INTERACTIONS BETWEEN CALANOID COPEPOD HOSTS AND THEIR ASSOCIATED MICROBIOTA

By

Amalia Aruda Almada

B.S., Georgetown University, Biological Sciences (2009)

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

February 2015

Signature of Author…….....

Joint Program in Oceanography/Applied Ocean Science and Engineering
Massachusetts Institute of Technology and Woods Hole Oceanographic Institution
December 18, 2014

Certified by…….

Dr. Ann M. Tarrant
Associate Scientist in Biology
Woods Hole Oceanographic Institution
Thesis Supervisor

Accepted by…….....

Professor Martin F. Polz
Chair, Joint Committee for Biological Oceanography
Massachusetts Institute of Technology and Woods Hole Oceanographic Institution

INTERACTIONS BETWEEN CALANOID COPEPOD HOSTS
AND THEIR ASSOCIATED MICROBIOTA

by

Amalia Aruda Almada

Submitted to the MIT/WHOI Joint Program in Oceanography/Applied Ocean Science and Engineering on December 18, 2014 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Oceanography

ABSTRACT

Zooplankton, such as copepods, are highly abundant environmental reservoirs of many bacterial pathogens. Although copepods are known to support diverse and productive bacterial communities, little is understood about whether copepods are affected by bacterial attachment and whether they can regulate these associations through mechanisms such as the innate immune response. This thesis investigates the potential role that copepod physiology may play in regulating *Vibrio* association and the community structure of its microbiome. To this end, the intrinsic ability of oceanic copepod hosts to transcriptionally respond to mild stressors was first investigated. Specifically, the transcriptional regulation of several heat shock proteins (Hsps), a highly conserved superfamily of molecular chaperones, in the copepod *Calanus finmarchicus* was examined and demonstrated that Hsps are a conserved element of the copepod's transcriptional response to stressful conditions and diapause regulation. To then investigate whether copepod hosts respond to and regulate their microbiota, the transcriptomic response of an estuarine copepod *Eurytemora affinis* to two distinct *Vibrio* species, a free-living strain (*V. ordalii 12B09*) and a zooplankton specialist (*V. sp. F10 9ZB36*), was examined with RNA-Seq. Our findings provide evidence that the copepod *E. affinis* does distinctly recognize and respond to colonizing vibrios via transcriptional regulation of innate immune response elements and transcripts involved in maintaining cuticle integrity. Our work also suggests that association with *E. affinis* can significantly impact the physiology of *Vibrio* colonists. Finally, the inter-individual variability of the *C. finmarchicus* microbiome was examined to identify how specifically and predictably bacterial communities assemble on copepods and whether host physiology influences the bacterial community structure. Our findings suggest that copepods have a predictable "core microbiome" that persists throughout the host's entrance into diapause, a dormancy period characterized by dramatic physiological changes in the host. However, diapausing and active populations harbor distinct flexible microbiomes which may be driven by factors such including the copepod's feeding history, body size, and bacterial interactions. This thesis work highlights the role of copepods as dynamic reservoirs of diverse bacterial communities and implicates copepod host physiology as an important contributor to the activity, abundance, and community structure of its microbiome.

Thesis Supervisor: Dr. Ann M. Tarrant
Title: Associate Scientist in Biology, Woods Hole Oceanographic Institution

Thesis Advisor

Dr. Ann M. Tarrant
Department of Biology
Woods Hole Oceanographic Institution

Thesis Committee

Professor Rita R. Colwell
Center for Bioinformatics and Computational Biology
University of Maryland

Dr. Mark F. Baumgartner
Department of Biology
Woods Hole Oceanographic Institution

Dr. Tracy J. Mincer
Department of Marine Chemistry & Geochemistry
Woods Hole Oceanographic Institution

Professor Martin F. Polz
Department of Civil & Environmental Engineering
Massachusetts Institute of Technology

Thesis Defense Chair

Dr. Mark E. Hahn
Department of Biology
Woods Hole Oceanographic Institution

ACKNOWLEDGEMENTS:

This work has been supported by a National Science Foundation Graduate Research Fellowship, an Environmental Protection Agency Science to Achieve Results Fellowship, National Science Foundation Grants (OCE-1132567 to AM Tarrant and MF Baumgartner), student awards from the WHOI Ocean Venture Fund, and private donation from Richard and Susan Hill.

What a long, strange (and amazing) trip it's been. Reaching the end of this journey has only been made possible with the help of my incredible mentors, colleagues, friends, and family. Those thanked below represent only some of those who have been an integral part of this experience.

To my advisor Ann, who has guided my scientific growth ever since I was a freshman in college. Thank you for your mentorship and for your confidence in me as I took the big plunge into microbiology. Thank you for giving me space to learn how to be an independent scientist and for always throwing me a rope when I was getting sucked down a rabbit hole. Thank you for encouraging me to pursue my interests in science policy. I have been blessed to have had you as an advisor all of these years.

To the members of my thesis committee, Rita Colwell, Mark Baumgartner, Tracy Mincer, and Martin Polz, for their dedication to my research projects and my personal growth as a scientist. Thank you for your thoughtful input and attention, which have helped these projects succeed. I feel honored to have had the opportunity to work with you all these past 5.5 years. Special thanks to Dr. Rita Colwell, whose research was the inspiration to my own thesis work. Thank you for your mentorship and for taking a personal interest in my research and career development. You have been an incredible role model to myself and to many other aspiring female scientists.

To members of the Tarrant lab, past and present, for your moral and scientific support. Thank you to Luisa Villamil Diaz for her ideas and assistance with the copepod-*Vibrio* exposure experiments. Thank you to Amy Maas for her guidance with the RNA-Seq analysis. Thank you to Maja Edenius for the needed study break walks around Eel pond.

To the members of the Polz lab, for welcoming me as a part of the lab group while I completed my 16S rRNA sequencing libraries. Thank you for your friendship, advice, and encouragement, especially after long days in the PCR hood. Special thanks to Michael Cutler for his help in keeping the library preparations running smoothly. Special thanks also to Manoshi S. Datta, who has been a wonderful collaborator and friend and whose help with the 16S rRNA analysis was a crucial piece to the copepod microbiome work. Thank you to Sarah Preheim for her help with the sequencing analysis and continued encouragement and interest in my copepod-*Vibrio* work. Thank you to Otto Cordero for the invigorating scientific discussions that spurred me onward with these challenging projects and for setting me on the right path with the copepod microbiome work. I wish all of you the best in your future scientific endeavors.

To my cohort, Kristen, Santiago, Liz, Britta, Jill, and Kalina, for your friendship. We have all grown so much these past 5.5 years. Thank you for making the growing process more fun – often by eating a lot of fondue cheese and chocolate together.

To my MIT and Woods Hole Christian community, for all of the prayers and fellowship that gave me the courage to keep pushing forward.

To my dad, my first scientific mentor and my editor-in-chief. Thank you for helping me reach my childhood dream of becoming a marine biologist. To my mom, for your constant support and for the great example of how to continuously love on others even when enormously busy. To my brothers, for their encouragement and sound advice.

To my husband Albert, who has been the best partner I could have ever hoped for throughout this whole process. Thank you for helping me grow to be a better scientist and person. I could not have done this without you.

TABLE OF CONTENTS

Chapter One

Introduction

INTRODUCTION:

We inhabit a microbial world; dynamic bacterial communities live in, on, and all around us. Once thought to be rare occurrences, bacterial associations with eukaryotic organisms are now known to be ubiquitous and to have dramatic implications for host biology (1, 2). The collective microbial community that inhabits an organism (i.e., the microbiome) can exert remarkable influence on the development, behavior, metabolism, and immunity of a broad diversity of animal hosts (1-3). The microbiome of animals can often be highly specific (4-6) due to the maternal transmission of bacterial symbionts and/or the interplay of host-microbe and inter-specific microbial interactions that sculpt the bacterial community structure (7-10). To enrich for specific bacterial symbionts and prevent invasion by harmful bacteria, animals have developed diverse mechanisms to initiate and regulate their associations with bacteria (3, 11). Those bacteria that successfully associate with animal hosts can in turn receive benefits including access to nutrient-rich environments (12), protection against environmental stressors (13), and enhanced persistence in the environment (14). Overall, our knowledge of the impact of the microbiome on animal biology and the mechanisms by which host- specific microbiomes form is still in its infancy, but has already transformed our existing understanding of the extent and function of the microbial biosphere.

Microbial contributions to animal biology

Animals are often considered 'superorganisms' whose biology represents a synergistic blend of bacterial and animal activity (15, 16) in light of the predominance in sheer numbers of bacterial cells relative to host cells (17, 18), and the contribution of microbes to host physiology (1). Research continues to uncover that fundamental host functions, including development (1, 2), metabolism (19), immunity (20), and behavior (3), are not autonomous and often depend on the activity of the host's microbiota. For example, the bobtail squid *Euprymna scolopes* depends on a light organ in order to camouflage itself, and the proper development of this light organ is only induced in the presence of their light-generating *Vibrio fischeri* symbiont. Colonization by *V. fischeri* leads to significant remodeling of the epithelial tissue in the light organ, and without exposure to their symbiont at the proper time, the organ's development is irrevocably stalled

(21). The juvenile development of the giant tubeworm *Riftia pachyptila* is also intimately linked to the presence of its obligate symbionts. The tubeworm symbionts pass through the skin of their hosts during early juvenile stages as the symbiont-housing organ (i.e., the trophosome) is established in the host's mesodermal tissues (22). As *Riftia* enters later juvenile stages, the epidermal and muscle tissues surrounding the trophosome undergo massive apopotosis, ultimately enclosing the symbionts within the trophosome (22). Even in systems without symbioses as intimate as those of the bobtail squid and *Riftia*, bacteria can play essential roles in host development. For example, recent studies have shown that surface-bound signals produced by bacterial biofilms trigger larval settlement and development in a number of marine invertebrates including biofouling barnacles and the tube-dwelling polychaete worm *Hydroides elegans* (23, 24). In vertebrate animals such as mice, it has been shown that the normal gut microbiota can even affect brain development by altering neuroendocrine system signaling pathways (25).

Beyond their influence on the normal development of animal hosts, microbiota can also drive host behavior in a manner that has significant implications for host ecology and evolution. For example, a recent study suggests that the microbiome of the fruit fly *Drosophila melanogaster* can influence mating preference by altering the host's production of cuticular hydrocarbon sex pheromones (26). In another study, the overall diversity of the bacterial community and the presence of certain bacteria on human skin were found to correlate with an individual's 'attractiveness' to hungry female *Anopheles gambiae* mosquitos (27). It was shown that humans that are poor attractants to *A. gambiae* have surface bacterial communities with high abundances of *Pseudomonas* and *Variovorax*, which may produce mosquito-repelling volatile compounds or even mask attractive human-produced volatiles. Beyond altering the host's volatile signals, a growing body of literature shows that the gut microbiome can also affect the host's intracellular signaling via neural, endocrine, and immune pathways, which can ultimately profoundly influence brain processing (28). Alteration of this 'gut-microbiome-brain axis' has been tied to changes in vertebrate emotional and stress responses. For instance, researchers found that feeding mice a probiotic of the lactic acid bacteria *Lactobacillus rhamnosus* reduced their overall anxiety as measured by changes in both mice behavior and brain chemistry (29).

Symbiotic associations with bacteria also provide important nutritional advantages that can enable invertebrate hosts to occupy novel nutritional niches (12). Such nutritional niches can be physical environments where food is scarce or diets that consist of nutrient-poor or difficult to digest food sources (30-32). Despite the presence of harsh chemicals and high temperatures, hydrothermal vents are highly productive ecosystems largely fueled by the activity of chemoautotrophic bacteria, which are often found in association with animal hosts (33, 34). The adult *Riftia* relies entirely on its intracellular sulfur-oxidizing symbiont for nutrition, which is physically demonstrated by the adult tubeworm's lack of a gut and mouthparts (35). Hydrothermal vent shrimp also depend on symbiotic chemosynthetic bacteria as a nutritional source in this system by feeding directly on their epibionts and other free-living bacteria (36). In other environments, invertebrates can rely on obligate symbioses to provide the necessary nutrients that are lacking in their unique food sources (12). For example, the tsetse fly *Glossina morsitans* strictly feeds on vertebrate blood, a food source lacking in essential vitamins, which are supplemented by the obligate bacterial symbiont *Wigglesworthia*. This proposed role of *Wigglesworthia* is supported by the fact that the symbionts retain a large proportion of genes involved in the synthesis of vitamins despite having a streamlined genome (37), and the finding that compromised hosts lacking symbionts partially regain their normal physiology when they are supplemented with B-complex vitamins (38). The pea aphid *Acyrthosiphon pisum* similarly feeds on a nutritionally deficient food source (plant phloem) and relies on its obligate symbiont *Buchnera* for provision of nitrogenous nutrients such as amino acids (39, 40). Bacteria can also be essential players in digestive symbioses in which the host requires help to digest complex substrates like the lignocellulose found in wood and leaf matter (12, 41). Cellulolytic and fermentive bacterial symbionts are found in many insects, including the scarab beetle (42), the crane fly (43), and termites, which harbor complex microbial communities comprised of bacteria, archaea, and protists in their hindgut (32, 44-46).

A diversity of invertebrates harbor bacterial symbionts that convey enhanced immunity to their hosts against invading pathogens by directly inhibiting the pathogens or by modulating the host's immune activity (20, 47). In the case of the symbiotic bacteria associated with the embryos of the shrimp *Palaemon macrodactylus* (48) and the lobster *Homarus americanus* (49), the symbionts produce antifungal metabolites that protect the hosts against infection by pathogenic fungi. In the fruitfly *Drosophila* (50) and the bumble bee *Bombus terristris* (51),

individuals with a normal gut microbiome were less susceptible to parasitic infection than were germ-free individuals. In *B. terristris* it was observed that successful prevention of parasitic infection was more dependent on the type of microbes colonizing the gut than on the host's genotype (51), highlighting the dramatic role that microbiota can play in enhancing the host's immunocapacity (52). Studies that examined how the microbiome enhances host immunity suggest that commensal gut bacteria prime the immune system of hosts by eliciting a basal immune response that aids in fending off further pathogen infection (53, 54). Beyond priming the immune system, the obligate symbiont of the tsetse fly *G. morsitans* is thought to help direct the development of the cellular and humoral immune responses (31, 55). Larval tsetse flies that lack the bacterial symbiont *Wigglesworthia* demonstrate an underdeveloped immune system as adults (55) and an increased susceptibility to parasitic infection (56).

Specificity of host associations with bacteria

Researchers are continually discovering that a vast diversity of animals, from sea anemones (5), to corals (4), to whales (6), to copepods (57, Chapter 3) demonstrate highly consistent species-specific bacterial communities. This is particularly true for hosts which have evolved to depend on their microbiota for survival (i.e., obligate symbioses). For many of the well-studied obligate symbioses (e.g., *Riftia* and chemosynthetic endosymbionts, pea aphid and *Buchnera*, tsete fly and *Wigglesworthia*, which are all described in the 'Microbial contributions to animal biology' section above), the necessity of the association derives from the host's dependence on the symbiont for nutritional provision (12). These obligate bacterial symbionts are often passed down from the mother (i.e., vertical transmission) to eliminate the risks inherent in acquiring the symbiont from the surrounding environment (10). Interestingly, the intracellular chemosynthetic symbiont of *Riftia* is not passed from the mother but is acquired by juveniles from the environment (i.e., horizontal transmission) (35). As previously discussed, *Riftia* juveniles are colonized by their symbionts through the skin before the symbionts are ultimately enclosed in the host's trophosome (22). Such a strategy requires tightly regulated host control mechanisms to ensure the acquisition of the specific obligate symbiont from a bacteria-rich environment. There are many other cases of horizontally transmitted symbionts including *V. fischeri* and chemoautotrophic bacteria, which are acquired by the bobtail squid and marine

nematodes, respectively. One potential advantage of adopting such a strategy is that the symbiont population is derived from a genetically diverse free-living population, which can reduce the potential for random accumulation and persistence of deleterious alleles in the symbiont genome (10). However, acquisition and maintenance of symbionts from environments populated with potential microbial invaders requires a sensitive and highly coordinated host response (see 'Host Regulation of the microbiome' section below).

Although many recent studies demonstrate the importance of host activity in structuring species-specific bacterial communities, bacterial growth dynamics, inter-specific interactions, and stochastic processes may also play important roles in structuring these highly stable associations (7-9). *In vivo* studies demonstrate that successional changes in the bacterial members colonizing particles can be highly consistent (58, 59). Furthermore, a highly predictable linkage between the population structure of coastal ocean *Vibrionaceae* and their associated habitat has been observed in several studies (7, 60, 61). This consistency may derive from predictable growth dynamics in which the dominant bacterial members on a surface are characterized by an ability to rapidly colonize and form biofilms, while the less dominant members form more transient associations (62). The trade-off is that the transiently associated members can more flexibly and rapidly respond to new nutrient hotspots than the tightly associated, dominant members (62). Microbial dynamics on surfaces may be further modulated by antagonistic inter-specific interactions mediating competition between co-habitating bacterial populations through the production of antibiotics (63, 64) or more indirectly via environment modification due to rapid nutrient consumption by fast-growing populations (65). The competitive advantage of symbionts over other colonizing bacteria can also derive from their unique genetic arsenal, including genes that enable bacterial symbionts to specifically colonize their hosts (66, 67). Other studies have observed positive interactions between bacteria colonizing a surface in which a certain bacterial species promotes the growth of others (68) potentially via hydrolytic activity that releases nutrients otherwise unattainable by other bacteria (8). Alternatively, some recent studies have implicated stochastic birth-death and immigration processes as important drivers of observed *Vibrio* community assembly patterns on some invertebrate hosts (7, 9). Further study is needed to gain a better understanding of how the mechanisms of host activity, inter-specific bacterial interactions, and stochastic processes interact to shape the observed structure of bacterial communities that form on animal hosts.

Although a large diversity of organisms demonstrate species-specific microbiomes, when examined at a finer phylogenetic resolution it becomes apparent that individuals host unique assemblages of bacteria (69, 70). What factors drive the individual variability in microbiome structure is largely unknown, although a few recent studies suggest that dietary history (71, 72) and maternal effects (73) may be important contributors. In humans, the high variability in the relative abundance of bacterial members and overall diversity of the microbiome across individuals was not well explained by factors such as gender, temperature, and blood pressure (74). Further research is needed to explore how bacterial responses to host genetics, extrinsic environmental factors, and stochastic processes contribute to the high inter-individual variability of animal microbiomes largely results from. The majority of the inter-individual microbiome variability studies have focused on human and mouse models, and few studies examining inter-individual variability in invertebrates exist (7, 75, 76, Chapter 3). Exploration of these questions in invertebrate systems is particularly significant in light of the abundance and role of invertebrate hosts as environmental reservoirs of pathogenic bacteria (77). Furthermore, the tractability of invertebrate host systems for experimental manipulation and the preponderance of well-characterized bacterial symbioses in these systems could enable a deeper understanding of the mechanisms which guide bacterial community assembly on animal hosts.

Host regulation of the microbiome

There is growing recognition of the important role that host action, via mechanisms such as physical barriers, behavior, and innate immunity, plays in symbiont acquisition and maintenance (3, 11). Often the first line of defense for many groups of invertebrates, the hard, chitinous exoskeleton provides a physical and chemical barrier against pathogen attachment and invasion (78, 79). A potentially vulnerable point of entry into the host, the gut is lined with the peritrophic matrix, which acts like a sieve that surrounds and prevents bacteria, bacterial toxins, and hard food fragments from contacting the intestinal epithelium (11, 80-82). When the thickness and permeability of the peritrophic matrix is compromised in *Drosophila*, there is higher susceptibility to infection by pathogenic bacteria or mortality from bacterial toxins (83). Furthermore, ingestion of bacteria elicits a stronger immune response in individuals with a compromised peritrophic matrix, demonstrating the important role that this barrier defense

contributes to host immunity (83). In cases where the peritrophic matrix is breached and the gut epithelium is damaged, repair mechanisms to restore tissue integrity are crucial to the host's ability to endure infection (11, 53). In those organisms without a hard exoskeleton barrier, mucous secretions can be an important method to simultaneously reduce encounters with unwanted bacteria and enrich for symbionts (2, 84). For example, in response to the presence of the bacterial cell wall component peptidoglycan in the surrounding seawater, the bobtail squid begins shedding 'toxic' mucus from the surface of its light organ (85). The mucous secretions contain antimicrobial molecules including reactive oxygen species (2) that inhibit the growth of most bacteria but enable the symbiont *V. fischeri* to densely aggregate and dominate the bacterial community in the mucus (86). *V. fischeri* is known to aggregate towards *N*-acetylneuraminic acid, another component of the squid mucus (87), further demonstrating that selection of the specific *Vibrio* symbiont commences before the symbiont directly contacts the host surface (2).

Host behavior can also be an important mechanism to establish and regulate associations with microbiota in a diverse set of animals. In the case of the bobtail squid, a majority of the symbiont *V. fischeri* cells are predictably ejected from the light organ before sunrise to winnow down the symbionts and prevent their potential overgrowth and metabolic cost to the host (88). Symbiont-free juvenile kudzu bugs must ingest symbiont capsules left behind by the mother in order to acquire their bacterial symbionts. The young will search for capsules if they are not in close proximity or if they are removed from the nest (89). It is also thought that infant iguanas attempt to pre-colonize their guts with the fermentative microbiota required for their herbivorous material diet by solely consuming the soil and feces in their nest chamber for the first few weeks of life (90).

While physical barriers and behavioral adaptations can help the host to reduce the prevalence of non-specific bacterial associations, the innate immune system is also essential to the host's ability to initiate and regulate its association with microbes, including symbionts selected from the environment (11). The innate immune system is typically thought of as a generic response to pathogens, but recent studies have revealed a deeper level of specificity and complexity than previously realized. For example, after repeated exposure to bacterial pathogens, the bumblebee *Bombus terrestris* demonstrates unexpectedly specific and prolonged resistance to future infections (91), and the freshwater crustacean *Daphnia* appears to pass this protection to

the next generation, as the offspring of mothers exposed to a particular bacterial strain had enhanced fitness upon challenge with the same strain (92).

Elements of the innate immune system, such as C-type lectins, antimicrobial peptides (AMPs) and prophenoloxidase (proPO), enable invertebrate hosts to select for specific bacterial associates in addition to inhibiting growth of undesirable foreigners (5, 93, 94). C-type lectins can function as pattern recognition proteins, which enable the invertebrate host to recognize colonizing bacteria in order to acquire specific bacterial symbionts from the environment (10, 93, 95). The marine nematode *Laxus oneistus* produces a mucus-secreted C-type lectin that mediates symbiont association with the cuticle by inducing symbiont aggregation and by directly binding to the symbiont's bacterial antigens (93). C-type lectins can also function by internally inhibiting the proliferation of endogeneous bacteria by modulating the expression of AMPs (96) or by directly binding to bacteria and acting as antimicrobial agents (97). The proPO cascade is induced when host recognition proteins are activated by microbial compounds including bacterial surface attachment proteins and cell wall components (98). Activation of the proPO system is thought to be an early and rapid response to microbial invaders that results in the production of cytotoxic compounds and melanin, a pigment that inactivates and physically encapsulates microbial invaders (99, 100). The activation of the proPO pathway initiates the conversion of the inactive prophenoloxidase into catalytically active phenoloxidase (101, 102), triggering the production of cytotoxic compounds and encapsulation of the microbial invaders (99, 100). AMPs, easily and rapidly synthesized, disrupt bacterial membranes (103) and can bind to bacterial cells to promote their phagocytosis (104). In the early life stages of invertebrates, AMP production can be an important method by which hosts modulate bacterial colonization (105, 106). In the cnidarian *Hydra*, successive changes in expression patterns of perlucin family AMPs during embryogenesis mediate the formation of the microbiome (105). Furthermore, species-specific composition and expression patterns of arminin AMPs in adult *Hydra* help shape their highly stable, species-specific bacterial communities (5). Therefore, in *Hydra* AMPs are not unselective bacteriocides, but rather specific host regulators of microbiome community composition.

Impacts of animal host association on bacteria

Association with animal hosts can convey a multitude of benefits to bacteria that can ultimately increase bacterial growth and productivity. One well-characterized benefit of host association is the provision of nutrients either directly or indirectly from the host (107). In the pea aphid-*Buchnera* symbiosis, the host directly provides the symbiont with the non-essential amino acid glutamate, which is then processed by the symbiont to form the essential amino acids required by the host (108). Other more indirect methods by which the host provides nutrients include sloppy feeding or host excretions wherein nutrients leaked by the host are absorbed by epibiotic bacteria (109-111). Host movement across chemical gradients present in habitats such as vent plumes and anoxic sediments can also provide chemoautotrophic symbionts with predictable access to essential inorganic compounds that would otherwise be difficult to obtain (36). Another important role that hosts can play is as refuge from environmental stressors and open ocean environments that are not permissive to certain bacteria such as strict anaerobes (13, 14, 112, 113). As a whole, host-associated bacteria often demonstrate increased growth rates (36, 114), productivity (115), and fitness (116) relative to their free-living counterparts.

Association with animal hosts can also potentially provide long-term benefits that have important implications for microbial evolution, including enhanced persistence in the environment and increased genetic diversity (14, 117). Attachment to motile organisms (118) or hosts extensively dispersed by ocean currents (119) permits widespread dispersal of "hitchhiking" bacteria to otherwise unattainable new environments. Animal hosts can also serve as inocula to free-living microbial communities due to the periodic release of the host-associated symbionts into the natural environment (36, 88). In hydrothermal vent systems, this constant release of symbionts into the environment enables these bacterial species to dominate the free-living community and outcompete other bacterial species for resources (36). Attachment to hosts such as zooplankton may also enhance persistence of bacteria by triggering a viable but nonculturable (VBNC) state, which is thought to enhance bacterial survival during unfavorable environmental conditions such as dramatic shifts in salinity and temperature (14, 120-122). In addition to enhancing the persistence of bacteria, host association can expose bacteria to novel DNA. The host surface can serve as a forum where closely-related bacteria that would otherwise be separated in the open environment can convene and exchange genetic material via horizontal

gene transfer (HGT) (1, 8). Such genetic exchanges can perpetuate the specificity of the host-associated lifestyle in the case of the bacteria *V. fischeri* and *Xenorhabdus nemtaophila* whose genetic ability to initiate colonization of their specific hosts, bobtail squid and nematodes, respectively, may have been acquired via HGT (66, 67). Overall, association with animal hosts has profound implications for microbial ecology and evolution as well as on biogeochemical cycling, through direct and indirect effects that increase the growth, productivity, persistence, and genetic potential of the bacterial symbionts.

GOALS OF THESIS:

Roughly ten years ago, Dr. Rita Colwell and her team discovered an easy way to dramatically reduce the spread of cholera in Bangladesh. Simple filtration of drinking water through folded sari clothes reduced the cholera infection rate of Bangladeshi residents by 48% (123). The filtration removed copepods harboring high amounts of the cholera bacteria *Vibrio cholerae*, which can heavily colonize the oral region and the egg sac of copepods (124). Although *V. cholerae* can attach to other chitinous zooplankton, the relationship between *V. cholerae* and copepods is the most extensively studied. Associations with copepods provide protection to *V. cholerae*, suggested by the bacteria's enhanced survival and resistance to stressors in the presence of live copepods (124). The dispersal of copepods and their eggs by ocean currents is also a major mode of extended cholera distribution (119), underscoring the potential consequences of copepod-*Vibrio* associations for human health. We now know that many types of *Vibrio* bacteria, including pathogens, are prevalent on copepods (14, 125-130) and that colonization of copepods can have dramatic impacts on the proliferation, virulence, and physiology of many *Vibrio* species (122, 131-134).

Although the literature generally classifies the association of *V. cholerae* and other vibrios with copepods as commensal, the nature of the relationship between copepods and their associated bacteria is unknown (126). Copepods have often been considered passive vectors of bacteria with little attention to whether copepods are affected by bacterial attachment and whether they can regulate bacterial associations through mechanisms such as the innate immune response. However, host gene function is known to be important in enriching for beneficial

microbes in systems as diverse as plants selecting bacterial *Rhizobium* strains (135) and the bobtail squid *E. scolopes* selecting for luminous *V. fischeri* (136). Evidence that host factors can be essential to selecting and maintaining bacterial symbionts in a diversity of animals continues to mount (5, 10, 93, 95, 105, 137-139), yet it is still unknown whether copepods *actively* interact with their microbial communities and select for or against particular bacteria. Furthermore, although copepods harbor distinct bacterial communities (140, 141), it is not understood how stable these bacteria-copepod associations are and whether changes in copepod physiology may affect bacterial community assembly. In light of copepods' abundance, pervasiveness across aquatic habitats, and diverse microbiomes, copepod physiology could have an important influence on microbial ecology that has not yet been explored.

This thesis investigates the potential role that copepod physiology may play in regulating *Vibrio* association and the community structure of its microbiome. To this end, the intrinsic ability of oceanic copepod hosts to transcriptionally respond to mild stressors is first investigated (Chapter 2). Specifically, the transcriptional regulation of several heat shock proteins (Hsps), a highly conserved superfamily of molecular chaperones, was examined in the copepod *Calanus finmarchicus*. Several Hsps (Hsp21, Hsp22, and Hsp70A) were induced by handling stress and diapause (Hsp22), suggesting these genes are a conserved element of the copepod's transcriptional response to stressful conditions. To then investigate whether copepod hosts respond to and regulate their microbiota, the transcriptomic response of an estuarine copepod *Eurytemora affinis* to two distinct *Vibrio* species, a free-living (*V. ordalii 12B09*) and a zooplankton specialist (*V. sp. F10*), was examined using next-generation sequencing technologies (Chapter 3). Our findings provide evidence that the copepod *E. affinis* does distinctly recognize and respond to colonizing vibrios via transcriptional regulation of elements involved in the innate immune response and cuticle integrity. Our work also suggests that association with *E. affinis* can significantly impact the physiology of *Vibrio* colonists. Finally, the inter-individual variability of the *C. finmarchicus* microbiome was examined to identify how specifically and predictably bacterial communities assemble on copepods (Chapter 4). Our findings suggest that copepods have a predictable "core microbiome" that persists throughout the host's entrance into diapause, a dormancy period characterized by dramatic physiological changes in the host. Furthermore, the differences in the structure of the "flexible" microbiome in diapausing and active individuals appears to be partially driven by factors including the

copepod's feeding history, body size, and microbial interactions. The findings of this thesis work highlight the role of copepods as dynamic reservoirs of diverse microbial communities and implicate copepod host physiology and molecular responses as important contributors to the activity, abundance, and community structure of its microbiota (Chapter 5).

REFERENCES:

1. **McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Loso T, Douglas AE, Dubilier N, Eberl G, Fukami T, Gilbert SF, Hentschel U, King N, Kjelleberg S, Knoll AH, Kremer N, Mazmanian SK, Metcalf JL, Nealson K, Pierce NE, Rawls JF, Reid A, Ruby EG, Rumpho M, Sanders JG, Tautz D, Wernegreen JJ.** 2013. Animals in a bacterial world, a new imperative for the life sciences. Proc. Natl. Acad. Sci. U.S.A. **110:**3229-3236.
2. **McFall-Ngai MJ.** 2014. The importance of microbes in animal development: lessons from the squid-*Vibrio* symbiosis. Annu. Rev. Microbiol. **68:**177-194.
3. **Ezenwa VO, Gerardo NM, Inouye DW, Medina M, Xavier JB.** 2012. Animal behavior and the microbiome. Science **338:**198-199.
4. **Carlos C, Torres TT, Ottoboni LM.** 2013. Bacterial communities and species-specific associations with the mucus of Brazilian coral species. Sci. Rep. **3:**1624.
5. **Franzenburg S, Walter J, Künzel S, Wang J, Baines JF, Bosch TCG, Fraune S.** 2013. Distinct antimicrobial peptide expression determines host species-specific bacterial associations. Proc. Natl. Acad. Sci. U.S.A. **110:**E3730-E3738.
6. **Apprill A, Robbins J, Eren AM, Pack AA, Reveillaud J, Mattila D, Moore M, Niemeyer M, Moore KMT, Mincer TJ.** 2014. Humpback whale populations share a core skin bacterial ommunity: towards a health index for marine mammals? PLoS ONE **9:**e90785.
7. **Preheim SP, Boucher Y, Wildschutte H, David LA, Veneziano D, Alm EJ, Polz MF.** 2011. Metapopulation structure of *Vibrionaceae* among coastal marine invertebrates. Environ. Microbiol. **13:**265-275.
8. **Grossart H-P, Kiørboe T, Tang K, Ploug H.** 2003. Bacterial colonization of particles: growth and interactions. Appl. Environ. Microbiol. **69:**3500-3509.
9. **Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP.** 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environ. Microbiol. **8:**732-740.
10. **Bright M, Bulgheresi S.** 2010. A complex journey: transmission of microbial symbionts. Nat. Rev. Microbiol. **8:**218-230.
11. **Buchon N, Broderick NA, Lemaitre B.** 2013. Gut homeostasis in a microbial world: insights from *Drosophila melanogaster*. Nat. Rev. Microbiol. **11:**615-626.
12. **Douglas AE.** 2009. The microbial dimension in insect nutritional ecology. Func. Ecol. **23:**38-47.
13. **Chowdhury MAR, Huq A, Xu B, Madeira FJB, Colwell RR.** 1997. Effect of alum on free-living and copepod-associated *Vibrio cholerae* O1 and O139. Appl. Environ. Microbiol. **63:**3323-3326.
14. **Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR.** 1983. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Appl. Environ. Microbiol. **45:**275-283.
15. **Gilbert SF, Sapp J, Tauber AI.** 2012. A symbiotic view of life: we have never been individuals. Q. Rev. Biol. **87:**325-341.
16. **Moran NA.** 2007. Symbiosis as an adaptive process and source of phenotypic complexity. Proc. Natl. Acad. Sci. U.S.A. **1:**8627-8633.
17. **[Editorial].** 2011. Microbiology by numbers. Nat. Rev. Microbiol. **9:**628-628.
18. **Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA.** 2005. Diversity of the human intestinal microbial flora. Science **308:**1635-1638.
19. **Zientz E, Dandekar T, Gross R.** 2004. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiol. Mol. Biol. Rev. **68:**745-770.
20. **Feldhaar H.** 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. Ecol. Entomol. **36:**533-543.
21. **Montgomery MK, McFall-Ngai M.** 1994. Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. Development **120:**1719-1729.
22. **Nussbaumer AD, Fisher CR, Bright M.** 2006. Horizontal endosymbiont transmission in hydrothermal vent tubeworms. Nature **441:**345-348.
23. **Hadfield MG.** 2011. Biofilms and marine invertebrate larvae: what bacteria produce that larvae use to choose settlement sites. Ann. Rev. Mar. Sci. **3:**453-470.
24. **Huang Y, Callahan S, Hadfield MG.** 2012. Recruitment in the sea: bacterial genes required for inducing larval settlement in a polychaete worm. Sci. Rep. **2:**228.

25. **Heijtz RD, Wang S, Anuar F, Qian Y, Björkholm B, Samuelsson A, Hibberd ML, Forssberg H, Pettersson S.** 2011. Normal gut microbiota modulates brain development and behavior. Proc. Natl. Acad. Sci. U.S.A. **108:**3047-3052.
26. **Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E.** 2010. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. **107:**20051-20056.
27. **Verhulst NO, Qiu YT, Beijleveld H, Maliepaard C, Knights D, Schulz S, Berg-Lyons D, Lauber CL, Verduijn W, Haasnoot GW, Mumm R, Bouwmeester HJ, Claas FHJ, Dicke M, van Loon JJA, Takken W, Knight R, Smallegange RC.** 2011. Composition of human skin microbiota affects attractiveness to malaria mosquitoes. PLoS ONE **6:**e28991.
28. **Cryan JF, Dinan TG.** 2012. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat. Rev. Neurosci. **13:**701-712.
29. **Bravo JA, Forsythe P, Chew MV, Escaravage E, Savignac HlnM, Dinan TG, Bienenstock J, Cryan JF.** 2011. Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. Proc. Natl. Acad. Sci. U.S.A. **108:**16050-16055.
30. **Dubilier N, Bergin C, Lott C.** 2008. Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. Nat. Rev. Microbiol. **6:**725-740.
31. **Wang JW, Wu YN, Yang GX, Aksoy S.** 2009. Interactions between mutant *Wigglesworthia* and the tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission Am. J. Trop. Med. Hyg. **81:**291-291.
32. **Brune A, Ohkuma M.** 2011. Role of the termite gut microbiota in symbiotic digestion, p. 439-475. *In* Bignell DE, Roisin Y, Lo N (ed.), Biology of Termites: A Modern Synthesis.
33. **Felbeck H, Somero GN.** 1982. Primary production in deep-sea hydrothermal vent organisms: roles of sulfide-oxidizing bacteria. Trends Biochem. Sci. **7:**201-204.
34. **Zierenberg RA, Adams MWW, Arp AJ.** 2000. Life in extreme environments: hydrothermal vents. Proc. Natl. Acad. Sci. U.S.A. **97:**12961-12962.
35. **Stewart FJ, Cavanaugh CM.** 2006. Symbiosis of thioautotrophic bacteria with *Riftia pachyptila*. Prog. Mol. Subcell. Biol. **41:**197-225.
36. **Polz MF, Ott, J.A., Bright, M., Cavanaugh, C.M. .** 2000. When bacteria hitch a ride. ASM News **66:**531-539.
37. **Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, Hattori M, Aksoy S.** 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. Nat. Genet. **32:**402-407.
38. **Nogge G.** 1981. Significance of symbionts for the maintenance of an optional nutritional state for successful reproduction in hematophagous arthropods. Parasitology **82:**101-104.
39. **Wilson ACC, Ashton PD, Calevro F, Charles H, Colella S, Febvay G, Jander G, Kushlan PF, Macdonald SJ, Schwartz JF, Thomas GH, Douglas AE.** 2010. Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. Insect Mol. Biol. **19:**249-258.
40. **Douglas AE, Prosser WA.** 1992. Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrthosiphon pisum*) symbiosis. J. Insect Physiol. **38:**565-568.
41. **Martin MM, Jones CG, Bernays EA.** 1991. The evolution of cellulose digestion in insects. Philos. Trans. R. Soc. Lond., B, Biol. Sci. **333:**281-288.
42. **Huang S, Sheng P, Zhang H.** 2012. Isolation and identification of cellulolytic bacteria from the gut of *Holotrichia parallela* larvae (Coleoptera: Scarabaeidae). Int. J. Mol. Sci. **13:**2563-2577.
43. **Cook DM.** 2010. Bacterial communities associated with the hindgut of *Tipula abdominalis* larvae (Diptera: Tipulidae), a natural biorefinery. Ph.D. Thesis. University of Georgia, Athens,GA.
44. **Ohkuma M.** 2003. Termite symbiotic systems: efficient bio-recycling of lignocellulose. Appl. Microbiol. Biotechnol. **61:**1-9.
45. **Cleveland LR.** 1923. Symbiosis between termites and their intestinal protozoa. Proc. Natl. Acad. Sci. U.S.A. **9:**424-428.
46. **Purdy KJ.** 2007. The distribution and diversity of *Euryarchaeota* in termite guts, p. 63-80. *In* Laskin AI, Sariaslani S, Gadd GM (ed.), Advances in Applied Microbiology, vol. 62.
47. **Dillon RJ, Dillon VM.** 2004. The gut bacteria of insects: nonpathogenic interactions. Annu. Rev. Entomol. **49:**71-92.

48. **Gil-Turnes MS, Hay ME, Fenical W.** 1989. Symbiotic marine bacteria chemically defend crustacean embryos from a pathogenic fungus. Science **246:**116-118.
49. **Gil-Turnes MS, Fenical W.** 1992. Embryos of *Homarus americanus* are protected by epibiotic bacteria. Biol. Bull. **182:**105-108.
50. **Blum JE, Fischer CN, Miles J, Handelsman J.** 2013. Frequent replenishment sustains the beneficial microbiome of *Drosophila melanogaster*. mBio **4**(6)**:**10.1128/mBio.00860-00813.
51. **Koch H, Schmid-Hempel P.** 2012. Gut microbiota instead of host genotype drive the specificity in the interaction of a natural host-parasite system. Ecol. Lett. **15:**1095-1103.
52. **Koch H, Schmid-Hempel P.** 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. Proc. Natl. Acad. Sci. U.S.A. **108:**19288-19292.
53. **Buchon N, Broderick NA, Chakrabarti S, Lemaitre B.** 2009. Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in *Drosophila*. Genes Dev. **23:**2333-2344.
54. **Rodrigues J, Brayner FA, Alves LC, Dixit R, Barillas-Mury C.** 2010. Hemocyte differentiation mediates innate immune memory in *Anopheles gambiae* mosquitoes. Science **329:**1353-1355.
55. **Weiss BL, Wang J, Aksoy S.** 2011. Tsetse immune system maturation requires the presence of obligate symbionts in larvae. PLoS Biol. **9:**e1000619.
56. **Pais R, Lohs C, Wu Y, Wang J, Aksoy S.** 2008. The obligate mutualist *Wigglesworthia glossinidia* influences reproduction, digestion, and immunity processes of its host, the tsetse fly. Appl. Environ. Microbiol. **74:**5965-5974.
57. **Grossart HP, Dziallas C, Tang KW.** 2009. Bacterial diversity associated with freshwater zooplankton. Environ. Microbiol. Rep. **1:**50-55.
58. **Dang HY, Lovell CR.** 2000. Bacterial primary colonization and early succession on surfaces in marine waters as determined by amplified rRNA gene restriction analysis and sequence analysis of 16S rRNA genes. Appl. Environ. Microbiol. **66:**467-475.
59. **Banning NC, Gleeson DB, Grigg AH, Grant CD, Andersen GL, Brodie EL, Murphy DV.** 2011. Soil microbial community successional patterns during forest ecosystem restoration. Appl. Environ. Microbiol. **77:**6158-6164.
60. **Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ, Polz MF.** 2013. Reproducibility of *Vibrionaceae* population structure in coastal bacterioplankton. ISME J **7:**509-519.
61. **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.
62. **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. U.S.A. **111:**5622-5627.
63. **Cordero OX, Wildschutte H, Kirkup B, Proehl S, Ngo L, Hussain F, Le Roux F, Mincer T, Polz MF.** 2012. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science **337:**1228-1231.
64. **Tait K, Sutherland IW.** 2002. Antagonistic interactions amongst bacteriocin-producing enteric bacteria in dual species biofilms. J. Appl. Microbiol. **93:**345-352.
65. **Rao D, Webb JS, Kjelleberg S.** 2005. Competitive interactions in mixed-species biofilms containing the marine bacterium *Pseudoalteromonas tunicata*. Appl. Environ. Microbiol. **71:**1729-1736.
66. **Cowles CE, Goodrich-Blair H.** 2008. The *Xenorhabdus nematophila* nilABC genes confer the ability of *Xenorhabdus spp*. to colonize *Steinernema carpocapsae* nematodes. J. Bacteriol. **190:**4121-4128.
67. **Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG.** 2009. A single regulatory gene is sufficient to alter bacterial host range. Nature **458:**215-U217.
68. **Burmølle M, Webb JS, Rao D, Hansen LH, Sørensen SJ, Kjelleberg S.** 2006. Enhanced biofilm formation and increased resistance to antimicrobial agents and bacterial invasion are caused by synergistic interactions in multispecies biofilms. Appl. Environ. Microbiol. **72:**3916-3923.
69. **Lazarevic V, Whiteson K, Hernandez D, Francois P, Schrenzel J.** 2011. Study of inter- and intra-individual variations in the salivary microbiota. BMC Genomics **11:**1471-2164.
70. **Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, Gordon JI, Knight R.** 2011. Moving pictures of the human microbiome. Genome Biol. **12:**2011-2012.
71. **Davenport ER, Mizrahi-Man O, Michelini K, Barreiro LB, Ober C, Gilad Y.** 2014. Seasonal variation in human gut microbiome composition. PLoS ONE **9:**e90731.

72. **Wang J, Linnenbrink M, Kunzel S, Fernandes R, Nadeau MJ, Rosenstiel P, Baines JF.** 2014. Dietary history contributes to enterotype-like clustering and functional metagenomic content in the intestinal microbiome of wild mice. Proc. Natl. Acad. Sci. U.S.A. **111:**E2703-E2710.
73. **Spor A, Koren O, Ley R.** 2011. Unravelling the effects of the environment and host genotype on the gut microbiome. Nat. Rev. Microbiol. **9:**279-290.
74. **The Human Microbiome Project Consortium.** 2012. Structure, function and diversity of the healthy human microbiome, p. 207-214, Nature, vol. 486.
75. **King GM, Judd C, Kuske CR, Smith C.** 2012. Analysis of stomach and gut microbiomes of the eastern oyster *Crassostrea virginica* from coastal Louisiana, USA. PLoS ONE **7:**e51475.
76. **Aksoy E, Telleria EL, Echodu R, Wu Y, Okedi LM, Weiss BL, Aksoy S, Caccone A.** 2014. Analysis of multiple tsetse fly populations in Uganda reveals limited diversity and species-specific gut microbiota. Appl. Environ. Microbiol. **80:**4301-4312.
77. **Waterfield NR, Wren BW, French-Constant RH.** 2004. Invertebrates as a source of emerging human pathogens. Nat. Rev. Microbiol. **2:**833-841.
78. **Lemaitre B, Hoffmann J.** 2007. The host defense of *Drosophila melanogaster*. Annu. Rev. Immunol. **25:**697-743.
79. **Vallet-Gely I, Lemaitre B, Boccard F.** 2008. Bacterial strategies to overcome insect defenses. Nat. Rev. Microbiol. **6:**302-313.
80. **Lehane MJ.** 1997. Peritrophic matrix structure and function. Annu. Rev. Entomol. **42:**525-550.
81. **Yoshikoshi K, Ko Y.** 1988. Structure and function of the peritrophic membranes of copepods. Nippon Suisan Gakkai Shi **54:**1077-1082.
82. **Buchon N, Broderick NA, Poidevin M, Pradervand S, Lemaitre B.** 2009. *Drosophila* intestinal response to bacterial infection: activation of host defense and stem cell proliferation. Cell Host Microbe **5:**200-211.
83. **Kuraishi T, Binggeli O, Opota O, Buchon N, Lemaitre B.** 2011. Genetic evidence for a protective role of the peritrophic matrix against intestinal bacterial infection in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. **108:**15966-15971.
84. **Sonnenburg JL, Angenent LT, Gordon JI.** 2004. Getting a grip on things: how do communities of bacterial symbionts become established in our intestine? Nat. Immunol. **5:**569-573.
85. **Nyholm SV, McFall-Ngai M.** 2004. The winnowing: establishing the squid-vibrio symbiosis. Nat. Rev. Microbiol. **2:**632-642.
86. **Nyholm SV, Deplancke B, Gaskins HR, Apicella MA, McFall-Ngai MJ.** 2002. Roles of *Vibrio fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont colonization of the *Euprymna scolopes* light organ. Appl. Environ. Microbiol. **68:**5113-5122.
87. **DeLoney-Marino CR, Wolfe AJ, Visick KL.** 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.
88. **Boettcher KJ, Ruby EG, McFall-Ngai MJ.** 1996. Bioluminescence in the symbiotic squid *Euprymna scolopes* is controlled by a daily biological rhythm. J. Comp. Physiol. A **179:**65-73.
89. **Hosokawa T, Kikuchi Y, Shimada M, Fukatsu T.** 2008. Symbiont acquisition alters behaviour of stinkbug nymphs. Biol. Lett. **4:**45-48.
90. **Troyer K.** 1984. Behavioral acquisition of the hindgut fermentation system by hatchling *Iguana iguana*. Behav. Ecol. Sociobiol. **14:**189-193.
91. **Sadd BM, Schmid-Hempel P.** 2006. Insect immunity shows specificity in protection upon secondary pathogen exposure. Curr. Biol. **16:**1206-1210.
92. **Little TJ, O'Connor B, Colegrave N, Watt K, Read AF.** 2003. Maternal transfer of strain-specific immunity in an invertebrate. Curr. Biol. **13:**489-492.
93. **Bulgheresi S, Schabussova I, Chen T, Mullin NP, Maizels RM, Ott JA.** 2006. A new c-type lectin similar to the human immunoreceptor DC-SIGN mediates symbiont acquisition by a marine nematode. Appl. Environ. Microbiol. **72:**2950-2956.
94. **Binggeli O, Neyen C, Poidevin M, Lemaitre B.** 2014. Prophenoloxidase activation is required for survival to microbial infections in *Drosophila*. PLoS Pathog. **10:**e1004067.
95. **Kvennefors ECE, Leggat W, Hoegh-Guldberg O, Degnan BM, Barnes AC.** 2008. An ancient and variable mannose-binding lectin from the coral *Acropora millepora* binds both pathogens and symbionts. Dev. Comp. Immunol. **32:**1582-1592.

