28, 2015. *Shewanella hsp*60 strains were obtained from the same environmental strain collection as the Vibrionaceae.



Supplementary Figure 1: Taxonomic composition of 16S rRNA and hsp60 gene clusters with different clustering strategies. Within each panel, each column represents the taxonomic composition of at least one sequence cluster for the given gene and clustering strategy. Within a column, a colored bar indicates that sequence(s) from the corresponding taxon in the phylogeny are present in that sequence cluster composition. Phylogeny is based on 52 concatenated ribosomal proteins.

Chapter 4 References

Biller SJ, Berube PM, Lindell D, Chisholm SW. (2015). Prochlorococcus: the structure and function of collective diversity. *Nat Rev Microbiol* **13**: 13–27.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* **108**: 4516–4522.

Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC, *et al.* (2010). Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proc Natl Acad Sci* **107**: 2383–2390.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, *et al.* (2013). Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* **4**: 1111–1119.

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081–1085.

Preheim SP. (2010). Ecology and population structure of Vibrionaceae in the coastal ocean. Massachusetts Institute of Technology http://dspace.mit.edu/handle/1721.1/58184 (Accessed November 22, 2015).

Preheim SP, Boucher Y, Wildschutte H, David LA, Veneziano D, Alm EJ, *et al.* (2011). Metapopulation structure of Vibrionaceae among coastal marine invertebrates. *Environ Microbiol* **13**: 265–275.

Preheim SP, Perrotta AR, Martin-Platero AM, Gupta A, Alm EJ. (2013). Distribution-Based Clustering: Using Ecology To Refine the Operational Taxonomic Unit. *Appl Environ Microbiol* **79**: 6593–6603.

Preheim SP, Timberlake S, Polz MF. (2011). Merging taxonomy with ecological population prediction in a case study of Vibrionaceae. *Appl Environ Microbiol* **77**: 7195–7206.

Schmidt VT, Reveillaud J, Zettler E, Mincer TJ, Murphy L, Amaral-Zettler LA. (2014). Oligotyping reveals community level habitat selection within the genus Vibrio. *Front Microbiol* **5**. doi:10.3389/fmicb.2014.00563.

Shapiro BJ, Polz MF. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol* **22**: 235–247.

Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ, *et al.* (2012). Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. *ISME J* **7**: 509–519.

Takemura AF, Chien DM, Polz MF. (2014). Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Aquat Microbiol* **5**: 38.

Tringe SG, Hugenholtz P. (2008). A renaissance for the pioneering 16S rRNA gene. *Curr Opin Microbiol* **11**: 442–446.