96. **Wang X-W, Xu J-D, Zhao X-F, Vasta GR, Wang J-X.** 2014. A shrimp c-type lectin inhibits proliferation of the hemolymph microbiota by maintaining the expression of antimicrobial peptides. J. Biochem. **289:**11779-11790.
97. **Cash HL, Whitham CV, Behrendt CL, Hooper LV.** 2006. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science **313:**1126-1130.
98. **Medzhitov R.** 2007. Recognition of microorganisms and activation of the immune response. Nature **449:**819-826.
99. **Cerenius L, Lee BL, Söderhäll K.** 2008. The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol. **29:**263-271.
100. **Rowley AF, Powell A.** 2007. Invertebrate immune systems-specific, quasi-specific, or nonspecific? J. Immunol. **179:**7209-7214.
101. **Cerenius L, Söderhäll K.** 2004. The prophenoloxidase-activating system in invertebrates. Immunol. Rev. **198:**116-126.
102. **Jiravanichpaisal P, Lee BL, Soderhall K.** 2006. Cell-mediated immunity in arthropods: Hematopoiesis, coagulation, melanization and opsonization. Immunobiology **211:**213-236.
103. **Tincu JA, Taylor SW.** 2004. Antimicrobial peptides from marine invertebrates. Antimicrob. Agents Chemother. **48:**3645-3654.
104. **Bachere E, Gueguen Y, Gonzalez M, de Lorgeril J, Garnier J, Romestand B.** 2004. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas*. Immunol. Rev. **198:**149-168.
105. **Fraune S, Augustin R, Anton-Erxleben F, Wittlieb J, Gelhaus C, Klimovich VB, Samoilovich MP, Bosch TCG.** 2010. In an early branching metazoan, bacterial colonization of the embryo is controlled by maternal antimicrobial peptides. Proc. Natl. Acad. Sci. U.S.A. **107:**18067-18072.
106. **Gorman MJ, Kankanala P, Kanost MR.** 2004. Bacterial challenge stimulates innate immune responses in extra-embryonic tissues of tobacco hornworm eggs. Insect Mol. Biol. **13:**19-24.
107. **Douglas AE.** 2009. The microbial dimension in insect nutritional ecology. Functional Ecology **23:**38-47.
108. **Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H.** 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera sp. APS*. Nature **407:**81-86.
109. **Eckert EM, Pernthaler J.** 2014. Bacterial epibionts of *Daphnia*: a potential route for the transfer of dissolved organic carbon in freshwater food webs. ISME J **8:**1808-1819.
110. **Möller EF.** 2005. Sloppy feeding in marine copepods: prey-size-dependent production of dissolved organic carbon. J. Plankton Res. **27:**27-35.
111. **Carman KR.** 1994. Stimulation of marine free-living and epibiotic bacterial activity by copepod excretions. FEMS Microbiol. Ecol. **14:**255-261.
112. **Braun ST, Proctor LM, Zani S, Mellon MT, Zehr JP.** 1999. Molecular evidence for zooplankton-associated nitrogen-fixing anaerobes based on amplification of the nifH gene. FEMS Microbiol. Ecol. **28:**273-279.
113. **Proctor LM.** 1997. Nitrogen-fixing, photosynthetic, anaerobic bacteria associated with pelagic copepods. Aquat. Microb. Ecol. **12:**105-113.
114. **Tang K.** 2005. Copepods as microbial hotspots in the ocean: Effects of host feeding activities on attached bacteria. Aquat. Microb. Ecol. **38:**31-40.
115. **Griffith PC, Douglas DJ, Wainright SC.** 1990. Metabolic activity of size-fractionated microbial plankton in estuarine, nearshore, and continental shelf waters of Georgia. Mar. Ecol. Prog. Ser. **59:**263-270.
116. **Wollenberg MS, Ruby EG.** 2012. Phylogeny and fitness of *Vibrio fischeri* from the light organs of *Euprymna scolopes* in two Oahu, Hawaii populations. ISME J **6:**352-362.
117. **Garcia JR, Gerardo NM.** 2014. The symbiont side of symbiosis: do microbes really benefit? Front. Microbiol. **5**.
118. **Grossart H-P, Dziallas C, Leunert F, Tang KW.** 2010. Bacteria dispersal by hitchhiking on zooplankton. Proc. Natl. Acad. Sci. U.S.A. **107:**11959-11964.
119. **Lipp EK, Huq A, Colwell RR.** 2002. Effects of global climate on infectious disease: the cholera model. Clin. Microbiol. Rev. **15:**757-770.
120. **Thomas KU, Joseph N, Raveendran O, Nair S.** 2006. Salinity-induced survival strategy of *Vibrio cholerae* associated with copepods in Cochin backwaters. Mar. Pollut. Bull. **52:**1425-1430.
121. **Signoretto C, Burlacchini G, Pruzzo C, Canepari P.** 2005. Persistence of *Enterococcus faecalis* in aquatic environments via surface interactions with copepods. Appl. Environ. Microbiol. **71:**2756-2761.

122. **Epstein SS, Colwell R.** 2009. Viable but Not Cultivable Bacteria, p. 121-129, Uncultivated Microorganisms, vol. 10. Springer Berlin Heidelberg.
123. **Colwell RR, Huq A, Islam MS, Aziz KMA, Yunus M, Khan NH, Mahmud A, Sack RB, Nair GB, Chakraborty J, Sack DA, Russek-Cohen E.** 2003. Reduction of cholera in Bangladeshi villages by simple filtration. Proc. Natl. Acad. Sci. U.S.A. **100:**1051-1055.
124. **Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR.** 1983. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Applied and Environmental Microbiology **45:**275-283.
125. **Heidelberg JF, Heidelberg KB, Colwell RR.** 2002. Bacteria of the gamma-subclass *Proteobacteria* associated with zooplankton in Chesapeake Bay. Appl. Environ. Microbiol. **68:**5498-5507.
126. **Rawlings TK, Ruiz GM, Colwell RR.** 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.
127. **Tamplin ML, Gauzens AL, Huq A, Sack DA, Colwell RR.** 1990. Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. Appl. Environ. Microbiol. **56:**1977-1980.
128. **Martinelli JE, Lopes RM, Rivera ING, Colwell RR.** 2011. *Vibrio cholerae* O1 detection in estuarine and coastal zooplankton. J. Plankton Res. **33:**51-62.
129. **Maugeri TL, Carbone M, Fera MT, Irrera GP, Gugliandolo C.** 2004. Distribution of potentially pathogenic bacteria as free living and plankton associated in a marine coastal zone. J. Appl. Microbiol. **97:**354-361.
130. **Gugliandolo C, Irrera GP, Lentini V, Maugeri TL.** 2008. Pathogenic *Vibrio*, *Aeromonas* and *Arcobacter spp*. associated with copepods in the Straits of Messina (Italy). Mar. Pollut. Bull. **56:**600-606.
131. **Tang KW, Turk V, Grossart H-P.** 2010. Linkage between crustacean zooplankton and aquatic bacteria. Aquat. Microb. Ecol. **61:**261-277.
132. **Kirn TJ, Jude BA, Taylor RK.** 2005. A colonization factor links *Vibrio cholerae* environmental survival and human infection. Nature **438:**863-866.
133. **Hunt DE, Gevers D, Vahora NM, Polz MF.** 2008. Conservation of the chitin utilization pathway in the *Vibrionaceae*. Appl. Environ. Microbiol. **74:**44-51.
134. **Huq A, Colwell RR, Rahman R, Ali A, Chowdhury MAR, Parveen S, Sack DA, Russekcohen E.** 1990. Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and cell cultures Appl. Environ. Microbiol. **56:**2370-2373.
135. **Smith KP, Goodman RM.** 1999. Host variation for interactions with beneficial plant-associated microbes. Annu Rev Phytopathol **37:**473-491.
136. **Thompson FL, Iida T, Swings J.** 2004. Biodiversity of *Vibrios*. Microbiol. Mol. Biol. Rev. **68:**403-431.
137. **Douglas A.** 2011. Lessons from studying insect symbioses. Cell Host Microbe **10:**359-367.
138. **Heath-Heckman EAC, Gillette AA, Augustin R, Gillette MX, Goldman WE, McFall-Ngai MJ.** 2014. Shaping the microenvironment: evidence for the influence of a host galaxin on symbiont acquisition and maintenance in the squid-vibrio symbiosis. Environ. Microbiol.**:**Epub.
139. **Login FH, Balmand S, Vallier A, Vincent-Monégat C, Vigneron A, Weiss-Gayet M, Rochat D, Heddi A.** 2011. Antimicrobial peptides keep insect endosymbionts under control. Science **334:**362-365.
140. **Turner JW, Good B, Cole D, Lipp EK.** 2009. Plankton composition and environmental factors contribute to Vibrio seasonality. ISME J **3:**1082-1092.
141. **Dziallas C, Grossart H-P, Tang KW, Nielsen TG.** 2013. Distinct communities of free-living and copepod-associated microorganisms along a salinity gradient in Godthabsfjord, West Greenland. Arctic, Antarctic, and Alpine Research **45:**471-480.

Chapter Two

Heat shock protein expression during stress and diapause in the marine copepod *Calanus finmarchicus*[1]

[1] Published in the Journal of Insect Physiology (2011) with Mark F. Baumgartner, Adam M. Reitzel, and Ann M. Tarrant

ABSTRACT:

Calanoid copepods, such as *Calanus finmarchicus*, are a key component of marine food webs. *C. finmarchicus* undergoes a facultative diapause during juvenile development, which profoundly affects their seasonal distribution and availability to their predators. The current ignorance of how copepod diapause is regulated limits understanding of copepod population dynamics, distribution, and ecosystem interactions. Heat shock proteins (*Hsp*s) are a superfamily of molecular chaperones characteristically upregulated in response to stress conditions and frequently associated with diapause in other taxa. In this study, 8 heat shock proteins were identified in *C. finmarchicus* C5 copepodids (*Hsp21, Hsp22*, *p26, Hsp90,* and 4 forms of *Hsp70*), and expression of these transcripts was characterized in response to handling stress and in association with diapause. *Hsp21, Hsp22,* and *Hsp70A* (cytosolic subfamily) were induced by handling stress. Expression of *Hsp70A* was also elevated in shallow active copepodids relative to deep diapausing copepodids, which may reflect induction of this gene by varied stressors in active animals. In contrast, expression of *Hsp22* was elevated in deep diapausing animals; *Hsp22* may play a role both in short-term stress responses and in protecting proteins from degradation during diapause. Expression of most of the *Hsp*s examined did not vary in response to diapause, perhaps because the diapause of *C. finmarchicus* is not associated with the extreme environmental conditions (e.g., freezing, desiccation) experienced by many other taxa, such as overwintering insects or *Artemia* cysts.

INTRODUCTION:

Marine food webs depend on herbivorous zooplankton such as copepods to process and repackage energy harnessed by photosynthetic primary producers. Representing often more than half of the zooplankton biomass in the temperate North Atlantic (Planque and Batten, 2000; Williams et al., 1994), the calanoid copepod *Calanus finmarchicus* provides an essential route of energy transfer to higher trophic levels either via direct predation or trophic links. The ecological success of *C. finmarchicus* is facilitated by its ability to avoid adverse seasonal conditions and high predation risk (Kaartvedt, 1996) by vertically migrating to depth and entering a facultative diapause during the last juvenile stages (typically stage C5) (Hirche, 1996). In the Gulf of Maine, a portion of the *C. finmarchicus* population enters into this diapause period during the warm spring and summer months and exits during mid to late winter to molt into adults; however, some C5 juveniles skip diapause and proceed to molt into adults, reproduce, and spawn another generation (Durbin et al., 2000; Durbin et al., 1997). The appropriate translation of environmental cues into the physiological changes required for diapause in *C. finmarchicus* suggests the involvement of complex but flexible internal regulatory processes. Despite the ecological implications of seasonal dormancy, shockingly little is understood about the factors that regulate the adaptive diapause response in calanoid copepods.

True diapause, an endogeneously regulated process during which organisms undergo progressive physiological changes over several successive phases, is distinct from quiescence, an immediate response to changes in limiting environmental factors that is not restricted to a specific ontogenetic stage (Kostál, 2006). Diapause is characterized by a persistent reduction of metabolism, increased stress resistance, and an arrest of development at a specific life stage (Kostál, 2006). Many aspects of true diapause are observed in the physiology of dormant *C. finmarchicus*, including a preparatory phase that precedes unfavorable environmental conditions (Dahms, 1995; Hirche, 1996), a dormancy phase characterized by an endogeneous arrest in development, reduction of metabolism and respiration (Hirche, 1983; Ingvarsdóttir et al., 1999), and gene expression patterns consistent with increased stress resistance (Tarrant et al., 2008), and a distinct post-dormancy phase when development resumes (Hirche, 1996). Therefore, we refer here to the *C. finmarchicus* dormancy as a facultative diapause; however, we recognize that not all individuals or populations of *C. finmarchicus* appear to enter a diapause state at the same time or of similar intensity (Hirche, 1996), which is dissimilar from the true diapause observed in

other well-characterized organisms (e.g. *Artemia* cysts, overwintering insects).

Diapause in *C. finmarchicus* is identified by a combination of classic behavioral, morphological, and biochemical characteristics: diapausing C5 copepodids accumulate at depths below 200 to 300 m in oceanic waters (Heath et al., 2004; Miller et al., 1991; Sameoto and Herman, 1990), have empty guts with thin epithelia (Bonnet et al., 2007; Hirche, 1983), and have large oil sacs (Miller et al. 2000). Indicative of arrested development, diapausing copepods have reduced transcriptional activity (i.e. low RNA:DNA ratios) (Wagner et al., 1998), low ecdysteroid levels (Johnson, 2004), and delayed molt progression (i.e. cessation of tooth formation) (Miller et al., 1991). Molecular markers of diapause include low expression of genes related to lipid synthesis, transport, and storage (*ELOV*, *FABP*, *RDH*) and high expression of *ferritin* and *ecdysteroid receptor (EcR)* in diapausing copepods (Tarrant et al., 2008). Further exploration of *C. finmarchicus* diapause using molecular techniques is required to understand the mechanisms regulating the physiological changes that characterize the diapause response.
Heat shock proteins (*Hsp*s) are a highly conserved superfamily of molecular chaperones that facilitate proper protein folding and localization while preventing protein aggregation (Feder and Hofmann, 1999; Hartl and Hayer-Hartl, 2002). Furthermore, previous research has shown that *Hsp*s play a major role in diapause regulation in a wide range of organisms (Denlinger et al., 2001; MacRae, 2010; Qiu, 2008; Yuan et al., 1996). Induction of *Hsp*s by protein denaturating stressors (e.g., heat, toxins) appears to be an ancient and universal response within all examined taxa ranging from bacteria to plants, flies, and human beings (Hansen et al., 2008; Lindquist, 1986; Sorensen et al., 2003). In addition to their well-known roles in stress tolerance, *Hsp*s are integral to normal cell growth and development. Through protein-protein interactions, *Hsp*s help to regulate fundamental cellular processes such as protein turnover, mitochondrial and endoplasmic reticulum trafficking, cell cycle progression, and steroid signaling (Beato and Klug, 2000; Helmbrecht et al., 2000; Pratt, 1997; Taipale et al., 2010). During diapause, *Hsp*s are thought to contribute to cell cycle arrest and increased stress (e.g. cold) resistance (Denlinger et al., 2001; MacRae, 2010; Rinehart et al., 2007).

Unlike the classic stress response which is characterized by transient and universal upregulation of a wide range of *Hsp*s, *Hsp* expression patterns observed during dormancy may be prolonged and highly variable among species and *Hsp* types (Denlinger et al., 2001). Comparison of *Hsp70* regulation during the diapause periods of several insect species

demonstrates how the participation of a given class of *Hsp* in the diapause response can differ significantly; for example, *Hsp70* is highly expressed during the pupal diapause of the flesh fly *Sarcophaga crassipalpis* (Rinehart and Denlinger, 2000), but expression is low during the adult diapause of the Colorado potato beetle *Leptinotarsa decemlineata* (Yocum, 2001) and the larval diapause of the bamboo borer *Omphisa fuscidentalis* (Tungjitwitayakul et al., 2008). In addition, different classes of *Hsp*s can play distinct roles in dormancy *within* a species, as indicated by discordant expression patterns: in the pupal diapause of *S. crassipalpis, Hsp90* is downregulated, while *Hsp70* and several small *Hsp*s are upregulated (Rinehart and Denlinger, 2000; Rinehart et al., 2007; Rinehart et al., 2000). However in some species, such as in the fruit fly *Drosophila triauraria*, *Hsp*s do not appear to participate in diapause at all (Goto and Kimura, 2004; Goto et al., 1998).

In light of the involvement (albeit varied) of *Hsp*s in the diapause response of many organisms, we hypothesized that *Hsp*s play a role in regulating *C. finmarchicus* diapause. In addition to their role in the stress response, *Hsp70* and *Hsp90* expression varies during the crustacean molt cycle (Cesar and Yang, 2007; Spees et al., 2003) and could therefore help regulate the developmental delay associated with diapause in *C. finmarchicus*. Production of *Hsp70* and *Hsp90* transcripts in *C. finmarchicus* has been previously examined in response to stressors such as increased temperature and exposure to toxins (Hansen et al., 2007; Hansen et al., 2008; Voznesensky et al., 2004), but not as a factor regulating diapause. Small *Hsp*s have been shown to play an important role in stress tolerance during diapause in the crustacean *Artemia franciscana* (Clegg et al., 1999; Qiu, 2008; Qiu and MacRae, 2008), but have yet to be investigated in calanoid copepods. In this study, we examined the expression patterns of several large and small *Hsp*s (*Hsp90*, 4 forms of *Hsp70*, *Hsp21*, *Hsp22*, *p26*) in individual diapausing and active *C. finmarchicus* C5 copepodids by quantitative real-time PCR (qRT-PCR). We also characterized expression of these genes during exposure to handling stress to identify inducible forms of *Hsp*s and to confirm that diapause-associated patterns of *Hsp* expression were not a byproduct of incidental stressors. We further examined the phylogenetic relationships among the large *Hsp*s to facilitate comparisons with inducible or diapause-associated *Hsp*s in other taxa. This study represents the first characterization of *Hsp* expression in association with diapause in a calanoid copepod species and expands the current understanding of the molecular regulation of diapause.

MATERIALS AND METHODS:

Identification and cloning of C. finmarchicus Hsp genes

Material for initial cloning of *Hsp*s was obtained from bulk samples of *C. finmarchicus* C5 copepodids that were collected in 2005, preserved in RNAlater (Ambion) at -80°C, and described previously (Tarrant et al., 2008). We searched the NCBI Expressed Sequence Tag (EST) database for *C. finmarchicus Hsp* sequences using the tblastn algorithm with selected crustacean *Hsp21*, *Hsp22*, *p26*, *Hsp70* and *Hsp90* sequences (Table 1). Specific primers were designed and commercially synthesized (Eurofins MWG Operon) to amplify partial *Hsp* sequences for cloning. *Hsp*s were amplified from pooled *C. finmarchicus* cDNA using 0.25 µl Amplitaq gold polymerase per 50 µl reaction. The PCR conditions were as follows: 94°C/10 min; 40 cycles of 94°C/15 sec, 60-67°C/30 sec, 68°C/7 min; 68°C/7 min; hold at 4°C. Products were visualized on 1% agarose gels, excised, and purified using the MinElute gel extraction kit (Qiagen). PCR products were cloned into pGEM-T Easy (Promega) and sequenced.

Phylogenetic analysis

The phylogenetic relationships among the *C. finmarchicus Hsp70* and *Hsp90* partial sequences and *Hsp*s from a broad selection of taxa were determined using maximum likelihood analyses. We retrieved representative *Hsp70* sequences from yeast, bacteria, red algae, human, insects, and crustaceans (Table B.1), including two *Hsp70* sequences that had previously been reported from *C. finmarchicus* (Hansen et al., 2008; Voznesensky et al., 2004). The selected *Hsp70* sequences encompass all four monophyletic groups of eukaryotic *Hsp*70s as defined by intracellular localization (i.e., cytosol, endoplasmic reticulum, mitochondria, and chloroplast) (Boorstein et al., 1994; Daugaard et al., 2007; Rhee et al., 2009). Similarly, for phylogenetic analysis of *Hsp90* we retrieved sequences from yeast, plants, human, nematodes, insects, and crustaceans that are representative of the four eukaryotic *Hsp90* subfamilies: cytosolic (*Hsp90A*), endoplasmic reticulum (*Hsp90B*), mitochondrial (*TRAP*), and chloroplast (*Hsp90C*) (Chen et al., 2006) (Table B.2). This dataset also included one *Hsp90* EST previously examined in *C. finmarchicus* (Hansen et al., 2007). Using the multiple sequence alignment program MUSCLE (Edgar, 2004), we aligned the *Hsp* sequences and then trimmed the variable 5' and 3' regions for a total sequence length of ~ 650 amino acids for each *Hsp* family. We first conducted analyses of the *Hsp70* and *Hsp90* sequences in order to classify the *Hsp* subfamily types represented by

our cloned sequences. Maximum likelihood analyses were run using RaxML (v7.0.4, Stamatakis, 2006) under a RTREV+G model of protein evolution (selected by AIC with ProtTestv1.4, Abascal et al., 2005). In a second analysis we retrieved additional insect *Hsp70* and *Hsp90* sequences that had been studied for their role in diapause (reviewed by MacRae, 2010). We aligned these sequences with representative *Hsp*s from each subfamily and conducted maximum likelihood analyses as described above. Support for nodes was assessed as a proportion of 1000 bootstrap replicates and the most likely trees were constructed and visualized in FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/) with bootstrap values > 50% reported.

Handling stress effects on gene expression

To test whether handling stress affects *Hsp* expression in *C. finmarchicus*, shallow C5 copepodids were collected in the southwestern Gulf of Maine at a station 25.5 km east/northeast of Chatham, Massachusetts, USA (41° 46' N/ 69° 38' W) on 18 May 2010 during a cruise aboard the NOAA Ship *Delaware II*. Zooplankton were collected between 13:10 and 13:30 local time from 0-20 m using a 70-cm diameter ring outfitted with a 150 μm conical mesh net. Once the net was recovered, the contents of the cod end were poured onto a 150 μm mesh sieve and 10 ml of copepods were added to each of three, covered, black, 7.6-liter ice-chilled containers of ambient filtered seawater stored in several closed ice chests. To mimic the stress that may be experienced during extended waiting times associated with the processing of collected copepod samples, each 7.6-liter ice-chilled container was left for a specific time (t = 0, 2, 3 hours) before its contents were gently sieved and transferred to an ice-chilled Petri dish. From the Petri dish, live *C. finmarchicus* C5 copepodids were individually captured using a wide-bore glass Pasteur pipette, mounted on a depression slide, photographed on a stereomicroscope equipped with a digital camera (see below), preserved in microcentrifuge tubes with 500 μl RNAlater (Ambion), and stored at -20°C until analysis. For this study, we pooled three C5 copepodids per tube to increase sample RNA yield for 10 (t = 0 h and 2 h) and 9 (t = 3 h) total samples. For each of the three time points, we used qRT-PCR (see below) to measure expression of six selected *Hsp* transcripts (i.e. *Hsp70A, Hsp70B, Hsp70D*, *Hsp90*, *Hsp21*, and *Hsp22;* in a pilot study *Hsp70C* was not reliably amplified and *p26* did not appear to be induced) and three previously identified molecular markers of diapause (*ELOV*, *RDH*, *ferritin*) (Tarrant et al., 2008).

Sampling of deep and shallow C. finmarchicus

To assess the relationship between diapause and *Hsp* expression, *C. finmarchicus* C5 copepodids were collected in the southwestern Gulf of Maine during a cruise in 2006 aboard the NOAA Ship *Albatross IV* using a 1 m2 Multiple Opening-Closing Net and Environmental Sensing System (MOCNESS) (Wiebe et al., 1976) outfitted with 333 □m mesh nets. Zooplankton were collected in two depth strata: 169 to 208 m and 0 to 50 m (hereafter referred to as the deep and shallow samples, respectively). On 20 May 2006 the deep sample was collected at a station in Franklin Basin just to the north of Georges Bank (41° 54' N/68° 16' W) between 11:15 and 11:50 local time. The shallow sample was collected at a station 83 km to the west of the Franklin Basin station in southern Wilkinson Basin (41° 53' N/69° 17' W) between 17:08 and 17:21 local time on the same day. Water depths at the stations where the deep and shallow samples were collected were 220 m and 198 m, respectively. Upon recovery of the MOCNESS, the contents of the cod end were immediately poured into a transparent, ice-chilled 1.5-liter container and stored in a closed ice chest. Live *C. finmarchicus* were periodically transferred from this container to an ice-chilled Petri dish with a 44-ml Pasteur pipette. From the Petri dish they were individually captured using a wide-bore glass Pasteur pipette, mounted on a depression slide, photographed, individually preserved in microcentrifuge tubes with 500 □l RNAlater (Ambion), and frozen (-20°C). Observations of gut contents or fecal pellet production were noted while viewing the live animals. Photographs of single animals were taken with a Canon EOS-20D digital camera mounted on a Zeiss Stemi 2000C stereomicroscope, and all measurements were calibrated with digital photographs of a stage micrometer taken just prior to sampling. The length, width, oil sac volume, and fractional fullness (after Miller et al., 2000) of each copepodid were estimated from these photographs as described previously (Tarrant et al., 2008).

RNA extraction

Total RNA was extracted from preserved individual or pooled C5 copepodids using the Aurum Total RNA Fatty and Fibrous Tissue Kit (Bio-Rad) with slight modification. C5 copepodids were homogenized in 1 ml PureZOL using a teflon homogenizer. The homogenate was added to pre-spun (16,000 x g/ 30s) Phase Lock Gel Heavy 2 ml tubes (5 PRIME), mixed with 200 µl of chloroform and centrifuged at 14,000 x g for 5 min at 4°C. The upper aqueous phase was mixed (1:1) with isopropanol, added to the extraction columns, and processed

according to the manufacturer's protocol, including on-column DNase digestion. RNA yield and purity were quantified using a Nanodrop ND-1000 spectrophotometer. RNA quality was visualized for selected individual samples on a denaturing agarose gel.

Quantitative real-time polymerase chain reaction (qRT-PCR)

ELOV, *ferritin*, *RDH*, and *16S* qRT-PCR primers and assay conditions have been described previously (Tarrant et al., 2008). Assays were developed to measure expression of *C. finmarchicus Hsp*s by qRT-PCR. Oligonucleotide primers were designed against the cloned *Hsp* sequences to target 75-150 bp amplicons (Table 2). For the handling stress assays, total RNA was extracted from pooled (3 individuals/tube) shallow C5 copepodids collected at three time treatments (0, 2 and 3 hours in an ice-chilled bucket) and used to prepare cDNA (300 ng RNA per 20 µl reaction). Expression of *ferritin*, *RDH*, *16S* and selected *Hsp*s was measured by qRTPCR using SsoFast EvaGreen Supermix (Bio-Rad) from 10 (t = 0 h and t = 2 h) or 9 (t = 3 h) pooled samples. To compare expression between deep and shallow animals, total RNA was extracted from individual deep and shallow C5 copepodids and used to prepare cDNA (200 ng RNA per 20 µl reaction). Expression of the eight *Hsp*s, *ELOV*, *ferritin*, *RDH*, and *16S* (housekeeping gene) was quantified via qRT-PCR from 21 individual deep and 21 individual shallow C5 copepodids. Expression of *Hsp21*, *Hsp22*, and *RDH* was measured using SsoFast EvaGreen Supermix, and the other genes were measured using iQ SYBR Green Supermix (Bio-Rad). In all *Hsp* assays, plasmid standards were run in duplicate on the same plate. For each gene, all samples were run in duplicate wells on a single plate.

All qRT-PCR reactions were run in an iCycler iQ Real-Time PCR detection system (Bio-Rad). The iQ SYBR Green PCR mixture consisted of 11 µl molecular biology grade distilled water, 12.5 µl iQ SYBR Green Supermix, 0.25 µl 5'-primer (10 µM), 0.25 µl 3' primer (10 µM), and 1 µl cDNA. The PCR conditions were: 95°C/30 min; 40 cycles of 95°C/15 s, 64-66°C/45 s. The SsoFast EvaGreen PCR mixture consisted of 9 µl molecular biology grade distilled water 12.5 EvaFast, 1.25 µl 5'-primer (10 µM), 1.25 µl 3' primer (10 µM), and 1 µl cDNA. The PCR conditions were: 95°C/30 min; 40 cycles of 95°C/5 s, 62-64°C/45 s. After amplification by either procedure, PCR products from each reaction were subjected to melt-curve analysis to ensure that only a single product was amplified. Selected products were also visualized on 15% TBE gels and consistently yielded single bands.

Normalization and analysis of qRT-PCR expression data

For both the handling experiment and the diapause study, *16S* ribosomal RNA was used as a housekeeping gene (Tarrant et al., 2008). Expression was normalized using the Pfaffl method (Pfaffl, 2001), a relative quantification approach that allowed for consistent comparison of gene expression without requiring plasmid standards ('Expression' = $E_{target}^{\Delta Ct\ target(sample-calibrator)}/E_{16S}^{\Delta Ct\ 16S\ (sample-calibrator)}$). For each of the *Hsp*s the amplification efficiency was calculated from a standard curve generated by amplification of serially diluted plasmid standards. The amplification efficiencies for *ELOV*, *RDH*, and *ferritin* were previously calculated from a relative standard curve of serially diluted cDNA (Tarrant et al., 2008). To calculate the 'Expression' of each *Hsp*, the mean threshold cycle (Ct) of the 0-hour treatment samples (handling experiment) or of the shallow samples (diapause study) was used as the 'calibrator' (also known as the reference sample). The Ct values for duplicate wells were averaged and relative expression values were base-10 log-transformed for statistical analysis.

Statistical analysis

Two-sample, two-tailed t-tests for the morphometric parameters, RNA:DNA ratios, and expression of *ELOV*, *RDH*, and *ferritin* were performed to confirm that deep and shallow samples represented diapausing and active copepods, respectively. *16S* and *Hsp* expression in deep and shallow samples were also evaluated with two-sample, two-tailed t-tests. One-way ANOVAs were used to compare qRT-PCR expression between handling stress treatments (0, 2, and 3 hour). Planned posthoc comparisons (Dunnett test) in genes with significant ANOVA results compared the 0-hour treatment mean (control) with the 2- and 3- hour treatment sample means.

RESULTS:

Identification of C. finmarchicus Hsp genes and phylogenetic analyses

Through a search of the *C. finmarchicus* ESTs at NCBI, we identified eight different putative *Hsp* sequences (*Hsp90*; 4 forms of *Hsp70* designated *Hsp70A*, *Hsp70B*, *Hsp70C* and *Hsp70D*; *Hsp21*; *Hsp22*; and *p26*; Table 1). The *Hsp70* forms and *Hsp90* were highly conserved, with Expect values (E-values) less than 10-40 and about 50% identity with *Hsp* proteins previously annotated from *Artemia*. In comparison, the small *Hsp*s (*Hsp21*, *Hsp22*, *p26*) were

more variable, with larger E-values and about 30% identity with amino acid sequences from *Artemia*. The partial *Hsp* sequences we amplified ranged in size from 361 to 637 bp (Table 1). Alignments of our *C. finmarchicus* clones with several crustacean and insect *Hsp* sequences demonstrated that the cloned *Hsp70A* and *Hsp90* sequences encode the 3' ends of the predicted proteins, while *Hsp70* (B, C, and D), *Hsp21*, *Hsp22*, and *p26* all encode the 5' ends (Fig. A.1-5, respectively).

Phylogenetic analyses based on maximum likelihood criteria confirm that the *C. finmarchicus* sequences are members of the *Hsp70* and *Hsp90* families (Fig. 1 and 2, respectively). In the *Hsp70* analysis we considered the sequences of bacterial origin (i.e., the monophyletic group of bacteria, plastid, and mitochondrial sequences) as the outgroup. Similarly, bacterial *Hsp90* homologs (high-temperature protein G, HTPG; Chen et al., 2006) were used to root the *Hsp90* tree. All three monophyletic groups of the *Hsp70* family that have been described in animals are represented by our four partial *Hsp70* sequences: one form is closely related to cytosolic forms of *Hsp70* (i.e., *Hsp70A*), two to mitochondrial forms (*Hsp*70*B,* and *Hsp70C*), and one to endoplasmic reticulum-associated forms (*Hsp70D*). The *Hsp70* sequences that were identified in our study are distinct from those previously examined in *C. finmarchicus* (Hansen et al., 2007; Voznesensky et al., 2004) (Fig. 1).The *Hsp70* described by Voznesensky (2004) falls into a clade with *Hsp70A* from our study and an *Hsp70* from the intertidal copepod *Tigriopus japonicus* (Rhee et al., 2009) although bootstrap support for this clade is weak (< 50%). The two *C. finmarchicus* sequences in this clade share only 58% identity, suggesting that they represent distinct genes. The *C. finmarchicus Hsp70* described by Hansen et al. (2008) falls into a clade distinct from the subfamilies associated with the cytosol, endoplasmic reticulum, and bacterial origin. Sequences within this clade include divergent human and *Drosophila Hsp70*-like genes (e.g., human *Hsp70_14* and *Hsp70_4*) that have been demonstrated in human to be functionally distinct from other forms of *Hsp70* (Kaneko et al., 1997; Wan et al., 2004). The *Hsp90* that we have identified in *C. finmarchicus* is identical to an EST sequence previously reported by Hansen et al. (2007) and it falls into a well-supported clade (100% bootstrap support) with cytosolic (*Hsp90A*) sequences from other organisms including crustaceans, insects, and human (Fig. 2).

Handling experiments

In a pilot study (data not shown), *p26* was not as strongly induced as other small *Hsp*s, and our assay did not reliably amplify *Hsp70C* (melt curves indicated non-specific products); therefore, expression of *p26* and *Hsp70C* were not measured in the handling experiment. Of the six examined *Hsp*s (*Hsp90*, *Hsp70A*, *Hsp70B*, *Hsp70D*, *Hsp21* and *Hsp22*), expression of three of these was significantly induced in shallow animals by a handling stress of increased waiting time before sampling: *Hsp70A* (one-way ANOVA; $F = 4.38$, $p = 0.023$), *Hsp21* ($F = 4.99$, $p = 0.015$) and *Hsp22* ($F = 4.22$, $p = 0.027$). Expression of *Hsp70A*, *Hsp21*, and *Hsp22* was significantly higher in the 3-hour treatment relative to the control treatment at hour 0 (Dunnett's test: $D = 2.34, 2.34, 2.35$, respectively; Fig. 3). The median expression in the 3-hour treatment for *Hsp70A*, *Hsp21*, and *Hsp22* was 2.95, 2.11, and 1.82 times higher than in the 0-hour treatment, respectively. Expression of these three *Hsp* transcripts was not significantly different between the 2-hour and 0-hour treatments. Genes previously identified as molecular markers of diapause (*ELOV*, *ferritin, RDH*) (Tarrant et al., 2008) showed no significant change in expression with handling stress (one-way ANOVA; $F = 2.087$, $p = 0.15$ for *ELOV*, $F = 0.21$, $p = 0.81$ for *ferritin*, $F = 3.13$, $p = 0.061$ for *RDH*; there was suggestive, but inconclusive, evidence that expression of *RDH* was lower during later sampling periods when compared to hour zero).

Outliers

Scatterplots of all morphological and gene-expression data were used to identify outliers (not shown). Two individual shallow *C. finmarchicus* C5 copepodids collected in 2006 for the diapause study had degraded oil sacs when photographed (thereby eliminating the possibility of estimating oil sac volume and fractional fullness) and were therefore excluded from these morphometric analyses. Two individual deep *C. finmarchicus* C5 copepodids collected in 2006 were also excluded from further analysis based on very low expression of the housekeeping gene and every other gene examined, likely caused by mRNA degradation. Finally, a few individual wells produced poor melt curves during amplification of *p26* and *Hsp70D* so these wells and their replicates were eliminated from further analysis of *p26* and *Hsp70D*. Removal of all outliers due to low expression or poor melt curves yielded individual deep copepod samples of $n = 18$ (*p26*), $n = 16$ (*Hsp70D*), and $n = 19$ (all other genes), and $n = 21$ individual shallow copepod samples for all genes in the diapause study.

Hallmarks of diapause

For all examined physiological, biochemical, and molecular indicators of diapause, the shallow and deep samples followed patterns expected of active and diapausing copepods, respectively (Table 3). Morphometric analysis demonstrated that animals collected from deep water had significantly larger oil sac volumes and greater oil sac fractional fullness, as would be expected of diapausing copepods. A majority of the shallow animals had food in their guts while the guts of all examined individual deep copepodids were empty. Higher RNA to DNA ratios, reflective of increased transcriptional activity, were observed in shallow animals. Moreover, expression of genes related to lipid synthesis (*ELOV* and *RDH*) were significantly lower while *ferritin* expression was significantly higher in deep animals (Fig. 4a-c). As expected, expression of the housekeeping gene, *16S*, did not significantly differ between the shallow and deep samples (two-sample, two-tailed t-test; $t = 0.48$, $p = 0.93$; Fig. 4d).

Hsp expression in deep and shallow samples

Expression of the eight cloned *Hsp*s was quantified in deep and shallow copepod samples (Fig. 5); *Hsp70C* expression was below the sensitivity of our assay and was not analyzed further. Two of the *Hsp*s (*Hsp70A* and *Hsp22*) exhibited significantly different expression between the deep and shallow samples (two-sample, two-tailed t-tests; *Hsp70A*: $t = -5.86$, $p < 0.0001$; *Hsp22*: $t = 3.15$, $p = 0.0032$). The median expression of *Hsp70A* was 7.52 times higher in the shallow samples than in the deep samples (95% CI: 3.74-15.09), while the median expression of *Hsp22* was 2.19 times higher in the deep samples than in the shallow samples (95% CI: 1.32-3.62). There was suggestive, but inconclusive, evidence that *p26* expression was also higher in the deep samples ($t = 1.80$, $p = 0.081$).

DISCUSSION:

In this study we isolated 8 *Hsp* transcripts (*Hsp90*, *Hsp70A, Hsp70B, Hsp70C, Hsp70D*, *Hsp21*, *Hsp22*, *p26*) and examined their expression in association with diapause and handling stress in the marine copepod *Calanus finmarchicus*. The cloned *C. finmarchicus Hsp*s showed high sequence similarity to corresponding genes from shrimp and *Artemia*, with the greatest conservation among *Hsp90* and *Hsp70* forms and relatively lower conservation among small *Hsp*s. These findings are consistent with observations that small *Hsp*s have diversified within

many lineages and generally exhibit a lower degree of conservation (reviewed by Denlinger et al., 2001; Haslbeck et al., 2005). Phylogenetic analyses based on maximum likelihood trees indicate that the *C. finmarchicus Hsp70* and *Hsp90* transcripts represent cytosolic (*Hsp70A*, *Hsp90*), mitochondrial (*Hsp70B* and *Hsp70C*) and endoplasmic reticulum (*Hsp70D*) forms. Cytosolic forms of *Hsp70* and *Hsp90* are of particular interest for this study because these forms are typically induced in response to stress (Chen et al., 2006; Daugaard et al., 2007; Taipale et al., 2010). While *Hsp* mRNA expression has been shown in many studies to be an accurate indicator of stress responsiveness and/or preparation for diapause in a variety of species (e.g. Tungjitwitayakul et al., 2008; Yocum, 2001; Zhang and Denlinger, 2009), future studies may benefit from measuring both mRNA and protein expression, as these two can indicate different regulatory processes and timescales of response.

Exposure to a handling stress of increased waiting time before sampling induced two small *Hsp*s (*Hsp21* and *Hsp22*), as well as *Hsp70A*. Induction of small *Hsp*s in *C. finmarchicus* was not surprising, as they play a direct role in stress tolerance in arthropods (Clegg et al., 1999; MacRae, 2010; Qiu and MacRae, 2008). The cytosolic *Hsp70* subfamily includes a mixture of constitutively expressed (cognate) and inducible forms (Daugaard et al., 2007), with the inducible genes varying among taxa (e.g., in mammals the cognate and inducible forms of cytosolic *Hsp70* resulted from lineage-specific duplication and subsequent diversification). Induction of *Hsp70* has been reported following thermal stress in *C. finmarchicus* (Voznesensky et al., 2004) and following thermal stress or metal exposure in the intertidal copepod *Tigriopus japonicus* (Rhee et al., 2009). Our phylogenetic analysis showed that the sequences from Voznesensky et al. (2004) and Rhee et al. (2009) fall within the cytosolic *Hsp70* subfamily and form a clade with *C. finmarchicus Hsp70A*, which was induced by handling in our study (Fig. 1). In contrast, *Hsp70B*, *Hsp70C*, and *Hsp70D* clustered with constitutively expressed mitochondrial and endoplasmic reticulum forms, so we expected that they would not be induced by stress. Our results confirm this, as none of these transcripts were significantly induced by handling stress. Our phylogenetic analysis also clustered the *C. finmarchicus Hsp70* sequence previously identified by Hansen et al. (2008) with human *Hsp70*-like proteins (*Hsp70_14*, *Hsp70_*4) that are known to share low sequence and functional similarity with cytosolic *Hsp70* forms (Kaneko et al., 1997; Wan et al., 2004). Therefore, previously reported differences in *C. finmarchicus Hsp70* inducibility by stressors (Hansen et al., 2008; Voznesensky et al., 2004) may be attributed, at

least in part, to differences in the '*Hsp70'* transcripts examined. Cytosolic forms of *Hsp90* are also sometimes induced in response to stress (Chen et al., 2006; Taipale et al., 2010); however, this is not observed in all species. Cytosolic *Hsp90* was not significantly induced by thermal stress in *T. japonicus* (Rhee et al., 2009) or by naphthalene exposure in *C. finmarchicus* (Hansen et al., 2008), although Hansen et al. (2007) suggested that *Hsp90* may be induced by combined chemical and thermal stress in *C. finmarchicus*. *Hsp90* was not induced by handling in our study. Unlike the small *Hsp*s, *Hsp*90 does not play a direct protective role in stabilizing denatured proteins (reviewed by MacRae, 2010) but acts as a molecular chaperone to enable the proper folding of a wide range of cellular proteins (Taipale et al., 2010), which may help explain our findings.

Two of the genes induced by handling stress in shallow samples, *Hsp70A* and *Hsp22*, were also differentially expressed between active and diapausing copepods. Diapause was classified using both traditional and molecular markers (Table 3), and associations between these two markers were identical to those observed by Tarrant et al. (2008). Expression of *Hsp70A* was both elevated in active animals and induced upon handling in shallow samples. We speculate that the observed differential expression of this gene reflects greater exposure of copepods to stressors while active in the upper ocean, such as temperature gradients encountered during diel vertical migrations, visual predators, solar radiation (including ultraviolet radiation; Wold and Norrin, 2004), turbulence, and starvation. In contrast, diapausing copepods likely experience a more stable environment at depth with far fewer stressors. Like *Hsp70A*, *Hsp22* was induced upon handling in shallow samples; however, *Hsp22* expression was elevated in diapausing animals. These observations could be explained by differential induction of *Hsp22* during the process of collecting active and diapausing samples (i.e., if *Hsp22* is more inducible in diapausing copepods); however, we consider this unlikely since diapausing copepods are expected to be in a state of torpor with reduced responsiveness to external stimuli (Hirche, 1983). Instead, we speculate that *Hsp22* plays a role in both stress response and diapause in *C. finmarchicus*. Qiu and MacRae (2008) found that in *Artemia, Hsp22* is similarly induced both in response to thermal stress and in preparation for diapause.

The role of *Hsp*70 and *Hsp*90 in mediating diapause varies greatly among insect taxa. For example, while changes in *Hsp*70 expression are not a component of the larval diapause of the blowfly (Tachibana et al., 2005), the pupal diapause of the corn earworm (Zhang and Denlinger,

2009) or the adult diapause of a fruit fly (Goto et al., 1998), *Hsp*70 is highly induced in the pupal diapause of the solitary bee (Yocum et al., 2005) and the onion maggot (Chen et al., 2006). Conversely, *Hsp*70 is down-regulated in the larval diapause of the corn stalk borer (Gkouvitsas et al., 2009) and the bamboo borer (Tungjitwitayakul et al., 2008). *Hsp*90 demonstrates similar discrepancies in diapause-related expression patterns, including down-regulation in the pupal diapause of the flesh fly (Rinehart and Denlinger, 2000), up-regulation at the termination of the larval diapause of the blow fly (Tachibana et al., 2005), and constant expression in the pupal diapause of the solitary bee (Yocum et al., 2005).

The variation in diapause-associated expression patterns of *Hsp*70 and *Hsp*90 among taxa cannot be attributed to differences in the developmental stage at which diapause occurs. In addition, our phylogenetic analyses demonstrate that all of the genes described in the studies cited in the preceding paragraph belong to the cytosolic subfamilies (Fig. B.1 and B.2); therefore, variability in *Hsp*70 and *Hsp*90 expression patterns among taxa cannot be attributed to different subfamily membership. However, within the *Hsp* subfamilies, more recent duplication events have given rise to multiple isoforms of cytosolic *Hsp70* and *Hsp90* genes (Chen et al., 2006; Daugaard et al., 2007; Taipale et al., 2010). In our analyses, non-inducible and inducible (both positively and negatively by diapause) *Hsp70* and *Hsp90* isoforms were distributed throughout the clusters of the cytosolic subfamily, which highlights the need for studies of the inducibility and function of individual *Hsp* isoforms.

In conclusion, this study represents the first characterization of *Hsp* expression in association with diapause in a calanoid copepod and expands the current understanding of the molecular regulation of diapause. We have identified a small *Hsp (Hsp22)* that is upregulated during diapause. We have also identified three *Hsp*s (*Hsp70A*, *Hsp21*, and *Hsp22*) that were induced by handling stress. We now know that diapause in *C. finmarchicus* is marked by changes in lipid metabolism (*FABP*, *ELOV*, *RDH*), endocrine signaling (*EcR*), and protection from cellular stress and protein degradation (*Hsp22, ferritin*) (Tarrant et al., 2008, current study). Enhanced stress tolerance is a particularly common feature of diapause, regardless of the developmental stage during which diapause occurs and/or differences in the degree of metabolic arrest (MacRae, 2010; Qiu, 2008; Rinehart et al., 2007; Sonoda et al., 2006). In the case of *C. finmarchicus,* the environmental conditions experienced during diapause are relatively minor (i.e., small variation in temperature, not subject to freezing or desiccation) compared to those

experienced by overwintering insects or *Artemia* cysts, so the induction of *ferritin* and *Hsp*22 may be sufficient for protection of proteins and other cellular components. Future work will be needed to identify the full complement of genes and protein products associated with diapause, and particularly the mechanisms that regulate preparation for and emergence from diapause.

ACKNOWLEDGEMENTS:

We are grateful for assistance at sea by Nadine Lysiak, Sarah Mussoline, Melissa Patrician, and Christopher Tremblay, the officers and crew of the NOAA Ships 'Albatross IV' and 'Delaware II', and chief scientists Fred Wenzel and Lisa Conger of the NOAA Northeast Fisheries Science Center's Protected Species Branch. We also thank Peter Wiebe for the loan of his MOCNESS, and Kristen Hunter Cevera for technical assistance. Helpful criticisms were provided by Mark Ohman and an anonymous reviewer. Funding for AMA was provided by the WHOI Summer Student Fellowship Program and an EPA STAR fellowship.

REFERENCES:

Abascal F., Zardoya R., Posada D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics, 21, 2104-2105.

Beato M., Klug J., 2000. Steroid hormone receptors: an update. Hum Reprod Update, 6, 225-236.

Bonnet D., Harris R.P., Hay S., Ingvarsdottir A., Simon O., 2007. Histological changes of the digestive epithelium in *Calanus finmarchicus*: an index for diapause? Mar Biol, 151, 313-326.

Boorstein W.R., Ziegelhoffer T., Craig E.A., 1994. Molecular evolution of the HSP70 multigene family. J Mol Evol, 38, 1-17.

Cesar J.R.O., Yang J., 2007. Expression patterns of ubiquitin, heat shock protein 70, α-actin and β-actin over the molt cycle in the abdominal muscle of marine shrimp *Litopenaeus vannamei*. Molecular Reproduction and Development, 74, 554-559.

Chen B., Zhong D., Monteiro A., 2006. Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. BMC Genomics, 7, 156.

Clegg J., WIillsie J.K., Jackson S., 1999. Adaptive Significance of a Small Heat Shock Alpha-Crystallin Protein (p26) in Encysted Embryos of the Brine Shrimp, *Artemia franciscana*. American Zoologist, 39, 836-847.

Dahms H.-U., 1995. Dormancy in the Copepoda — an overview. Hydrobiologia, 306, 199-211.

Daugaard M., Rohde M., Jaattela M., 2007. The heat shock protein 70 family: Highly homologous proteins with overlapping and distinct functions. FEBS Letters, 581, 3702-3710.

Denlinger D.L., Rinehart J.P., Yocum G.D., Denlinger D.L., Giebultowicz J.M., Saunders

D.S.,2001 Stress proteins: A role in insect diapause? Insect Timing: Circadian Rhythmicity to Seasonality. Elsevier Science B.V., Amsterdam, pp 155-171

Durbin E.G., Garrahan P.R., Casas M.C., 2000. Abundance and distribution of *Calanus finmarchicus* on the Georges Bank during 1995 and 1996. ICES Journal of Marine Science: Journal du Conseil, 57, 1664-1685.

Durbin E.G., Runge J.A., Campbell R.G., Garrahan P.R., Casas M.C., Plourde S. , 1997. Late fall-early winter recruitment of *Calanus finmarchicus* on Georges Bank Marine Ecology Progress Series, 151, 103.

Edgar R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32, 1792-1797.

Feder M.E., Hofmann G.E., 1999. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology, 61, 243-282.

Gkouvitsas T., Kontogiannatos D., Kourti A., 2009. Cognate Hsp70 gene is induced during deep larval diapause in the moth *Sesamia nonagriodes*. Insect Molecular Biology, 18, 253-264.

Goto S.G., Kimura M.T., 2004. Heat-shock-responsive genes are not involved in the adult diapause of *Drosophila triauraria*. Gene, 326, 117-122.

Goto S.G., Yoshida K.M., Kimura M.T., 1998. Accumulation of Hsp70 mRNA under environmental stresses in diapausing and nondiapausing adults of *Drosophila triauraria*. Journal of Insect Physiology, 44, 1009-1015.

Hansen B.H., Altin D., Nordtug T., Olsen A.J., 2007. Suppression subtractive hybridization library prepared from the copepod *Calanus finmarchicus* exposed to a sublethal mixture of environmental stressors. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 2, 250-256.

Hansen B.H., Altin D., Vang S.H., Nordtug T., Olsen A.J., 2008. Effects of naphthalene on gene transcription in *Calanus finmarchicus* (Crustacea : Copepoda). Aquatic Toxicology, 86, 157-165.

Hartl F.U., Hayer-Hartl M., 2002. Protein folding - Molecular chaperones in the cytosol: from nascent chain to folded protein. Science, 295, 1852-1858.

Haslbeck M., Franzmann T., Weinfurtner D., Buchner J., 2005. Some like it hot: the structure and function of small heat-shock proteins. Nature Structural & Molecular Biology, 12, 842-846.

Heath M.R., Boyle P.R., Gislason A., Gurney W.S.C., Hay S.J., Head E.J.H., Holmes S., Ingvarsdóttir A., Jónasdóttir S.H., Lindeque P., Pollard R.T., Rasmussen J., Richards K., Richardson K., Smerdon G., Speirs D., 2004. Comparative ecology of over-wintering *Calanus finmarchicus* in the northern North Atlantic, and implications for life-cycle patterns. ICES Journal of Marine Science: Journal du Conseil, 61, 698-708.

Helmbrecht K., Zeise E., Rensing L., 2000. Chaperones in cell cycle regulation and mitogenic signal transduction: a review. Cellular Proliferation, 33, 341-365.

Hirche H.J., 1983. Overwintering of *Calanus finmarchicus* and *Calanus helgolandicus*. Marine Ecology Progress Series, 11, 281-290.

Hirche H.J., 1996. Diapause in the marine copepod, *Calanus finmarchicus* - A review. Ophelia,44, 129-143.

Ingvarsdóttir A., Houlihan D.F., Heath M.R., Hay S.J., 1999. Seasonal changes in respiration rates of copepodite stage V *Calanus finmarchicus* (Gunnerus). Fisheries Oceanography, 8, 73-83.

Johnson C.L., 2004. Seasonal variation in the molt status of an oceanic copepod. Progress in Oceanography, 62, 15-32.

Kaartvedt S., 1996. Habitat preference during overwintering and timing of seasonal vertical migration of *Calanus finmarchicus*. Ophelia, 44, 145-156.

Kaneko Y., Kimura T., Kishishita M., Noda Y., Fujita J., 1997. Cloning of apg-2 encoding a novel member of heat shock protein 110 family. Gene, 189, 19-24.

Kostál V., 2006. Eco-physiological phases of insect diapause. Journal of Insect Physiology, 52,113-127.

Lindquist S., 1986. The heat-shock response Annual Review of Biochemistry, 55, 1151-1191.

MacRae T.H., 2010. Gene expression, metabolic regulation and stress tolerance during diapause. Cellular and Molecular Life Science, 67, 2405-2424.
Miller C., Cowles T., Wiebe P., Copley N., Grigg H., 1991. Phenology in *Calanus finmarchicus*; hypotheses about control mechanisms. Marine Ecology Progress Series, 72, 79-91.
Miller C.B., Crain J.A., Morgan C.A., 2000. Oil storage variability in *Calanus finmarchicus*. ICES Journal of Marine Science: Journal du Conseil, 57, 1786-1799.
Pfaffl M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research, 29, e45.
Planque B., Batten S.D., 2000. *Calanus finmarchicus* in the North Atlantic: the year of *Calanus* in the context of interdecadal change. ICES Journal of Marine Science: Journal du Conseil, 57,1528-1535.
Pratt W.B., 1997. The role of the hsp90-based chaperone system in signal transduction by nuclear receptors and receptors signaling via MAP kinase. Annual Review of Pharmacology and Toxicology, 37, 297-326.
Qiu Z., MacRae, Thomas H., 2008. ArHsp21, a developmentally regulated small heat-shock protein synthesized in diapausing embryos of *Artemia franciscana* Biochemistry Journal, 411,605-611.
Qiu Z.J., MacRae T.H., 2008. ArHsp22, a developmentally regulated small heat shock protein produced in diapause-destined *Artemia* embryos, is stress inducible in adults. FEBS Journal, 275,3556-3566.
Rhee J.-S., Raisuddin S., Lee K.-W., Seo J.S., Ki J.-S., Kim I.-C., Park H.G., Lee J.-S., 2009. Heat shock protein (Hsp) gene responses of the intertidal copepod *Tigriopus japonicus* to environmental toxicants. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 149, 104-112.
Rinehart J.P., Denlinger D.L., 2000. Heat-shock protein 90 is down-regulated during pupal diapause in the flesh fly, *Sarcophaga crassipalpis*, but remains responsive to thermal stress. Insect Molecular Biology, 9, 641-645.
Rinehart J.P., Li A., Yocum G.D., Robich R.M., Hayward S.A.L., Denlinger D.L., 2007. Upregulation of heat shock proteins is essential for cold survival during insect diapause. Proceedings of the National Academy of Science U S A, 104, 11130-11137.
Rinehart J.P., Yocum G.D., Denlinger D.L., 2000. Developmental upregulation of inducible hsp70 transcripts, but not the cognate form, during pupal diapause in the flesh fly, *Sarcophaga crassipalpis*. Insect Biochemistry and Molecular Biology, 30, 515-521.
Sameoto D.D., Herman A.W., 1990. Life-cycle and distribution of *Calanus finmarchicus* in deep basins on the Nova-Scotia shelf and seasonal changes in *Calanus spp.* Marine Ecology Progress Series, 66, 225-237.
Sonoda S., Fukumoto K., Izumi Y., Yoshida H., Tsumuki H., 2006. Cloning of heat shock protein genes (hsp90 and hsc70) and their expression during larval diapause and cold tolerance acquisition in the rice stem borer, *Chilo suppressalis* Walker. Archives of Insect Biochemistry and Physiology, 63, 36-47.
Sorensen J.G., Kristensen T.N., Loeschcke V., 2003. The evolutionary and ecological role of heat shock proteins. Ecology Letters, 6, 1025-1037.
Spees J.L., Chang S.A., Mykles D.L., Snyder M.J., Chang E.S., 2003. Molt cycle-dependent molecular chaperone and polyubiquitin gene expression in lobster. Cell Stress Chaperones, 8, 258-264.
Stamatakis A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22, 2688-2690.
Tachibana S.I., Numata H., Goto S.G., 2005. Gene expression of heat-shock proteins (Hsp23, Hsp70 and Hsp90) during and after larval diapause in the blow fly *Lucilia sericata*. Journal of Insect Physiology, 51, 641-647.
Taipale M., Jarosz D.F., Lindquist S., 2010. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nature Reviews Molecular Cell Biology, 11, 515-528.
Tarrant A.M., Baumgartner M.F., Verslycke T., Johnson C.L., 2008. Differential gene expression in diapausing and active *Calanus finmarchicus* (Copepoda). Marine Ecology Progress Series, 355, 193-207.
Tungjitwitayakul J., Tatun N., Singtripop T., Sakurai S., 2008. Characteristic expression of three heat shock-responsive genes during larval diapause in the bamboo borer *Omphisa fuscidentalis*. Zoological Science, 25, 321-333.
Voznesensky M., Lenz P.H., Spanings-Pierrot C., Towle D.W., 2004. Genomic approaches to detecting thermal stress in *Calanus finmarchicus* (Copepoda: Calanoida). Journal of Experimental Marine Biology and Ecology, 311, 37-46.
Wagner M., Durbin E., Buckley L., 1998. RNA : DNA ratios as indicators of nutritional condition in the copepod *Calanus finmarchicus*. Marine Ecology Progress Series, 162, 173-181.
Wan T., Zhou X., Chen G., An H., Chen T., Zhang W., Liu S., Jiang Y., Yang F., Wu Y., Cao X., 2004. Novel heat shock protein Hsp70L1 activates dendritic cells and acts as a Th1 polarizing adjuvant. Blood, 103, 1747-1754.
Wiebe P.H., Burt K.H., Boyd S.H., Morton A.W., 1976. Multiple opening-closing net and environmental sensing system for sampling zooplankton Journal of Marine Research, 34, 313-326.
Williams R., Conway D.V.P., Hunt H.G., 1994. The role of copepods in the planktonic ecosystems of mixed and stratified waters of the European shelf seas. Hydrobiologia, 292-293, 521-530.
Wold A., Norrin F., 2004. Vertical migration as a response to UVR stress in *Calanus finmarchicus* females and nauplii. Polar Research, 23, 27-34.
Yocum G.D., 2001. Differential expression of two HSP70 transcripts in response to cold shock, thermoperiod, and adult diapause in the Colorado potato beetle. Journal of Insect Physiology, 47,1139-1145.
Yocum G.D., Kemp W.P., Bosch J., Knoblett J.N., 2005. Temporal variation in overwintering gene expression and respiration in the solitary bee *Megachile rotundata*. Journal of Insect Physiology, 51, 621-629.

Yuan Y., Crane D., Barry C., 3rd, 1996. Stationary phase-associated protein expression in *Mycobacterium tuberculosis*: function of the mycobacterial alpha-crystallin homolog. Journal of Bacteriology, 178, 4484-4492.
Zhang Q.R., Denlinger D.L., 2009. Molecular characterization of heat shock protein 90, 70 and 70 cognate cDNAs and their expression patterns during thermal stress and pupal diapause in the corn earworm. Journal of Insect Physiology, 56, 138-150.

Figure legends:

Fig. 1 : Maximum likelihood analysis of *Hsp70* from *Calanus finmarchicus* and representative *Hsp70* sequences. Bootstrap percentages of 1000 replicates are indicated above branches when they are greater than 50%. Sequences of bacterial origin (i.e., the monophyletic group of bacteria, plastid, and mitochondrial sequences) were used as the outgroup. Sequences selected for these analyses are principally from Rhee et al. (2009), Boorstein et al. (1994), and Daugaard et al. (2007) with some additions (Table B.1 for full list of sequences and accession numbers). '*C. finmarchicus 70 (V)*' and '*C. finmarchicus 70 EST*' represent *Hsp70* sequences previously characterized by Voznesensky et al. (2004) and Hansen et al. (2008), respectively. Symbols denote *Hsp70* forms that are non-inducible [ ] and inducible [(+)] by stress as described in this study as well as in studies listed in Table B.1. Distance bar at the bottom of the tree indicates branch length scale, or the number of substitutions per amino acid site.

Fig. 2: Maximum likelihood analysis of *Hsp90* from *Calanus finmarchicus* and representative cytosolic (*Hsp90A*), endoplasmic reticulum (*Hsp90B*), chloroplast (*Hsp90C*) and mitochondrial (TRAP) *Hsp90* genes. Bootstrap percentages of 1000 replicates are indicated above branches when they are greater than 50%. Bacterial *Hsp*90 homologs (i.e., high-temperature protein G, HTPG) sequences were used to root the tree. Sequences are primarily from Chen et al. (2005; 2006) with some additions (see Table B.2 for full list of sequences and accession numbers). Symbols denote copepod *Hsp90* forms that are non-inducible [ ] and inducible [(+)] by stress as described in this study and Rhee et al. (2009). Distance bar at the bottom of the tree indicates branch length scale, or the number of substitutions per amino acid site.

Fig. 3 (a-f). Sample means (± 95% CI) and jitter plots of base-10 log-transformed gene expression of pooled (3 individuals/tube) shallow *Calanus finmarchicus* C5 copepodids exposed to handling stress. Expression was calculated relative to mean expression at t = 0 hour and normalized to *16S* rRNA expression. Expression of *Hsp70A*, *Hsp21*, and *Hsp22* was significantly higher in the 3-hour treatment than in the 0-hour treatment (Dunnett test values represented by 'D' = 2.34, 2.34, 2.35, respectively; $p < 0.05$).

Fig. 4 (a-d). Sample means (± 95% CI) and jitter plots of base-10 log-transformed gene expression of *ELOV*, *RDH*, *ferritin*, and *16S* in individual deep ( ) and shallow ( ) *Calanus finmarchicus* C5 copepodids. Expression values are calculated relative to mean expression in shallow samples and normalized to *16S* rRNA expression (for *ELOV*, *RDH*, and *ferritin*).

Fig. 5 (a-g). Sample means (± 95% CI) and jitter plots of base-10 log-transformed *Hsp* expression in individual deep ( ) and shallow ( ) *Calanus finmarchicus* C5 copepodids. Expression was calculated relative to mean expression in shallow samples and normalized to *16S* rRNA.

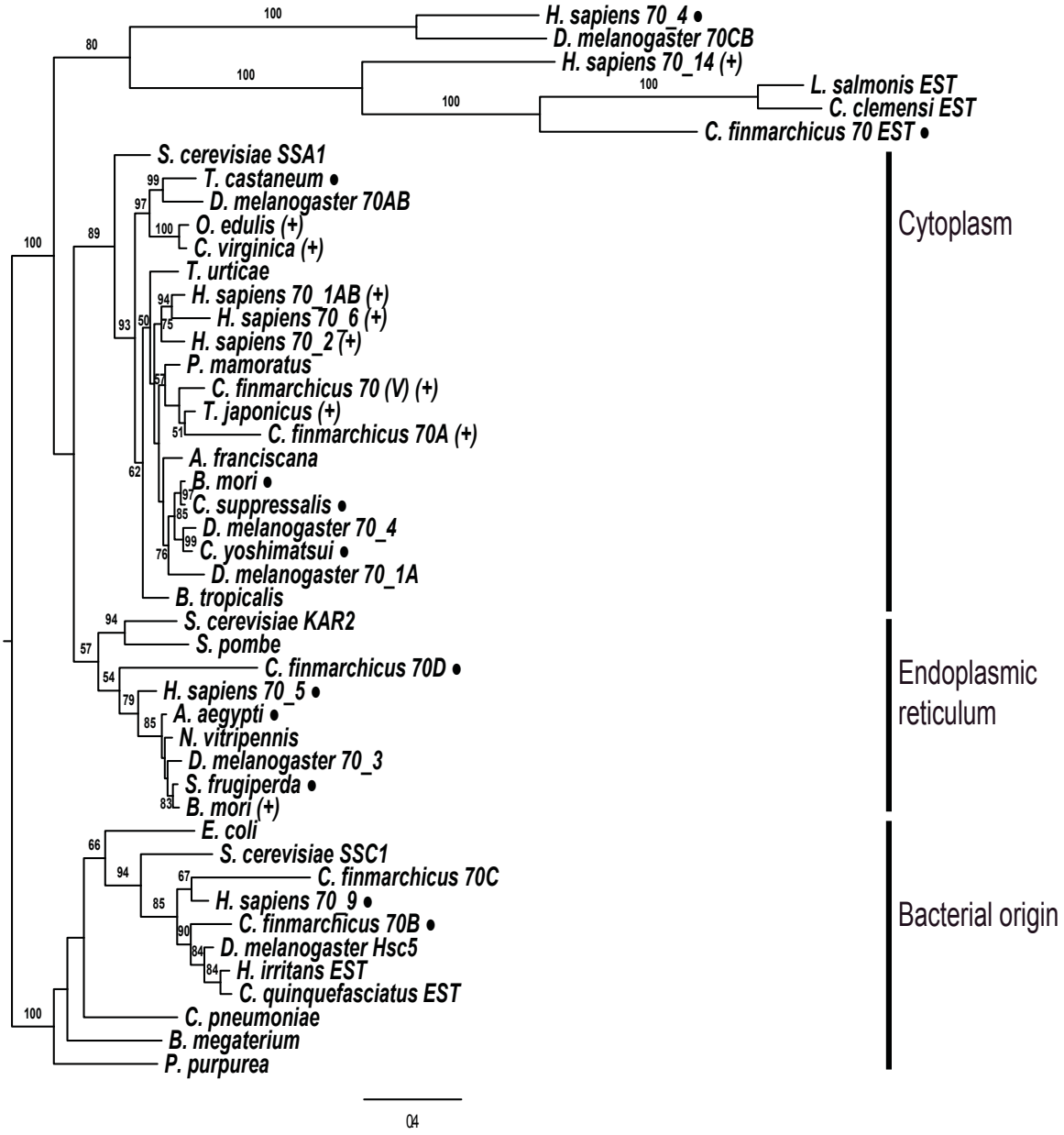

H. sapiens 70_4 •
D. melanogaster 70CB
H. sapiens 70_14 (+)
L. salmonis EST
C. clemensi EST
C. finmarchicus 70 EST •
S. cerevisiae SSA1
T. castaneum •
D. melanogaster 70AB
O. edulis (+)
C. virginica (+)
T. urticae
H. sapiens 70_1AB (+)
H. sapiens 70_6 (+)
H. sapiens 70_2 (+)
P. mamoratus
C. finmarchicus 70 (V) (+)
T. japonicus (+)
C. finmarchicus 70A (+)
A. franciscana
B. mori •
C. suppressalis •
D. melanogaster 70_4
C. yoshimatsui •
D. melanogaster 70_1A
B. tropicalis
S. cerevisiae KAR2
S. pombe
C. finmarchicus 70D •
H. sapiens 70_5 •
A. aegypti •
N. vitripennis
D. melanogaster 70_3
S. frugiperda •
B. mori (+)
E. coli
S. cerevisiae SSC1
C. finmarchicus 70C
H. sapiens 70_9 •
C. finmarchicus 70B •
D. melanogaster Hsc5
H. irritans EST
C. quinquefasciatus EST
C. pneumoniae
B. megaterium
P. purpurea
Cytoplasm
Endoplasmic reticulum
Bacterial origin
04

Figure 1

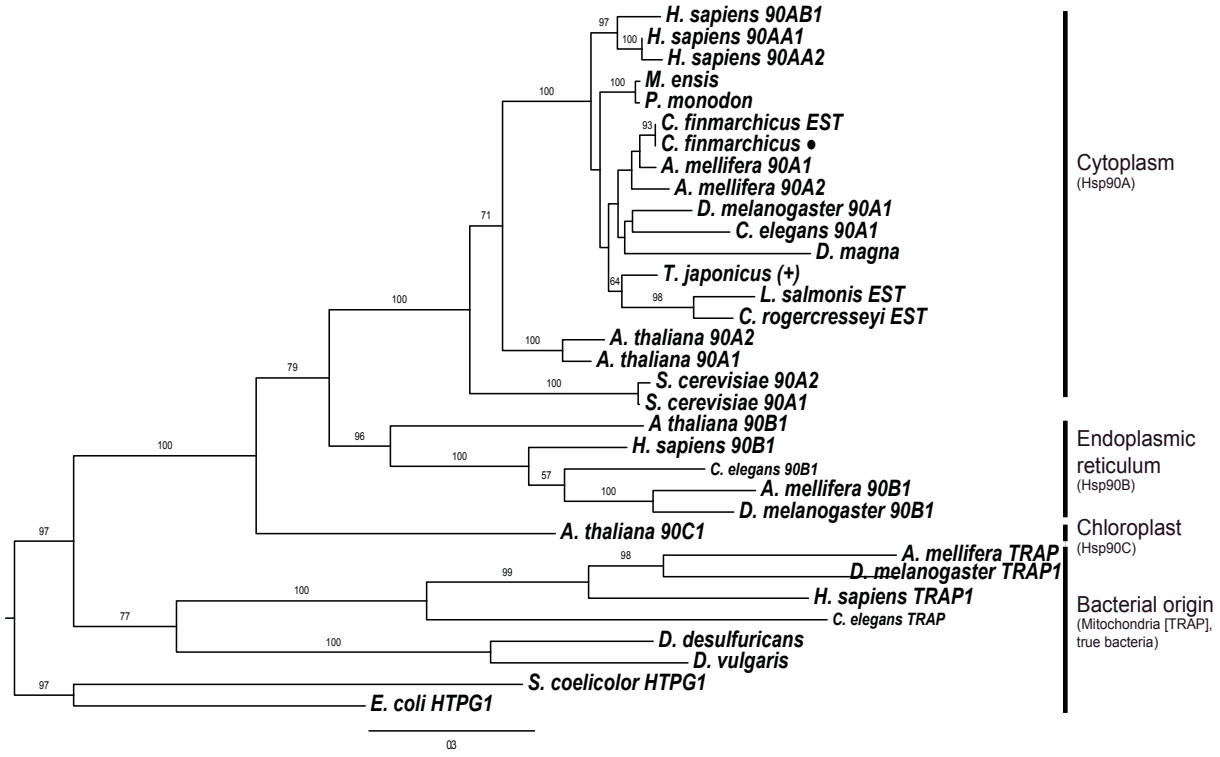

H. sapiens 90AB1
H. sapiens 90AA1
H. sapiens 90AA2
M. ensis
P. monodon
C. finmarchicus EST
C. finmarchicus •
A. mellifera 90A1
A. mellifera 90A2
D. melanogaster 90A1
C. elegans 90A1
D. magna
T. japonicus (+)
L. salmonis EST
C. rogercresseyi EST
A. thaliana 90A2
A. thaliana 90A1
S. cerevisiae 90A2
S. cerevisiae 90A1
A thaliana 90B1
H. sapiens 90B1
C. elegans 90B1
A. mellifera 90B1
D. melanogaster 90B1
A. thaliana 90C1
A. mellifera TRAP
D. melanogaster TRAP1
H. sapiens TRAP1
C. elegans TRAP
D. desulfuricans
D. vulgaris
S. coelicolor HTPG1
E. coli HTPG1
03
Cytoplasm
(Hsp90A)
Endoplasmic reticulum
(Hsp90B)
Chloroplast
(Hsp90C)
Bacterial origin
(Mitochondria [TRAP], true bacteria)

Figure 2

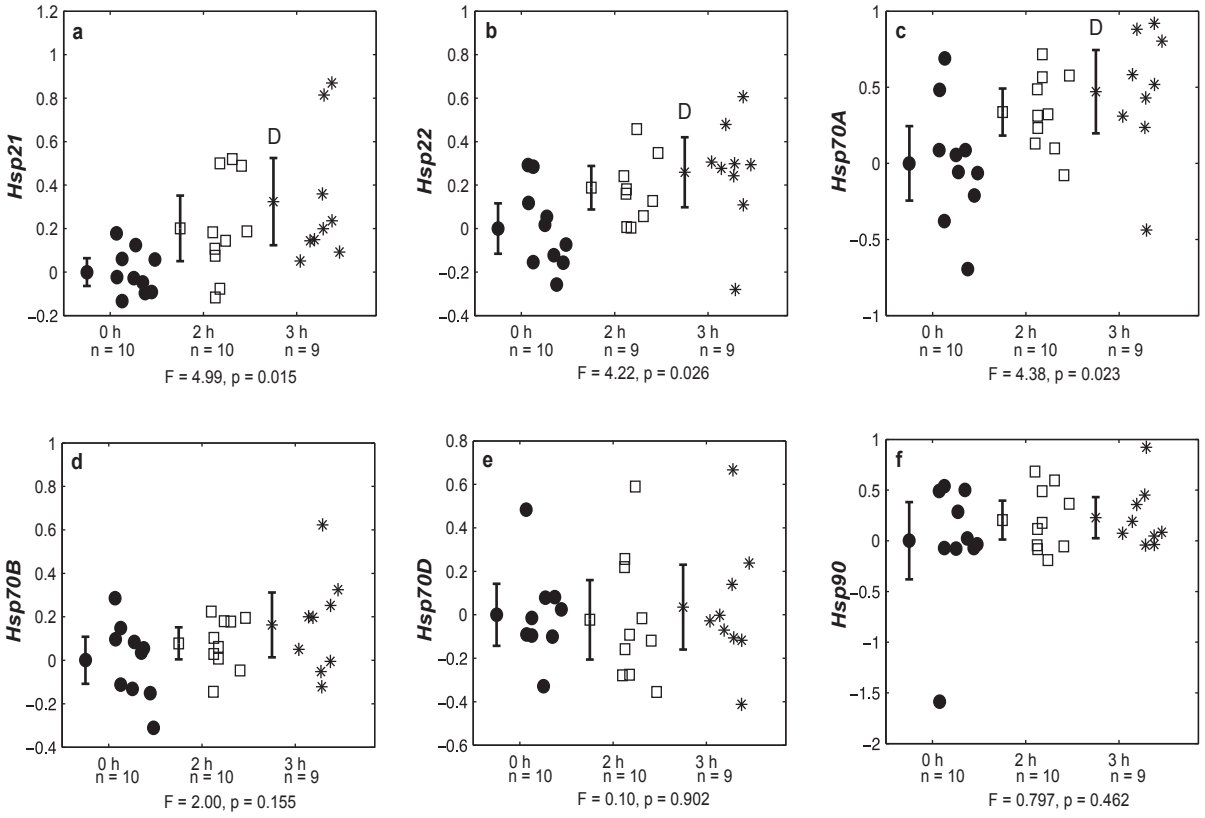

a
Hsp21
D
0 h
n = 10
2 h
n = 10
3 h
n = 9
F = 4.99, p = 0.015
b
Hsp22
D
0 h
n = 10
2 h
n = 9
3 h
n = 9
F = 4.22, p = 0.026
c
Hsp70A
D
0 h
n = 10
2 h
n = 10
3 h
n = 9
F = 4.38, p = 0.023
d
Hsp70B
0 h
n = 10
2 h
n = 10
3 h
n = 9
F = 2.00, p = 0.155
e
Hsp70D
0 h
n = 10
2 h
n = 9
3 h
n = 9
F = 0.10, p = 0.902
f
Hsp90
0 h
n = 10
2 h
n = 10
3 h
n = 9
F = 0.797, p = 0.462

Figure 3

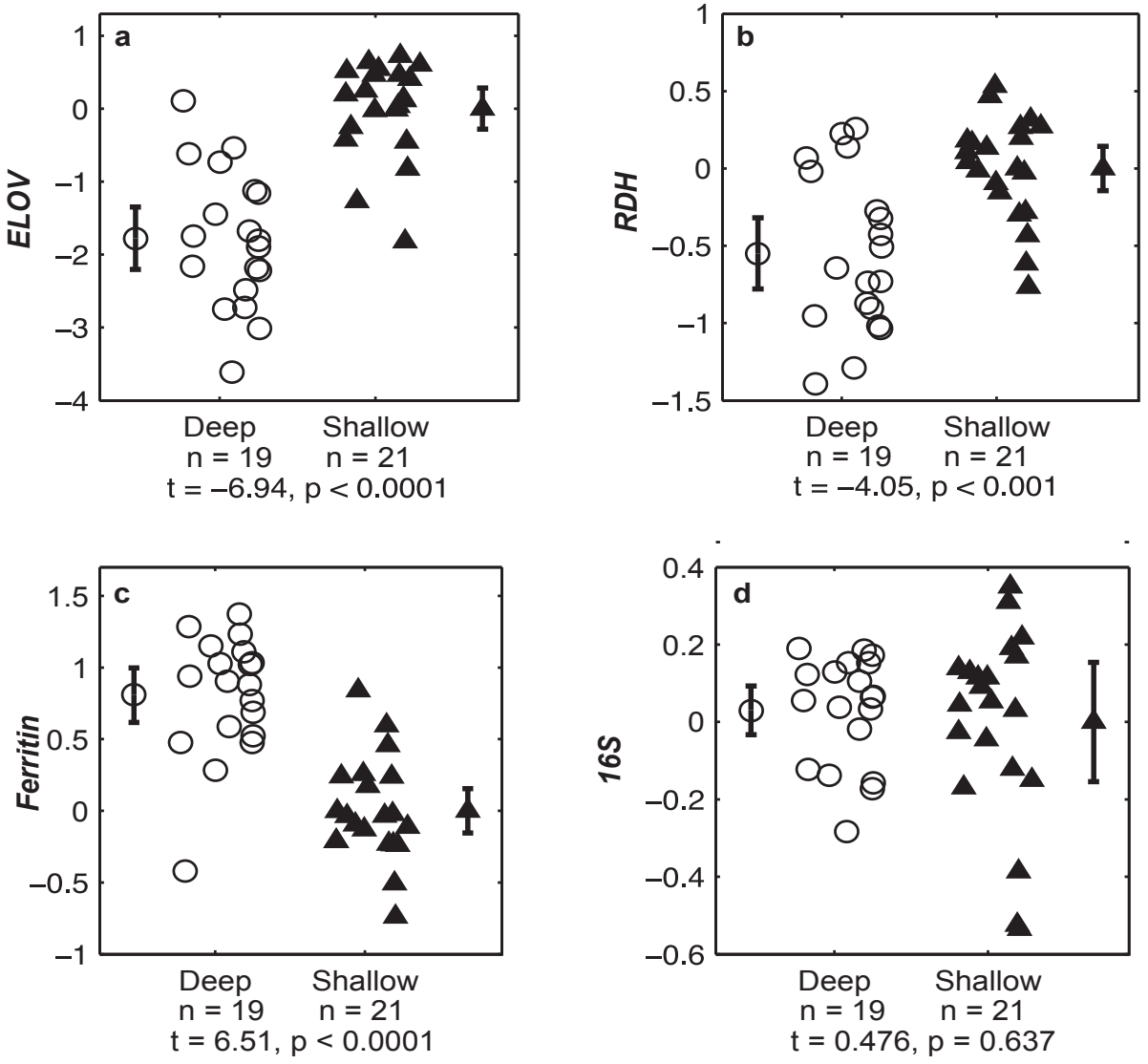

a
ELOV
1
0
−1
−2
−3
−4
Deep
Shallow
n = 19
n = 21
t = −6.94, p < 0.0001
b
RDH
1
0.5
0
−0.5
−1
−1.5
Deep
Shallow
n = 19
n = 21
t = −4.05, p < 0.001
c
Ferritin
1.5
1
0.5
0
−0.5
−1
Deep
Shallow
n = 19
n = 21
t = 6.51, p < 0.0001
d
16S
0.4
0.2
0
−0.2
−0.4
−0.6
Deep
Shallow
n = 19
n = 21
t = 0.476, p = 0.637

Figure 4

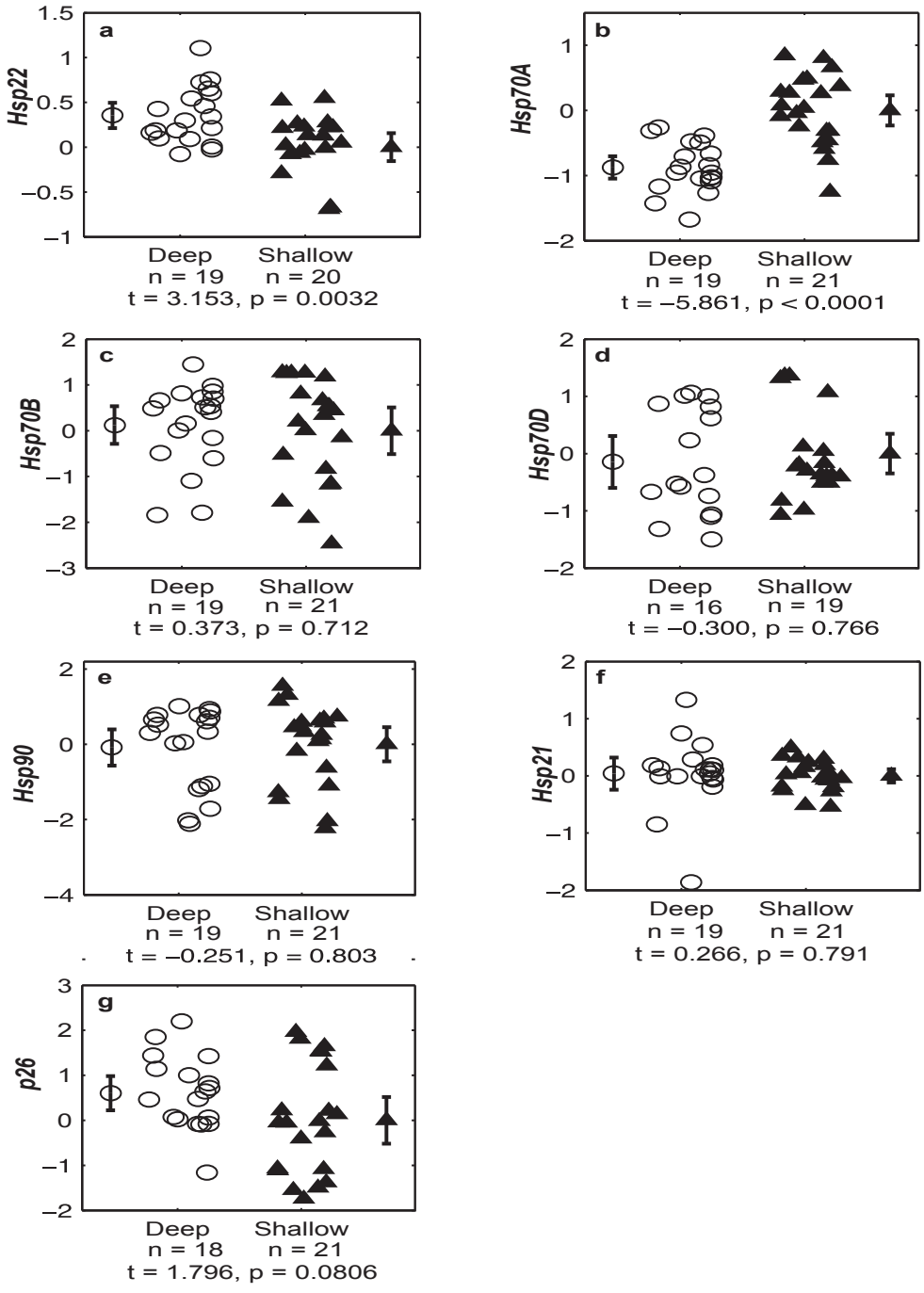

a
Hsp22
Deep
Shallow
n = 19
n = 20
t = 3.153, p = 0.0032
b
Hsp70A
Deep
Shallow
n = 19
n = 21
t = –5.861, p < 0.0001
c
Hsp70B
Deep
Shallow
n = 19
n = 21
t = 0.373, p = 0.712
d
Hsp70D
Deep
Shallow
n = 16
n = 19
t = –0.300, p = 0.766
e
Hsp90
Deep
Shallow
n = 19
n = 21
t = –0.251, p = 0.803
f
Hsp21
Deep
Shallow
n = 19
n = 21
t = 0.266, p = 0.791
g
p26
Deep
Shallow
n = 18
n = 21
t = 1.796, p = 0.0806

Figure 5

Table 1. Annotation of cloned *Calanus finmarchicus Hsp*s. Accession numbers are given for the selected crustacean *Hsps* used to search the EST database ("BLAST input") and thesubsequent *C. finmarchicus* EST hits. Portions of each EST sequence were cloned in the present study. The percent identity was calculated as the percentage of identical amino acids in the input sequence relative to the *C. finmarchicus* EST.

Gene Name	BLAST input	*Calanus* EST hit	EST length (bp)	[a]E-value	Input/EST identity (% identity)	Cloned sequence length (bp)
Hsp90	*Metapenaeus ensis* (ABR66910)	ES414827	413	7.00E-41	104/128 (81)	361
Hsp70A	*Artemia franciscana* (AAL27404)	EL965576	667	1.00E-71	152/221 (53)	630
Hsp70B	*Artemia franciscana* (AAL27404)	EH666605	650	6.00E-57	110/205 (53)	586
Hsp70C	*Artemia franciscana* (AAL27404)	ES237720	665	1.00E-55	110/211 (52)	637
Hsp70D	*Artemia franciscana* (AAL27404)	FG342764	496	5.00E-52	102/164 (62)	463
Hsp21	*Artemia franciscana* (ABD19712)	EH667182	700	4.00E-07	30/102 (29)	536
Hsp22	*Artemia franciscana* (ABD19713)	FK041659	665	1.00E-09	30/87 (34)	502
p26	*Artemia franciscana* (ABC41138)	EH666286	634	3.00E-13	34/97 (35)	551

[a]E-values based on BLAST search on 5 August 2010.

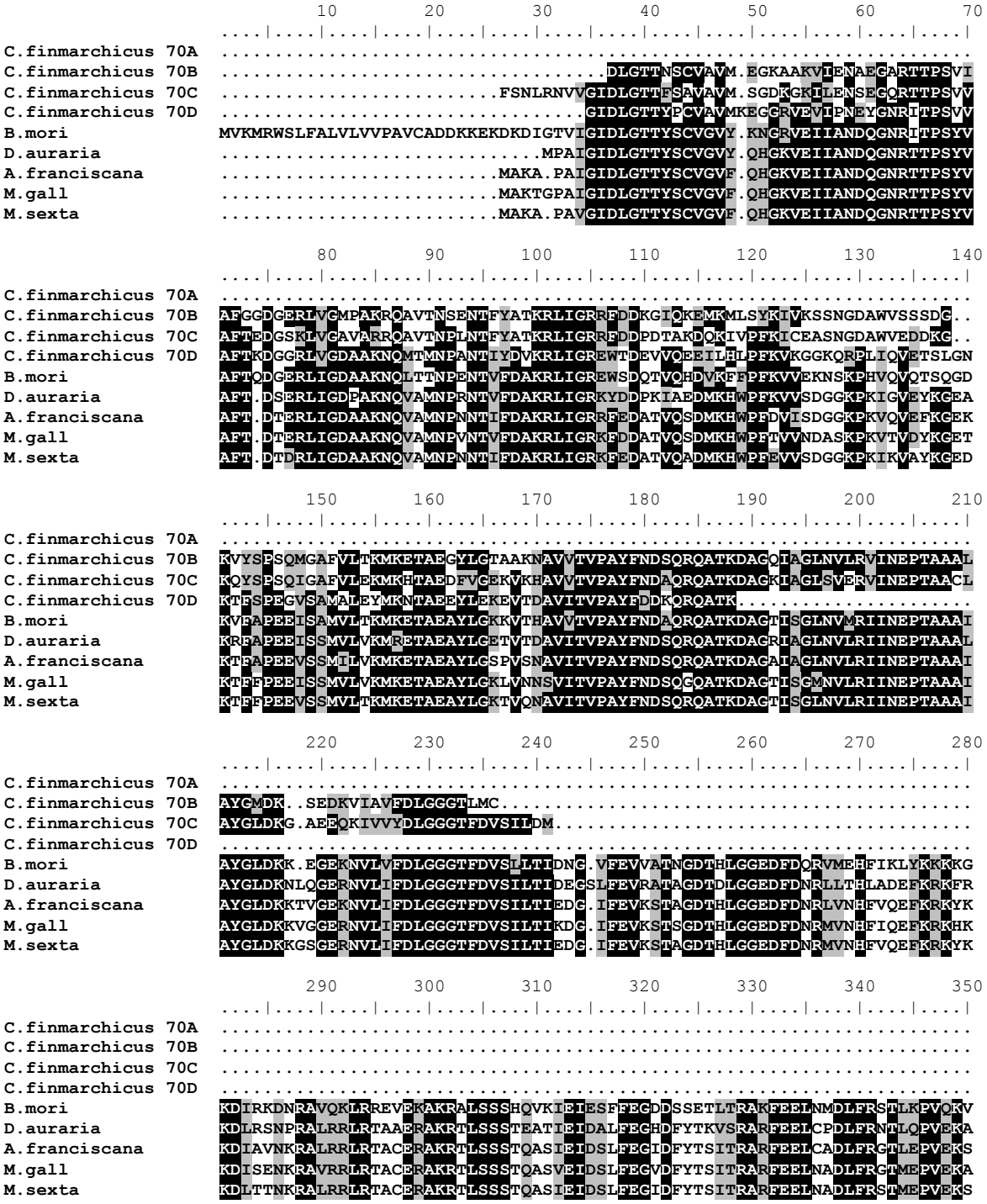
Table 2. Oligonucleotide primer sequences and annealing temperatures (T_m) used in qPCR assays. IQMix and EvaFast are distinct qPCR reagents (see text for additional details).

Gene	Primer Sequence	Tm IQMix	Tm EvaFast
Hsp90	F: 5'-TCATCCGGATTCAGCTTGGAG-3'	64	60
	R: 5'-GGTGGCATGTCGCTGTCATC-3'	64	60
Hsp70A	F: 5'-CGAAACAGCAGGAGGAGTGATG-3'	64	60
	R: 5'-TGACAGCAGGTTGGTTGTCTTG-3'	64	60
Hsp70B	F: 5'-TGGAGGGAAAGGCAGCTAAAG-3'	66	60
	R: 5'-CATCGCTGGAACTAACCCAAGC-3'	66	60
Hsp70D	F: 5'-GGGTGGAGGTGATCCCTAATG-3'	66	60
	R: 5'-TGCACCACTTCATCAGTCCAC-3'	66	60
Hsp22	F: 5'-GGCTACAAGCCAAGTGAGCTG-3'	—	64
	R: 5'-GAGACCATGGTGTGGCCTTC-3'	—	64
Hsp21	F: 5'-TGCAAACACAGCAACAAGCTG-3'	—	62
	R: 5'-GCCTCGGAAAGAGCATTCTTC-3'	—	62
p26	F: 5'-CTTGCCAAGCATGAGACCAAG-3'	64	60
	R: 5'-GGATTGACCCCAGATGGTAATG-3'	64	60
ELOV	F: 5'-GTCTGGTGGTGTTTCTTCTCC-3'	64	60
	R: 5'-CACATGCAGAGAGGTAAGTTGG-3'	64	60
RDH	F: 5'-CTAGCCAGGTTGCTGATGAAG-3'	—	64
	R: 5'-TCTTGGAGATGGTGAGGTCTG-3'	—	64
ferritin	F: 5'-AATATCAGACCAAGCGTGGAG-3'	64	60
	R: 5'-AGCTTCCATTGCCTGAATAGG-3'	64	60
16S	F: 5'-AAGCTCCTCTAGGGATAACAGC-3'	64	62
	R: 5'-CGTCTCTTCTAAGCTCCTGCAC-3'	64	62

Table 3. Evidence for diapause in deep copepod samples. Two-sample, two-tailed t-tests were conducted using reported expectations as the alternative hypothesis. Sample sizes were as follows: n = 21 shallow, n = 19 deep. Gene expression reported as base-10 log-transformed differences relative to the average shallow sample expression. RNA:DNA ratios are also base-10 log transformed. 95% confidence intervals are provided for means, and asterisks indicate significance of the t-test as follows: ** indicates $p < 0.0001$, * indicates $p = 0.0002$.

Indicator of diapause	**Expectation**	**Deep**	**Shallow**	**t**
ELOV expression	Deep < Shallow	-1.77 ± 0.43	0.00 ± 0.28	-6.94**
RDH expression	Deep < Shallow	-0.55 ± 0.23	0.00 ± 0.14	-4.05*
Ferritin expression	Deep > Shallow	0.81 ± 0.19	0.00 ± 0.15	6.51**
Oil sac volume (mm^3)	Deep > Shallow	0.40 ± 0.053	0.14 ± 0.052	6.97**
Oil sac fractional fullness (mm^3)	Deep > Shallow	0.78 ± 0.077	0.33 ± 0.098	7.43**
RNA:DNA ratio	Deep < Shallow	0.30 ± 0.12	0.67 ± 0.13	-4.23*
Empty gut (%)	Deep > Shallow	100%	9.5%	N/A

Figure A.1: Alignment of translated *Hsp70* sequences from *Calanus finmarchicus* (this study), *Bombyx mori* (BAA32395), *Drosophila auraria* (CAA04699), *Artemia franciscana* (AAL27404), *Mytilus galloprovincialis* (M.gall; AAW52766), and *Manduca sexta* (AAF09496). The *Hsp70A* sequence identified in this study is 210 amino acids and aligns at the 3' end of the approximately 640 amino acid long *Hsp70* protein while the *Hsp70B*, *Hsp70C* and *Hsp70D* clones range in size from 150-200 bp and align at the 5' end.



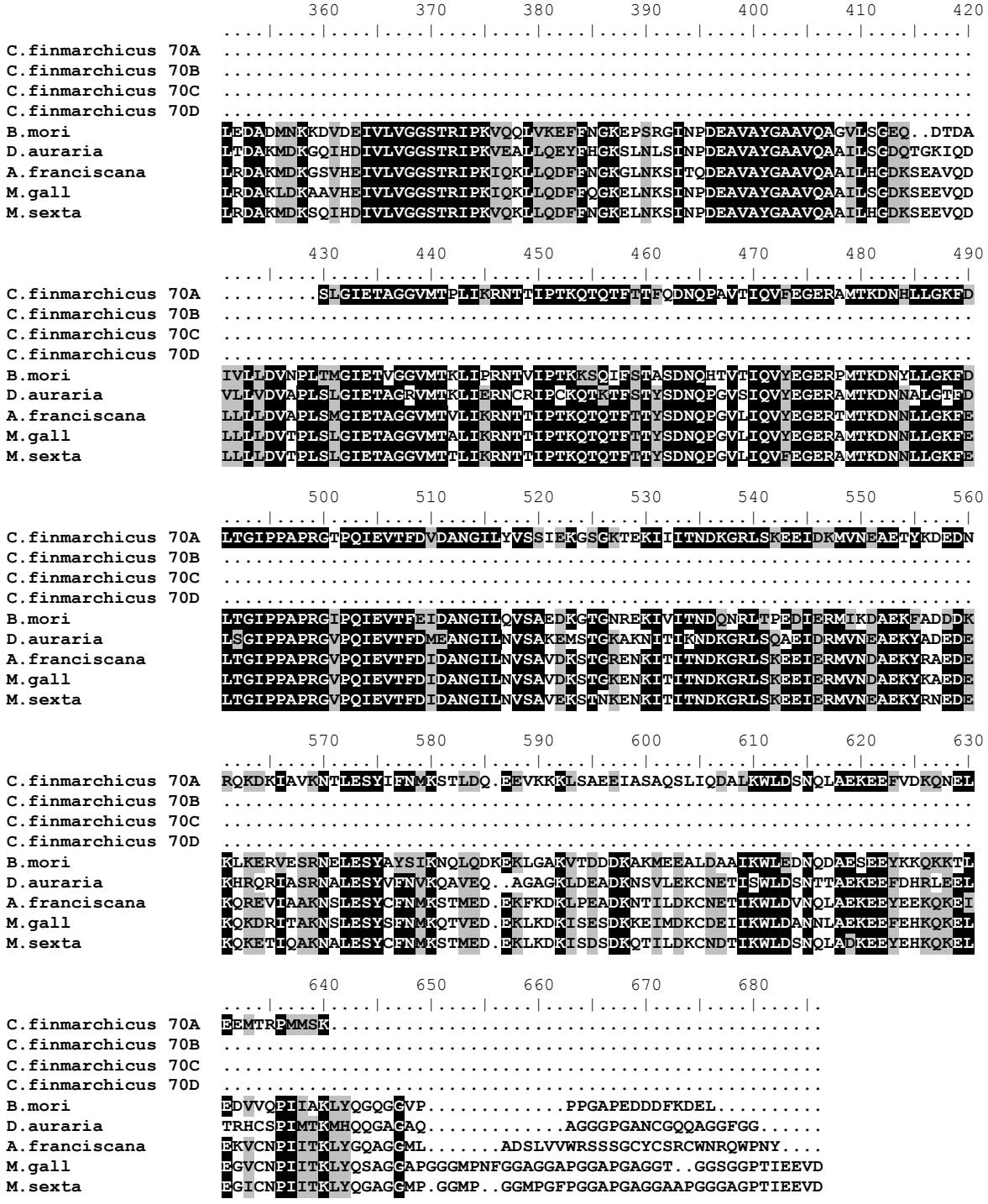
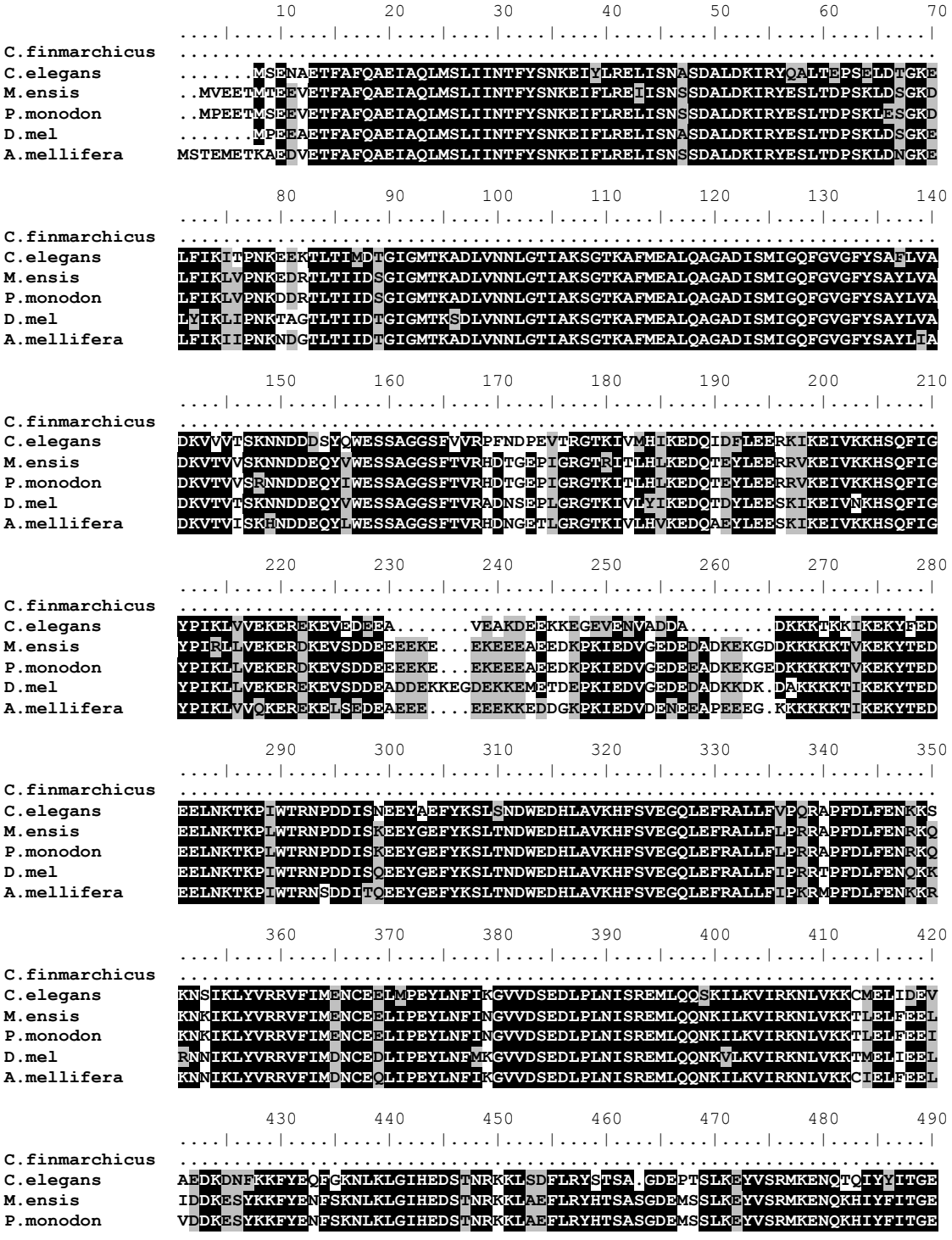

360 370 380 390 400 410 420
C.finmarchicus 70A
C.finmarchicus 70B
C.finmarchicus 70C
C.finmarchicus 70D
B.mori LEDADMNKKDVDEIVLVGGSTRIPKVQQLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGEQ..DTDA
D.auraria LTDAKMDKGQIHDIVLVGGSTRIPKVEALLQEYFHGKSLNLSINPDEAVAYGAAVQAAILSGDQTGKIQD
A.franciscana LRDAKMDKGSVHEIVLVGGSTRIPKIQKLLQDFFNGKGLNKSITQDEAVAYGAAVQAAILHGDKSEAVQD
M.gall LRDAKLDKAAVHEIVLVGGSTRIPKIQKLLQDFFQGKELNKSINPDEAVAYGAAVQAAILSGDKSEEVQD
M.sexta LRDAKMDKSQIHDIVLVGGSTRIPKVQKLLQDFFNGKELNKSINPDEAVAYGAAVQAAILHGDKSEEVQD
430 440 450 460 470 480 490
C.finmarchicus 70A .........SLGIETAGGVMTPLIKRNTTIPTKQTQTFTTFQDNQPAVTIQVFEGERAMTKDNHLLGKFD
C.finmarchicus 70B
C.finmarchicus 70C
C.finmarchicus 70D
B.mori IVLLDVNPLTMGIETVGGVMTKLIPRNTVIPTKKSQIFSTASDNQHTVTIQVYEGERPMTKDNYLLGKFD
D.auraria VLLVDVAPLSLGIETAGRVMTKLIERNCRIPCKQTKTFSTYSDNQPGVSIQVYEGERAMTKDNNALGTFD
A.franciscana LLLLDVAPLSMGIETAGGVMTVLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERTMTKDNNLLGKFE
M.gall LLLLDVTPLSLGIETAGGVMTALIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLGKFE
M.sexta LLLLDVTPLSLGIETAGGVMTTLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVFEGERAMTKDNNLLGKFE
500 510 520 530 540 550 560
C.finmarchicus 70A LTGIPPAPRGTPQIEVTFDVDANGILYVSSIEKGSGKTEKIIITNDKGRLSKEEIDKMVNEAETYKDEDN
C.finmarchicus 70B
C.finmarchicus 70C
C.finmarchicus 70D
B.mori LTGIPPAPRGIPQIEVTFEIDANGILQVSAEDKGTGNREKIVITNDQNRLTPEDIERMIKDAEKFADDDK
D.auraria LSGIPPAPRGVPQIEVTFDMEANGILNVSAKEMSTGKAKNITIKNDKGRLSQAEIDRMVNEAEKYADEDE
A.franciscana LTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGRENKITITNDKGRLSKEEIERMVNDAEKYRAEDE
M.gall LTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGKENKITITNDKGRLSKEEIERMVNDAEKYKAEDE
M.sexta LTGIPPAPRGVPQIEVTFDIDANGILNVSAVEKSTNKENKITITNDKGRLSKEEIERMVNEAEKYRNEDE
570 580 590 600 610 620 630
C.finmarchicus 70A RQKDKIAVKNTLESYIFNMKSTLDQ.EEVKKKLSAEEIASAQSLIQDALKWLDSNQLAEKEEFVDKQNEL
C.finmarchicus 70B
C.finmarchicus 70C
C.finmarchicus 70D
B.mori KLKERVESRNELESYAYSIKNQLQDKEKLGAKVTDDDKAKMEEALDAAIKWLEDNQDAESEEYKKQKKTL
D.auraria KHRQRIASRNALESYVFNVKQAVEQ..AGAGKLDEADKNSVLEKCNETISWLDSNTTAEKEEFDHRLEEL
A.franciscana KQREVIAAKNSLESYCFNMKSTMED.EKFKDKLPEADKNTILDKCNETIKWLDVNQLAEKEEYEEKQKEI
M.gall KQKDRITAKNSLESYSFNMKQTVED.EKLKDKISESDKKEIMDKCDEIIKWLDANNLAEKEEFEHKQKEL
M.sexta KQKETIQAKNALESYCFNMKSTMED.EKLKDKISDSDKQTILDKCNDTIKWLDSNQLADKEEYEHKQKEL
640 650 660 670 680
C.finmarchicus 70A EEMTRPMMSK.....
C.finmarchicus 70B
C.finmarchicus 70C
C.finmarchicus 70D
B.mori EDVVQPIIAKLYQGQGGVP.....PPGAPEDDDFKDEL.....
D.auraria TRHCSPIMTKMHQQGAGAQ.....AGGGPGANCGQQAGGFGG.....
A.franciscana EKVCNPIITKLYGQAGGML.....ADSLVVWRSSSGCYCSRCWNRQWPNY....
M.gall EGVCNPIITKLYQSAGGAPGGGMPNFGGAGGAPGGAPGAGGT..GGSGGPTIEEVD
M.sexta EGICNPIITKLYQGAGGMP.GGMP..GGMPGFPGGAPGAGGAAPGGGAGPTIEEVD

Figure A.2: Alignment of translated *Hsp90* sequences from *Caenorhabditis elegans* (NP_506626.1), *Calanus finmarchicus* (this study), *Metapenaeus ensis* (ABR66910.1), *Penaeus monodon* (ABM54577.1), *Drosophila melanogaster* (D.mel; NP_523899.1), and *Apis mellifera* (NP_001153536.1). The partial *Hsp90* sequence identified in this study represents 119 amino acids and aligns at the 3' end of *Hsp*90A (~720 amino acids).



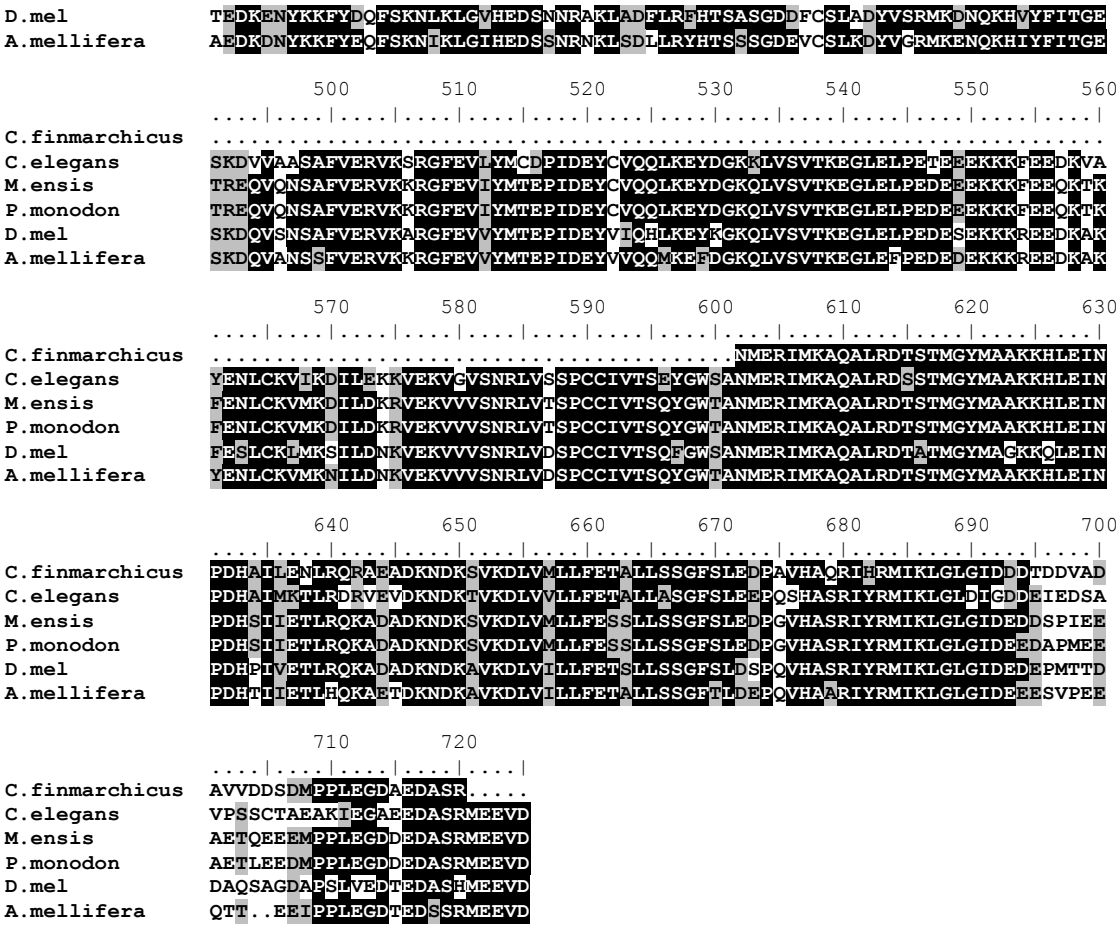

D.mel
A.mellifera
500 510 520 530 540 550 560
C.finmarchicus
C.elegans
M.ensis
P.monodon
D.mel
A.mellifera
570 580 590 600 610 620 630
C.finmarchicus
C.elegans
M.ensis
P.monodon
D.mel
A.mellifera
640 650 660 670 680 690 700
C.finmarchicus
C.elegans
M.ensis
P.monodon
D.mel
A.mellifera
710 720
C.finmarchicus
C.elegans
M.ensis
P.monodon
D.mel
A.mellifera

Figure A.3: Alignment of translated *Hsp21* sequences from *Calanus finmarchicus* (this study), *Artemia franciscana* (ABD19712), and *Bombyx mori* (BAD74197.1). The *Hsp21*-like sequence identified in this study represents 181 amino acids aligned towards the middle of an approximately 200 amino acid long Hsp21 protein.

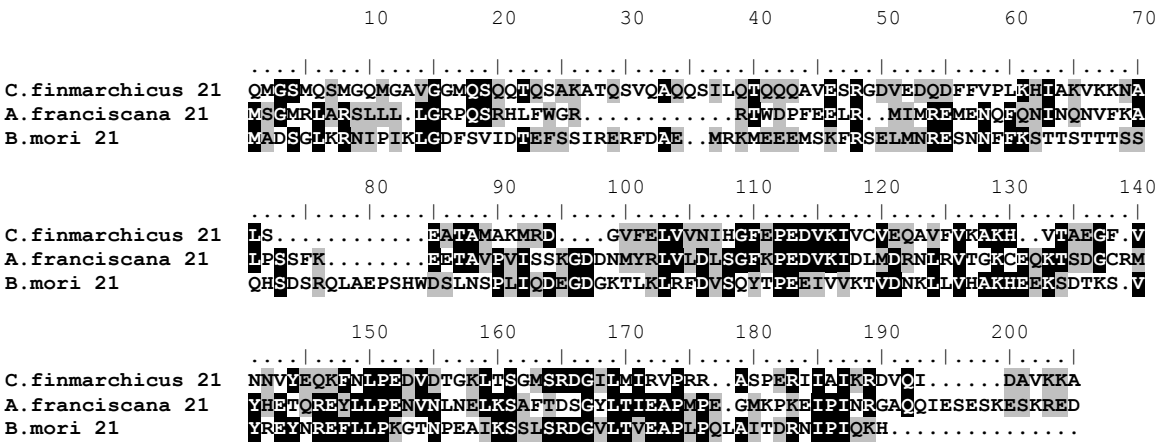


Figure A.4: Alignment of translated *Hsp22* sequences from *Calanus finmarchicus* (this study), *Artemia franciscana* (ABD19713), and *Culex quinquefasciatus* (XP_001847194). The *Hsp22*-like sequence identified in this study represents 168 amino acids aligned towards the 5' end of an approximately 200 amino acid long Hsp22 protein.

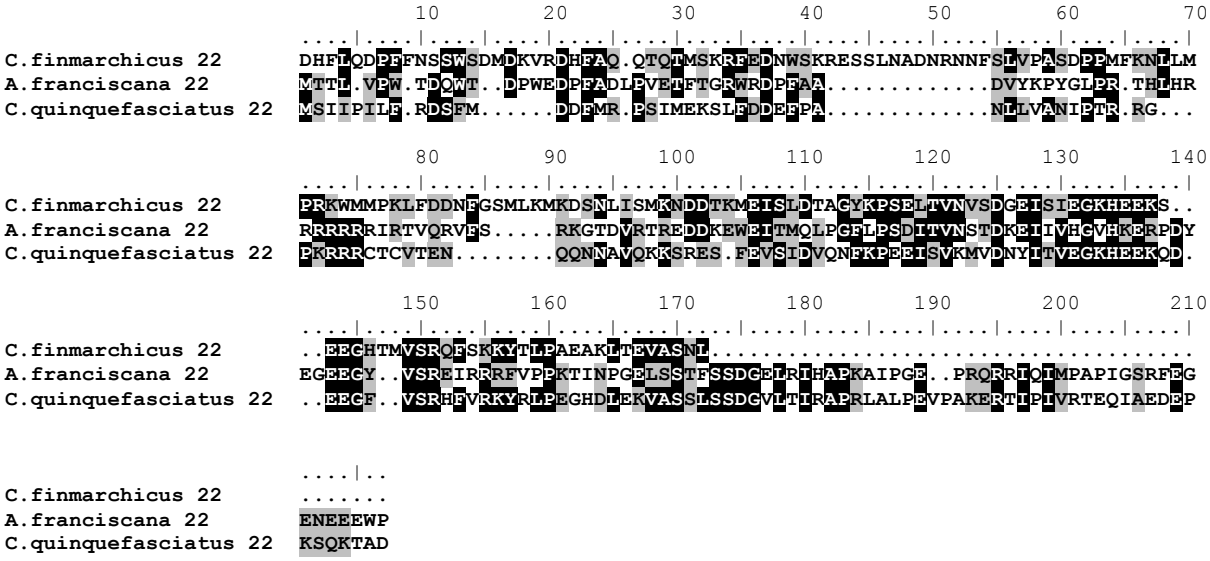
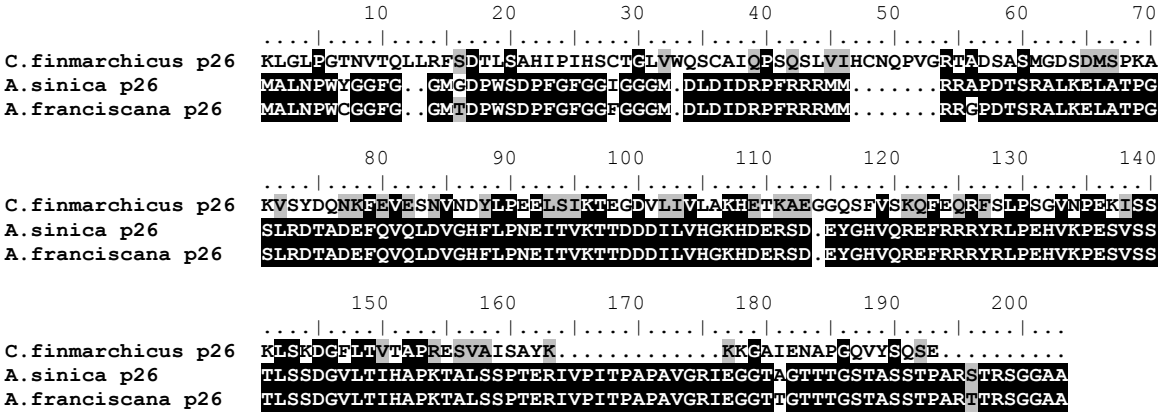


Figure A.5: Alignment of translated *p26* sequences from *Calanus finmarchicus* (this study), *Artemia sinica* (ABC41137.1), and *Artemia franciscana* (ABC41138). The *p26*-like sequence identified in this study represents 180 amino acids aligned towards the 5' end of an approximately 200 amino acid long p26 protein.



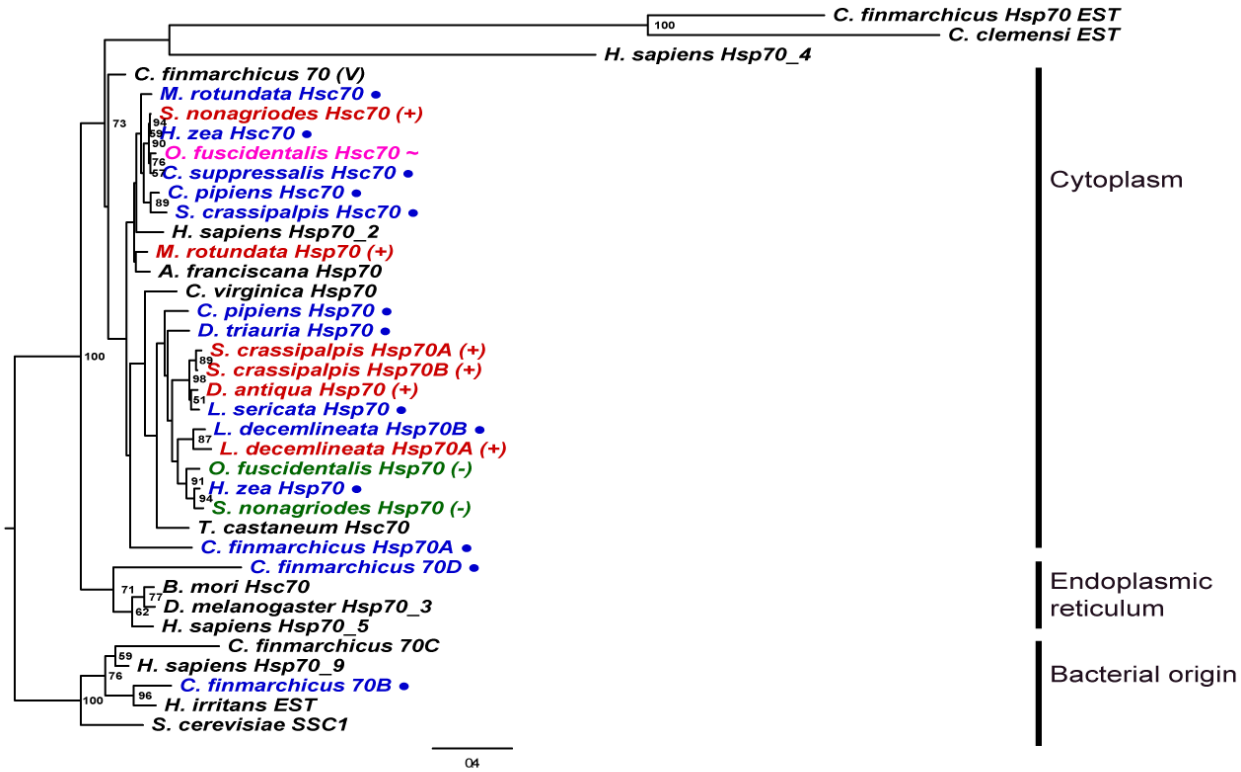


Fig. B.1: Maximum likelihood analysis of *Hsp70* from *Calanus finmarchicus*, several diapausing insect species, and representative cytosolic, endoplasmic reticulum, mitochondrial, plastid, and bacterial *Hsp70* members. Bootstrap percentages of 1000 replicates are indicated above branches when they are greater than 50%. Sequences of bacterial origin were used as the outgroup. Sequences selected for these analyses are principally from Rhee et al. (2009), Boorstein et al. (1994), Daugaard et al. (2007), and MacRae (2010) with some additions (see Table B.1 for full list of sequences and accession numbers). Symbols and colors denote *Hsp70* forms that are non-inducible [●], inducible [*(+)*], down-regulated [*(-)*], or show variable expression in response to diapause [~] according to findings in our study and those reviewed by MacRae (2010). Distance bar at the bottom of the tree indicates branch length scale, or the number of substitutions per amino acid site.

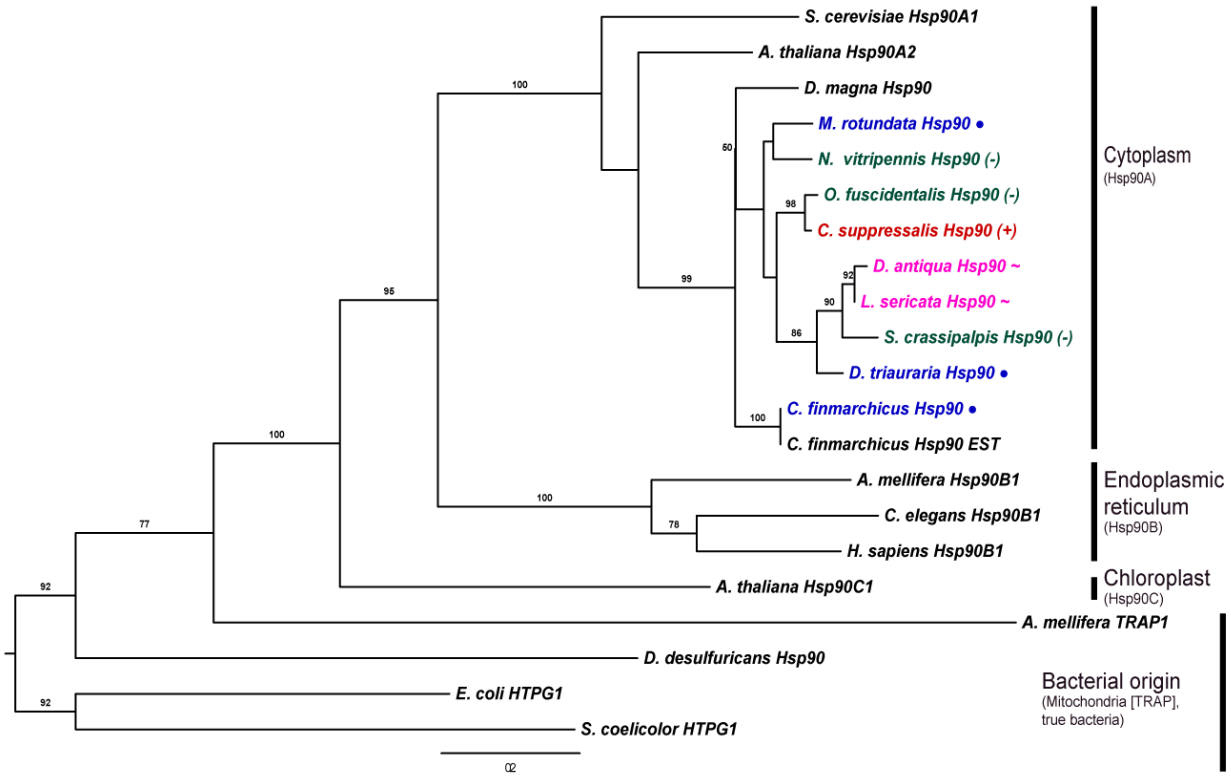


Fig. B.2: Maximum likelihood analysis of *Hsp90* from *Calanus finmarchicus*, several diapausing insect species, and representative cytosolic (*Hsp90A*), endoplasmic reticulum (*Hsp90B*), chloroplast (*Hsp90C*), and mitochondrial (TRAP) *Hsp90* members. Bootstrap percentages of 1000 replicates are indicated above branches when they are greater than 50%. Bacterial *Hsp90* homologs (i.e., high-temperature protein G, HTPG) were used to root the *Hsp90* tree. Sequences selected for these analyses are primarily from Chen et al. (2005; 2006) with some additions (see Table B.2 for full list of sequences and accession numbers). Symbols and colors denote *Hsp90* forms that are non-inducible [•], inducible [(+)], down-regulated [(-)], or show variable expression in response to diapause [~] according to findings in our study and those reviewed by MacRae (2010). Distance bar at the bottom of the tree indicates branch lengths or the number of substitutions per amino acid site.

Table B.1: *Hsp70* sequences used in phylogenetic analyses. *Hsp70* sequences and information regarding their cellular localization and responsiveness to stress are principally from Rhee et al. (2009), Boorstein et al. (1994), Daugaard et al. (2007), and MacRae (2010) as noted, with some additions.

Species	**Common name**	**Localization**	**[c]Stress/ [d]diapause response**	**[e]Amino acid**	**Reference**
*Bacillus megaterium Hsp*70	Bacteria			P05646	Boorstein et al., 1994
*Chlamydophila pneumoniae Hsp*70	Bacteria			AAA23121.1	Boorstein et al., 1994
Culex quinquefasciatus EST	Mosquito	[a]Mitochondria		*FF223480*	
*Drosophila melanogaster Hsc*5	Fruit fly	[a]Bacterial origin	[c]Not inducible	NP_523741	
*Escherichia coli Hsp*70	Bacteria			P04475	Boorstein et al., 1994
Haematobia irritans EST	Horn fly	[a]Mitochondria		*FD461791*	
*Homo sapiens Hsp*70_9	Human	Mitochondria	[c]Not inducible	NP_004125.3	Daugaard et al., 2007
*Porphyra purpurea Hsp*70	Red algae	Plastid		CAA44160.1	Boorstein et al., 1994
Saccharomyces cerevisiae SSC1	Baker's yeast	Mitochondria		AAA63792.1	Boorstein et al., 1994
*Aedes aegypti Hsc*70	Mosquito	[a]ER	[c]Not inducible	ABF18258	Rhee et al., 2009
*Bombyx mori Hsp*70	Silk moth	[a]ER	[c]Inducible	BAA32395	Rhee et al., 2009
*Drosophila melanogaster Hsp*70_3	Fruit fly	[a]ER	[c]Not inducible	NP_727563	
*Homo sapiens Hsp*70_5	Human	ER	[c]Not inducible	NP_005338	Daugaard et al., 2007
Nasonia vitripennis EST	Parasitic wasp	[a]ER		*XM_001606413*	
Saccharomyces cerevisiae KAR2	Baker's yeast	ER			Boorstein et al., 1994
Schizosaccharomyces pombe	Fission yeast	ER		CAA45762	Boorstein et al., 1994
*Spodoptera frugiperda Hsc*70	Fall armyworm	[a]ER	[c]Not inducible	AAN86047	Rhee et al., 2009
*Artemia franciscana Hsp*70	Brine shrimp	[a]Cytoplasm		AAL27404	Rhee et al., 2009
*Blomia tropicalis Hsp*70	Mite	[a]Cytoplasm		AAQ24552	Rhee et al., 2009

*Bombyx mori Hsc*70	Silk moth	[a]Cytoplasm	[c]Not inducible	BAB92074	Rhee et al., 2009
*Calanus finmarchicus Hsp*70 (V)	Marine copepod	[a]Cytoplasm		(not in NCBI)	Voznesensky et al. 2004
*Chilo suppressalis Hsc*70	Striped rice borer	[a]Cytoplasm	[c]Not inducible	BAE44308	Rhee et al., 2009
*Chironomus yoshimatsui Hsc*70	Midge larva	[a]Cytoplasm	[c]Not inducible	AAN14526	Rhee et al., 2009
*Crassostrea virginica Hsp*70	Eastern oyster	[a]Cytoplasm	[c]Inducible	CAB89802	Rhee et al., 2009
*Drosophila melanogaster Hsc*70_4	Fruit fly	[a]Cytoplasm	[c]Not inducible	NP_524356	
*Drosophila melanogaster Hsp*701A	Fruit fly	[a]Cytoplasm		NP_524063	
*Drosophila melanogaster Hsp*70AB	Fruit fly	[a]Cytoplasm		AAG26896	
*Homo sapiens Hsp*70_1AB	Human	[a]Cytoplasm		NP_005337	
*Homo sapiens Hsp*70_2	Human	Cytoplasm	[c]Not inducible	NP_068814	Daugaard et al., 2007
*Homo sapiens Hsp*70_6	Human	Cytoplasm	[c]Inducible	NP_002146	Daugaard et al., 2007
*Ostrea edulis Hsp*70	European oyster	[a]Cytoplasm	[c]Inducible	AAM46635	Rhee et al., 2009
*Pachygrapsus mamoratus Hsp*70	Marbled crab	[a]Cytoplasm		ABA02164	Rhee et al., 2009
Saccharomyces cerevisiae SSA1	Baker's yeast	Cytoplasm		CAA31393	Boorstein et al., 1994
*Tetranychus urticae Hsc*70	Spider mite	[a]Cytoplasm		ABC33921	Rhee et al., 2009
*Tigriopus japonicus Hsp*70	Marine copepod	[a]Cytoplasm	[c]Inducible	ABX89903	Rhee et al., 2009
*Tribolium castaneum Hsc*70	Red flour beetle	[a]Cytoplasm	[c]Not inducible	XP_973521	Rhee et al., 2009
Calanus finmarchicus 70 EST	Marine copepod		[b,c] Not inducible	*EH666705*	Hansen et al., 2008
Caligus clemensi EST	Copepod sea louse			*GO396551*	
*Drosophila melanogaster Hsc*70CB	Fruit fly			NP_648687	
*Homo sapiens Hsp*70_14	Human		[c]Not inducible	NP_057383	Wan et al., 2004
*Homo sapiens Hsp*70_4	Human		[c]Not inducible	NP_002145	Kaneko et al., 1997
Lepeophtheirus salmonis EST	Copepod sea louse			FK926374	
*Chilo suppressalis Hsc*70	Rice stem borer	[a]Cytoplasm	[d]Constant	BAE44308	MacRae, 2010
*Culex pipiens Hsp*70	Mosquito	[a]Cytoplasm	[d]Down	AAX84696.1	MacRae, 2010

*Culex pipiens Hsc*70	Mosquito	[a]Cytoplasm	[d]Down	XP_001850527.1	MacRae, 2010
*Delia antiqua Hsp*70	Onion maggot	[a]Cytoplasm	[d]Up	AAY28732.1	MacRae, 2010
*Drosophila triauraria Hsp*70	Fruit fly	[a]Cytoplasm	[d]Constant	BAC77410.1	MacRae, 2010
*Helicoverpa zea Hsc*70	Corn earworm	[a]Cytoplasm	[d]Constant	ACV32641.1	MacRae, 2010
*Helicoverpa zea Hsp*70	Corn earworm	[a]Cytoplasm	[d]Down	ACV32640.1	MacRae, 2010
*Leptinotarsa decemlineata Hsp*70A	CO potato beetle	[a]Cytoplasm	[d]Up	AAG01177.2	MacRae, 2010
*Leptinotarsa decemlineata Hsp*70B	CO potato beetle	[a]Cytoplasm	[d]Constant	AAG42838.1	MacRae, 2010
*Lucilia sericata Hsp*70	Blow fly	[a]Cytoplasm	[d]Constant	BAD12047.1	MacRae, 2010
*Megachile rotundata Hsc*70	Leafcutter bee	[a]Cytoplasm	[d]Constant	AAS57865.1	MacRae, 2010
*Megachile rotundata Hsp*70	Leafcutter bee	[a]Cytoplasm	[d]Up	AAS57864.1	MacRae, 2010
*Omphisa fuscidentalis Hsc*70	Bamboo borer	[a]Cytoplasm	[d]Variable	ABP93403.1	MacRae, 2010
*Omphisa fuscidentalis Hsp*70	Bamboo borer	[a]Cytoplasm	[d]Down	ABP93405.1	MacRae, 2010
*Sarcophaga crassipalpis Hsc*70	Flesh fly	[a]Cytoplasm	[d]Constant	AAD17996.1	MacRae, 2010
*Sarcophaga crassipalpis Hsp*70A	Flesh fly	[a]Cytoplasm	[d]Up	AAD17995.2	MacRae, 2010
*Sarcophaga crassipalpis Hsp*70B	Flesh fly	[a]Cytoplasm	[d]Up	ABL06944.1	MacRae, 2010
*Sesamia nonagrioides Hsc*70	Corn stalk borer	[a]Cytoplasm	[d]Up	AAY26452.2	MacRae, 2010
*Sesamia nonagrioides Hsp*70	Corn stalk borer	[a]Cytoplasm	[d]Down	ABZ10939.1	MacRae, 2010

[a] denotes inferred cellular localization from our phylogenetic analyses

[b]Hansen et al., 2008 saw modest but insignificant increases of *Hsp*70 in response to naphthalene exposure

[c] denotes response to stress

[d] denotes response to diapause. Diapausing insect species and indicated diapause response are from MacRae, 2010 and internal citations.

[e] Accession numbers in italics represent nucleotide sequences

Table B.2: *Hsp90* sequences used in phylogenetic analyses. *Hsp90* sequences are principally from Chen et al. (2006; 2005) and MacRae (2010) as noted, with some additions.

Species	**Common name**	**[a]Stress/[b]diapause response**	**[c]Acession no.**	**Reference**
*Apis mellifera Hsp*90A1	Honey bee		NP_001153536.1	Chen et al., 2006
*Apis mellifera Hsp*90A2	Honey bee		XP_395168.3	Chen et al., 2006
*Apis mellifera Hsp*90B1	Honey bee		XP_395614.3	Chen et al., 2006
Apis mellifera TRAP1	Honey bee		XP_623366.1	Chen et al., 2006
*Arabidopisis thaliana Hsp*90A1	Thale cress		NP_200076.1	Chen et al., 2006
*Arabidopisis thaliana Hsp*90A2	Thale cress		NP_200414.1	Chen et al., 2006
*Arabidopisis thaliana Hsp*90B1	Thale cress		NP_194150.1	Chen et al., 2006
*Arabidopisis thaliana Hsp*90C1	Thale cress		NP_178487.1	Chen et al., 2006
*Caenorhabditis elegans Hsp*90A1	Nematode		NP_506626.1	Chen et al., 2006
*Caenorhabditis elegans Hsp*90B1	Nematode		NP_502080.1	Chen et al., 2006
Caenorhabditis elegans TRAP1	Nematode		NP_741219.2	Chen et al., 2006
*Calanus finmarchicus Hsp*90 (EST)	Marine copepod	[a]Inducible	*ES414827.1*	
Caligus rogercresseyi (EST)	Copepod sea louse		*FK890997.1*	
*Desulfovibrio desulfuricans Hsp*90	Bacteria		YP_387643.1	
*Daphnia magna Hsp*90	Water flea		ABI35831.1	
*Drosophila melanogaster Hsp*90A1	Fruit fly		NP_523899.1	Chen et al., 2006
*Drosophila melanogaster Hsp*90B1	Fruit fly		NP_651601.1	Chen et al., 2006
Drosophila melanogaster TRAP1	Fruit fly		NP_477439.2	Chen et al., 2006
*Desulfovibrio vulgaris Hsp*90	Bacteria		YP_011855	
Escherichia coli HTPG1	Bacteria		NP_415006.1	Chen et al., 2006
Homo sapiens Hsp90AA1	Human		NP_005339.2	Chen et al., 2005
*Homo sapiens Hsp*90AA2	Human		AAA36024.1	Chen et al., 2005

*Homo sapiens Hsp*90AB1	Human		NP_031381.2	Chen et al., 2005
*Homo sapiens Hsp*90B1	Human		NP_003290.1	Chen et al., 2005
Homo sapiens TRAP1	Human		NP_057376.1	Chen et al., 2005
Lepeophtheirus salmonis (EST)	Copepod sea louse		FK920985.1	
*Metapenaeus ensis Hsp*90	Sand shrimp		ABR66910.1	
*Penaeus monodon Hsp*90	Giant tiger prawn		ABM54577.1	
*Saccharomyces cerevisiae Hsp*90A1	Baker's yeast		AAA02743.1	Chen et al., 2006
*Saccharomyces cerevisiae Hsc*90A2	Baker's yeast		AAA02813.1	Chen et al., 2006
Streptomyces coelicolor HTPG1	Bacteria		NP_631561	
*Tigriopus japonicus Hsp*90	Marine copepod	[a]Inducible	ACA03524.1	
Lucilia sericata Hsp90	Blow fly	[b]Variable	BAD12048.1	MacRae et al., 2010
*Chilo suppressalis Hsp*90	Rice stem borer	[b]Up	BAE44307.1	MacRae et al., 2010
*Delia antiqua Hsp*90	Onion maggot	[b]Variable	CAI64494.1	MacRae et al., 2010
*Drosophila triauraria Hsp*90	Fruit fly	[b]Constant	BAC77528.1	MacRae et al., 2010
*Helicoverpa zea Hsp*90	Corn earworm	[b]Down	ACV32639.1	MacRae et al., 2010
*Megachile rotundata Hsp*90	Solitary bee	[b]Constant	AAS57866.1	MacRae et al., 2010
*Nasonia vitripennis Hsp*90	Parasitic wasp	[b]Down	XP_001605191.1	MacRae et al., 2010
*Omphisia fuscidentalis Hsp*90	Bamboo borer	[b]Down	ABP93404.1	MacRae et al., 2010
*Sarcophaga crassipalpis Hsp*90	Flesh fly	[b]Down	AAF69019.1	MacRae et al., 2010

[a] denotes response to stress

[b] denotes response to diapause. Diapausing insect species and indicated diapause response are reviewed by MacRae (2010)

[c]Accession numbers in italics are nucleotide sequences

Chapter Three

The copepod *Eurytemora affinis* mounts a targeted transcriptional response to colonization by *Vibrio* bacteria[1]

[1] Preliminary author list: Almada AA, Tarrant AM
Proposed journal: Applied and Environmental Microbiology

ABSTRACT:

Living zooplankton, such as copepods, are highly abundant environmental reservoirs of many bacterial pathogens. Although copepods are known to support diverse and productive bacterial communities, little is understood about whether copepods actively interact with or regulate their bacterial associations. Here we report that the estuarine copepod *Eurytemora affinis* mounts a specific response to colonizing *Vibrio* species. We exposed adult female *E. affinis* to a putative copepod symbiont (*Vibrio sp. F10 9ZB36*) or a free-living *Vibrio* (*V. ordalii 12B09*) for 24 hours and compared the *E. affinis* transcriptomic response using high-throughput sequencing. We identified a total of 78 genes differentially expressed with *Vibrio* exposure, with many of these genes associated with the innate immune response, cuticle integrity, and the general stress response. The magnitude of the *E. affinis* transcriptomic response varied greatly between the two *Vibrio* treatments, with the majority of the genes differentially expressed with *V. sp. F10*-exposure and very few with *V. ordalii*-exposure (3 genes down-regulated). We further profiled the expression of six genes up-regulated upon *V. sp. F10* association by quantitative PCR, and we suggest that the expression of these genes (3 C-type lectin-like, 2 chitin-binding, 1 saposin-like transcript) may represent a mechanism by which *E. affinis* recognizes and maintains symbiotic *Vibrio* bacteria. We further report that *Vibrio* differ in their ability to attach to *E. affinis* and that colonization of *E. affinis* can specifically alter *Vibrio* culturability. These findings suggest that rather than merely being passive vectors, copepods can actively regulate their *Vibrio* colonizers, which may ultimately impact the abundance, activity, and community structure of copepod-associated vibrios.

INTRODUCTION:

Beyond the classically described contribution of zooplankton to aquatic food webs, their role as environmental reservoirs of a diversity of microbes, including many bacterial pathogens, is increasingly recognized (1-4). The dynamic associations with living zooplankton convey multiple benefits to their associated bacterial communities including a steady supply of nutrients (5-7) and protection from environmental stressors (2), leading to increased growth rates (8) and production (9) relative to those bacteria free-living in the surrounding seawater. Attachment to migrating zooplankton permits widespread dispersal of "hitchhiking" bacteria to new environments (1) and provides access to unique microenvironments (10, 11), which is reflected in the distinct microbial communities found on zooplankton relative to the ambient environment (12, 13). Colonization of the copepod exoskeleton by *Vibrio* bacteria is a relatively well-studied zooplankton-bacteria interaction due to the prevalence of pathogenic *Vibrio*s (e.g., *V. cholerae*, *V. parahaemolyticus*) on these abundant chitinous organisms (3, 4, 12, 14-17) and the dramatic impacts of these associations on the proliferation, virulence, and physiology of many *Vibrio* species (2, 18-21). Despite evidence that living zooplankton in particular harbor unique and distinct *Vibrio* communities (22, 23), little work has been done to explore whether the specificity of these communities may derive from direct action by the zooplankton host. A growing body of literature has demonstrated the importance of invertebrate host factors in selecting and maintaining obligate bacterial symbionts (24-31), yet whether copepod hosts are able to actively select for or against particular *Vibrio* colonizers is still unknown.

Crustaceous zooplankton, such as copepods, are constantly immersed in an environment enriched with potential microbial invaders. Crustaceans' first line of defense against unwanted microbial invasion is a hard, rigid exoskeleton that provides a physical and chemical barrier against pathogen attachment and invasion (32, 33). The gut is also a vulnerable point of entry into the host, but the thin peritrophic membrane, composed of proteins and chitin, lines the mid-gut of arthropods to protect against abrasion by hard food fragments and to serve as a barrier against microbial invasion (34-36). When the dynamic protection of the cuticle fails, activation of the innate immune defense quickly follows. Although crustaceans lack an adaptive immune system equipped with specificity and memory, recent studies have revealed a deeper level of complexity of the innate immune system than previously appreciated. For example, copepods demonstrated specific memory upon repeated encounters with cestode parasites (37), and the

offspring of *Daphnia* mothers exposed to a particular bacterial strain had enhanced fitness upon challenge with the same strain (38). The complex activity of the innate immune system depends on dynamic mechanisms, such as the prophenoloxidase (proPO) cascade and antimicrobial peptide (AMP) production. Activation of the proPO system is thought to be an early and rapid response to microbial invaders that results in the production of cytotoxic compounds and melanin, a pigment that inactivates and physically encapsulates microbial invaders (39, 40). AMPs are key effectors in the elimination of bacteria, including *Vibrio* (41), and they are easily and rapidly synthesized due to their small size (42). AMPs act by forming holes in microbial membranes to disrupt cellular integrity and function (42) and by binding to bacterial cells to enhance downstream phagocytosis of these cells (43). Recent work has shown that in addition to inhibiting growth of undesirable foreigners, the production of AMPs enables invertebrate hosts to select for specific microbial associates (26, 29-31).

Despite the growing understanding of the complexity of the innate immune response and the potential for dynamic interactions between invertebrate hosts and their microbiota (26, 27, 29-31), little attention has been paid to whether copepods *actively* interact with their microbial communities. In this study, we addressed the question of whether copepods differentially respond to distinct *Vibrio* colonizers. We chose as our model the copepod *Eurytemora affinis,* an invasive and abundant species (44, 45) that naturally associates with a diversity of pathogenic vibrios (46) and has been consistently used in the few laboratory studies examining copepod-*Vibrio* interactions (14, 15, 47). The objective of this study was to quantitatively measure the colonization densities of two ecologically distinct *Vibrio* species (*V. sp F10* and *V. ordalii*) on the copepod *E. affinis* and the subsequent transcriptomic response of *E. affinis* to the *Vibrio* colonizers. The two *Vibrio* species selected inhabit similar coastal environments to *E. affinis* (14, 48, 49) and possess distinct physical characteristics and ecological specializations: *V. sp. F10* is classified as a putative copepod symbiont (49) that lacks the ability to degrade chitin (50), while *V. ordalii 12B09* is classically described as a 'free-living' *Vibrio* (48). *V. ordalii* has been implicated in vibriosis of fish (51) and the strain *V. ordalii 12B09* also demonstrates pathogenicity towards the marine invertebrate *Artemia* (T. Mincer, *personal communication*) through the production of the signaling chemical indole (52). We found that *V. ordalii* elicited virtually no transcriptomic response in *E. affinis*, whereas *V. sp. F10* elicited a highly consistent and targeted immune response in the copepod host. Furthermore, we observed that colonization

of *E. affinis* triggered dramatic changes in the culturability of *V. sp. F10*. Our findings provide evidence that the copepod *E. affinis* distinctly recognizes and responds to colonizing vibrios and suggest that association with *E. affinis* can significantly impact the physiology of *Vibrio* colonists.

MATERIALS AND METHODS:

Bacterial cultures

Several *Vibrio* species (Table S1) were screened for their pathogenicity towards and ability to colonize *Eurytenora affinis*. Ultimately, *V. ordalii 12B09 WT* and *V. sp. F10 9ZB36*, were selected for the *E. affinis* transcriptome profiling experiments based on their high colonization densities of *E. affinis* and their distinct ecological specializations and pathogenicities towards *E. affinis*. *Vibrio* growth was measured over 24 hours under different temperature (15 °C, 18 °C, 30 °C) and salinity (15 PSU, 30 PSU) regimes to confirm their ability to survive and grow under the *E. affinis* exposure conditions (15 PSU, 18 °C) (Figure S1).

To prepare *Vibrio* cultures for *E. affinis* exposure experiments, glycerol stocks of *Vibrio* cultures (stored at -80 °C) were streaked onto seawater complete (SWC) agar plates containing 15 PSU artificial seawater (ASW), peptone, yeast extract, and glycerol before a 24 hour incubation at room temperature. Several colonies were then transferred into 10 mL of SWC liquid media, shaken at 200 rpm, and incubated for 19 hours at 28 °C (18 °C for *V. sp. F10* based on growth curve results; Figure S1). For RNA-Seq and qPCR experiments profiling the transcriptomic response of *E. affinis* to *Vibrio*, 1 mL of the overnight liquid culture was transferred to 100 mL of SWC liquid media (15 PSU) and incubated for 19 hours at 100 rpm (*V. ordalii* at 28 °C and *V. sp F10* at 18 °C). For all experiments, cultures were pelleted at 5,500 x g for 5 minutes, and rinsed twice with 0.22-µm sterile filtered artificial seawater (15 PSU) before diluting to the desired cell density (2×10^7 CFU mL^{-1} for the Illumina expression profiling and qPCR studies).

To test whether *V. sp. F10* and *V. ordalii* secrete extracellular chitinases, overnight cultures were grown in SWC media as described above and spread onto plates comprised of approximately 2% (w/v) colloidal chitin in 1x marine agar (2216). Colloidal chitin was prepared

from crab shell chitin flakes (Sigma) (53) and dyed with Remazol Brilliant Violet (54). When extracellular chitinases hydrolyze the chitin substrate and covalently linked dye, a clear halo is left surrounding the chitinase-producing culture. Those cultures that do not secrete chitinases will successfully grow on the plate but will not produce a clear halo.

Gnotobiotic cultures of the estuarine copepod Eurytemora affinis

Eurytemora affinis cultures that originated from the Baie de L'isle Verte in the St. Lawrence estuary were generously provided by Carol Lee (University of Wisconsin). The copepod cultures were maintained at 12 °C and 15 PSU on a 14 h light/10 h dark cycle with moderate air bubbling (1-2 bubbles/second). The cultures were fed with the red cryptophyte *Rhodomonas lens* three times a week at a concentration of 1 x 10^6 cells mL^{-1}.

At the beginning of the *Vibrio* exposure experiments, *E. affinis* were treated for 24 hours with an antibiotic mixture of ampicillin (0.3 mg mL^{-1}), streptomycin (0.1 mg mL^{-1}), and chloramphenicol (0.05 mg mL^{-1}) in autoclaved seawater (15 °C, 15 PSU). The antibiotic treatment was performed at 15 °C in order to gradually acclimate the copepods to the higher temperature used in the *Vibrio* exposure experiments (18 °C) that is more amenable to *Vibrio* growth. During the antibiotic treatment, copepods were aerated with moderate bubbling (~2 bubbles/sec) and fed 1 x 10^6 cells mL^{-1} of *R. lens*. The efficacy of the antibiotic treatment was tested by placing individual whole copepods, homogenized copepods, or 400 µL of seawater from flasks containing either antibiotic-treated or untreated copepods into 2 mL of Marine Broth 2216 (Difco). The presence of bacteria was estimated by measuring the absorbance at 600 nm of each Marine Broth treatment after 48 hours of incubation at 22 °C. We found that the antibiotic cocktail virtually eliminated the natural microbiota of *E. affinis* individuals (Figure S2) and had no observable impact on copepod survival, mobility, or feeding.

We also tested the potential impact of residual antibiotics in the copepods on *Vibrio* growth. A pool of five antibiotic-treated mature, adult female copepods were rinsed in sterile artificial seawater, homogenized in 200 µL of sterile artificial seawater, and passed through a 0.22-µm filter to remove copepod parts and any remaining microbiota, while retaining any residual antibiotics. We consider 'mature, adult females' to include ovigerous females and non-

ovigerous females with enlarged oviducts full of large oocytes, as previously defined (55). The filtered copepod solution was diluted 1:3 in sterile ASW before addition to equal volumes of SWC liquid media and *V. sp. F10* culture (1:1:1). Antibiotic control samples consisted of SWC liquid media added to equal volumes of sterile ASW and *V. sp. F10* culture, while the negative control consisted of SWC liquid media. All treatments were aliquoted into triplicate wells of one 96-well plate and incubated at 18 °C for 24 hours before reading the absorbance at 600 nm. The sterile-filtered, homogenized copepod sample had 1.5-fold higher growth than the antibiotic control samples, demonstrating no harmful impact of residual antibiotics on *V. sp. F10* growth. The increased growth in the antibiotic treatment is likely due to the nutrients released by homogenizing and sterile-filtering the copepods.

E. affinis- Vibrio exposure experiments

Copepods were fed 1 x 10^6 cells mL^{-1} of *R. lens* 24 and 48 hours before *Vibrio* exposure. After 24 hours of treatment with an antibiotic cocktail (as described above), copepods were rinsed with autoclaved seawater (15 °C, 15 PSU) onto an autoclaved 400-µm sieve and into sterile Petri dishes. To validate the effectiveness of the antibiotic treatment within each experiment, three antibiotic-treated copepods were placed into individual 600 µL aliquots of SWC media and incubated for 24 hours at room temperature before reading the absorbance at 600 nm. Any experiments with significant growth in the SWC treatments by 24 hours were terminated.

For survival and attachment experiments, 5 adult copepods (males and females) were captured with a transfer pipette for each treatment replicate (n = 2). For the transcriptomic profiling experiment, 20 mature, adult females were captured for each treatment replicate (n = 5). Follow-up qPCR experiments were performed with pools (n = 10 - 20 per replicate) of mature, adult females. Once copepods were placed into autoclaved 50 mL glass flasks containing autoclaved seawater (15 °C, 15 PSU), the flasks were inoculated with the desired titer of *Vibrio*. The samples were then moved to an 18 °C incubator (14 h light/10 h dark cycle), where they were aerated with moderate bubbling for 24 hours. Copepod samples used for gene expression studies were gently rinsed onto autoclaved 333-µm mesh, transferred using plastic pipettors into

1 mL of PureZOL (Bio-Rad), and stored at -80 °C until RNA extraction (Figure S3). Follow-up qPCR analysis also tested whether the *E. affinis* response to *V. ordalii* was influenced by the production of indole, an exotoxin known to be produced by wild type *V. ordalii* as a grazing deterrent (52). Specifically, pools of mature, adult female *E. affinis* (n = 10 per replicate) were exposed to either *V. sp. F10 9ZB36, V. ordalii 12B09* Wild Type ("WT"), or *V. ordalii 12B09* Δ*tnaA* ("K/O", no indole production) (52) in two independent experiments before qPCR analysis.

To quantify the colonization density of *Vibrio*s on *E. affinis* after 24 hours, copepods were rinsed onto an autoclaved 333-μm mesh sieve with sterile ASW (18 °C, 15 PSU). After noting any mortality, pools of copepods (n = 3-5 per replicate) were transferred to 1.5 mL microcentrifuge tubes containing 200 μL of filter-sterilized ASW and homogenized with sterile plastic darts. Homogenized copepods were then serially diluted, plated onto SWC agar, and incubated for 20 hours at room temperature before counting colony-forming units (CFU). Follow-up experiments examined the growth of *V. sp. F10* or *V. ordalii WT* recovered from *E. affinis* and plated on thiosulfate-citrate-bile salts-sucrose (TCBS) agar, a culture medium that is highly selective for *Vibrio* species. Direct counts of the copepod-associated *Vibrio*s were also performed by fixing the remaining serial dilutions of homogenized copepods in formalin (1%), staining with DAPI (10%), and filtering onto a 0.22-μm black polycarbonate filter. Filters were mounted in Citifluor (Citifluor Ltd.), and cells were counted with an epifluorescence scope under blue light excitation.

RNA extractions and library sequencing

Total RNA was extracted from *E. affinis* samples using the Aurum Total RNA Fatty and Fibrous Tissue Kit (Bio-Rad). Samples were homogenized in 1 ml PureZOL using a teflon homogenizer and processed according to the manufacturer's protocol, with final elution from columns in 40 μL of warmed elution buffer (Tris buffer). For the qPCR study, residual genomic DNA was removed with on-column DNase digestion, and the final elution was followed by re-precipitation with 1 μL of Glycoblue (Life Technologies), 2.5x volume of 100% ethanol, and 0.15x volume of 3.25M ammonium acetate to increase the purity of extracted samples. RNA yield and purity were quantified using a Nanodrop ND-1000 spectrophotometer, and RNA

quality was visualized on a denaturing agarose gel. Quality of RNA samples submitted for Illumina sequencing was further assessed using a Bioanalyzer. The *E. affinis* samples, like many other arthropods, yield one sharp peak on the Bioanalyzer due to a hidden break in their 28S rRNA that causes it to run at about the same size as the 18S rRNA.

Directional, polyA-enriched RNA libraries were built by the Hudson Alpha Genomic Services Laboratory with the NEBNext® Ultra Directionality Kit (New England BioLabs) from 1 μg of total RNA from each sample. The average fragment size of each library was approximately 300 bp. Libraries were quantified by quantitative PCR using a SYBR Fast Illumina Library Quantification Kit (Kapa Biosystems) before sequencing on a HiSeq2000 machine. For transcriptome assembly, a library was constructed from a sample of pooled RNA made by combining approximately 200 ng from each sample [Control (n=4), *V. sp. F10*-exposed (n=4), *V. ordalii*-exposed (n=4)]. The library constructed from this pooled sample was sequenced using a paired-end approach with 100 bp reads at a total sequencing depth of 50 million reads. The libraries constructed from each of the twelve individual samples were multiplexed and sequenced across two lanes of the HiSeq2000 using a paired-end approach with 50 bp reads at a total depth of 25 million reads per sample for differential expression analysis.

De novo transcriptome assembly and post-assembly analysis

A library constructed from pooled adult female *E.affinis* RNA and sequenced on the Illumina HiSeq2000 platform produced 111 million 100-bp, paired-end, directional reads. Trimmomatic software (56) was used in paired-end mode to remove adaptor sequences, low quality sequences (phred score < 20 bp), and the first 12 bp of the 5' end of the read, which often contain a biased nucleotide composition due to nonrandom hexamer priming (57). Reads greater than 50 bp in length after quality trimming were retained for assembly, resulting in a total of 102 million reads for assembly. An *E. affinis* transcriptome was assembled *de novo* with the RNA-seq assembler Trinity (version r2013-08-14) using default parameters for paired-end, directional reads (58). The assembled transcriptome consisted of 138,581 contiguous consensus sequences (contigs) that were grouped into 82,891 Trinity components ('genes'). The size range of the transcripts was 201-23,627 bp with an N50 (weighted median) of 2,087 bp. The *E. affinis*

assembly is qualitatively similar to other recently reported copepod and amphipod transcriptomes (Table S5). The assembled *E. affinis* transcriptome is accessible through the Transcriptome Shotgun Assembly database (TSA, Bioproject PRJNA242763).

Trinity-supported protocols and scripts for downstream analyses were followed using default parameters (59) to align reads associated with each library to the assembled transcriptome and to estimate abundances of the assembled transcripts (RSEM). Abundance counts of genes were TMM- (trimmed-mean of M-values) and FPKM- (fragments per kilobase per million reads mapped) normalized to account for differences in RNA production across samples (60) and gene length, respectively. Principal components analysis (PCA) of the TMM- and FPKM- normalized abundance counts of all biological replicates across the three treatments identified one outlier in the control treatment that was subsequently dropped from further analysis (Figure S4). Analysis of differentially expressed genes across the three treatments was performed with edgeR software (61) with a minimum 2-fold difference in expression and a p-value cutoff for false discovery rate (FDR) of 0.05. We chose a 2.0-fold threshold in light of previous findings that known modulators of host-microbiota interactions are regulated within this range (62).

Representative sequences corresponding to the differentially expressed genes were provisionally annotated using blastx against the NCBI non-redundant (nr) database with a threshold e-value of 10^{-4}. The *Eurytemora affinis* genome was released in the midst of our analysis, so a blast search against the genome with a threshold e-value of 10^{-10} was performed to validate the origin of the differentially expressed genes as belonging to *E. affinis*. The remainder of the transcriptome was annotated by blastx against the Swissprot database. Blast2GO (63) was also used to gain further information about the gene ontology (GO) terms and conserved protein domains associated with the genes of interest. In addition, *Vibrio* sequences captured within the transcriptome assembly were identified through blastx searches against a *Vibrio* nr database with a threshold of 10^{-5}.

Cloning and Quantitative PCR (qPCR)

To confirm the predicted sequences of the genes of interest and to generate standards for qPCR, 205-790 bp regions were cloned and sequenced as described previously (64). All primer sequences are provided in Tables S2 and S3. Material for cloning was obtained from mature, adult *E. affinis* females preserved in PureZOL at -80 °C. Complementary cDNA (cDNA) was synthesized from 1 µg of total RNA per 20 µL reaction using the I-Script cDNA-synthesis kit (Bio-Rad) according to the manufacturer's instructions. PCR products were cloned into pGEM-T Easy (Promega) and sequenced. For qPCR experiments, cDNA was synthesized from 450 ng of total *E. affinis* RNA in a 20 µL reaction. The 20 µL cDNA reaction was diluted with molecular biology grade water, such that each microliter of diluted cDNA contained 10 ng of total RNA.

Gene expression was measured using SsoFast EvaGreen Supermix (Bio-Rad) on an iCycler iQ real-time PCR detection system (Bio-Rad). The 20 µL EvaGreen reaction mixture contained 10 µL master mix, 8 µL molecular biology grade water, 1 µL diluted cDNA and 1 µL of 10 µM primers. The PCR conditions were: 95 °C for 2 min followed by 40 cycles of 95 °C for 5 s and 62 °C - 64 °C for 10 s. All samples and standards were run in duplicate wells on the same plate for each gene of interest. After amplification, PCR products from each reaction were subjected to melt-curve analysis to ensure that only a single product was amplified. Selected products were also visualized on 1.5% agarose gels and consistently yielded single bands.

Gene expression was calculated relative to a standard curve of serially diluted plasmid standards encompassing the amplicon of interest and then base-2 log-transformed. A normalization factor equal to the geometric mean of the selected housekeeping genes (65) was subtracted from the gene expression values. Three housekeeping genes were chosen from the Illumina data based on their moderate expression and low coefficient of variation between samples (i.e., thioredoxin domain-containing protein 5 (comp52262_c0), thyroid adenoma-associated protein homolog (comp59254_c0), human leucine-rich repeat neuronal protein 2-like (comp53361_c0)). The housekeeping genes exhibited stable expression throughout the study except for one *V. sp. F10*-exposed sample that exhibited very low expression of all three housekeeping genes and was subsequently removed from further analysis. Results from three independent exposure experiments were combined to give a total of 9 biological replicates in the *V. sp. F10* and *V. ordalii* WT-exposed treatments, 10 in the control treatment, and 7 in the *V. ordalii* ΔtnaA- exposed treatment. One-way ANOVAs were used to compare gene expression

among *Vibrio*-exposure treatments. For those genes that demonstrated unequal variance in expression across treatments, Welch ANOVAs were used. Unplanned post-hoc comparisons (Tukey's test) in genes with significant ANOVA results compared all possible pairs of treatment means.

RESULTS:

E. affinis survival and Vibrio attachment experiments

The copepod *E. affinis* had high survival rates across a range of inoculation densities, incubation durations, and purportedly pathogenic *Vibrio* species tested (Table S1). The majority of the *Vibrio* species tested (i.e., *V. splendidus RSK9*, *V. splendidus LGP32*, *V. nigripulchritudo Sfn27)* had no influence on *E. affinis* survival. *V. campbelli HY01*, *V. penaecidae AQ115*, and *V. ordalii 12B09* demonstrated weak and variable pathogenicity towards *E. affinis* at inoculation titers in the range of 2-7 x 10^7 CFU mL^{-1}. When we further screened *V. ordalii 12B09* at high inoculation titers (1 x 10^8 CFU mL^{-1}) and exposure times (48 hours), we observed an average mortality of 35% (± 15%) across three independent experiments. Exposure to *V. sp. F10* did not cause *E. affinis* mortality at any of the inoculation densities tested (2 - 6 x 10^7 CFU mL^{-1}).

The colonization density of *E. affinis* differed by *Vibrio* species and inoculation titer of *Vibrio*. The *Vibrio* attachment experiments demonstrated that as the inoculation titers of *V. ordalii WT* increased, the abundance of culturable *V. ordalii WT* recovered from copepods increased exponentially (Figure 1; Table S4). At low inoculation titers, *V. ordalii WT* did not colonize *E. affinis* as much as other *Vibrio* strains; *V. nigripulchritudo Sfn27* and *V. penaecidae AQ115* colonized *E. affinis* 10-fold more than *V. ordalii WT* at an inoculation titer of 2 x 10^7 CFU mL^{-1} (Figure 1). Interestingly, the culturability of *V. sp. F10* on SWC agar was consistently below detectable levels (below 30 colonies per pooled copepod sample) for all inoculation titers tested.

The abundance estimates of *Vibrio* associated with copepods obtained from direct and culturable counts (on SWC agar) were highly consistent for both *V. ordalii* and *V. penaecidae* (Figure 2; Table S4). Conversely, there was great discrepancy (10^6 fold difference) between the

direct and culturable counts (on SWC agar) for *V. sp. F10* across all inoculation titers tested (Figure 2; Table S4). The culturability of *V. sp. F10* appears to rapidly decrease upon association with copepods, as we observed a 300-fold decrease in culturability between 6 and 24 hours of *E. affinis* colonization (Figure S5). In contrast, the culturability of *V. ordalii WT* on SWC media changed only slightly during this time period (Figure S5). We observed that the *V. sp. F10* free-living in the incubation flasks did not have reduced culturability on SWC agar, suggesting that the change in the *V. sp. F10* culturability is specific to association with copepods. The change in culturability of *Vibrio* upon copepod colonization is also dependent on the *Vibrio* species, as *V. ordalii WT* colonizing copepods had similar direct and culturable counts (TCBS and SWC media) (Figure 3). Interestingly, copepod-associated *V. sp. F10* demonstrated highly consistent culturable and direct counts when the samples were plated on TCBS agar (Figure 3). We also observed that colonies of *V. sp. F10* recovered from copepods were yellow on TCBS media, which indicates sucrose metabolism, while those of *V. ordalii WT* were not.

Further characterization of the chitin-degrading ability of *V. sp. F10* and *V. ordalii* using the Remazol Brilliant Violet-dyed colloidal chitin plates suggested that *V. ordalii WT* and *V. ordalii* Δ*tnaA* do secrete exogeneous chitinase based on the presence of clear halos surrounding the colonies. The *V. sp. F10 9ZB36* colonies did not clear the labeled chitin, suggesting that this strain does not degrade colloidal chitin (Figure S6).

RNA-Seq differential expression analysis

Overall, the global gene expression pattern of the *V. sp. F10*-exposed treatment was most distinct from the control treatment, demonstrated by the separation of the treatments on a PCA plot (Figure 4A). Conversely, the global expression pattern of the *V. ordalii*-exposed treatment was very similar to the control treatment. A heat map of the genes differentially expressed between treatments also suggests high similarity in expression patterns between the *V. ordalii*-exposed and the control treatments, whereas the *V. sp. F10*-exposed replicates are distinct from the other treatments and highly congruent within the *V. sp. F10* treatment (Figure 4B).

A total of 78 genes were differentially expressed with a fold change > 2 and an FDR > 0.05 across the three *Vibrio*-exposure treatments (Figure S7; Table S6). Among these, the 38

genes with annotation information represent diverse organismal functions including cell signaling, immune function, maintenance of cuticle integrity, cellular transport, metabolism, and the stress response (Figure 5). Gene categories involved in the host response to microbiota, namely cuticle integrity, immune response, and stress response, comprised approximately half of the annotated differentially expressed genes (Figure 5).

The majority of the 78 differentially expressed transcripts originated from the *V. sp. F10*-exposed treatment. The 47 genes up-regulated by *V. sp. F10*-exposure are primarily involved in the stress response, cuticle integrity (chitin metabolism, chitin binding) and the innate immune response (C-type lectins, saposin-like) (Figure 5, Table S6). *V. sp. F10*-exposure also induced mild up-regulation of several cell transport and cell signaling genes and mild down-regulation of several cell signaling, metabolism, stress response, and immune elements (14 total down-regulated genes). Exposure to *V. ordalii WT* induced few transcriptional changes in *E. affinis*, with a total of two transcripts of unknown function strongly down-regulated (6-8 fold change in expression) and one stress response element mildly down-regulated. A total of 53 genes were differentially expressed between the *V. sp. F10*- and *V. ordalii*-exposed treatments, 16 of which were unique to this comparison (Table S8) and were primarily up-regulated in the *V. sp. F10*-exposed treatment. The majority of these genes were of unknown function with a few involved in cell signaling and maintenance of cuticle integrity (Table S8).

Interestingly, two genes were similarly regulated in direction and magnitude in the *V. sp. F10*- and *V. ordalii*-exposed treatments (Figure 5, Table S7). These two transcripts had no significant match to the nr or InterProScan databases, although a BLAT (BLAST-like Alignment Tool) search against the *E. affinis* genome confirmed their origin as *Eurytemora* (99-100% nucleotide match to *E. affinis* genome; data not shown). The two commonly regulated transcripts could represent candidate markers of *E. affinis* exposure to *Vibrio* given that they are both strongly down-regulated regardless of whether a slight (*V. ordalii*-exposed) or more moderate (*V. sp. F10*-exposed) transcriptional response was mounted.

We also identified a few *Vibrio* transcripts in our transcriptome assembly that were detected at low levels only in the *V. sp. F10* treatments. Although these transcripts were all of unknown function, these genes could serve as important targets for further studies of the

mechanisms of *V. sp. F10* colonization of copepods and the mechanisms that trigger the change in culturability of *V. sp. F10* upon association with *E. affinis* (Table S9).

We also briefly surveyed the genomes of *V. sp. F10 9ZB36* and *V. ordalii 12B09* to identify whether several pathways known to be involved in regulating host interactions, namely sialic acid synthesis, toxin production, and quorum sensing, were present. We found that *V. sp. F10* and *V. ordalii* both have the HapR quorum-sensing regulator of virulence and hemolysin. *V. ordalii* has several other toxin-encoding genes including those that code for the Zona occludens toxin, the pore-forming toxin HlyA, a 'Death on curing' toxin, and the ParE toxin. *V. sp. F10* did not have as many toxin-encoding genes, but did have a gene coding for a YafQ toxin. *V. ordalii* also has a gene involved in sialic acid metabolism (Sialic acid utilization regulator, RpiR family).

E. affinis gene expression profiling via qPCR

Eight genes were selected for further qPCR profiling based on their known function in the innate immune response to microbes. Six of these genes were highly differentially expressed within the transcriptome profiling study (three C-type lectin-like transcripts, a saposin-like transcript, and 2 chitin-binding transcripts). The three C-type lectin-like transcripts selected for further study are all predicted via InterPro (66) to have mannose-binding domains, and two of the transcripts (comp49674_c0, comp46353_c1) are predicted by SignalP (67) to have signal peptides, suggesting they may be secreted. Mannose-binding C-type lectins can function as pattern recognition proteins to initiate acquisition of bacterial symbionts from the environment (24, 25, 28), but can also internally inhibit the proliferation of endogeneous bacteria by modulating the expression of AMPs (68) or directly binding to bacteria and acting as antimicrobial agents (69). The saposin-like transcript further profiled in this study is also predicted by SignalP to have a signal peptide and to be secreted. Saposin-like proteins with antimicrobial activity are found in diverse organisms (e.g., amoebae, nematodes, and humans (70-72)). Saposins can act as pore-forming AMPs in invertebrates (73, 74), and their expression can be induced following microbial infection (75). Finally, the two chitin-binding transcripts selected (comp43891_c0, comp47090_c0) are both predicted to have chitin-binding domains (InterPro), which are often found in genes involved in maintaining the integrity of the arthropod gut to prevent against invasion of microbes and their toxins (36, 76).

Two genes that were not differentially expressed in the transcriptome analysis, prophenoloxidase and catalase, were selected for further study in light of their highly conserved and significant role in the innate immune response. The proPO cascade is induced when host recognition proteins are activated by microbial compounds including bacterial surface attachment proteins and cell wall components such as lipopolysaccharides and peptidoglycans (77). The activation of the proPO pathway initiates the conversion of the inactive prophenoloxidase into catalytically active phenoloxidase (78, 79), triggering production of cytotoxic compounds and encapsulation of the microbial invaders (39, 40). Catalases are essential proteins that decompose reactive oxygen species, specifically hydrogen peroxide, produced as part of the innate immune response (80-82).

The Illumina expression analysis was strongly supported by the qPCR studies, with consistency in the magnitude and direction of induction of the target genes across the *E. affinis* treatments (Table 1).Expression of the targeted innate immune elements (three C-type lectin-like transcripts, saposin-like transcript) was significantly induced in the *V. sp. F10*-exposed copepods compared to the control, *V. ordalii* WT, and *V. ordalii* ΔtnaA- exposed treatments (Figure 6). While the basal expression of the immune resistance genes (C-type lectin-like and saposin-like transcripts) was variable across the independent replicate control samples, their expression was highly and uniformly up-regulated across *V. sp. F10*-exposed biological replicates, suggesting a highly conserved and targeted immune response to *V. sp. F10*-exposure. Conversely, the expression of these genes in *V. ordalii WT*-exposed copepods was more variable than in the control or *V. ordalii* treatments. In follow-up experiments with the *V. ordalii* knockout strain unable to produce indole (*V. ordalii* ΔtnaA), we observed more variable expression of the target genes in the *V. ordalii WT* treatment than in the *V. ordalii* ΔtnaA-exposed or control treatments, suggesting that there may be some individual variability in the host's response to indole. However, there was no statistically significant difference between the control and the *V. ordalii*-exposed treatments for any of the immune transcripts examined (Figure 6). The chitin-binding transcripts were more subtly and variably up-regulated in the *V. sp. F10*-exposed treatment (Figure 6), implying that these genes may be regulated within a smaller range than the other examined immune elements under these *Vibrio* exposure conditions. In accordance with the Illumina sequencing results, prophenoloxidase and catalase were not differentially expressed across the different *Vibrio*-exposure treatments (Figure 7).

DISCUSSION:

In this study, we demonstrate the potential for the estuarine copepod *E. affinis* to discriminately interact with colonizing *Vibrio* species. In response to colonization by *V. sp. F10*, *E. affinis* consistently up-regulated a targeted set of genes, many of which are involved in cuticle integrity and immune elements, including C-type lectin-like transcripts and a saposin-like transcript. Conversely, there was virtually no transcriptomic response of *E. affinis* to *V. ordalii.* Our work also shows that *Vibrio* species vary in their colonization of and physiological response to *E. affinis.* A putative copepod symbiont, *V. sp. F10,* strongly associates with *E. affinis* but does not detectably grow on SWC agar plates after 24 hours of *E. affinis* colonization.

Although it has been observed that commensal and pathogenic colonizing bacteria often elicit similar immune responses in invertebrate hosts, the magnitude of the response is significantly lower in commensal associations (83, 84). Therefore, the moderate number of genes induced by *V. sp. F10* association with *E. affinis* may reflect the ecological specialization of *V. sp. F10* as a commensal symbiont of living zooplankton that lacks the ability to degrade chitin (49, 50, 85-87). The up-regulation of genes with chitin-binding properties upon *V. sp. F10* exposure may reflect the renewal of the peritrophic membrane to restrict the bacteria from invading the host through the gut (83). The immune response genes up-regulated by *V. sp. F10* association, specifically saposin-like genes and C-type lectins, also belong to families that are characteristically involved in symbiont acquisition from the environment (24, 25, 28) and in symbiont maintenance (26, 27, 29, 30). Furthermore, components of highly conserved innate immune pathways such as the Toll and IMD signaling pathways (31, 88) and the prophenoloxidase cascade (39, 78, 89) were not up-regulated by *V. sp. F10* exposure, which lends additional support to the hypothesis that the response elicited by *V. sp. F10* is one of recognition and maintenance of a commensal symbiont rather than elimination of a harmful invader.

E. affinis exposure to *V. ordalii* induced a limited transcriptomic response despite the observation that *V. ordalii 12B09* can digest chitin (this study), contains several toxin-encoding genes, and can be pathogenic towards some marine invertebrates (this study; T. Mincer, *personal communication*). However, we found that *E. affinis* was resilient to *V. ordalii* and a suite of other purportedly pathogenic *Vibrio* species. Our study shows that although *V. ordalii*

has been classified as predominantly free-living (48), it can colonize *E. affinis* at a density typical of copepods' natural microbiota (1 x 10^5 cells copepod^{-1}) (2, 8, 90, 91). Yet our findings suggest that significant colonization by *Vibrio* does not necessarily elicit a significant molecular response in the *E. affinis* host. If further localization studies were to demonstrate that *V. ordalii* loosely and transiently attaches to the copepod cuticle, it might imply that *E. affinis* only recognizes tightly and stably associated microbiota like *V. sp. F10*. Another hypothesis to explain the limited *E. affinis* transcriptomic response is that our exposure timeframe may have missed the full *E. affinis* response to *V. ordalii* colonization. Two of the genes that were strongly down-regulated by *V. ordalii* exposure were similarly down-regulated in the *V. sp. F10* treatment, suggesting that these unknown transcripts may be candidate markers of *Vibrio* exposure (Figure 5). Characterization of the function of these two transcripts and examination of their expression patterns upon copepod exposure to other *Vibrio* species warrant further study.

Attachment to *E. affinis* alters the metabolism of *V. sp. F10*, quickly and dramatically reducing its culturability on SWC agar to below the detection limit. This phenomenon is not observed in the free-living *V. sp. F10* collected from the ambient seawater, suggesting that this process is specific to the surface of the copepod and is not likely caused by a broadly secreted factor. The association of viable but non-culturable (VBNC) bacteria with copepods and other zooplankton has been previously observed in environmental samples that harbored bacteria non-culturable on standard media but detectable by immunological or PCR-based methods (14, 21, 92, 93). The VBNC phenomenon is thought to enhance bacterial survival during unfavorable environmental conditions including dramatic shifts in salinity and temperature (14, 21, 92, 93). Although many previous studies describe VBNC vibrios as non-culturable on TCBS agar (93-95), a highly selective medium often used for isolation and enumeration of vibrios, in our study the *V. sp. F10* associated with copepods is culturable on TCBS agar but non-culturable on SWC agar. Further study is needed to identify which components unique to TCBS media, including sucrose and bile salts, lead to the observed differences in the culturability of copepod-associated *V. sp. F10* on SWC and TCBS agar plates. Bile salts are known to affect the physiology of many bacteria (96) and can serve as stimuli for biofilm formation, increased motility, and activation of virulence genes in several vibrio species (97-99). Even upon entering the VBNC state vibrio species can be highly sensitive to bile salts (100), leading to one hypothesis that the non-culturable copepod-associated *V. sp. F10* could have undergone a physiological and/or

transcriptional change leading to restored culturability upon exposure to the bile salts present in the TCBS agar. In light of *V. sp. F10*'s extensive colonization of living zooplankton in the natural environment, future work should also investigate whether the change in culturability in *V. sp. F10* triggered by association with copepods confers a fitness advantage to *V. sp. F10*.

In summary, this study demonstrates that the estuarine copepod *E. affinis* dynamically and discriminately interacts with colonizing *Vibrio* species. Specifically, we have shown that *E. affinis* can distinctly respond to colonizing *Vibrio* through up-regulation of immune elements that may be involved in recognition and maintenance of symbiotic *Vibrio* associates. Further transcriptomic studies of other ecologically relevant copepod – *Vibrio* interactions (e.g., *V. cholerae*) in the context of our results could identify whether colonization by such naturally associating, chitinolytic *Vibrio*s elicits a stronger immune response in *E. affinis* than do non-chitinolytic symbionts (i.e., *V. sp. F10*). The effect of *E. affinis* association on *V. sp. F10* culturability highlights our limited understanding of the physiological impacts of copepod colonization on other ecologically relevant *Vibrio* species. Furthermore, the observed variability between *Vibrio* species in their colonization of *E. affinis* raises the question of whether this pattern is driven by the copepod host's differential innate immune response to the colonizing *Vibrio*. We propose that continued study of the dynamics of copepod-*Vibrio* interactions may reveal that copepods exert significant control over *Vibrio* activity, abundance, and community structure in the natural environment.

ACKNOWLEDGEMENTS:

This work was supported by grant number OCE-1132567 from the National Science Foundation to AMT and by the WHOI Ocean Venture Fund. Funding for AAA was provided by the EPA STAR Fellowship and the NSF Graduate Research Fellowship Program (GRFP). High-performance computing was provided by the Pittsburgh Supercomputing Center's Blacklight system through the National Science Foundation's Extreme Science and Engineering Discovery Environment (XSEDE) program. We are grateful for the provision of the *Eurytemora affinis* cultures by Carol Lee and the provision of the labeled colloidal chitin agar plates by Manoshi Datta. We thank Tracy Mincer and Martin Polz for the *V. ordalii WT, V. ordalii* Δ*tnaA*, and *V. sp. F10* cultures, Frederique Le Roux for the *V. nigripulchritudo Sfn27*, *V. penaecidae AQ115*, *V. splendidus LGP32*, and *V. splendidus RSK9* cultures, and Janelle Thompson for the *V. campbelli HY01* and *V. harveyi PSU3316* cultures. We also thank Amy Maas, Albert Almada, and Luisa Villamil Diaz for technical assistance and Phil Alatalo, Dag Altin, Kristen Hunter-Cevera, and Meredith White for copepod culturing assistance.

REFERENCES:

1. **Grossart H-P, Dziallas C, Leunert F, Tang KW.** 2010. Bacteria dispersal by hitchhiking on zooplankton. Proc. Natl. Acad. Sci. U.S.A. **107:**11959-11964.
2. **Tang KW, Turk V, Grossart H-P.** 2010. Linkage between crustacean zooplankton and aquatic bacteria. Aquat. Microb. Ecol. **61:**261-277.
3. **Martinelli JE, Lopes RM, Rivera ING, Colwell RR.** 2011. *Vibrio cholerae* O1 detection in estuarine and coastal zooplankton. J. Plankton Res. **33:**51-62.
4. **Gugliandolo C, Irrera GP, Lentini V, Maugeri TL.** 2008. Pathogenic *Vibrio*, *Aeromonas* and *Arcobacter spp*. associated with copepods in the Straits of Messina (Italy). Mar. Pollut. Bull. **56:**600-606.
5. **Eckert EM, Pernthaler J.** 2014. Bacterial epibionts of *Daphnia*: a potential route for the transfer of dissolved organic carbon in freshwater food webs. ISME J. **8:**1808-1819.
6. **Möller EF.** 2005. Sloppy feeding in marine copepods: prey-size-dependent production of dissolved organic carbon. J. Plankton Res. **27:**27-35.
7. **Carman KR.** 1994. Stimulation of marine free-living and epibiotic bacterial activity by copepod excretions. FEMS Microbiol. Ecol. **14:**255-261.
8. **Tang K.** 2005. Copepods as microbial hotspots in the ocean: Effects of host feeding activities on attached bacteria. Aquat. Microb. Ecol. **38:**31-40.
9. **Griffith PC, Douglas DJ, Wainright SC.** 1990. Metabolic activity of size-fractionated microbial plankton in estuarine, nearshore, and continental shelf waters of Georgia. Mar. Ecol. Prog. Ser. **59:**263-270.
10. **Braun ST, Proctor LM, Zani S, Mellon MT, Zehr JP.** 1999. Molecular evidence for zooplankton-associated nitrogen-fixing anaerobes based on amplification of the nifH gene. FEMS Microbiol. Ecol. **28:**273-279.
11. **Proctor LM.** 1997. Nitrogen-fixing, photosynthetic, anaerobic bacteria associated with pelagic copepods. Aquat. Microb. Ecol. **12:**105-113.
12. **Heidelberg JF, Heidelberg KB, Colwell RR.** 2002. Bacteria of the gamma-subclass *Proteobacteria* associated with zooplankton in Chesapeake Bay. Appl. Environ. Microbiol. **68:**5498-5507.
13. **Parveen B, Reveilliez JP, Mary I, Ravet V, Bronner G, Mangot JF, Domaizon I, Debroas D.** 2011. Diversity and dynamics of free-living and particle-associated *Betaproteobacteria* and *Actinobacteria* in relation to phytoplankton and zooplankton communities. FEMS Microbiol. Ecol. **77:**461-476.
14. **Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR.** 1983. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Appl. Environ. Microbiol. **45:**275-283.
15. **Rawlings TK, Ruiz GM, Colwell RR.** 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.
16. **Tamplin ML, Gauzens AL, Huq A, Sack DA, Colwell RR.** 1990. Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. Appl. Environ. Microbiol. **56:**1977-1980.
17. **Maugeri TL, Carbone M, Fera MT, Irrera GP, Gugliandolo C.** 2004. Distribution of potentially pathogenic bacteria as free living and plankton associated in a marine coastal zone. J. Appl. Microbiol. **97:**354-361.
18. **Kirn TJ, Jude BA, Taylor RK.** 2005. A colonization factor links *Vibrio cholerae* environmental survival and human infection. Nature **438:**863-866.
19. **Hunt DE, Gevers D, Vahora NM, Polz MF.** 2008. Conservation of the chitin utilization pathway in the *Vibrionaceae*. Appl. Environ. Microbiol. **74:**44-51.

20. **Huq A, Colwell RR, Rahman R, Ali A, Chowdhury MAR, Parveen S, Sack DA, Russekcohen E.** 1990. Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and cell cultures Appl. Environ. Microbiol. **56:**2370-2373.
21. **Epstein SS, Colwell R.** 2009. Viable but Not Cultivable Bacteria, p. 121-129, Uncultivated Microorganisms, vol. 10. Springer Berlin Heidelberg.
22. **Mueller RS, McDougald D, Cusumano D, Sodhi N, Kjelleberg S, Azam F, Bartlett DH.** 2007. *Vibrio cholerae* strains possess multiple strategies for abiotic and biotic surface colonization. J. Bacteriol. **189:**5348-5360.
23. **Turner JW, Good B, Cole D, Lipp EK.** 2009. Plankton composition and environmental factors contribute to Vibrio seasonality. ISME J **3:**1082-1092.
24. **Bright M, Bulgheresi S.** 2010. A complex journey: transmission of microbial symbionts. Nat. Rev. Microbiol. **8:**218-230.
25. **Bulgheresi S, Schabussova I, Chen T, Mullin NP, Maizels RM, Ott JA.** 2006. A new c-type lectin similar to the human immunoreceptor DC-SIGN mediates symbiont acquisition by a marine nematode. Appl. Environ. Microbiol. **72:**2950-2956.
26. **Fraune S, Augustin R, Anton-Erxleben F, Wittlieb J, Gelhaus C, Klimovich VB, Samoilovich MP, Bosch TCG.** 2010. In an early branching metazoan, bacterial colonization of the embryo is controlled by maternal antimicrobial peptides. Proc. Natl. Acad. Sci. U.S.A. **107:**18067-18072.
27. **Douglas A.** 2011. Lessons from Studying Insect Symbioses. Cell Host & Microbe **10:**359-367.
28. **Kvennefors ECE, Leggat W, Hoegh-Guldberg O, Degnan BM, Barnes AC.** 2008. An ancient and variable mannose-binding lectin from the coral *Acropora millepora* binds both pathogens and symbionts. Dev. Comp. Immunol. **32:**1582-1592.
29. **Heath-Heckman EA, Gillette AA, Augustin R, Gillette MX, Goldman WE, McFall-Ngai MJ.** 2014. Shaping the microenvironment: evidence for the influence of a host galaxin on symbiont acquisition and maintenance in the squid-vibrio symbiosis. Environ. Microbiol. **2014:**1462-2920.
30. **Login FH, Balmand S, Vallier A, Vincent-Monégat C, Vigneron A, Weiss-Gayet M, Rochat D, Heddi A.** 2011. Antimicrobial Peptides Keep Insect Endosymbionts Under Control. Science **334:**362-365.
31. **Franzenburg S, Walter J, Künzel S, Wang J, Baines JF, Bosch TCG, Fraune S.** 2013. Distinct antimicrobial peptide expression determines host species-specific bacterial associations. Proc. Natl. Acad. Sci. U.S.A. **110:**E3730-E3738.
32. **Lemaitre B, Hoffmann J.** 2007. The host defense of *Drosophila melanogaster*. Annu. Rev. Immunol. **25:**697-743.
33. **Vallet-Gely I, Lemaitre B, Boccard F.** 2008. Bacterial strategies to overcome insect defenses. Nat. Rev. Microbiol. **6:**302-313.
34. **Lehane MJ.** 1997. Peritrophic matrix structure and function. Annu. Rev. Entomol. **42:**525-550.
35. **Yoshikoshi K, Ko Y.** 1988. Structure and function of the peritrophic membranes of copepods. Nippon Suisan Gakkai Shi **54:**1077-1082.
36. **Buchon N, Broderick NA, Poidevin M, Pradervand S, Lemaitre B.** 2009. *Drosophila* intestinal response to bacterial infection: activation of host defense and stem cell proliferation. Cell Host Microbe **5:**200-211.
37. **Kurtz J, Franz K.** 2003. Innate defence: evidence for memory in invertebrate immunity. Nature **425:**37-38.
38. **Little TJ, O'Connor B, Colegrave N, Watt K, Read AF.** 2003. Maternal transfer of strain-specific immunity in an invertebrate. Curr. Biol. **13:**489-492.
39. **Cerenius L, Lee BL, Söderhäll K.** 2008. The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol. **29:**263-271.
40. **Rowley AF, Powell A.** 2007. Invertebrate immune systems-specific, quasi-specific, or nonspecific? J. Immunol. **179:**7209-7214.

41. **Durai S, Pandian SK, Balamurugan K.** 2011. Changes in *Caenorhabditis elegans* exposed to *Vibrio parahaemolyticus*. J. Microbiol. Biotechnol. **21:**1026-1035.
42. **Tincu JA, Taylor SW.** 2004. Antimicrobial peptides from marine invertebrates. Antimicrob. Agents Chemother. **48:**3645-3654.
43. **Bachere E, Gueguen Y, Gonzalez M, de Lorgeril J, Garnier J, Romestand B.** 2004. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas*. Immunol. Rev. **198:**149-168.
44. **Lee CE.** 2000. Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate "populations". Evolution **54:**2014-2027.
45. **Winkler G, Dodson JJ, Lee CE.** 2008. Heterogeneity within the native range: population genetic analyses of sympatric invasive and noninvasive clades of the freshwater invading copepod *Eurytemora affinis*. Mol. Ecol. **17:**415-430.
46. **Zo Y-G, Chokesajjawatee N, Grim C, Arakawa E, Watanabe H, Colwell RR.** 2009. Diversity and Seasonality of Bioluminescent *Vibrio cholerae* Populations in Chesapeake Bay. Appl. Environ. Microbiol. **75:**135-146.
47. **Huq A, West PA, Small EB, Huq MI, Colwell RR.** 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.
48. **Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ, Polz MF.** 2013. Reproducibility of *Vibrionaceae* population structure in coastal bacterioplankton. ISME J **7:**509-519.
49. **Preheim SP, Boucher Y, Wildschutte H, David LA, Veneziano D, Alm EJ, Polz MF.** 2011. Metapopulation structure of *Vibrionaceae* among coastal marine invertebrates. Environ. Microbiol. **13:**265-275.
50. **Preheim SP.** 2010. Ecology and population structure of *Vibrionaceae* in the coastal ocean. Ph.D. Thesis. Massachusetts Institute of Technology, Cambridge, MA.
51. **Schiewe M, Trust T, Crosa J.** 1981. *Vibrio ordalii* sp. nov.: A causative agent of vibriosis in fish. Curr. Microbiol. **6:**343-348.
52. **Mincer TJ, Johnson M, Flynn-Carroll A, Sharma RS, Wildschutte H, Polz MF.** 2014. Indole as a mediator of protozoan grazing of bacteria: a new role for a multifaceted infochemical, American Society of Microbiology Annual Meeting, Boston, MA.
53. **Murphy N, Bleakeley B.** 2012. Simplified method of preparing colloidal chitin used for screening of chitinase-producing microorganisms. Internet J Microbiol **10**.
54. **Gomez Ramirez M, Rojas Avelizapa LI, Rojas Avelizapa NG, Cruz Camarillo R.** 2004. Colloidal chitin stained with Remazol Brilliant Blue, a useful substrate to select chitinolytic microorganisms and to evaluate chitinases. J. Microbiol. Methods **56:**213-219.
55. **Boulange-Lecomte C, Forget-Leray J, Xuereb B.** 2014. Sexual dimorphism in Grp78 and Hsp90A heat shock protein expression in the estuarine copepod *Eurytemora affinis*. Cell Stress Chaperones **19:**591-597.
56. **Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:**2114-2120.
57. **Hansen KD, Brenner SE, Dudoit S.** 2010. Biases in Illumina transcriptome sequencing caused by random hexamer priming. Nucleic Acids Res. **38:**e131.
58. **Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A.** 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. **29:**644-652.
59. **Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A.** 2013.

De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protocols **8:**1494-1512.

60. **Robinson MD, Oshlack A.** 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. **11**.
61. **Robinson MD, McCarthy DJ, Smyth GK.** 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26:**139-140.
62. **Broderick NA, Buchon N, Lemaitre B.** 2014. Microbiota-induced changes in *Drosophila melanogaster* host gene expression and gut morphology. mBio **5**(3)**:**e01117-01410.01128/mBio.01117-01114.
63. **Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M.** 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21:**3674-3676.
64. **Aruda AM, Baumgartner MF, Reitzel AM, Tarrant AM.** 2011. Heat shock protein expression during stress and diapause in the marine copepod *Calanus finmarchicus*. J. Insect Physiol. **57:**665-675.
65. **Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F.** 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. **3:**research0034.0031 - research0034.0011.
66. **Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, McMenamin C, Mi H, Mutowo-Muellenet P, Mulder N, Natale D, Orengo C, Pesseat S, Punta M, Quinn AF, Rivoire C, Sangrador-Vegas A, Selengut JD, Sigrist CJ, Scheremetjew M, Tate J, Thimmajanarthanan M, Thomas PD, Wu CH, Yeats C, Yong SY.** 2012. InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res. **40:**16.
67. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods **8:**785-786.
68. **Wang X-W, Xu J-D, Zhao X-F, Vasta GR, Wang J-X.** 2014. A shrimp c-type lectin inhibits proliferation of the hemolymph microbiota by maintaining the expression of antimicrobial peptides. J. Biochem. **289:**11779-11790.
69. **Cash HL, Whitham CV, Behrendt CL, Hooper LV.** 2006. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science **313:**1126-1130.
70. **Hoeckendorf A, Stanisak M, Leippe M.** 2012. The saposin-like protein SPP-12 is an antimicrobial polypeptide in the pharyngeal neurons of *Caenorhabditis elegans* and participates in defense against a natural bacterial pathogen. Biochem. J. **445:**205-212.
71. **Leippe M, Ebel S, Schoenberger OL, Horstmann RD, Műller -Eberhard HJ.** 1991. Pore-forming peptide of pathogenic *Entamoeba histolytica*. Proc. Natl. Acad. Sci. U.S.A. **88:**7659-7663.
72. **Pena SV, Krensky AM.** 1997. Granulysin, a new human cytolytic granule-associated protein with possible involvement in cell-mediated cytotoxicity. Semin. Immunol. **9:**117-125.
73. **Banyai L, Patthy L.** 1998. Amoebapore homologs of *Caenohabditis elegans*. Biochima et Biophysica Ata **1429:**259-264.
74. **Roeder T, Stanisak M, Gelhaus C, Bruchhaus I, Grotzinger J, Leippe M.** 2010. Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. Dev. Comp. Immunol. **34:**203-209.
75. **Aguilar R, Jedlicka AE, Mintz M, Mahairaki V, Scott AL, Dimopoulos G.** 2005. Global gene expression analysis of *Anopheles gambiae* responses to microbial challenge. Insect Biochem. Mol. Biol. **35:**709-719.

76. **Kuraishi T, Binggeli O, Opota O, Buchon N, Lemaitre B.** 2011. Genetic evidence for a protective role of the peritrophic matrix against intestinal bacterial infection in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. **108:**15966-15971.
77. **Medzhitov R.** 2007. Recognition of microorganisms and activation of the immune response. Nature **449:**819-826.
78. **Cerenius L, Söderhäll K.** 2004. The prophenoloxidase-activating system in invertebrates. Immunol. Rev. **198:**116-126.
79. **Jiravanichpaisal P, Lee BL, Soderhall K.** 2006. Cell-mediated immunity in arthropods: Hematopoiesis, coagulation, melanization and opsonization. Immunobiology **211:**213-236.
80. **Ha E-M, Oh C-T, Ryu J-H, Bae Y-S, Kang S-W, Jang I-h, Brey PT, Lee W-J.** 2005. An antioxidant system required for host protection against gut infection in *Drosophila*. Dev. Cell **8:**125-132.
81. **Wang C, Yue X, Lu X, Liu B.** 2013. The role of catalase in the immune response to oxidative stress and pathogen challenge in the clam *Meretrix meretrix*. Fish Shellfish Immunol. **34:**91-99.
82. **Mydlarz LD, Harvell CD.** 2007. Peroxidase activity and inducibility in the sea fan coral exposed to a fungal pathogen. Comp Biochem Physiol A Mol Integr Physiol **146:**54-62.
83. **Buchon N, Broderick NA, Lemaitre B.** 2013. Gut homeostasis in a microbial world: insights from *Drosophila melanogaster*. Nat. Rev. Microbiol. **11:**615-626.
84. **Buchon N, Broderick NA, Chakrabarti S, Lemaitre B.** 2009. Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in *Drosophila*. Genes Dev. **23:**2333-2344.
85. **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.
86. **Szabo G, Preheim SP, Kauffman KM, David LA, Shapiro J, Alm EJ, Polz MF.** 2012. Reproducibility of Vibrionaceae population structure in coastal bacterioplankton. ISME J **7:**509-519.
87. **Takemura AF, Chien DM, Polz MF.** 2014. Associations and dynamics of *Vibrionaceae* in the environment, from the genus to the population level. Front. Microbiol. **5:**2014.
88. **Valenzuela-Munoz V, Gallardo-Escarate C.** 2014. TLR and IMD signaling pathways from *Caligus rogercresseyi* (Crustacea: Copepoda): In silico gene expression and SNPs discovery. Fish Shellfish Immunol. **36:**428-434.
89. **Binggeli O, Neyen C, Poidevin M, Lemaitre B.** 2014. Prophenoloxidase activation is required for survival to microbial infections in *Drosophila*. PLoS Pathog. **10:**e1004067.
90. **Hansen B, Bech G.** 1996. Bacteria associated with a marine planktonic copepod in culture .1. Bacterial genera in seawater, body surface, intestines and fecal pellets and succession during fecal pellet degradation. J. Plankton Res. **18:**257-273.
91. **Möller EF, Riemann L, Sondergaard M.** 2007. Bacteria associated with copepods: abundance, activity and community composition. Aquat. Microb. Ecol. **47:**99-106.
92. **Thomas KU, Joseph N, Raveendran O, Nair S.** 2006. Salinity-induced survival strategy of *Vibrio cholerae* associated with copepods in Cochin backwaters. Mar. Pollut. Bull. **52:**1425-1430.
93. **Signoretto C, Burlacchini G, Pruzzo C, Canepari P.** 2005. Persistence of *Enterococcus faecalis* in aquatic environments via surface interactions with copepods. Appl. Environ. Microbiol. **71:**2756-2761.
94. **Halpern M, Landsberg O, Raats D, Rosenberg E.** 2007. Culturable and VBNC *Vibrio cholerae*: interactions with chironomid egg masses and their bacterial population. Microb. Ecol. **53:**285-293.
95. **Chowdhury MAR, Huq A, Xu B, Madeira FJB, Colwell RR.** 1997. Effect of alum on free-living and copepod-associated *Vibrio cholerae* O1 and O139. Appl. Environ. Microbiol. **63:**3323-3326.

96. **Begley M, Gahan CGM, Hill C.** 2005. The interaction between bacteria and bile. FEMS Microbiol. Rev. **29:**625-651.
97. **Gotoh K, Kodama T, Hiyoshi H, Izutsu K, Park K-S, Dryselius R, Akeda Y, Honda T, Iida T.** 2010. Bile acid-induced virulence gene expression of *Vibrio parahaemolyticus* reveals a novel therapeutic potential for bile acid sequestrants. PLoS ONE **5:**e13365.
98. **Hung DT, Zhu J, Sturtevant D, Mekalanos JJ.** 2006. Bile acids stimulate biofilm formation in *Vibrio cholerae*. Mol. Microbiol. **59:**193-201.
99. **Hay AJ, Zhu J.** 2014. Host intestinal signal-promoted biofilm dispersal induces *Vibrio cholerae* colonization. Infect. Immun.
100. **Su C-P, Jane W-N, Wong H-c.** 2013. Changes of ultrastructure and stress tolerance of *Vibrio parahaemolyticus* upon entering viable but nonculturable state. Int. J. Food Microbiol. **160:**360-366.

FIGURES AND TABLES:

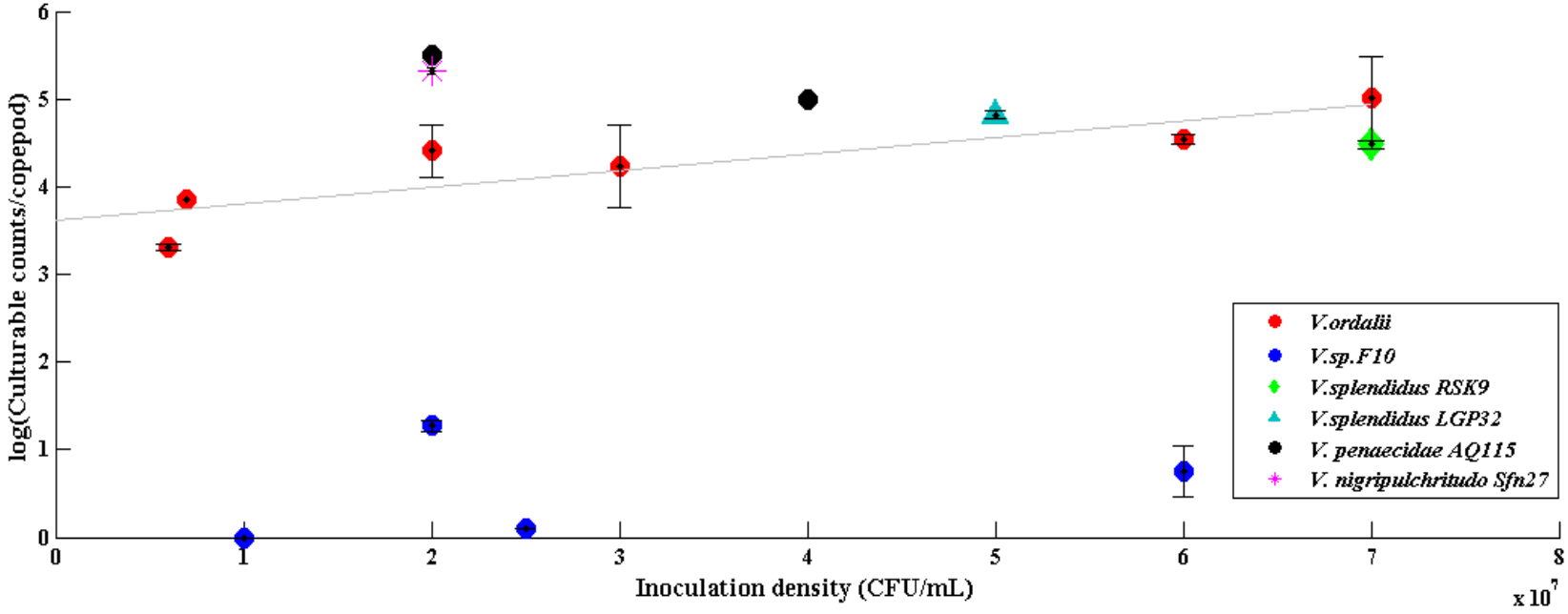


Figure 1: Colonization densities of the estuarine copepod *Eurytemora affinis* differs by *Vibrio* species and inoculation titer. After 24 hour exposure to a *Vibrio* species, pools of copepods (n = 5 per replicate) were rinsed with artificial seawater, homogenized, and plated on seawater complete agar (15 PSU). Plates were incubated for 24 hours at room temperature before counting. Colonization densities per copepod are shown as the base-10 log-transformed means ± 95% confidence interval. A best-fit line of the base-10 log-transformed *V. ordalii* colonization densities per copepod across a range of inoculation titers is shown ($R^2 = 0.76$; $y = 2 \times 10^{-8}x + 3.62$). All treatments, except for *V. penaecidae*, consisted of two biological replicates of pooled copepods.

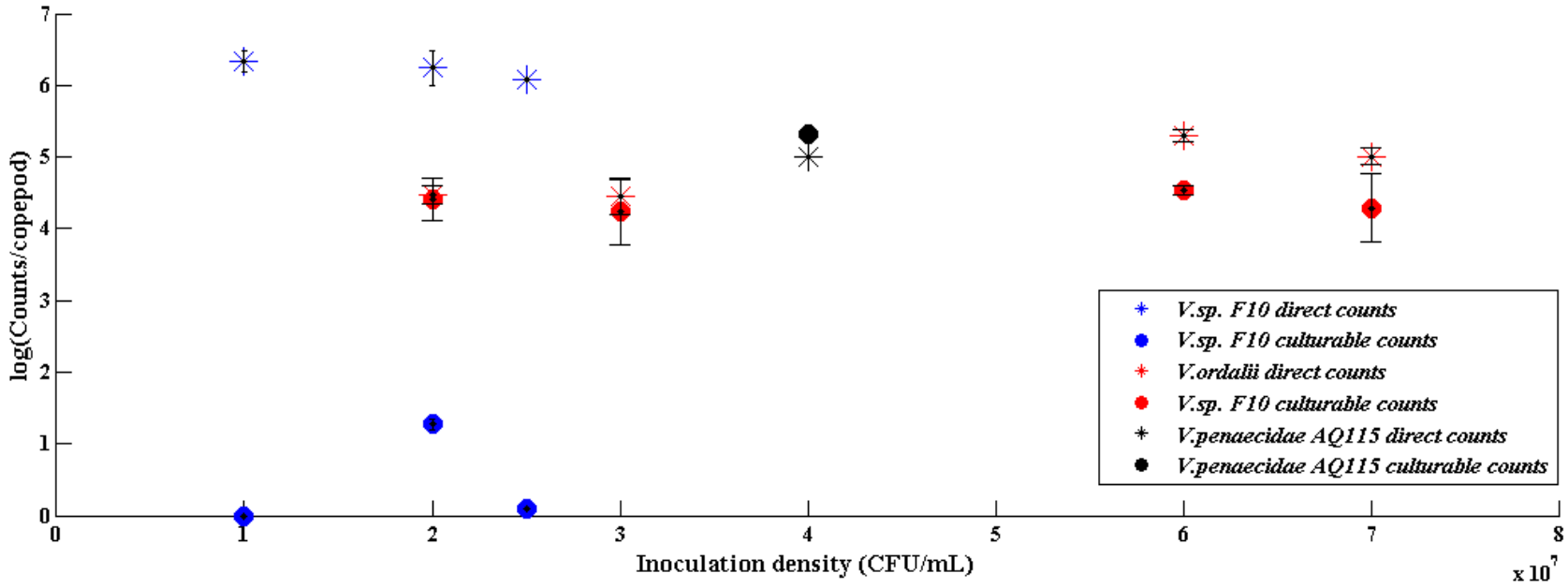


Figure 2: Association with the copepod *E. affinis* specifically alters the culturability of *V. sp. F10*. After 24 hour exposure to a *Vibrio* species, pools of copepods (n = 5 per replicate) were rinsed with artificial seawater, homogenized, and were either plated on seawater complete agar (15 PSU) (solid circles) or stained with DAPI (star symbol). Colonization densities are listed as the base-10 log-transformed means ± 95% confidence interval of two biological replicates. All treatments, except for *V. penaecidae*, had two biological replicates.

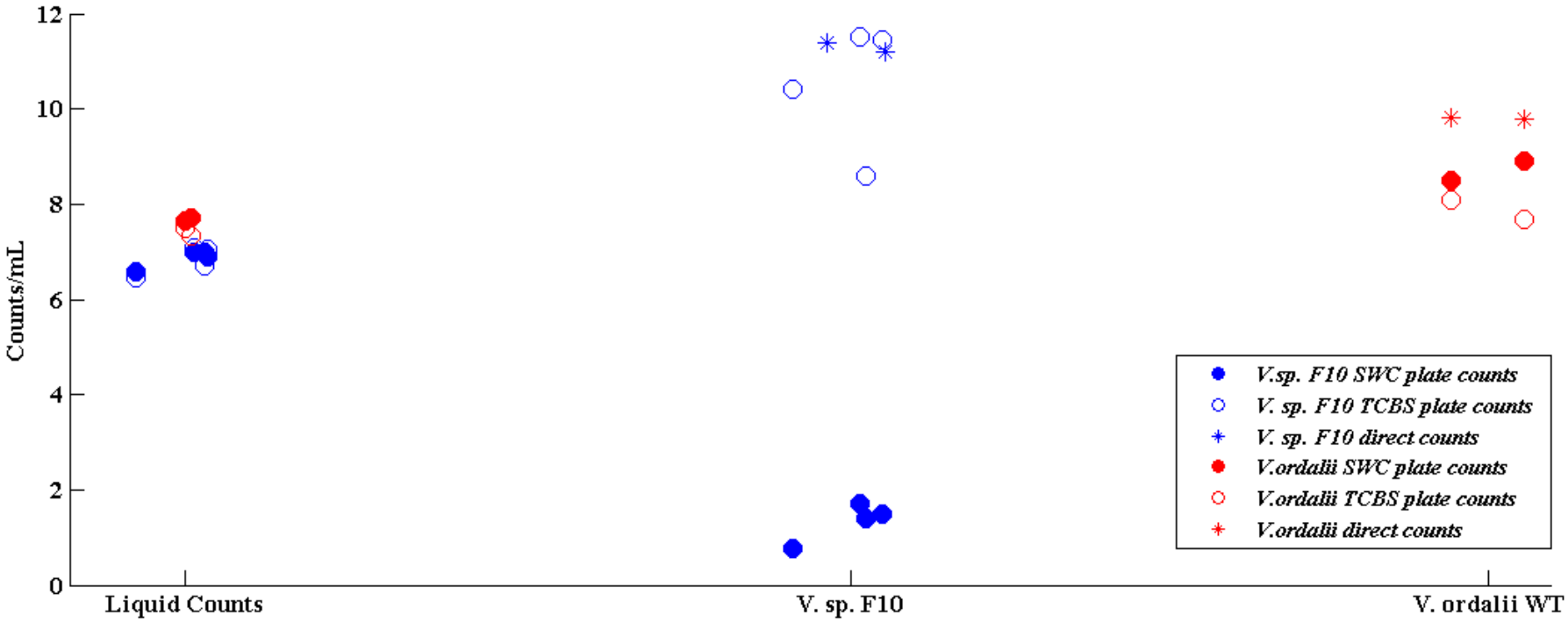


Figure 3: Culturability of *V. sp. F10* associated with *E. affinis* is restored using TCBS media. *V. sp. F10* and *V. ordalii WT* isolated from pools (n = 5) of homogenized adult, mature female *E. affinis* or from the ambient seawater were either directly stained with DAPI (star symbols), plated on seawater complete agar (filled circles) or plated on thiosulfate-citrate-bile salts-sucrose (TCBS) agar (hollow circles) and incubated at room temperature for 20 hours. Culturable counts were normalized by the approximate volume of *E. affinis* (~2.5 x 10^{-5} mL) in order to directly compare the bacterial counts in the ambient seawater and on the copepods.

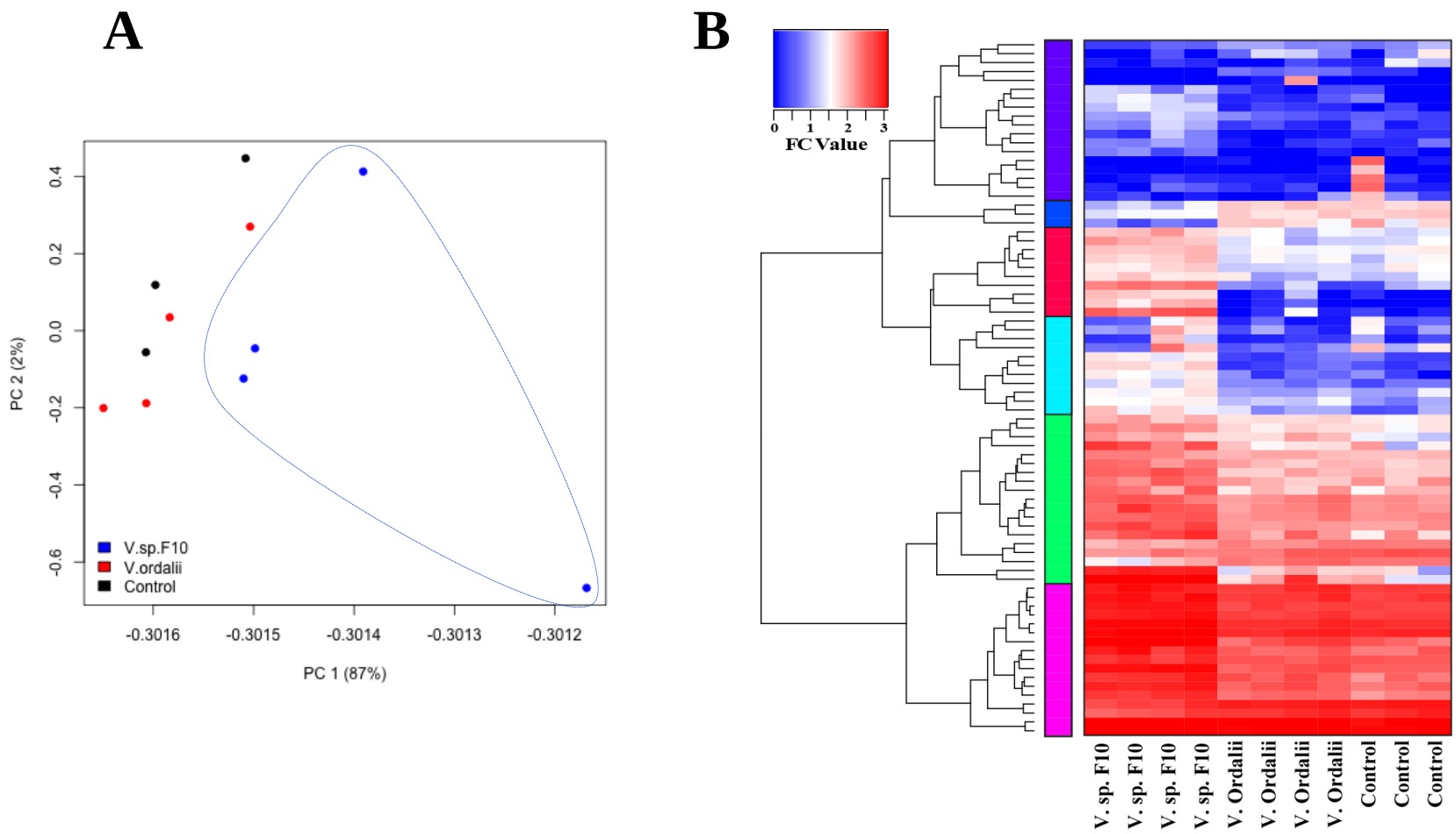
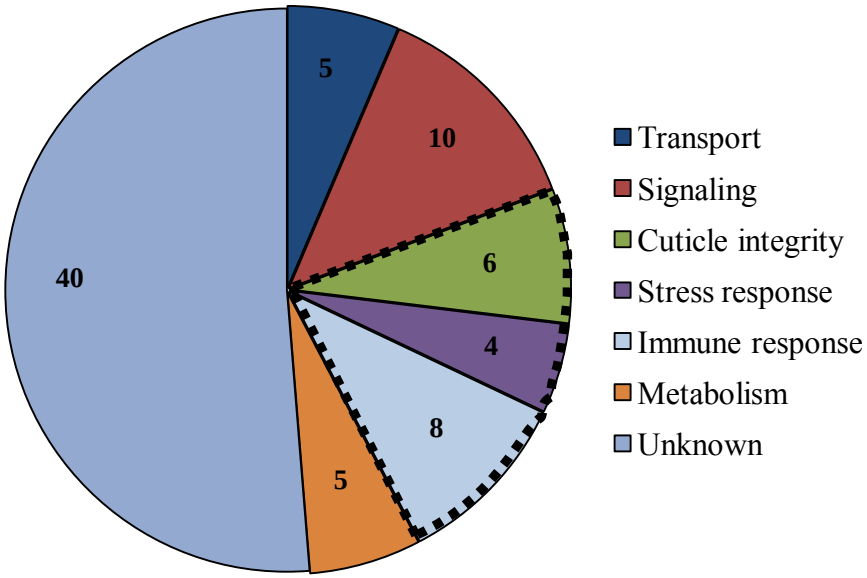


Figure 4: *Vibrio* exposure treatments show distinct transcriptional profiles. (A) Principal component analysis demonstrates strong distinction between the *V. sp. F10*-exposed and control treatments, whereas the *V. ordalii* and control treatments are similar. (B) Heat map representing the base-2 log-transformed FPKM expression values of the 78 differentially expressed genes (fold change > 2, FDR > 0.05) across the three *Vibrio* exposure treatments. Horizontal groupings indicate hierarchical clustering of expression values across all biological replicates. The *V. sp. F10*-exposed biological replicates demonstrate congruent expression patterns within the treatment that are distinct from the expression patterns of the *V. ordalii*-exposed and control treatments.

A



B

Function	Transcript ID	FC (*V. sp. F10*)	FC (*V. ordalii*)
RESPONSE TO STRESS			
Injury response	comp32809_c1	1.59	-
Detoxification	comp55690_c0	-1.50	-
Detoxification	comp46208_c1	-1.04	-
Inhibitory protein	comp44575_c0	-	-2.81
CUTICLE INTEGRITY			
Chitin metabolism	comp55805_c0	1.28	-
Chitin-binding	comp35157_c0	1.88	-
Chitin-binding	**comp43891_c0**	2.10	-
Chitin-binding	**comp47090_c0**	2.03	-
IMMUNE SYSTEM PROCESSES			
C-type lectin-like	comp43463_c0	-1.16	-
C-type lectin-like	comp50187_c1	-1.80	-
Saposin-like	**comp58868_c1**	4.52	-
C-type lectin-like	comp46353_c0	8.21	-
C-type lectin-like	**comp46353_c1**	7.64	-
C-type lectin-like	**comp49674_c0**	6.93	-
C-type lectin-like	**comp47544_c0**	4.61	-
C-type lectin-like	comp40027_c0	5.99	-
UNKNOWN			
Unknown	comp51822_c0	-5.53	-5.90
Unknown	comp40339_c0	-8.35	-7.66

Figure 5: Impact of *Vibrio* exposure on *E. affinis* gene expression.
(A) Functional gene ontology terms associated with the 78 differentially expressed genes identified by Illumina sequencing. The total gene number in each category is indicated on the pie chart. Those functional categories known to be involved in the transcriptional response to microbiota (i.e., cuticle integrity, stress response, immune response) are outlined with a dashed line. (B) Highlight of *E. affinis* genes that were most altered by exposure to *Vibrio*. Fold changes (FC) in gene expression for each *Vibrio* exposure condition relative to the control treatment. Positive and negative FC values reflect genes up-regulated and down-regulated, respectively, compared to the control treatment. Genes highlighted with bolder colors are more intensely altered by *Vibrio* exposure, with red hues indicating up-regulation and blue hues indicating down-regulation. Those genes further profiled by qPCR are in bold.

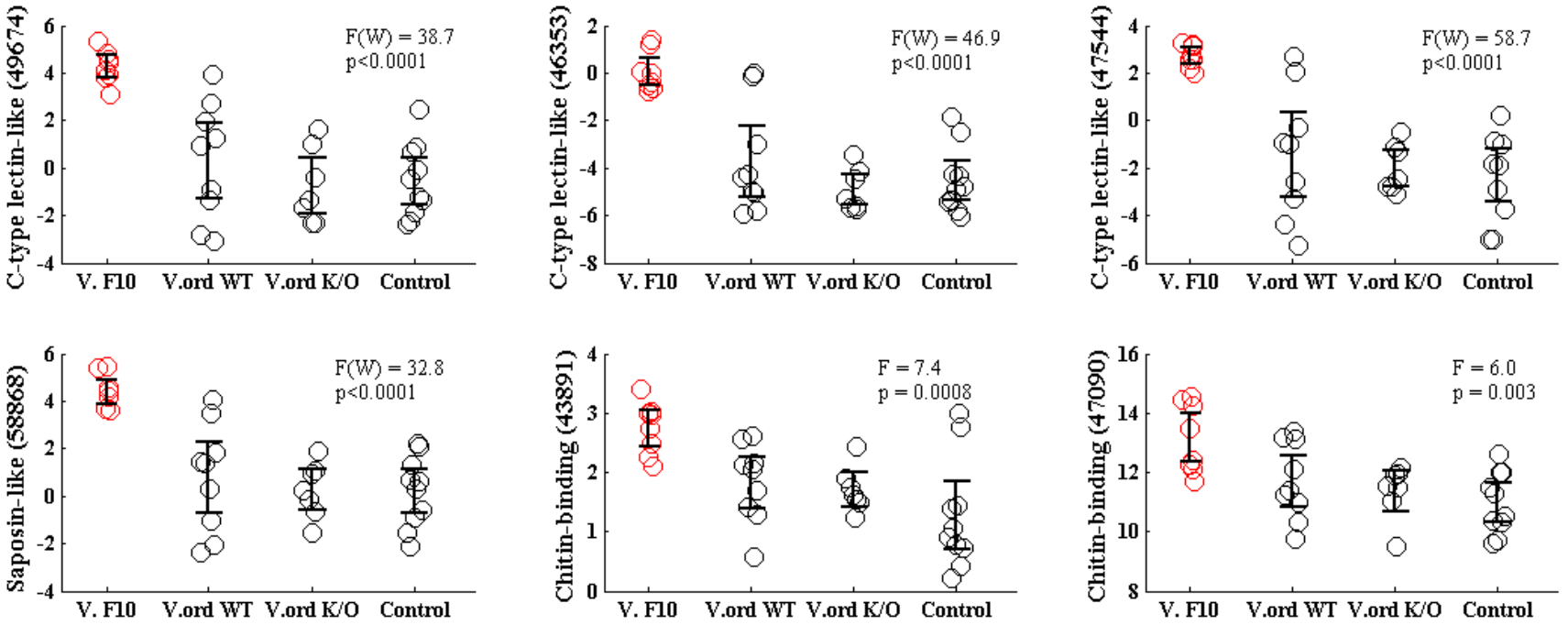


Figure 6: qPCR validation of RNA-Seq gene targets up-regulated in *V. sp. F10*-exposure treatment. Gene expression values listed were normalized to housekeeping genes and base-2 log-transformed. A 95% confidence interval is centered at the mean expression value for each treatment. The F-statistics and p-values from one-way ANOVAs or Welch ANOVAs ('F(W)') are listed for each profiled gene. Tukey's post-hoc comparisons demonstrated that the *V. sp. F10* treatment was different from all other treatments in all genes shown above (except for 'Chitin-binding (comp47090)') where the *V. sp. F10* treatment was only significantly different from the *V. ord* K/O and control treatments). Gene expression was measured in pools of *E. affinis* adult females in three independent experiments (sample sizes: control (n = 10 biological replicates), *V. ordalii WT* (n = 9), *V. ordalii* tnaA knockout ('K/O') (n= 7), *V. sp. F10* (n = 8)).

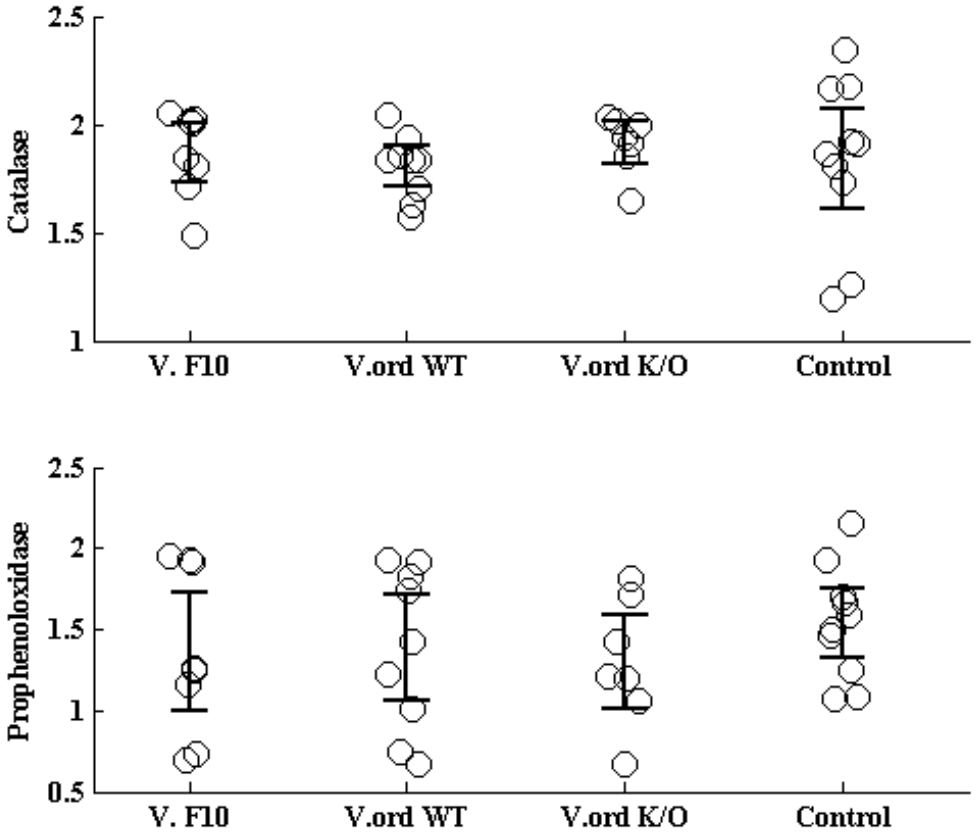


Figure 7: qPCR validation of RNA-Seq gene targets not up-regulated in the *V. sp. F10*-exposure treatment. Gene expression values listed were normalized to housekeeping genes and base-2 log-transformed. A 95% confidence interval is centered at the mean expression value for each treatment. Gene expression was measured in pools of adult female *E. affinis* in three independent experiments (Sample sizes: control (n = 10 biological replicates), *V. ordalii WT* (n = 9), *V. ordalii* ΔtnaA (n= 7), *V. sp. F10* (n = 8)).

Table 1: Expression patterns of genes identified in the Illumina study and further validated via qPCR are directly comparable. Base-2 log-transformed fold change (FC), Base-2 log-transformed counts per million (CPM), and false discovery rate (FDR) refer to the Illumina study and base-2 log-transformed (average) FC refers to the qPCR study comparing (A) *V. sp. F10*-exposed vs. control *E. affinis* and (B) *V. sp. F10*-exposed vs. *V. ordalii WT*-exposed *E. affinis*. Log-transformed FC is reported as a positive number for transcripts exhibiting higher expression in the (A) *V. sp. F10*-exposed or the (B) *V. ordalii*-exposed treatments.

A

		Illumina			qPCR
Transcript name	**Component name**	**log_2FC**	**log_2CPM**	**FDR**	**log_2FC**
Saposin-like	comp58868_c1	4.34	7.92	9.85E-73	3.68
Chitin-binding	comp43891_c0	2.11	2.94	1.15E-02	1.24
Chitin-binding	comp47090_c0	1.86	3.23	3.94E-10	2.29
C-type lectin-like	comp46353_c0	7.29	2.97	1.17E-56	4.00
C-type lectin-like	comp47544_c0	4.47	4.92	6.85E-36	4.27
C-type lectin-like	comp49674_c0	6.47	7.30	5.07E-41	3.97
Catalase	comp50873_c0	-0.25	0.16	1.00	-0.03
Prophenoloxidase	comp58098_c0	-0.01	0.91	1.00	-0.21

B

		Illumina			qPCR
Transcript name	**Component name**	**log_2FC**	**log_2CPM**	**FDR**	**log_2FC**
Saposin-like	comp58868_c1	-3.76	7.94	1.03E-35	-1.35
Chitin-binding	comp43891_c0	-1.35	3.07	5.12E-06	-0.76
Chitin-binding	comp47090_c0	-1.96	3.17	6.08E-10	-1.20
C-type lectin-like	comp46353_c0	-5.04	2.94	2.82E-07	-1.88
C-type lectin-like	comp47544_c0	-3.93	4.92	1.97E-20	-0.12
C-type lectin-like	comp49674_c0	-4.50	7.32	9.64E-08	-2.28
Catalase	comp50873_c0	0.21	5.09	1.00	-0.01
Prophenoloxidase	comp58098_c0	0.01	7.90	1.00	0.05

SUPPORTING FIGURES AND TABLES:

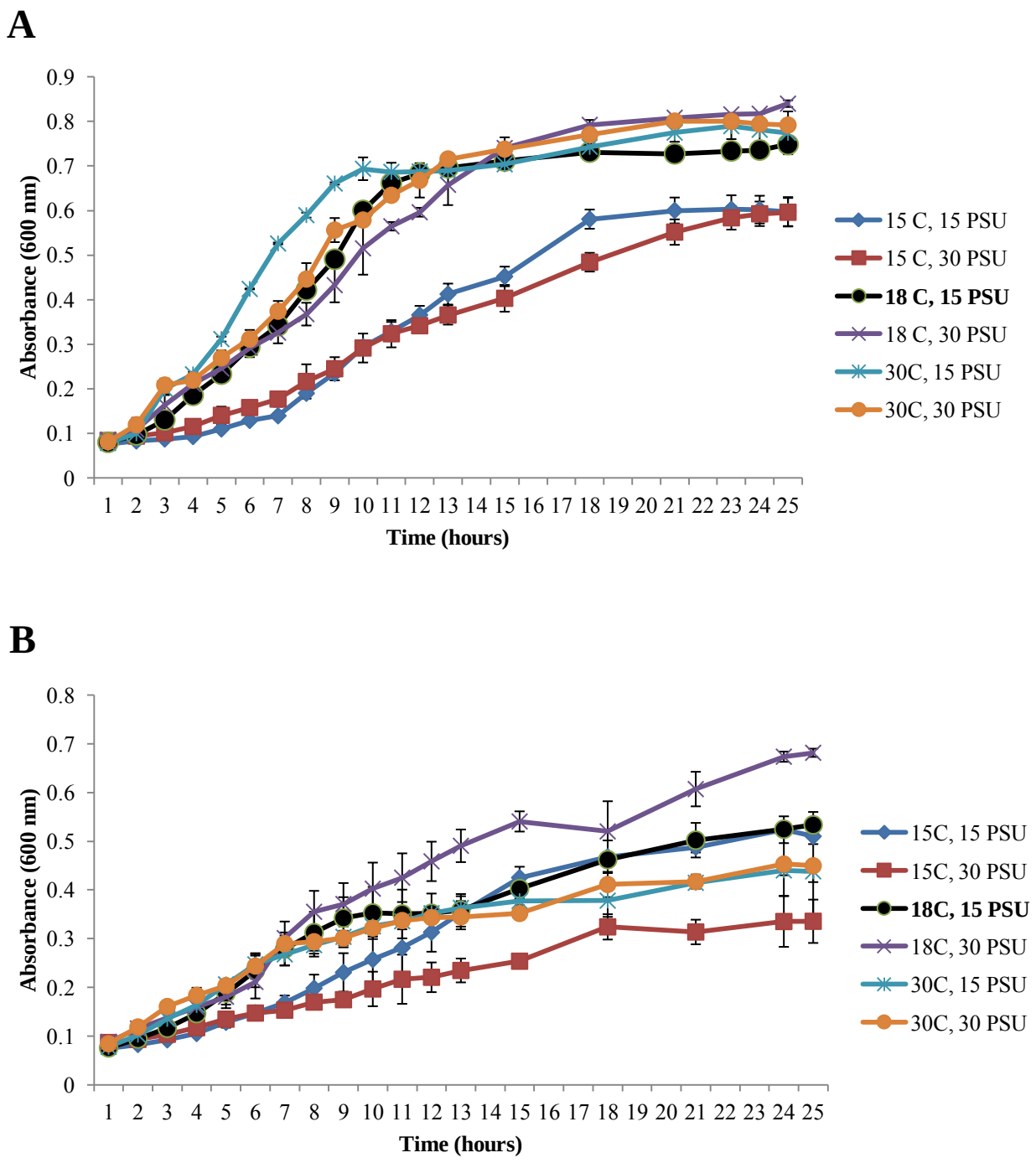


Figure S1: Growth curves of *Vibrio ordalii 12B09* (Figure 2A) and *Vibrio sp. F10 9ZB36* (Figure 2B) in seawater complete media (SWC) at different salinities (15 and 30 PSU) and temperatures (15 °C, 18 °C, 30 °C). The conditions represented in bold font and by black circles (18 °C, 15 PSU) were those used in the *E. affinis-Vibrio* exposure experiments. The results represent the mean ± SE of two experiments run in triplicate wells.

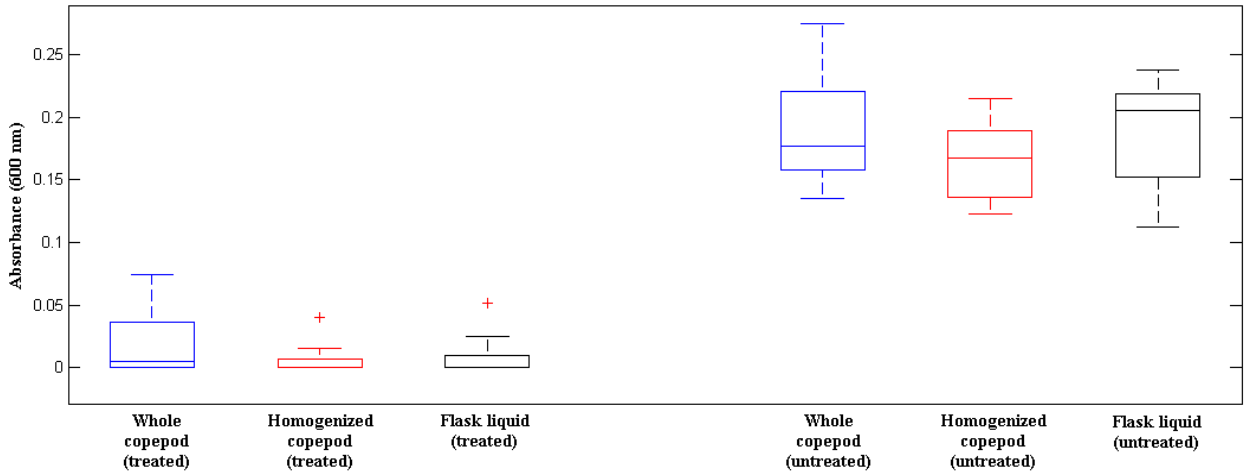


Figure S2: Validation of an antibiotic cocktail used to reduce the natural microbiota of *Eurytemora affinis*. Copepods were treated with a mixture of ampicillin (0.3 mg mL^{-1}), streptomycin (0.1 mg mL^{-1}), and chloramphenicol (0.05 mg mL^{-1}) for 24 hours. Individual whole copepods, homogenized copepods, or 400 µL of seawater from flasks containing either antibiotic treated or untreated copepods were placed into 2 mL of marine broth and the absorbance was measured after 48 hours of incubation at 22 °C. These boxplots represent 10 independent experiments.

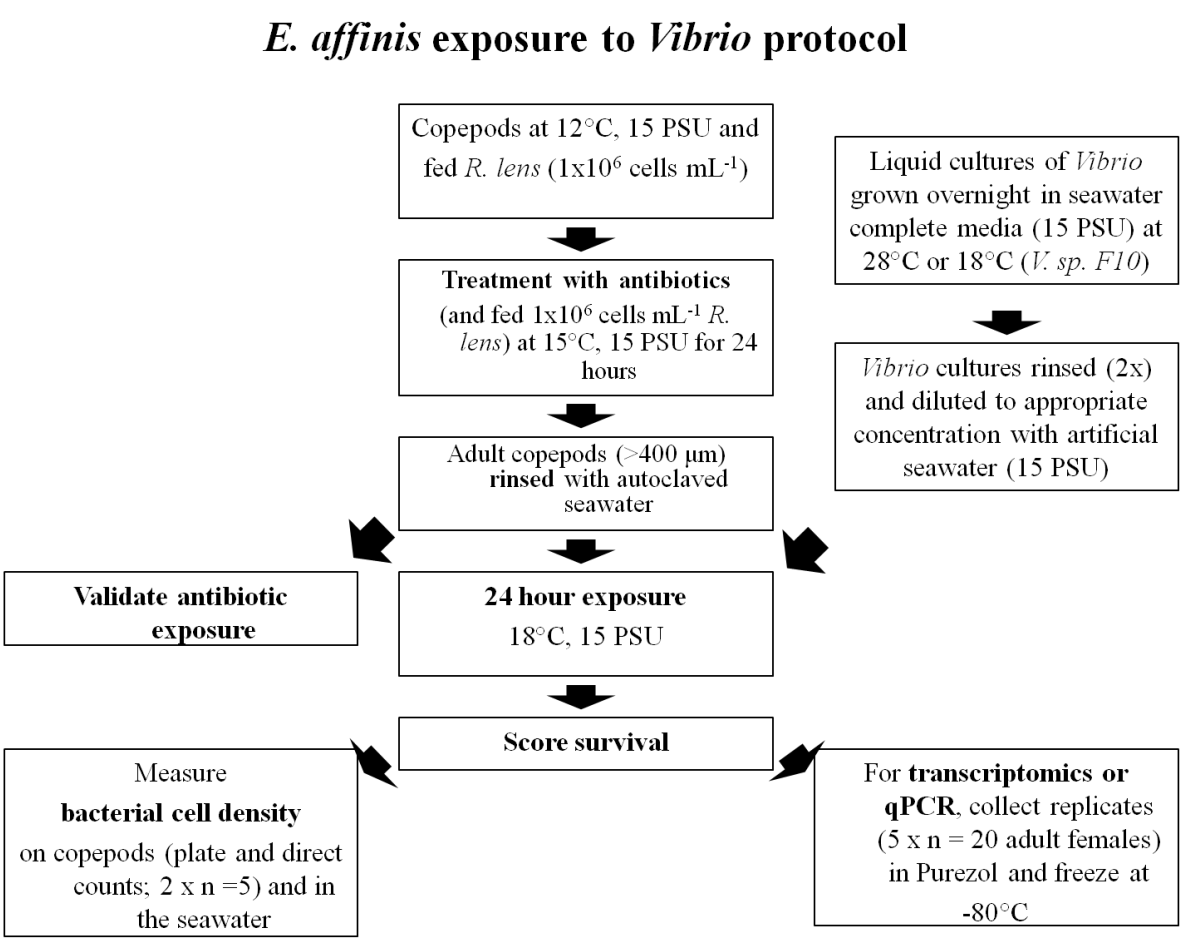


Figure S3: Workflow of *E. affinis* exposure to *Vibrio* species and downstream analyses. The antibiotic treatment was performed at 15 °C in order to gradually acclimate the copepods to the higher temperature used in the *Vibrio* exposure experiments (18 °C) that is more amenable to *Vibrio* growth.

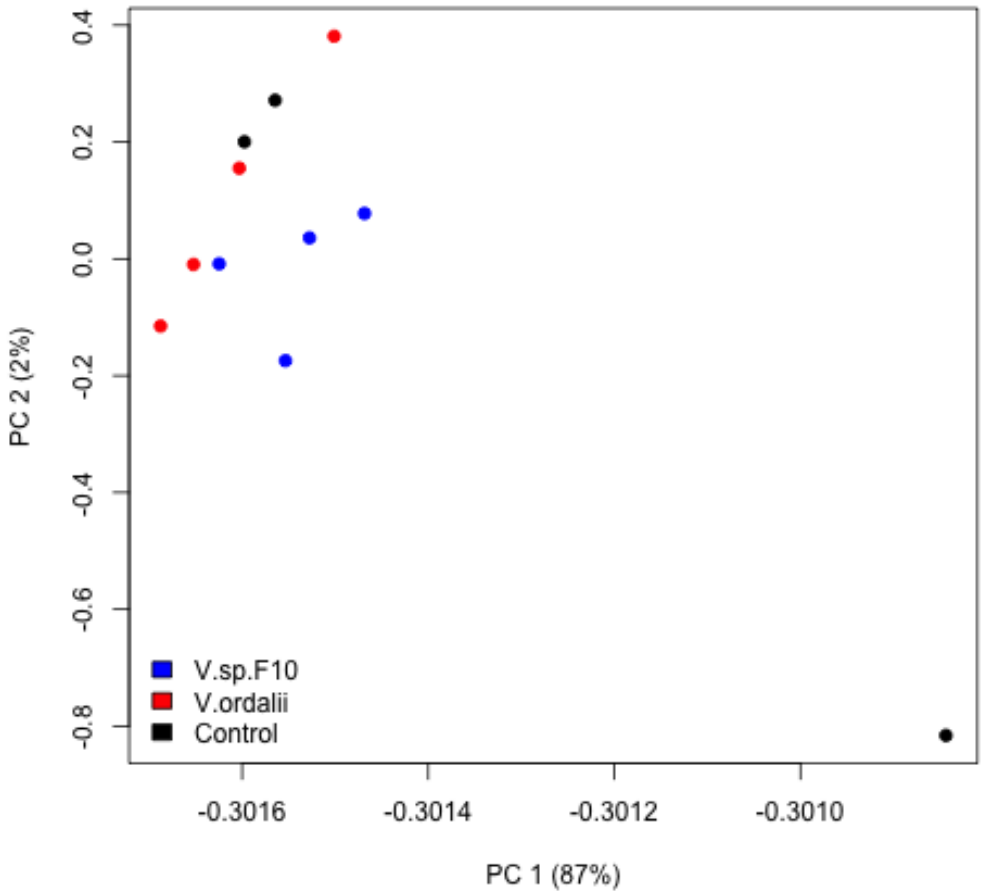


Figure S4: Principal component analysis of FPKM- and TMM-normalized Illumina gene expression data across all *Vibrio* samples suggests one biological replicate of the control treatment is an outlier. This biological replicate (plotted in black in the lower right hand corner) was subsequently dropped from further analysis.

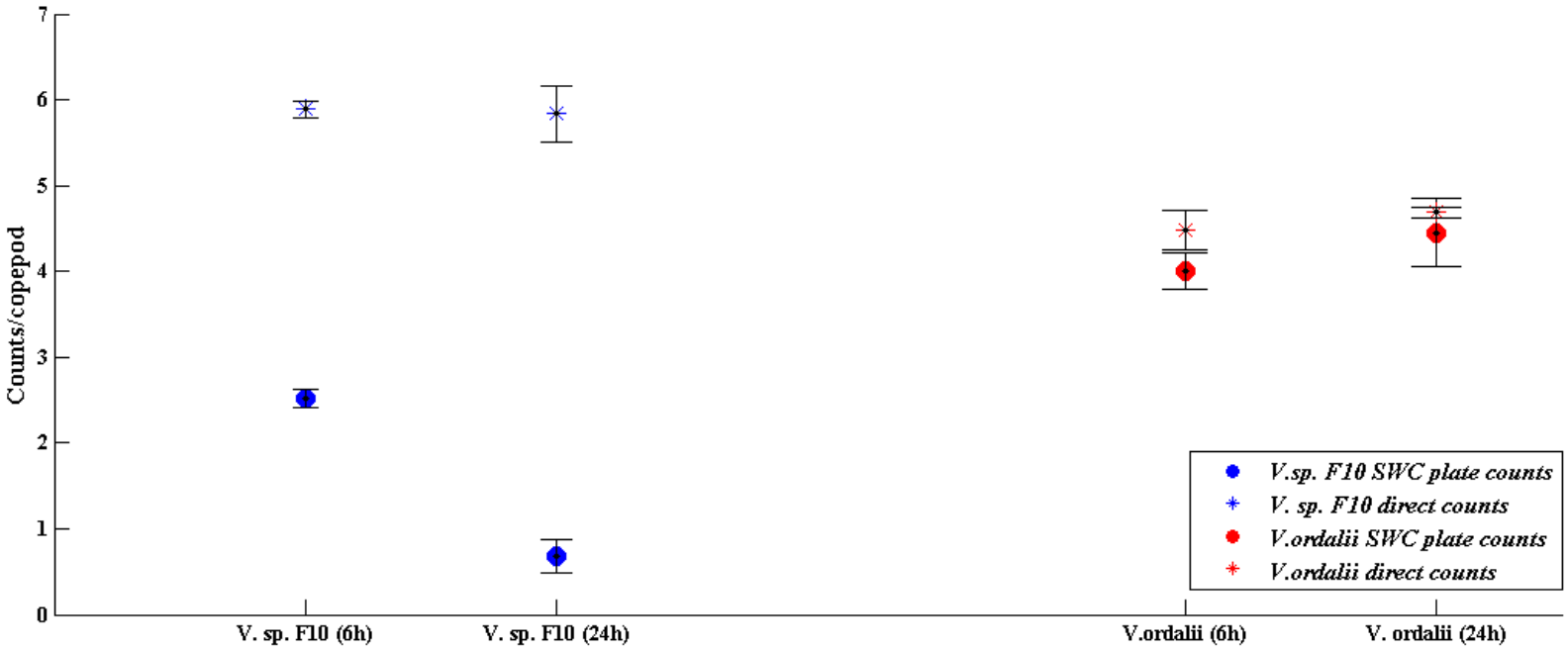


Figure S5: Association with the copepod *E. affinis* rapidly and specifically alters the culturability of *V. sp. F10*. After either 6 or 24 hour exposure to a *Vibrio* species, pools of copepods (n = 5 per replicate) were rinsed with artificial seawater, homogenized, and were either plated on seawater complete agar (15 PSU) (solid circles) or stained with DAPI (star symbol). Colonization densities are listed as the $\log_{10}$ of means ± 95% confidence interval of two biological replicates. All treatments had two biological replicates.

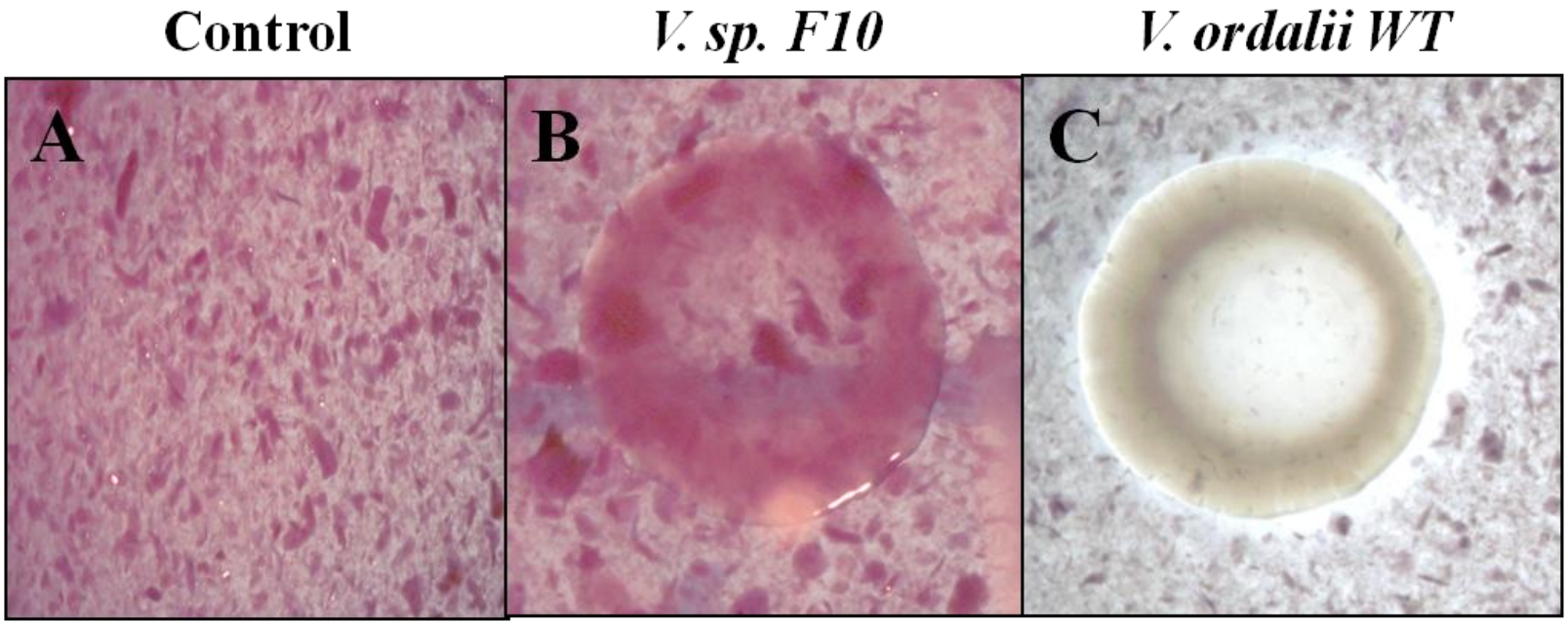


Figure S6: Exogeneous chitinase production of *Vibrio* strains was tested using a Remazol Brilliant Violet-labeled colloidal chitin agar plate. (A) The sterile plate control is purple due to the labeled chitin particles. **(B)** *Vibrio* cultures that do not produce exogeneous chitinase will grow on the marine agar plate, but those **(C)** *Vibrio* cultures that secrete chitinases will grow and also produce a clear halo surrounding the colony due to the cleavage of the chitin particles and the Remzaol dye. Our results suggest that *V. ordalii 12B09* wild type (C) and ΔtnaA (results not shown) produce an exogeneous chitinase, while *V. sp. F10 9ZB36* does not (B).

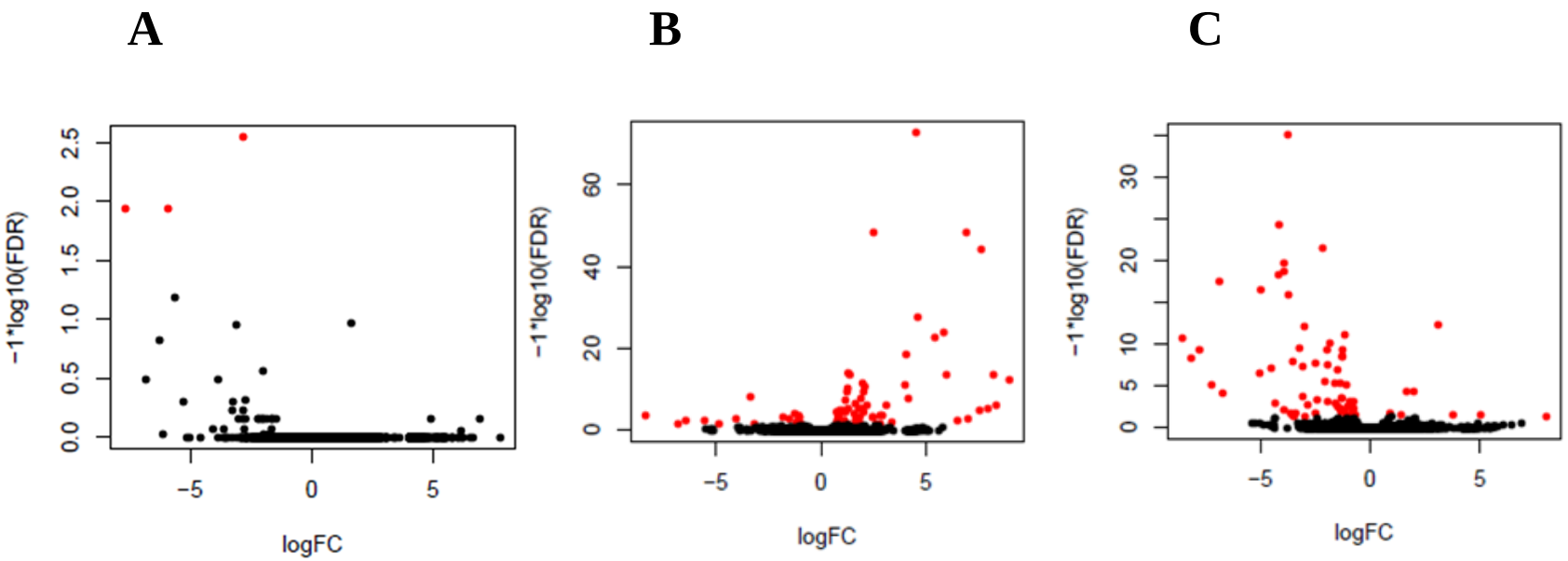


Figure S7: Volcano plot representing the differentially expressed genes present in the different *Vibrio* exposure treatments. Genes that are identified as significantly differentially expressed below a 0.1% false discovery rate (FDR) are indicated in red. **(A)** *V. ordalii* vs. control treatment (B) *V. sp. F10* vs. control treatment and (C) *V. ordalii* vs. *V. sp. F10* treatments. The base-10 log-transformed false discovery rate (FDR) as a function of the base-10 log-transformed fold change (FC) is plotted on the y- and x-axes, respectively. As demonstrated above, a majority of the differentially expressed genes are captured with our subsequently instated threshold of a FC > 2 and a FDR > 0.05.

Table S1: *E. affinis* mortality rates after exposure to different *Vibrio* species at 18 °C, 15 PSU for 24 or 48 hours.

Vibrio strain	Inoculation density (CFU mL^{-1})	Mortality (%)	Length of Exposure (h)	Sex of copepods in experiment	Purpose of exposure
V. sp. F10 9ZB36	1 x 10^7	0 ± 0 [a]	24	Males and females (>400 um)	Survival and attachment
V. sp. F10 9ZB36	2 x 10^7	0 ± 0 [a]	24	Ovigerous females	Transcriptome and attachment
V. sp. F10 9ZB36	2.5 x 10^7	0 ± 0 [a]	24	Males and females (>400 um)	Survival and Attachment
V. sp. F10 9ZB36	6 x 10^7	0 ± 0 [a]	24	Ovigerous females	Survival and attachment
V. ordalii 12B09 WT	6 x 10^6	0 ± 0	24	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	6 x 10^6	0 ± 0	48	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	7 x 10^6	5 ± 5	24	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	2 x10^7	0 ± 0	24	Ovigerous females	Transcriptome and attachment
V. ordalii 12B09 WT	3 x 10^7	10 ± 10 [a]	24	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	6 x 10^7	0 ± 0	24	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	6 x 10^7	0 ± 0	24	Ovigerous females	Survival and attachment
V. ordalii 12B09 WT	6 x 10^7	5 ± 5	48	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	7 x 10^7	5 ± 5	24	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	1 x 10^8	30 ± 30 [a]	48	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	1 x 10^8	13 ± 13 [a]	48	Males and females (>400 um)	Survival and attachment
V. ordalii 12B09 WT	1 x 10^8	63 ± 13 [a]	48	Males and females (>400 um)	Survival and attachment
*V. ordalii 12B09Δ*tnaA*	2 x10^7	0 ± 0	24	Ovigerous females	Transcriptome and attachment
V. splendidus RSK9	7 x 10^7	0 ± 0	24	Males and females (>400 um)	Survival and Attachment
V. splendidus LGP32	5 x 10^7	0 ± 0	24	Males and females (>400 um)	Survival and Attachment
V. nigripulchritudo Sfn27	2 x 10^7	0 ± 0 [a]	24	Males and females (>400 um)	Survival and Attachment
V. campbelli HY01	2.5 x10^7	0 ± 0 [a]	48	Males and females (>400 um)	Survival and Attachment
V. campbelli HY01	3 x 10^7	13 ± 13 [a]	48	Males and females (>400 um)	Survival and Attachment
V. penaecidae AQ115	2 x 10^7	0 ± 0	24	Males and females (>400 um)	Survival and Attachment
V. penaecidae AQ115	4 x 10^7	5 ± 5	24	Males and females (>400 um)	Survival and Attachment
V. penaecidae AQ115	5 x 10^7	0 ± 0	24	Males and females (>400 um)	Survival and Attachment

[a] indicates that 2 replicates of n = 5 individuals were tested. All other treatments tested 2 replicates of n = 10 individuals

Table S2: Cloning primer sequences and melting temperatures.

Transcript name	Transcript ID		Primer sequence	T_m (°C)	Cloned sequence length (bp)
C-type lectin-like	comp47544_c0	F	CGGATGTGTTTCTGTTGAGCA	63	407
		R	TTGCTGCAAGTTGAGAGAGC		
C-type lectin	comp49674_c0	F	TCTTCATGGCCAGGAGAAGG	64	505
		R	TGCTACATCATTCCAGAGTCCA		
C-type lectin-like	comp46353_c1	F	AGCATTGGTTCTATTTCTGGAGA	63	417
		R	AGGAGCATTAATGGCCCAGT		
Chitin-binding	comp47090_c0	F	CATCTACACCCACCTACAATACTAC	62.2	298
		R	CTACAATTCTACATTTCAGCTGG		
Chitin-binding	comp43891_c0	F	GCTGTTCCTCTTAGTCTCTCTC	63	205
		R	GTAGAGAGGTGGAGCGCAG		
Catalase	comp50873_c0	F	GATGCCGCAAACTACTCACC	65.5	543
		R	CTGGTTTGGTTTGGTCCTGAG		
Prophenoloxidase	comp58098_c0	F	CTGCAATGCGTGATCCTCTC	65	790
		R	CTTCTCACTCCGCTGCTG		
Thioredoxin domain-containing protein	comp52622_c0	F	CAAGTTCTACGCTCCCTGG	65	689
		R	GAGTTCGTCCTTCTCTGCC		
Thyroid adenoma-associated protein homolog	comp59254_c0	F	CTGCCTGAAGAAGCTCACTC	65.5	735
		R	CTTGAAACCGTGTAGCCGAG		
Leucine-Rich Neuronal protein	comp53361_c0	F	CTACTGTACCTTGACCTCAGC	65.5	588
		R	CGTGACGTCATTGATCCAGG		
Saposin-like	comp58868_c1	F	TACCCCGTCTTCCTTGAACC	60.5	590
		R	TCCATGCAAAGGTACAACAGT		

Table S3: qPCR primer sequences and melting temperatures.

Transcript name	Transcript ID		Primer sequence	T_m (°C)
C-type lectin-like	comp47544_c0	F	CGGATGTGTTTCTGTTGAGC	62
		R	CCCTCCATTCCTTCATCAGTAG	
C-type lectin	comp49674_c0	F	CTGATGAAGGTATGGAGGGTC	63
		R	GCTAGCTGATATCCATGGGTG	
C-type lectin-like	comp46353_c1	F	AGCTGTCTGACCAACTCCTTAG	63
		R	GGTTCATCTTGTTCTGTCTTGC	
Chitin-binding	comp47090_c0	F	GCTACATCTACTTCACCATCCTAC	64
		R	CTGTACTTGGATGGCAAGCTAC	
Chitin-binding	comp43891_c0	F	GCTGTTCCTCTTAGTCTCTCTC	62
		R	ACAGTCAAATGGATGAGGAAC	
Catalase	comp50873_c0	F	ACAGGCTCGGACCTAACTTTG	64
		R	CTGGTTTGGTTTGGTCCTGAG	
Prophenoloxidase	comp58098_c0	F	CATCACCAAGTCTCCGCTTC	64
		R	GGTAGAACCATTGTCTCAGGC	
Thioredoxin domain-containing protein	comp52622_c0	F	GATTGTACCGAGCATCAGTCC	64
		R	GCTCATTCACCCAGTCCTTG	
Thyroid adenoma-associated protein homolog	comp59254_c0	F	ACCTAGGCTTGTCACTGAGC	64
		R	TGAAGAACAGTCCCTCTCCG	
Leucine-Rich Neuronal protein	comp53361_c0	F	TGACTGGTCCAAGCTCTCTG	64
		R	CGTGACGTCATTGATCCAGG	
Saposin-like	comp58868_c1	F	CGTCTTCCTTGAACCTGAGG	63
		R	CAGCTCCTGTACATTCTTCAC	

Table S4: *Vibrio* bacteria colonization densities of the estuarine copepod *Eurytemora affinis* after 24 hour exposure at 18 °C, 15 PSU. After exposure, copepods (n = 5 per replicate) were rinsed with artificial seawater, homogenized, and plated on seawater complete agar (15 PSU) or stained with DAPI to obtain culturable and direct counts of *Vibrio*s attached to *E. affinis*, respectively. Counts values are listed as means ± standard error of two biological replicates. The *V. sp. F10* and *V. nigripulchritudo Sfn27* survival experiments were performed with 2 replicates of 5 pooled individuals. The *V. ordalii, V. splendidus RSK9 and LGP32,* and *V. penaecidae* survival experiments were performed with 2 replicates of 10 pooled individuals. qPCR and transcriptome experiments were performed with 2 replicates of 5 pooled individuals. [a] Indicates these counts are approximates given that there were fewer than 30 colonies. [b] Indicates there is one biological replicate.

***Vibrio* strain**	**Inoculation density (CFU mL^{-1})**	**Density of culturable bacteria attached to copepods (CFU/copepod)**	**Density of total bacteria attached to copepods (direct cell counts/copepod)**	**Density of culturable bacteria on controls (CFU/copepod)**	**Sex of copepods in experiment**	**Purpose of exposure**
V. sp. F10 9ZB36	1.0×10^7	0 ± 0 [a]	$2.3 \times 10^6 \pm 4.0 \times 10^5$	0.6 ± 0.6	Males and females (>400 µm)	Survival and attachment
V. sp. F10 9ZB36	2.0×10^7	18.5 ± 1.5 [a]	$2.0 \times 10^6 \pm 5.0 \times 10^5$	2.5 ± 1	Mature, adult females	RNA-Seq and attachment
V. sp. F10 9ZB36	2.0×10^7	4.9 ± 1.1 [a]	$7.4 \times 10^5 \pm 2.7 \times 10^5$	8.5 ± 1	Mature, adult females	qPCR and attachment
V. sp. F10 9ZB36	2.0×10^7	15.6 ± 9.9 [a]	$5.0 \times 10^6 \pm 1.1 \times 10^6$	0.75 ± 0.25	Mature, adult females	qPCR and attachment
V. sp. F10 9ZB36	2.5×10^7	0 ± 0 [a]	1.2×10^6 [b]	0 ± 0	Males and females (>400 µm)	Survival and Attachment
V. sp. F10 9ZB36	6.0×10^7	6.0 ± 2.0 [a]	-	0.75 ± 0.75	Mature, adult females	Survival and attachment
V. ordalii 12B09 WT	6.0×10^6	$2.1 \times 10^3 \pm 6.0 \times 10^2$	-	0.2 ± 0.2	Males and females (>400 µm)	Survival and attachment
V. ordalii 12B09 WT	7.0×10^6	7.0×10^3 [b]	-	2.8 ± 0.4	Males and females (>400 µm)	Survival and attachment
V. ordalii 12B09 WT	2.0×10^7	$2.7 \times 10^4 \pm 9.0 \times 10^3$	$3.1 \times 10^4 \pm 4.5 \times 10^3$	2.5 ± 1	Mature, adult females	RNA-Seq and attachment
V. ordalii 12B09 WT	2.0×10^7	$3.2 \times 10^4 \pm 1.4 \times 10^4$	$4.9 \times 10^4 \pm 3.5 \times 10^3$	8.5 ± 1	Mature, adult females	qPCR and attachment
V. ordalii 12B09 WT	2.0×10^7	9.0×10^4 [a]	3.1×10^4 [a]	0 ± 0	Mature, adult females	qPCR and attachment
V. ordalii 12B09 WT	2.0×10^7	$1.4 \times 10^4 \pm 6.0 \times 10^3$	$1.6 \times 10^5 \pm 1.5 \times 10^4$	0.75 ± 0.25	Mature, adult females	qPCR and attachment
V. ordalii 12B09 ΔtnaA	2.0×10^7	$4.0 \times 10^4 \pm 9.5 \times 10^3$	$9.0 \times 10^4 \pm 6.5 \times 10^3$	0.75 ± 0.25	Mature, adult females	qPCR and attachment
V. ordalii 12B09 ΔtnaA	2.0×10^7	$3.6 \times 10^4 \pm 5.0 \times 10^3$	$9.0 \times 10^4 \pm 6.5 \times 10^3$	0 ± 0	Mature, adult females	qPCR and attachment
V. ordalii 12B09 WT	3.0×10^7	$2.0 \times 10^4 \pm 1.0 \times 10^4$	$3.3 \times 10^4 \pm 8.5 \times 10^3$	1.25 ± 0.75	Males and females (>400 µm)	Survival and attachment
V. ordalii 12B09 WT	6.0×10^7	$3.4 \times 10^4 \pm 2.0 \times 10^3$	$2.0 \times 10^5 \pm 2.0 \times 10^4$	0.2 ± 0.2	Males and females (>400 µm)	Survival and attachment
V. ordalii 12B09 WT	6.0×10^7	$4.4 \times 10^4 \pm 1.0 \times 10^3$	-	0.75 ± 0.75	Mature, adult females	Survival and attachment
V. ordalii 12B09 WT	6.0×10^7	$5.4 \times 10^4 \pm 2.5 \times 10^3$	-	0 ± 0	Mature, adult females	Survival and attachment
V. ordalii 12B09 WT	7.0×10^7	$1.2 \times 10^5 \pm 6.1 \times 10^4$	$2.0 \times 10^4 \pm 2.5 \times 10^4$	2.8 ± 0.4	Males and females (>400 µm)	Survival and attachment
V. splendidus RSK9	7.0×10^7	$3.1 \times 10^4 \pm 1.5 \times 10^3$	-	0.8 ± 0.4	Males and females (>400 µm)	Survival and Attachment
V. splendidus LGP32	5.0×10^7	$6.7 \times 10^4 \pm 3.5 \times 10^3$	-	0.8 ± 0.4	Males and females (>400 µm)	Survival and Attachment
V. nigripulchritudo Sfn27	2.0×10^7	$2.1 \times 10^5 \pm 1 \times 10^4$	-	0 ± 0	Males and females (>400 µm)	Survival and Attachment
V. penaecidae AQ115	4.0×10^7	1e5 [b]	2.1e5 [b]	1.25 ± 0.75	Males and females (>400 um)	Survival and Attachment

Table S5: Comparison of the present study with recent studies utilizing next-generation sequencing technologies to assemble *de novo* transcriptomes of crustacean species.

Species	Description	Read length (bp)	Number of reads (million)	Contigs	'Genes'	N50	Contig length range	Platform	Assembler	Investigator
Eurytemora affinis	Estuarine copepod	100	100	138,581	82,891	2,087	201-23,627	Illumina	Trinity	Almada (current study)
Calanus finmarchicus	Marine copepod	100	80	241,778	124,618	987	201-25,048	Illumina	Trinity	Tarrant
Calanus finmarchicus	Marine copepod	100	400	206,041	96,090	1,418	300-23,068	Illumina	Trinity	Lenz et al. (2014) *PLoS ONE*
Tigriopus californicus	Intertidal copepod	384	0.6	22,262	42,473	(925: mean contig length)	8807 (max)	454	CLC Genomics Workbench	Barreto et al. (2011) *Mol Ecol*
Parhyale hawaiensis	Amphipod	400	3	89,664	25,735	1,510	~60-8,000	454	Newbler	Zeng (2011) *BMC Genomics*
Calanus sinicus	Marine copepod	380	1.5	56,809	~14,000	873	~100-3,500	454	Newbler	Ning et al. (2013) *PLoS ONE*

Table S6: *E. affinis* genes differentially expressed in the *V. sp. F10* exposure treatment, compared to the control samples. Abbreviations: 'FC' = fold change relative to the control treatment; 'FDR' = false discovery rate; 'GOs' = gene ontology terms. Blank entries reflect a lack of significant blast hits with associated GO terms at the set parameters (E-value < 1 x 10^{-4}). Differentially expressed genes that were further profiled via qPCR are in bold. Genes that have a red '#' are those that are differentially expressed in both *Vibrio* exposure treatments in comparison to the control samples.

Transcript Description	Transcript ID	FC	FDR	Top BLASTx Hit Species	Top Hit Accession Number	Min. E-Value	Mean similarity	GOs	InterProScan results
CELL SIGNALLING PROCESSES									
homeobox protein nkx	comp12937_c0	-3.18	3.34E-02	*Strongyloides ratti* (nematode)	CEF65008	1.50E-05	47.00%	F:DNA binding	IPR001356 (homeobox domain); IPR009057 (homeodomain-like domain)
a disintegrin and metalloproteinase with thrombospondin motifs partial	comp42146_c0	1.87	2.21E-02	*Stegodyphus mimosarum* (spider)	KFM61983	6.86E-73	72.00%	P:proteolysis; F:metalloendopeptidase activity	IPR001590 (peptidase_M12B domain); IPR024079 (metallopeptidase catalytic domain)
f-box kelch-repeat protein at2g44130-like	comp42229_c0	1.16	3.68E-08	*Pyrus x bretschneideri* (pear)	XP_009335865	1.28E-06	46.67%	-	signal peptide domain; transmembrane domain
cholesterol desaturase daf-36-like	comp36258_c0	1.80	4.63E-06	*Latimeria chalumnae* (coelacanth)	XP_006009329	6.46E-104	61.33%	F:2 iron, 2 sulfur cluster binding; F:oxidoreductase activity; P:oxidation-reduction process	IPR017941 (Rieske [2Fe-2S] iron-sulphur domain); PTHR21266 (iron-sulfur domain containing); transmembrane helix domain
phosphatidylethanolamine-binding protein	comp48058_c1	-1.26	9.38E-05	*Danaus plexippus* (butterfly)	EHJ71177	8.45E-18	47.67%	-	IPR008914, PTHR11362 (phosphatidylethanolamine-binding protein PEBP family); cytoplasmic domain; transmembrane helix domain
beta-crystallin a1	comp51193_c0	1.29	9.21E-15	*Lepeophtheirus salmonis* (copepod)	ADD38111	1.21E-35	54.33%	-	IPR001064 (Beta/gamma crystallin); signal peptide domain ; IPR011024 (Gamma-crystallin-related domain)
hypothetical protein	comp57629_c1	1.16	2.78E-05	*Daphnia pulex* (waterflea)	EFX79782	7.28E-91	39.67%	F:serine-type endopeptidase inhibitor activity	SSF54403 (cystatin/monellin family); IPR002223 (Proteinase inhibitor I2, Kunitz domain); IPR018073 (Proteinase inhibitor I25, cystatin, conserved site); signal peptide domain

METABOLISM									
hypothetical protein	comp45348_c0	-3.36	6.21E-09	*Ciona intestinalis* (tunicate)	XP_002121160	4.00E-42	56.33%	-	PTHR10366 (NAD dependent epimerase/dehydratase); IPR027417 (P-loop containing nucleoside triphosphate hydrolase); transmembrane helix domain
violaxanthin de-epoxidase	comp42733_c0	1.38	4.73E-14	*Physcomitrella patens* (moss)	XP_001773358	5.55E-13	40.00%	F:violaxanthin de-epoxidase activity; C:chloroplast; P:oxidation-reduction process	IPR012674 (calycin domain); IPR010788 (violaxanthin de-epoxidase) ; IPR011038 (calycin-like superfamily); signal peptide domain
hypothetical protein	comp53782_c0	1.12	5.86E-05	*Daphnia pulex* (waterflea)	EFX83386	1.13E-92	55.00%	F:hydrolase activity	IPR002018, IPR019826 (carboxylesterase, type B domain/active site); IPR029058 (alpha/Beta hydrolase fold domain); PTHR11559 (carboxylesterase family); signal peptide domain
aldehyde dehydrogenase family 3 member partial	comp56580_c0	-1.00	6.01E-04	*Stegodyphus mimosarum* (spider)	KFM66996	3.36E-175	69.33%	F:oxidoreductase activity; P:biological_process	IPR012394, PTHR11699 (Aldehyde dehydrogenase NAD(P)-dependent family); IPR016162 (Aldehyde dehydrogenase, N-terminal domain); IPR016163 (Aldehyde dehydrogenase, C-terminal domain); cytoplasmic domain; transmembrane domain
aldehyde oxidase 2-like	comp59156_c0	-1.07	3.99E-04	*Daphnia pulex* (waterflea)	EFX86357	0.00E+00	60.67%	F:molecular_function	IPR005107 (CO dehydrogenase flavoprotein, C-terminal domain); IPR000674 (aldehyde oxidase/xanthine dehydrogenase, a/b hammerhead domain); IPR016208 (Aldehyde oxidase/xanthine dehydrogenase family); IPR008274 (Aldehyde oxidase/xanthine dehydrogenase, molybdopterin binding domain)

RESPONSE TO STRESS									
inter-alpha-trypsin inhibitor heavy chain h4	comp32809_c1	1.59	1.53E-04	*Crassostrea gigas* (oyster)	EKC36390	6.50E-102	55.67%	-	no IPS match
Cytochrome P450	comp55690_c0	-1.50	2.38E-03	*Tigriopus japonicus* (copepod)	AIL94133	1.16E-87	53.67%	P:oxidation-reduction process; F:iron ion binding; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; F:heme binding	IPR001128 (cytochrome P450 family); IPR002401 (cytochrome P450, E-class, group I family); signal peptide domain
glutathione s-transferase mu 1	comp46208_c1	-1.04	4.77E-03	*Oryctolagus cuniculus* (rabbit)	NP_001075721	2.87E-37	51.33%	F:protein binding	IPR004046 (Glutathione S-transferase, C-terminal domain); IPR004045 (Glutathione S-transferase, N-terminal domain); IPR010987 (glutathione S-transferase, C-terminal-like domain)
CUTICLE INTEGRITY									
chitotriosidase	comp55805_c0	1.28	6.34E-11	*Daphnia pulex* (waterflea)	EFX90412	2.17E-134	66.00%	F:hydrolase activity, acting on glycosyl bonds; P:biological_process	IPR017853 (glycoside hydrolase, superfamily); IPR011583 (chitinase II domain); IPR002557 (chitin-binding domain); IPR029070 (chitinase insertion domain); PTHR11177 (chitinase family); signal peptide domain
chondroitin proteoglycan-2-like	comp35157_c0	1.88	4.63E-06	*Tribolium castaneum* (beetle)	XP_008192409	4.01E-08	60.33%	C:extracellular region; P:chitin metabolic process; F:chitin binding	IPR002557 (chitin-binding domain); PTHR23301 (chitin-binding peritrophin A family)
chitin-binding protein	**comp43891_c0**	**2.10**	**3.42E-11**	***Drosophila virilis* (fly)**	**XP_002048076**	**3.65E-05**	**57.67%**	**P:chitin metabolic process; C:extracellular region; F:chitin binding**	**chitin-binding domain (PFAM); signal peptide domain**
chondroitin proteoglycan-2-like	**comp47090_c0**	**2.03**	**5.71E-10**	***Tribolium castaneum* (beetle)**	**XP_008192409**	**1.86E-09**	**60.33%**	**F:chitin binding; P:chitin metabolic process; C:extracellular region**	**IPR002557 (chitin-binding domain)**

IMMUNE SYSTEM PROCESSES									
C-type lectin-like	**comp47544_c0**	**4.61**	**2.55E-28**	-	-	-	-	-	**IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold domain); signal peptide domain**
macrophage mannose receptor partial	comp50187_c1	-1.80	8.02E-04	*Chaetura pelagica* (bird)	KFU96626	1.50E-15	41.67%	F:carbohydrate binding	IPR001304 (c-type lectin domain); PTHR22803 (mannose, phospholipase, lectin receptor related family); IPR016187 (c-type lectin fold domain); signal peptide domain
hepatic lectin-like	**comp49674_c0**	**6.93**	**4.15E-49**	***Oreochromis niloticus* (fish)**	**XP_005459156**	**3.02E-05**	**37**	**F:carbohydrate binding**	**IPR001304 (c-type lectin domain); IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold); cytoplasmic domain; transmembrane helix domain**
C-type lectin-like	comp46353_c0	8.21	3.99E-14	-	-	-	-	-	IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold)
C-type lectin-like	**comp46353_c1**	**7.64**	**6.90E-45**	-	-	-	-	-	**IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold)**
C-type lectin-like	comp40027_c0	5.99	3.99E-14	-	-	-	-	-	IPR016186 (c-type lectin-like); IPR016187 (c-type lectin fold)
c-type mannose receptor 2- partial	comp43463_c0	-1.16	1.41E-02	*Saccoglossus kowalevskii* (worm)	XP_006825556	2.63E-18	44.67%	F:carbohydrate binding	IPR001304 (c-type lectin); IPR016186 (c-type lectin-like); PTHR22803 (mannose, phospholipase, lectin receptor related); IPR016187 (c-type lectin fold); signal peptide domain
Saposin-like	**comp58868_c1**	**4.52**	**2.49E-73**	-	-	-	-	-	**IPR011001 (saposin-like domain); IPR008139 (saposin B domain); signal peptide domain**

TRANSPORT									
sodium-dependent phosphate transporter 1-a-like	comp51144_c0	1.12	3.86E-03	*Metaseiulus occidentalis* (mite)	XP_003742817	1.38E-67	52.67%	F:inorganic phosphate transmembrane transporter activity; C:membrane; P:phosphate ion transport	IPR001204 (phosphate transporter family); cytoplasmic domain; transmembrane helix domain
peptide transporter family 1-like	comp56914_c0	1.64	2.73E-07	*Dendroctonus ponderosae* (beetle)	ENN73556	3.13E-159	59.33%	F:transporter activity; C:membrane; P:oligopeptide transport	IPR000109 (Proton-dependent oligopeptide transporter family); PTHR11654:SF96 (peptide transporter family 1); IPR018456 (PTR2 family proton/oligopeptide symporter, conserved site); IPR016196 (Major facilitator superfamily domain, general substrate transporter domain); transmembrane helix domain; cytoplasmic domain
hypothetical protein	comp57280_c0	1.04	1.96E-05	*Daphnia pulex* (waterflea)	EFX71591	1.72E-149	52.00%	-	IPR002035 (von Willebrand factor, type A domain); IPR013642 (Chloride channel calcium-activated); PTHR10579 (calcium-activated chlorine channel regulator); cytoplasmic domain; transmembrane domain
adp-ribosylation factor	comp45127_c0	2.03	2.68E-05	*Dugesia japonica* (flatworm)	P91924	1.69E-71	82.33%	P:response to stress; P:catabolic process; P:signal transduction; P:vesicle-mediated transport; P:transport; C:Golgi apparatus; F:ion binding	IPR006689 (Small GTPase superfamily, ARF/SAR type); IPR027417 (P-loop containing nucleoside triphosphate hydrolase domain); IPR005225 (Small GTP-binding protein domain)

UNKNOWN									
Unknown	comp62318_c0	2.83	3.06E-02	-	-	-	-	-	signal peptide domain
Unknown	comp16598_c0	2.65	2.13E-02	-	-	-	-	-	transmembrane helix domain; cytoplasmic domain
Unknown	comp16910_c0	4.16	2.58E-08	-	-	-	-	-	no IPS match
Unknown	comp17377_c0	2.48	1.09E-03	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp17945_c0	1.24	3.75E-10	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
hypothetical protein	comp18829_c0	1.90	1.49E-08	*Helobdella robusta* (leech)	XP_009029394	8.92E-04	44.00%	-	signal peptide domain
Unknown	comp58868_c2	4.00	9.54E-12	-	-	-	-	-	no IPS match
Unknown	comp58868_c3	4.07	5.23E-19	-	-	-	-	-	no IPS match
Unknown	comp56716_c0	1.26	2.04E-02	-	-	-	-	-	no IPS match
Unknown#	comp51822_c0	-5.53	6.05E-03	-	-	-	-	-	transmembrane helix domain
Unknown	comp52925_c1	2.50	4.15E-49	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp53341_c1	1.67	4.97E-02	-	-	-	-	-	IPR029469 (PAN-4 domain); G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp53341_c2	1.64	3.86E-03	-	-	-	-	-	IPR029469 (PAN-4 domain); G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp53492_c0	1.76	1.89E-02	-	-	-	-	-	cytoplasmic domain; transmembrane helix domain
Unknown	comp49776_c0	-6.42	3.61E-03	-	-	-	-	-	no IPS match
Unknown	comp50150_c0	1.08	4.43E-02	-	-	-	-	-	coiled-coil domain; transmembrane domain

Unknown	comp46444_c2	2.83	3.53E-04	-	-	-	-	-	signal peptide domain
Unknown	comp46722_c0	2.46	5.94E-04	-	-	-	-	-	no transmembrane domain
Unknown	comp47218_c0	1.06	3.34E-02	-	-	-	-	-	signal peptide domain; transmembrane helix domain
Unknown	comp46043_c0	2.93	3.99E-04	-	-	-	-	-	no IPS match
Unknown	comp44011_c0	3.11	9.29E-07	-	-	-	-	-	no IPS match
Unknown#	comp40339_c0	-8.35	3.53E-04	-	-	-	-	-	signal peptide domain; transmembrane domain
Unknown	comp40368_c0	1.99	4.09E-06	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); SSF57414 (hairpin loop containing domain-like superfamily)
Unknown	comp41942_c0	1.99	2.82E-12	-	-	-	-	-	no IPS match
Unknown	comp42970_c0	-4.86	4.83E-02	-	-	-	-	-	transmembrane domain
Unknown	comp43319_c0	2.19	1.33E-06	-	-	-	-	-	IPR003014 (PAN-1 domain); IPR003609 (apple-like domain) SSF57414 (hairpin loop containing domain-like superfamily); signal peptide domain
Unknown	comp33114_c0	-4.03	2.31E-03	-	-	-	-	-	no IPS match
Unknown	comp36118_c0	1.31	5.91E-06	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp36128_c0	1.86	6.66E-04	-	-	-	-	-	transmembrane, cytoplasmic domain
Unknown	comp39845_c0	3.38	1.48E-02	-	-	-	-	-	no IPS match

Table S7: *E. affinis* genes differentially expressed in the *V. ordalii* exposure treatment, compared to the control samples.
Abbreviations: 'FC' = fold change relative to the control treatment; 'FDR' = false discovery rate; 'GOs' = gene ontology terms. Blank entries reflect a lack of significant blast hits with associated GO terms at the set parameters (E-value < 1 x 10^{-4}). Genes that have a red '#' are those that are differentially expressed in both *Vibrio* exposure treatments in comparison to the control samples.

Transcript Description	Transcript ID	FC	FDR	Top BLASTx Hit Species	Top Hit Accession Number	Min. E-Value	Mean similarity	GOs	InterProScan results
RESPONSE TO STRESS									
Knottin-like inhibitory protein	comp44575_c0	-2.81	2.86E-03	-	-	-	-	P:defense response	IPR003614 (knottin, scorpion-toxin-like domain); signal peptide domain
UNKNOWN									
Unknown#	comp40339_c0	-7.66	1.17E-02	-	-	-	-	-	signal peptide domain; transmembrane domain
Unknown#	comp51822_c0	-5.90	1.17E-02	-	-	-	-	-	transmembrane helix domain

Table S8: *E. affinis* genes differentially expressed in the *V. ordalii* exposure treatment, compared to the *V. sp. F10* exposure treatment. Abbreviations: 'FC' = fold change relative to the control treatment; 'FDR' = false discovery rate; 'GOs' = gene ontology terms. Blank entries reflect a lack of significant blast hits with associated GO terms at the set parameters (E-value < 1 x 10^{-4}). Gene entries that are in bold are those that are not differentially expressed in any other treatment comparison.

Transcript Description	Transcript ID	FC	FDR	Top BLASTx Hit Species	Top Hit Accession Number	Min. E-Value	Mean similarity	GOs	InterProScan results
CELL SIGNALLING PROCESSES									
beta-crystallin a1	**comp45441_c0**	**-1.14**	**1.76E-02**	***Lepeophtheirus salmonis* (copepod)**	**ADD38111**	**8.62E-37**	**54.33%**	-	**G3DSA:2.60.20.10 (crystallin superfamily); IPR011024 (gamma-crystallin related domain); IPR001064 (Beta/gamma crystallin domain)**
a disintegrin and metalloproteinase with thrombospondin motifs partial	comp42146_c0	-1.36	9.78E-02	*Stegodyphus mimosarum* (spider)	KFM61983	6.86E-73	72.00%	P:proteolysis; F:metalloendopeptidase activity	IPR001590 (peptidase_M12B domain); IPR024079 (metallopeptidase catalytic domain)
f-box kelch-repeat protein at2g44130-like	comp42229_c0	-1.07	8.22E-06	*Pyrus x bretschneideri* (pear)	XP_009335865	1.28E-06	46.67%	-	signal peptide domain; transmembrane domain
f-box kelch-repeat protein at2g44130-like	**comp21522_c0**	**-1.30**	**8.33E-03**	***Pyrus x bretschneideri* (pear)**	**XP_009335865**	**4.15E-07**	**46.33%**	**F:protein binding**	**SSF117281 (kelch motif superfamily); IPR006652 (kelch repeat type 1); IPR015915 (kelch-type beta propeller domain)**
elongation factor 1-delta	**comp45173_c0**	**-2.97**	**4.22E-02**	***Artemia salina* (brine shrimp)**	**P32192**	**9.88E-26**	**65.00%**	**C:eukaryotic translation elongation factor 1 complex; P:translational elongation; F:translation elongation factor activity**	**IPR014038 (translation elongation factor EF1B, beta/delta subunit, guanine nucleotide exchange domain)**
beta-crystallin a1	comp51193_c0	-1.15	7.31E-12	*Lepeophtheirus salmonis* (copepod)	ADD38111	1.21E-35	54.33%	-	IPR001064 (Beta/gamma crystallin); signal peptide domain ; IPR011024 (Gamma-crystallin-related domain)

METABOLISM									
hypothetical protein	comp45348_c0	3.10	4.76E-13	*Ciona intestinalis* (tunicate)	XP_002121160	4.00E-42	56.33%	-	PTHR10366 (NAD dependent epimerase/dehydratase); IPR027417 (P-loop containing nucleoside triphosphate hydrolase); transmembrane helix domain
violaxanthin de-epoxidase	comp42733_c0	-1.26	4.46E-10	*Physcomitrella patens* (moss)	XP_001773358	5.55E-13	40.00%	F:violaxanthin de-epoxidase activity; C:chloroplast; P:oxidation-reduction process	IPR012674 (calycin domain); IPR010788 (violaxanthin de-epoxidase) ; IPR011038 (calycin-like superfamily); signal peptide domain
hypothetical protein	comp53782_c0	-0.96	3.65E-03	*Daphnia pulex* (waterflea)	EFX83386	1.13E-92	55.00%	F:hydrolase activity	IPR002018, IPR019826 (carboxylesterase, type B domain/active site); IPR029058 (alpha/Beta hydrolase fold domain); PTHR11559 (carboxylesterase family); signal peptide domain
aldehyde dehydrogenase family 3 member partial	comp56580_c0	0.82	9.10E-02	*Stegodyphus mimosarum* (spider)	KFM66996	3.36E-175	69.33%	F:oxidoreductase activity; P:biological_process	IPR012394, PTHR11699 (Aldehyde dehydrogenase NAD(P)-dependent family); IPR016162 (Aldehyde dehydrogenase, N-terminal domain); IPR016163 (Aldehyde dehydrogenase, C-terminal domain); cytoplasmic domain; transmembrane domain
aldehyde oxidase 2-like	comp59156_c0	-4.14	5.26E-25	*Daphnia pulex* (waterflea)	EFX86357	0.00E+00	60.67%	F:molecular_function	IPR005107 (CO dehydrogenase flavoprotein, C-terminal domain); IPR000674 (aldehyde oxidase/xanthine dehydrogenase, a/b hammerhead domain); IPR016208 (Aldehyde oxidase/xanthine dehydrogenase family); IPR008274 (Aldehyde oxidase/xanthine dehydrogenase, molybdopterin binding domain)

RESPONSE TO STRESS									
Knottin-like inhibitory protein	comp44575_c0	-3.54	2.77E-02	-	-	-	-	P:defense response	IPR003614 (knottin, scorpion-toxin-like domain); signal peptide domain
inter-alpha-trypsin inhibitor heavy chain h4	comp32809_c1	-1.34	1.12E-02	*Crassostrea gigas* (oyster)	EKC36390	6.50E-102	55.67%	-	no IPS match
Cytochrome P450	comp55690_c0	1.67	4.54E-05	*Tigriopus japonicus* (copepod)	AIL94133	1.16E-87	53.67%	P:oxidation-reduction process; F:iron ion binding; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; F:heme binding	IPR001128 (cytochrome P450 family); IPR002401 (cytochrome P450, E-class, group I family); signal peptide domain
glutathione s-transferase mu 1	comp46208_c1	0.90	4.84E-02	*Oryctolagus cuniculus* (rabbit)	NP_001075721	2.87E-37	51.33%	F:protein binding	IPR004046 (Glutathione S-transferase, C-terminal domain); IPR004045 (Glutathione S-transferase, N-terminal domain); IPR010987 (glutathione S-transferase, C-terminal-like domain)
CUTICLE INTEGRITY									
Chitotriosidase	**comp33461_c0**	**-1.30**	**3.23E-04**	***Daphnia pulex* (waterflea)**	**EFX90412**	**8.52E-80**	**73.33%**	**F:hydrolase activity, acting on glycosyl bonds; P:biological_process; P:carbohydrate metabolic process**	**IPR017853 (Glycoside hydrolase, superfamily); PTHR11177 (chitinase family); IPR011583 (chitinase II domain); IPR001579 (Glycoside hydrolase, chitinase active site); signal peptide domain**
hypothetical protein	**comp32479_c0**	**-1.58**	**1.25E-03**	***Daphnia pulex* (waterflea)**	**EFX90414**	**1.20E-45**	**63.33%**	**F:hydrolase activity, acting on glycosyl bonds; P:biological_process**	**IPR017853 (Glycoside hydrolase, superfamily domain); IPR001223 (Glycoside hydrolase, family 18, catalytic domain); IPR029070 (chitinase insertion domain); PTHR11177 (chitinase family)**
chitotriosidase	comp55805_c0	-1.29	2.82E-09	*Daphnia pulex* (waterflea)	EFX90412	2.17E-134	66.00%	F:hydrolase activity, acting on glycosyl bonds; P:biological_process	IPR017853 (glycoside hydrolase, superfamily); IPR011583 (chitinase II domain); IPR002557 (chitin-binding domain); IPR029070 (chitinase insertion domain); PTHR11177 (chitinase family); signal peptide domain
chondroitin proteoglycan-2-like	comp35157_c0	-1.60	4.70E-06	*Tribolium castaneum* (beetle)	XP_008192409	4.01E-08	60.33%	C:extracellular region; P:chitin metabolic process; F:chitin binding	IPR002557 (chitin-binding domain); PTHR23301 (chitin-binding peritrophin A family)

chitin-binding protein	comp43891_c0	-1.35	5.12E-06	*Drosophila virilis* (fly)	XP_002048076	3.65E-05	57.67%	P:chitin metabolic process; C:extracellular region; F:chitin binding	chitin-binding domain (PFAM); signal peptide domain
chondroitin proteoglycan-2-like	comp47090_c0	-1.96	6.08E-10	*Tribolium castaneum* (beetle)	XP_008192409	1.86E-09	60.33%	F:chitin binding; P:chitin metabolic process; C:extracellular region	IPR002557 (chitin-binding domain)
IMMUNE SYSTEM PROCESSES									
C-type lectin-like	comp47544_c0	-3.93	1.97E-20	-	-	-	-	-	IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold domain); signal peptide domain
macrophage mannose receptor partial	comp50187_c1	1.99	6.07E-05	*Chaetura pelagica* (bird)	KFU96626	1.50E-15	41.67%	F:carbohydrate binding	IPR001304 (c-type lectin domain); PTHR22803 (mannose, phospholipase, lectin receptor related family); IPR016187 (c-type lectin fold domain); signal peptide domain
hepatic lectin-like	comp49674_c0	-4.50	9.64E-08	*Oreochromis niloticus* (fish)	XP_005459156	3.02E-05	37	F:carbohydrate binding	IPR001304 (c-type lectin domain); IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold); cytoplasmic domain; transmembrane helix domain
C-type lectin-like	comp46353_c0	-4.98	2.64E-17	-	-	-	-	-	IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold)
C-type lectin-like	comp46353_c1	-5.04	2.82E-07	-	-	-	-	-	IPR016186 (c-type lectin-like domain); IPR016187 (c-type lectin fold)
C-type lectin-like	comp40027_c0	-3.53	1.33E-08	-	-	-	-	-	IPR016186 (c-type lectin-like); IPR016187 (c-type lectin fold)
c-type mannose receptor 2-partial	comp43463_c0	0.98	5.84E-02	*Saccoglossus kowalevskii* (worm)	XP_006825556	2.63E-18	44.67%	F:carbohydrate binding	IPR001304 (c-type lectin); IPR016186 (c-type lectin-like); PTHR22803 (mannose, phospholipase, lectin receptor related); IPR016187 (c-type lectin fold); signal peptide domain
Saposin-like	comp58868_c1	-3.76	1.03E-35	-	-	-	-	-	IPR011001 (saposin-like domain); IPR008139 (saposin B domain); signal peptide domain

TRANSPORT									
sodium-dependent phosphate transporter 1-a-like	comp51144_c0	-0.83	1.47E-01	*Metaseiulus occidentalis* (mite)	XP_003742817	1.38E-67	52.67%	F:inorganic phosphate transmembrane transporter activity; C:membrane; P:phosphate ion transport	IPR001204 (phosphate transporter family); cytoplasmic domain; transmembrane helix domain
sodium-dependent nutrient amino acid transporter 1-like	**comp12362_c0**	**-3.61**	**2.30E-02**	***Bombus terrestris* (bumblebee)**	**XP_003400703**	**2.47E-49**	**60.33%**	**P:neurotransmitter transport; F:neurotransmitter:sodium symporter activity; C:integral to membrane**	**IPR000175 (Sodium:neurotransmitter symporter family); SSF161070 (SNF-like superfamily); transmembrane helix domain; cytoplasmic domaiin**
peptide transporter family 1-like	comp56914_c0	-1.29	3.37E-04	*Dendroctonus ponderosae* (beetle)	ENN73556	3.13E-159	59.33%	F:transporter activity; C:membrane; P:oligopeptide transport	IPR000109 (Proton-dependent oligopeptide transporter family); PTHR11654:SF96 (peptide transporter family 1); IPR018456 (PTR2 family proton/oligopeptide symporter, conserved site); IPR016196 (Major facilitator superfamily domain, general substrate transporter domain); transmembrane helix domain; cytoplasmic domain
hypothetical protein	comp57280_c0	-0.92	6.55E-04	*Daphnia pulex* (waterflea)	EFX71591	1.72E-149	52.00%	-	IPR002035 (von Willebrand factor, type A domain); IPR013642 (Chloride channel calcium-activated); PTHR10579 (calcium-activated chlorine channel regulator); cytoplasmic domain; transmembrane domain
UNKNOWN									
Unknown	comp62318_c0	-3.06	2.01E-04	-	-	-	-	-	signal peptide domain
Unknown	comp16910_c0	-3.22	3.16E-10	-	-	-	-	-	no IPS match
Unknown	comp17945_c0	-1.25	3.84E-09	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
hypothetical protein	comp18829_c0	-1.85	8.61E-11	*Helobdella robusta* (leech)	XP_009029394	8.92E-04	44.00%	-	signal peptide domain

Unknown	comp58868_c2	-3.73	1.56E-16	-	-	-	-	-	no IPS match
Unknown	comp58868_c3	-4.14	5.26E-25	-	-	-	-	-	no IPS match
Unknown	comp56716_c0	-1.16	4.82E-01	-	-	-	-	-	no IPS match
Unknown	comp52925_c1	-2.18	2.85E-22	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp53341_c1	-1.47	1.65E-01	-	-	-	-	-	IPR029469 (PAN-4 domain); G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp53341_c2	-1.47	2.96E-03	-	-	-	-	-	IPR029469 (PAN-4 domain); G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp53492_c0	-1.46	7.38E-02	-	-	-	-	-	cytoplasmic domain; transmembrane helix domain
Unknown	comp46444_c2	-2.99	8.67E-13	-	-	-	-	-	signal peptide domain
Unknown	comp46722_c0	-1.93	6.30E-04	-	-	-	-	-	no transmembrane domain
Unknown	comp46043_c0	-3.07	5.77E-08	-	-	-	-	-	no IPS match
Unknown	comp44011_c0	-2.49	2.18E-08	-	-	-	-	-	no IPS match
Unknown	comp40368_c0	-1.16	1.09E-02	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); SSF57414 (hairpin loop containing domain-like superfamily)
Unknown	comp41942_c0	-1.48	1.33E-07	-	-	-	-	-	no IPS match
Unknown	comp43319_c0	-1.95	3.69E-08	-	-	-	-	-	IPR003014 (PAN-1 domain); IPR003609 (apple-like domain) SSF57414 (hairpin loop containing domain-like superfamily); signal peptide domain

Unknown	comp36118_c0	-1.02	3.57E-03	-	-	-	-	-	G3DSA:3.50.4.10 (hepatocyte growth factor superfamily); signal peptide domain
Unknown	comp36128_c0	-2.07	2.75E-06	-	-	-	-	-	transmembrane, cytoplasmic domain
Unknown	comp39845_c0	-2.85	2.28E-03	-	-	-	-	-	no IPS match
Unknown	**comp73005_c0**	**-3.92**	**7.86E-03**	**-**	**-**	**-**	**-**	**-**	**no IPS match**
Unknown	**comp63041_c0**	**3.79**	**2.40E-02**	**-**	**-**	**-**	**-**	**-**	**coiled coil domain**
Unknown	**comp60209_c0**	**-1.39**	**4.23E-03**	**-**	**-**	**-**	**-**	**-**	**signal peptide domain**
Unknown	**comp57815_c2**	**1.41**	**2.74E-02**	**-**	**-**	**-**	**-**	**-**	**signal peptide domain; transmembrane helix domain**
Unknown	**comp48674_c0**	**-3.48**	**4.49E-02**	**-**	**-**	**-**	**-**	**-**	**no IPS match**
hypothetical protein	**comp43699_c0**	**8.02**	**3.92E-02**	***Acartia pacifica* (copepod)**	**AGN29688**	**9.37E-48**	**69.67%**	**-**	**no IPS match**
Unknown	**comp41891_c0**	**-3.40**	**2.30E-02**	**-**	**-**	**-**	**-**	**-**	**signal peptide domain**
Unknown	**comp40961_c0**	**-2.49**	**2.30E-02**	**-**	**-**	**-**	**-**	**-**	**transmembrane helix domain**
Unknown	**comp39791_c0**	**-2.42**	**3.85E-04**	**-**	**-**	**-**	**-**	**-**	**coiled-coil domain**
Unknown	**comp16303_c0**	**5.03**	**2.46E-02**	**-**	**-**	**-**	**-**	**-**	**signal peptide domain; transmembrane helix domain**

Table S9: *Vibrio* transcripts identified in the assembled *E. affinis* transcriptome that are unique to the *V. sp. F10*-exposed treatment. The top blastx hit, GenBank accession number, and e-values are listed as well as the TMM- and FPKM- normalized expression values in the *V. sp. F10*-exposed treatment.

Transcript	BLASTx Top Hit	GenBank ID	E-value	Normalized expression
comp67439_c0	hypothetical protein [*Vibrio alginolyticus*]	WP_005380312.1	1.70E-14	F10: 2.58
comp26848_c0	hypothetical protein [*Vibrio sp. F10]*	WP_017036182.1	1.20E-11	F10: 4.4
comp18425_c0	hypothetical protein [*V. sp. F10*]	WP_017035998.1	9.70E-19	F10: 1.8
comp23751_c0	hypothetical protein [*Vibrio anguillarum*]	WP_019282965.1	1.80E-31	F10: 6.7

Chapter Four

Specificity of the bacterial communities associated with the copepod *Calanus finmarchicus*[1]

[1] To be submitted with: Manoshi S. Datta (co-first author for her contributions to the 16S rRNA sequencing analysis), Mark Baumgartner, Otto Cordero, Tracy Mincer, Ann Tarrant, Martin Polz
Proposed journal: Journal of the International Society for Microbial Ecology

ABSTRACT:

Animal microbiomes represent complex systems structured by dynamic interactions between host and microbial cells. Here, we investigate how specifically and reproducibly microbial communities assemble on ecologically relevant, marine invertebrate hosts and whether changes in host physiology can act as a selective force on host-associated microbial communities. We compared the abundance and composition of the bacterial communities on individual active and diapausing *Calanus finmarchicus* copepods collected from Trondheimfjord, Norway and examined the predictive power of host morphometrics on observed abundance patterns of microbial members. Our findings suggest that *C. finmarchicus* copepods have a predictable "core microbiome" that persists throughout the host's entrance into diapause, a dormancy period characterized by significant physiological changes in the host. Furthermore, we observe that the differences in the structure of the "flexible" microbiome in diapausing and active individuals appear to be at least partially driven by factors including the copepod's feeding history, copepod body size, and competitive microbial interactions. Ultimately our study provides a framework to better understand the processes that underlie microbial community assembly on *C. finmarchicus* copepods and suggests that host physiology may be an important selective force on copepod-associated bacterial community structure.

INTRODUCTION:

Microbial community composition can profoundly impact ecosystem biodiversity and function, yet the underlying forces that shape microbial community assembly are not well understood (1-3). Host-associated microbial communities provide unique systems in which to study the mechanisms that shape microbial community structure. Living hosts can often exert a strong selective force on associated microbiota, which is reflected in the frequent observation of species-specific (4-7) and even host site-specific (8, 9) microbiomes. The specificity of these complex microbial assemblages can be observed in terms of conserved taxonomy at high phylogenetic levels and/or function of the microbial community members (10-12). Identification of the predominant and consistent members of the host-associated community ('the core microbiome') can provide insight into how reliably the microbiomes assemble and the potential functional roles selected for by the host (13).

Although species-specific core microbiomes have been described in a variety of host organsisms, there can also be dynamic variability in microbial community composition across individuals (12, 14-16) and time (17, 18). The mechanisms that lead to variability in microbiome structure are poorly understood, although several studies suggest that inter-microbial interactions may be important contributors (8, 19, 20). In the human nasal cavity, for example, the presence of two species of resident bacteria (*Corynebacterium accolens*, *C. pseudodiphtheriticum*) was observed to be an important determinant of the occurrence of *Staphylococcus aureus* (19, 20), an opportunistic pathogen found in the nasal cavities of 20-30% of humans (21). In vitro co-cultivation experiments demonstrate that *C. accolens* supports the growth of *S. aureus* and that *C. pseudodiphtheriticum* inhibits the growth of *S. aureus*, suggesting that inter-species interactions contribute to the host variability in *S. aureus* presence in the nasal cavity (19). Host physiology and host life history are also thought to function as important selective forces on microbial community assembly and inter-individual variability in microbiome structure (16, 22). Research in vertebrates has shown that factors including dietary history (23, 24), maternal effects (25), genetics (25, 26), and gender (27) can be important determinants of host-derived inter-individual microbiome variability, but few such studies exist in invertebrate systems (e.g., 28). Alternatively, stochastic birth-death and immigration processes have also been implicated as

important drivers of assembly patterns of certain phylogenetic taxa on some invertebrate hosts (29).

The purpose of this study was to identify how specifically and reproducibly bacterial communities assemble on ecologically relevant, marine invertebrate hosts and whether dramatic changes in the physiology of a host population can act as a selective force on host-associated bacterial communities. We were also interested in examining whether inter-individual variability in the host microbiome could be attributed to bacterial interactions and/or inter-individual variability in host physiology. We examined the bacterial communities of marine copepods, abundant crustaceans known to cause dramatic impacts on the proliferation, virulence, and physiology of many bacterial species, including several pathogens (30-34). Several studies have suggested that copepod-associated bacterial communities are distinct from free-living communities in the water column (35, 36), but the mechanisms that drive the community assembly on these organisms have not yet been explored. We specifically studied the bacterial communities of the dominant large copepod in the North Atlantic, *Calanus finmarchicus*, which initiates a facultative diapause period typically during the last juvenile developmental stage (the fifth copepodite stage) (37) to avoid adverse seasonal conditions and high predation risk (38). This diapause period is characterized by a vertical migration to depth, arrested development, greatly reduced activity and metabolic rate, and cessation in feeding and molting (39, 40). The dramatic physiological changes associated with diapause in *C. finmarchicus* provide a unique system to study how a sustained change in the physiology of an invertebrate host may alter the abundance and community structure of its microbiome.

We collected numerous individual diapausing and active *C. finmarchicus* from Trondheimfjord, Norway to quantify bacterial abundance and community composition in light of morphometric data collected from each copepod. This first comprehensive analysis of the microbiome of individual copepods suggests that copepods have "core" bacterial members that persist throughout the host's entry into diapause. Furthermore, the distinct community structure of the "flexible" microbiome (i.e., the non-'core' members) of active and diapausing copepod populations appears to be partially driven by host factors including the copepod's feeding history and body size and by inter-specific bacterial interactions on the copepod surface. Overall, we observe that specific and predictable bacterial communities form on *C. finmarchicus* hosts which

may ultimately be driven by selective forces imposed by the host environment and/or host genetics in addition to inter-specific bacterial interactions.

METHODS:

Sampling of surface and deep C. finmarchicus

C. finmarchicus C5 copepodids were collected at Tröllet station near Trondheim, Norway over two sampling dates (6 June 2012, 11 June 2012) aboard the NTNU research vessel *R/V Gunnerus* (see Table S1 for details about sampling locations and times). Zooplankton were collected with a double trip close-open-close system using a 75 cm diameter ring outfitted with a 150 μm conical mesh net to capture samples in two depth strata: 250-350 m and 0-50 m (hereafter referred to as the deep and shallow samples, respectively). Depth was monitored in real time with an acoustic transmitter attached to the sampling net. Once the net was recovered, the contents of the cod end were poured into 10 gallon buckets filled with cold, sterile-filtered (0.22 μm) seawater. Before use, buckets were wiped with 10% bleach and thoroughly rinsed with MilliQ water and sterile-filtered (0.22 μm) seawater. Seawater samples were taken in tandem at representative depth intervals to capture the free-living bacterial communities (10 m intervals from surface to 50 m, 25 m intervals from 250-350 m). Water samples were taken with 5 L Niskin bottles, and temperature and salinity measurements were obtained with a SeaBird Conductivity, Temperature, and Depth (CTD) sensor.

Once back on shore, zooplankton and water samples were stored in the dark at 4°C for a few hours (< 6) until processing on the same day of collection. Live *C. finmarchicus* were periodically sieved (500 μm) from the larger bucket, rinsed with cold, sterile filtered seawater, and placed into an ice-chilled Petri dish. From the Petri dish they were individually captured using ethanol sterilized (95%) wide-bore glass Pasteur pipettes or forceps, mounted on autoclaved glass slides, photographed alive with a Canon EOS-20D camera attached to a Zeiss Stemi 2000C stereomicroscope, and placed in preservative. Copepod samples collected for bacterial community composition analysis were placed into microcentrifuge tubes containing 750 μl of RNALater (Ambion) and frozen (-20°C) until processing. Copepod samples collected for bacterial abundance counts were preserved in 1% formalin and kept at 4°C until processing.

Seawater samples were similarly preserved in 1% formalin for abundance counts or filtered (1 L) onto Sterivex and frozen at -80°C before DNA extractions.

Morphometrics

The prosome (i.e., the anterior portion of the copepod body) length and width, oil sac volume, and fractional fullness (41) of each copepodid were estimated from digital photographs of the live copepod samples as described previously (42). All measurements were calibrated with digital photographs of a stage micrometer taken repeatedly during sampling. Observations of gut contents were noted while viewing the live animals. The true prosome volume of *C. finmarchicus* is thought to be well approximated by numerical integrations of lateral views of the prosome as previously described (41). Prosome volume of a subset of copepod samples was calculated using three commonly used methods that are listed in order of increasing accuracy and complexity: the cubic function of body length (30), ellipsoid approximation using prosome length and prosome width (43), and numerical integrations of lateral views of the prosome (41). There was high consistency between the ellipsoid approximation method and the numerical integration method (Figure S1), prompting our use of the simpler ellipsoid approximation method to calculate prosome volume as a proxy for copepod body volume. Two-sample, two-tailed t-tests of the various morphometric variables were performed to compare the deep and shallow copepod populations for each sampling date in order to validate trends expected of diapausing and active populations, respectively. The relationships between the morphometric parameters (body length, body width, body volume, oil sac volume, oil sac fractional fullness, and food in gut) within and across copepod populations were evaluated by correlation analysis. The relationship between presence of food in the gut with the other morphometrics was not examined in the deep copepod population due to the low number of deep copepods with food in their guts.

Bacterial abundance

Staining with SYBR-Gold nucleic acid stain (Life Technologies) has been shown to provide greater contrast between bacterial cells and autofluorescent chitinous exoskeletons than

other traditional stains such as DAPI or SYBR Green (44-46). To enumerate the bacterial communities associated with individual *C. finmarchicus*, formalin-preserved copepods were first briefly homogenized within 1.5 mL Eppendorf tubes with a sterile plastic pestle and placed in a water bath sonicator for 3 mins to detach cells from the copepod carapace. Homogenized copepod samples were diluted (1:8) in 0.22-μm sterile-filtered artificial seawater and two replicate 2-mL aliquots of each diluted copepod sample were filtered onto a black polycarbonate 0.22 μm filter with vacuum pressure < 10 mm Hg. Blanks of 0.22-μm sterile-filtered artificial seawater were processed exactly as were copepod samples. For each water sample collected at the depths of copepods sampling, five replicate 1-mL aliquots were filtered onto a black polycarbonate 0.22-μm filter to enable enumeration of the free-living bacterial cells. Each polycarbonate filter was directly stained with 2x SYBR Gold (10,000x stock solution diluted in sterile-filtered TE buffer) for 15 minutes (47) and mounted in Citifluor (Citifluor Ltd.). Copepod and water filters were counted in a randomized order with an epifluorescence scope under blue light excitation. Cells on each replicate filter were counted with 50 random view fields and an average count was calculated across replicates to estimate the total bacterial cell abundance in each sample.

Bacterial abundance counts were pooled across the two sampling days (6/6/2012 and 6/11/2012) to increase the sample size from each depth (n = 20 individuals for each depth) and to minimize the effect of outliers. Bacterial cell density was calculated by dividing each individual's calculated bacterial cell abundance by its prosome volume. Two-sample, two-tailed t-tests for copepod bacterial cell counts (cells individual^{-1}) and bacterial cell density (cells mm^{-3}) were performed to examine differences in the abundance of bacteria associated with diapausing and active copepod populations. To investigate potential interactions between host physiology and bacterial load, relationships between the morphometric parameters (prosome length, prosome width, prosome volume, oil sac volume, oil sac fractional fullness) and the abundance counts on individuals within each copepod population (shallow, deep) were examined via correlation analysis. Two-sample, two-tailed t-tests were used to examine whether shallow copepods with food in their guts had higher bacterial cell counts (cells individual^{-1}) than those shallow copepods without food in their guts. Due to the low number of deep copepods with food in their guts as expected for diapausing copepods, this relationship was not examined in the deep population. Due to the strong correlation of depth with the examined morphometric parameters

(Table S2), the relationship between bacterial abundance counts and the various morphometric parameters was only examined within rather than across copepod populations.

DNA extraction

DNA was extracted following a modified protocol described previously (48). Sterivex filters were removed from their casing and cut using a sterile razor blade in a sterile Petri dish. The cut filters and individual copepod samples were each transferred with ethanol-flamed forceps into 2 mL screw-cap tubes containing a mixture of molecular biology grade 0.1 mm silica, 1.4 mm zirconium, and 4 mm silica beads (OPS Diagnostics PFMM 4000-100-28). To each tube, 400 µL of a DNA extraction buffer (0.1M Tris buffer (pH 8), 0.1 M NaEDTA (pH 8), 0.1 M phosphate buffer (pH 8), 1.5 M NaCl, 0.5% CTAB) and 100 µL of sodium dodecyl sulfate (10%) were added before 2.5 mins of beadbeating at top speed on a vortexer using a Vortex-Genie2® adapter (Mo-Bio). To each sample, 20 µL of lysozyme (10%) was added before incubation at 37°C for 30 mins. This was followed by a 30 min incubation with 20 µL of proteinase-K (10 mg mL^{-1}). The tubes were filled with 500 µL of phenol:chloroform:IAA (25:24:1, pH 8.0), slowly inverted on a rotating tube holder for 10 minutes, and spun at 12,000 x g for 5 mins. The supernatant was then extracted with 400 µL of chloroform and similarly inverted and spun down. DNA was precipitated from the samples with 1 µL of GlycoBlue (Life Technologies) and 0.6 x volume of isopropoanol (100%) at room temperature overnight. Precipitated DNA was spun down at 13,000 x g for 30 mins before rinsing twice with 70% ethanol. DNA was re-suspended in 30 µL of molecular biology grade water and stored at -20 °C. Copepod and seawater samples were extracted in random order, and blanks were processed simultaneously on each extraction day.

16S rRNA library preparation

Libraries for 16S rRNA paired-end sequencing were prepared following a previously described protocol (49) with some modifications. The PCR mixtures consisted of 0.25 µl of Q5 polymerase (New England Biolabs) with 1x Q5 buffer, 300 µM of dNTPs, 0.4 µg $µl^{-1}$ of bovine serum albumin (BSA), 0.3 µM of each primer, and 2 µl of DNA template for each 25 µl reaction.

To obtain adequate PCR product from the copepod samples, the addition of BSA was required, likely due to the presence of inhibitory substances in the copepod DNA as has been previously observed (35, 50). Quantitative PCR (qPCR) was performed before each PCR step to optimize the number of amplification cycles used. Selected qPCR products were also visualized on 2% TBE (tris-borate-EDTA) agarose gels to confirm the presence of the correct amplicon at each step.

Low quantities of DNA template in the copepod samples prevented proper amplification with the PE16S_V4_U515_F and PE16S_V4_E786_R primers (49), which target the V4 region of the 16S rRNA gene and contain a second-step priming site. Consequently, the copepod samples were first amplified with U515F (5'-GTGCCAGCMGCCGCGGTAA-3') and E786R (5'-GGACTACHVGGGTWTCTAAT-3') primers lacking the priming sites for 15 cycles to enrich for 16S rRNA V4 template. During the second amplification step, the samples were amplified for 10 cycles with the PE16S_V4_U515_F and PE16S_V4_E786_R primers (49). The Illumina-specific adapters (49) were added in the third and final amplification step following similar conditions as the previous PCR steps, except that 0.4 μM of the barcoded reverse primer, 0.4 μM of the PE-III-PCR-F primer and 4 μl of the AMPure-cleaned PCR product were used with 7 cycles of amplification. At each PCR step, samples were run in quadruplicate 25 μL reactions, then pooled and cleaned using Agencourt AMPure XP-PCR purification (Beckman Coulter, Brea, CA) according to the manufacturer's protocol. The seawater samples were diluted 1:100 in molecular biology grade water in order to use the same amplification scheme for both the copepod and seawater samples. Illumina libraries of DNA extraction blanks, no-template control (NTC) samples, and mock community samples consisting of 9 known microbial DNA templates (49) were prepared simultaneously with the copepod and seawater samples and sequenced.

After a final qPCR, all of the deep and shallow copepod libraries were multiplexed and gel extracted using the NucleoSpin® Gel and PCR Clean-up kit (Machery-Nagel) according to the manufacturer's instructions to eliminate a band identified by preliminary sequencing as the copepod 18S rRNA amplicon (data not shown). Another qPCR was performed before multiplexing the gel-extracted copepod samples and the seawater samples. The multiplexed copepod and seawater libraries were sequenced on a single lane using a paired-end approach

with 200 bp reads on the HiSeq 2000 Illumina sequencing machine at the BioMicro Center (Massachusetts Institute of Technology [MIT], Cambridge, MA).

Distribution-based Clustering (DBC)

Raw FASTQ files were merged and quality filtered using a combination of USEARCH (51), mothur (52), and custom Perl and Python script tools. Operational Taxonomic Units (OTUs) were clustered using distribution-based clustering (DBC), an OTU-calling algorithm that reduces the number of redundant OTUs by considering both genetic distance and the distribution of the sequences across environmental samples in calling an OTU (49). Therefore, this method is different than typical methods that use a fixed distance cut-off to estimate total species since a combination of sequence similarity and distribution is used to inform the clustering instead. To reduce sequencing errors, DBC was run on the quality-filtered sequences with the –k_fold 10 set initially at 90% sequence identity to clean up potential sequencing error by merging candidate sequences into OTUs if they were 10-fold less abundant than the OTU representative. This step resulted in a total of 15,060 OTUs across seawater and copepod samples. DBC was then run on the cleaned output with a –k_fold 0 and 95% identity, meaning that there was no abundance requirement for OTUs to be merged into an existing OTU representative with which it shared 95% identity. This step resulted in a total of 9,642 OTUs across all copepod and seawater samples. The sequential DBC steps reduced the redundancy in OTUs that were derived from the same organism or population by clustering sequences that were at least 95% similar and had highly correlated distributions across samples.

Sequence analysis

Initial hierarchical cluster analysis of the relative frequency of the most abundant OTUs across all sequenced samples was performed in R with a Euclidean distance metric and average-linkage clustering. The general trends observed were robust to the choice of clustering method (average- or complete-linkage) and distance metric (e.g., maximum, binary). Rarefaction curves and the Shannon index of diversity were calculated with the R package vegan. Further analysis focused on the most abundant OTUs (98 total) present at a mean relative abundance > 0.001

(reads/individual copepod sample) across all copepod individuals (excluding those identified as 'chloroplast') from both sampling dates (6/6/12, 6/11/12). The OTU relative abundances were base-10 log-transformed. To account for zeros in the data, 5.5×10^{-6} (equivalent to the smallest non-zero OTU relative frequency) was added to all OTUs with a relative frequency of 0. Copepod samples were ordered by average-linkage clustering of the Spearman correlation values between the log-transformed copepod communities. The top OTUs were also ordered by average-linkage clustering of the Spearman correlation values of their log-transformed abundance distribution across all copepod individuals. The "core" microbiome was defined as those OTUs present at non-zero relative abundance in at least 95% of the copepod samples collected on 6/6/12 and 6/11/12 (25 total). The "flexible" microbiome was defined as those remaining members of the most abundant 98 OTUs that were not in the core microbiome (73 total).

To explore the potential for positive or negative correlations between microbial members present on copepods, we calculated the Pearson correlation values between log-transformed OTU abundances with SparCC, a program designed to reduce the compositional effects inherent in correlation analysis of metagenomic data based on relative abundances (53). The default options were used for this analysis (-i 20, -x 10, -t 0.1). We focused on the top 102 OTUs found across all copepod individuals on 6/11/12 (excluding those identified as 'chloroplast'), as this was the sampling date where no deep copepods had food in their guts. Deep individuals with food in their guts observed on the earlier sampling date likely do not represent diapausing copepods, but rather diel vertical migrators (DVM), a phenomenon where active copepods descend to depth to avoid predators during the day. Therefore, we restricted the SparCC analysis to the 6/11/12 sampling date for which all deep copepods had no food in their guts.

Multivariate linear regression with depth, prosome volume, oil sac fractional fullness, and presence of food in gut was performed to examine which morphometric parameters significantly (stringent value of $p < 0.001$) predicted the log-transformed relative abundance patterns of each OTU found on copepod individuals sampled on 6/11/12 (102 total). A phylogenetic tree of the top 102 copepod OTUs was made by first aligning these OTU sequences against the SILVA 16S rRNA reference database (54). Based upon this alignment, FastTree (55) was used to prepare a tree based upon a generalized time reversible evolutionary model.

RESULTS:

Copepod morphometric analysis:

The shallow and deep samples followed patterns expected of active and diapausing copepods, respectively, in terms of the examined physiological indicators of diapause (Figure 1; Table S3). Morphometric analysis demonstrated that animals collected from deep water had significantly larger oil sac volumes, and oil sac fractional fullness, as would be expected of diapausing copepods. Additionally, the majority of shallow copepods had food in their guts, indicating recent feeding, while the majority of the deep copepods had no food in their guts, suggesting the cessation of active feeding, which is characteristic of diapausing copepods (Figure 1D; Table S3). Within each population, there was also notable inter-individual variability in the measured morphometric parameters (Figure 1). The deep population had higher variance in body and oil sac volume, while the shallow population had greater variance in oil sac fractional fullness. These overall trends were consistent across both sampling dates, although there were some differences in the morphometric variables between the two sampling dates. On the earlier sampling date (6/6/2012), the deep samples had slightly, but significantly smaller values than the later deep samples (6/11/12) for all measured morphometric parameters (prosome length/width/volume, oil sac volumes, oil sac fractional fullness) (Figure S2; Table S3). Conversely, the shallow samples collected on the later date (6/11/12) had slightly, but significantly larger values than the earlier date (6/6/12) for all parameters (except for oil sac fractional fullness which was not statistically different between the two dates) (Figure S2; Table S3).

Bacterial abundance counts:

Our results demonstrate that the bacterial density on individual *C. finmarchicus* is 10^3-10^4 times higher than that of an equivalent volume of ambient seawater, similar to findings in previous zooplankton studies (56) (Table S4; Figure S3). The bacterial load (cells individual^{-1}) and bacterial cell density (cells mm^{-3}) on individual copepods collected from shallow depths was approximately two times greater than those copepods collected at depth ($t = 5.016$, $p < 0.0001$) (Figure 2; Table S4). The inter-individual variability in the bacterial abundance counts of

shallow copepods (95% confidence interval: 5 x 10^4) was lower than in deep copepods (3.3 x 10^4) (Figure 2A). The average (± 95% confidence interval) bacterial load (cells/individual) observed in the shallow (4.5 x 10^5) and deep (2.8 x 10^5) copepod populations were also consistent with findings in other crustacean zooplankton studies when a similar approximation of body volume is used (Table S4). However, our findings demonstrate that the body volume approximation that uses a cubic function of prosome length (Figure S1), which is often used in copepod bacterial abundance studies (56), can underestimate body volume, leading to 2-fold overestimation of bacterial cell density in *C. finmarchicus* (Table S4).

Coupling the copepod morphometric analyses with the bacterial abundance counts, we observed that bacterial load (cells/individual) was unrelated to body volume in both the shallow ($p = 0.59$) and deep ($p = 0.35$) populations (Figure 2B-C). Bacterial load (cells/individual) was negatively correlated with oil sac fractional fullness in the shallow ($r = -0.47$; $p = 0.035$) population, but there was no significant relationship observed in the deep population ($p = 0.40$) (Figure 2D-E). Within the shallow and deep copepod populations, bacterial cell counts (counts/individual) were unrelated to prosome length, prosome width, and oil sac volume (Figure S4). The presence of food in the gut was unrelated to bacterial cell counts (cells/individual) ($p = 0.613$) (Figure S4-G) and bacterial cell density (cells/mm^3) ($p = 0.449$) in the shallow copepod population.

Community composition of copepods and ambient seawater:

The 16S rRNA amplicon sequencing of a total of 91 individual deep copepods, 108 individual shallow copepods, 49 seawater samples, and 12 control samples generated a total of 10,210,095 reads before quality filtering. OTU clustering with a distribution-based clustering algorithm (49) at the 95% sequence identity and distribution level resulted in approximately 9,000 OTUs. Rarefaction analysis showed that the copepod and seawater samples reach an inflection point (Figure S5), suggesting that our sequencing effort sufficiently sampled most members of the copepod and seawater microbial communities. Hierarchical clustering of all samples by microbial community composition demonstrated that the copepod and seawater samples clustered separately from the positive and negative controls (Figure S6). The positive

controls (Mix9) amplified on each plate also clustered together and had very short branch lengths compared to the seawater and copepod clades, suggesting that the amplification of samples across plates was highly consistent.

Our results demonstrate that copepod-associated microbial communities differ from ambient seawater communities in terms of overall diversity and structure of the microbial communities, as has been observed in previous studies (30, 36, 57) (Figure S6, S7). Deep and shallow copepod populations cluster into unique clades and demonstrate distinct relative abundance patterns of the microbial members, reflecting the significant differences in microbial community structure between active and diapausing copepods (Figure 3). Those few individuals from the deep population that clustered within the shallow population clade tended to have food in their guts (6 out of 11 deep individuals), a trait that is rare in diapausing copepods but common for active individuals, suggesting that these may be diapausing copepods that were awoken during collection or active copepods that are diel vertical migrators. The clade of deep copepods consisted of two distinct sub-groups, one of which (branches labeled in red on Figure 3) was comprised of individuals that were significantly larger than the rest of the deep copepod population in all examined morphometric parameters (Table S5). Deep and shallow copepods also exhibited differences in the overall diversity of the microbial communities, with the microbiomes of the shallow copepods demonstrating higher levels of diversity than those of the deep copepod population (Figure S7).

Through comparison of the microbial communities of numerous *C. finmarchicus* individuals, we identified a 'core' copepod microbiome persistent in nearly all samples. A core microbiome was present in both active and diapausing copepods, further highlighting the conservation of this microbial community structure despite the significant physiological changes that accompany entry into diapause (Figure 3). The 25 members of the core microbiome are from diverse phyla, including *Proteobacteria*, *Actinobacteria*, *Firmicutes*, and *Bacteroidetes* (Table S6). The top 4 most abundant OTUs within the core are classified at the genus level as *Vibrio* (seq16), *Propionibacterium* (seq13), and *Acinetobacter* (seq111, seq68), all of which have been previously isolated from copepod samples (57-59), and several of which have chitin-degrading members (60, 61).

We observed that the deep and shallow copepod populations exhibited characteristic community structure in their flexible microbiomes. In the deep population, one *Gammaproteobacteria* OTU (seq1) comprised up to 80% of the microbial community and was found at low abundance within the shallow population. The flexible microbiome of the deep population had higher relative abundance of Oceanospirillaceae, Pseudoalteromonadaceae, Vibrionaceae, Piscirickettsiaceae, Colwelliaceae, Streptococceae, and Pseudoalteromonadaceae (Table S7). The flexible microbiome of the shallow population had higher relative abundances of Thiotrichales, Alcaligenaceae, Hyphomonadaceae, Flavobacteriaceae, Puniceicoccaceae, Verrucomicrobiaceae, Planctomycetaceae, and Rhodobacteraceae (Table S8).

Correlation and multivariate linear regression analysis:

We further probed which specific mechanisms may be driving the characteristic community structure observed in the diapausing and active copepod populations by coupling correlation analysis of OTU occurrence with multivariate linear regression analysis of copepod morphometrics. This approach allowed us to begin to distinguish microbial co-occurrence patterns that may be driven by changes in host physiology from those that may reflect true microbial interactions. We identified several discrete groups of bacteria that were positively correlated with one another, representing the core, flexible deep, and flexible shallow microbiomes (Figure 4). Interestingly, bacteria that were highly positively correlated were often not closely related taxonomically, but their abundance patterns were frequently predicted by the same morphological characteristics in a multivariate linear regression model (Figure 4). We also observed several OTUs that were not related to any measured morphometric parameter, which may be due to a number of possibilities including the stringent p-value cutoff used, our focus on linear relationships between OTU occurrence and morphometric parameters, our omission of other important metrics of copepod physiology, or inter-specific bacterial interactions that encourage or prevent the co-occurrence of particular OTUs.

We observed that many OTUs that were enriched in the flexible microbiome of the deep population were well predicted by the depth of sample and body volume, while those of the shallow population were often predicted by depth of the sample, presence of food in the gut,

and/or oil sac fractional fullness. In several cases, we observed that the abundance pattern of an OTU was strongly predicted by several morphometric parameters in addition to depth of the sample, which shed light on the particular factors within each copepod population that were driving the characteristic flexible microbiome structure. For example, several members of the shallow flexible microbiome (e.g., *Flavobacteriaceae* (seq210)) were positively correlated with depth of the sample and presence of food in gut, suggesting that these OTUs are most abundant in the shallow copepod microbiome because they are driven by the presence of food in the gut. There are several examples of OTUs in both the deep and shallow flexible population for which abundance patterns were only predicted by depth of the sample (e.g., *Rhodobacteraceae* (seq317), *Staphylococceae* (seq462)). We hypothesize that this may be due to our inability to measure all of the physiological characteristics associated with the deep and shallow copepods that may be driving the observed microbiome structure. Several OTUs were strongly linked to one or more morphometric variables but not depth of the sample (e.g., *Flavobacteriaceae* (seq17), *Chitinophagaceae* (seq305)), which was primarily observed in those OTUs that were present across nearly all examined copepod individuals. For example, many of the members of the core copepod microbiome were positively correlated with each other and their abundance patterns were linked to prosome volume. This trend may be driven by the distinct subcluster within the deep copepod population that consists of individuals with significantly larger prosome volumes and lower relative abundances of the core microbiome members than the rest of the deep copepod population.

Although the members of the core and flexible shallow microbiome were often highly correlated to one another, we observed several negative correlations between OTUs characteristic of the deep flexible microbiome. A Pseudoalteromonadaceae (seq581) was negatively correlated to several other OTUs, namely a Colwelliaceae (seq67), Oceanspirillaceae (seq3), and a Pseudoalteromonadaceae (seq304). Another Oceanspirillaceae (seq387) was negatively correlated to a Vibrionaceae (seq546), Pseudoalteromonadaceae (seq788), Pseudomonadaceae (seq788), and a Streptococceae (seq1027). A Vibrionaceae (seq848) OTU was also negatively correlated with Pseudoalteromonadaceae (seq304) and Pseudomonadaceae (seq788) OTUs. A Streptococceae (seq372) OTU was negatively correlated with a Rhodobacteraceae (seq316) OTU as well.

DISCUSSION:

Each individual animal host can be viewed as a habitat "patch" colonized by microbial communities structured by dynamic forces including selection by the host or stochastic processes. Exploration of community assembly processes in host-associated microbiota is particularly important in light of the mounting evidence of considerable inter-individual variability in microbiomes (12, 14-16) and the significant implications of microbiome structure to host health (62). In this study, we analyzed the abundance and composition of the microbial communities of numerous individual *Calanus finmarchicus* copepods from Trondheimfjord to identify the specificity and reproducibility of microbial communities on these habitat "patches". By collecting both active and diapausing copepod individuals, we further tested whether the structure of the copepod microbiome may be influenced by the physiological changes associated with the host's entry into diapause. Our findings suggest that *C. finmarchicus* has a core microbiome that is common to both active and diapausing copepods, but that each copepod population harbors distinct flexible microbiomes, which may be driven by forces including host physiology and microbial interactions.

Bacterial abundance counts of individual copepods suggested that bacterial load was in fact influenced by host physiology. For example, we observed that active copepods have a higher bacterial load than do diapausing copepods. This trend may be explained by previous findings that the nutrients provided by "sloppy feeding" or excretions from metabolically active copepods enhance bacterial abundance and activity (30, 63, 64). Alternatively, since epibiotic bacterial community succession can be related to the host's age and time since molt (65, 66), the lower bacterial load on the diapausing copepods could be due to their greater age and a succession of their bacterial communities from transient founder species to less abundant constitutive species. Although the majority of the copepod morphometric factors examined did not significantly correlate with bacterial abundance within each copepod population, we did observe a significant negative correlation between bacterial abundance and oil sac fractional fullness within the active copepods. We hypothesize that the 'fuller' active copepods represent copepods that are closer to entering the diapause period and which may have consequently begun reducing metabolic rates and frequency of feedings, which may decrease the abundance of the gut or surface bacterial community. The abundance of bacterial communities within each copepod population was

independent of copepod body volume in accordance with previous studies (67). As the distribution of bacteria on the surface of copepods is known to be non-uniform and concentrated in nutrient-rich areas such as the mouthparts and gut (59), our results suggest that saturation of the copepod surface may be dependent on factors such as nutrient availability rather than space (67).

Our observation of a "core" group in the copepod microbiome builds on previous literature that identified core microbial members in numerous organisms, including plants (68), insects (69), and humans (9). However, a rare aspect of the core microbome that we have identified in this study is the conservation of microbial members at considerably high phylogenetic resolution (i.e., family level). In many other systems, a 'core' microbiome is only apparent at much lower phylogenetic resolution (i.e., phylum) or in consideration of the conserved function of microbial members (10-12, 62). Many members of the core *C. finmarchicus* microbiome, (e.g., *Vibrionaceae*, *Moraxellaceae*, *Chitinophagaceae*), and the flexible microbiome (e.g., *Flavobacteriaceae*, *Pseudoalteromonadaceae*, *Verrumicrobiaceae*) have been found on copepods including *C. finmarchicus* in previous studies (36, 57-59, 70), but our approach of examining physiologically distinct individual copepods enabled the identification of those microbial members that stably and reproducibly associate with *C. finmarchicus*. Interestingly, two of the most abundant OTUs represented in the core microbiome are identified as *Acinetobacter*, a genus which is frequently found in association with eukaryotic hosts including as mutualists of insects (71) and even as opportunistic pathogens of humans (72). The abundance of several genera known to contain several pathogenic species (i.e., *Vibrio, Actinobacter*) persistently associated with *C. finmarchicus* enforces the concept that copepods represent important environmental reservoirs of human pathogens (30, 73).

Although the specific members comprising the flexible microbiome differed between the active and diapausing copepod populations, both had communities enriched in surface-associated and opportunistically pathogenic taxa. For example, the active copepods had flexible microbiomes consisting of many representatives of Flavobacteriaceae, including the *Tenacibaculum*, *Lacinutrix*, and *Maribacter* genera, all of which are known to associate with diverse eukaryotic hosts (74-76). The flexible microbiome of diapausing copepods was dominated by Vibrionaceae, Streptococceae and Pseudomonadaceae, which consist of many

opportunistic pathogenic species (77-79). Furthermore, Rhodobacteraceae, known to be important primary surface colonizers in marine systems (80), were predominant members of the flexible microbiomes of both active and diapausing samples. These findings suggest that although the members of the flexible microbiome do not persist across the host's entry into diapause, that they may still be tightly associated copepod symbionts.

We observed that many members of the flexible microbiome of the diapausing copepods that were positively correlated to one another were also significantly predicted by copepod body volume. This trend may be driven by the sub-population of the deep copepods that have significantly larger body volumes and display lower relative abundances of many microbial members in the flexible microbiome. Body size of *C. finmarchicus* individuals is strongly related to environmental conditions, namely temperature and food availability, experienced during juvenile development (81). Therefore, differences in copepod life history could lead to the observed differences in microbial community composition of the deep sub-population if the diapausing individuals grew up at a time (i.e., earlier in the diapause season) or place where food availability and/or temperature were higher (82).

In the flexible microbiome of the shallow copepod population, we observed that the abundance patterns of many microbial members were strongly linked to the presence of food in the gut. These results make sense in light of previous findings that feeding history can influence the abundance and composition of microbial communities (23, 24, 83). Interestingly, we did not observe a significant relationship between the presence of food in the gut and the bacterial load of active copepods. Therefore, the active feeding of non-diapausing copepods may be an important mechanism of providing inocula for the flexible microbiome structure although inter-bacterial competition and/or host activity could control the overall abundance of bacteria found on individuals of this copepod population.

Positive and negative interactions between co-habitating bacterial populations are known to be important contributors to microbial community structure (84-86). Within the active and diapausing flexible microbiomes we observed several microbial members with highly correlated abundance patterns that were not significantly predicted by host morphometrics. These co-occurrence patterns could be due to our inability to accurately capture the causative physiological attributes of the hosts or due to true microbial interactions on the copepod. Within

the deep flexible microbiome, we observed several negative correlations between OTUs that may represent true antagonistic interactions as inferred from their phylogenetic associations. For example, several OTUs identified as *Pseudoalteromonas*, a genus characterized by species that synthesize a variety of biologically active molecules, were negatively correlated with OTUs from diverse taxa including Oceanspirillaceae, Vibrionaceae, Rhodobacteraceae and Colwelliaceae. Many of the diverse compounds synthesized by *Pseudoalteromonas* display anti-bacterial and bacteriolytic properties, which are thought to benefit the competitive ability of *Pseudoalteromonas* cells in colonizing surfaces and obtaining nutrients (87). An OTU classified as *Pseudomonas*, another genus known to produce antibiotic compounds (88), was also negatively associated with Vibrionaceae and Oceanspirillaceae OTUs. Furthermore, none of the pairs of OTUs described above had relative abundance patterns that were well predicted by host morphometrics. Overall, our findings suggest that competitive interactions may be especially important in structuring the flexible microbiome of deep copepods.

In summary, our study demonstrates that bacterial community assembly on similar ecological patches of *C. finmarchicus* invertebrates can be highly specific and reproducible. Our findings also suggest that physiological changes within the host can influence abundance and composition patterns of bacterial community members in accordance with previous findings in other invertebrate systems. However, there were several bacterial members for which their abundance patterns across active and diapausing copepod populations were not well predicted by copepod morphometrics, which suggests that inter-bacterial dynamics may also be important contributors to microbial community structure on copepods. Future studies should investigate whether the core members that we observe in our study are only associated with *C. finmarchicus* from Tronheimfjord, or whether they are more globally associated with marine copepods from diverse environments.

ACKNOWLEDGEMENTS:

This work was supported in part by grant number OCE-1132567 from the National Science Foundation to MFB and AMT. Funding for AAA provided by EPA STAR Fellowship, NSF GRFP, and the WHOI Ocean Venture Fund. We would like to thank Sarah Preheim for help with performing the distribution-based clustering, Michael Cutler for providing advice regarding 16S rRNA library preparation, Nadine Lysiak for assistance in the field, Morgan Rubanow for assistance with the morphometric analysis, and Krista Longnecker for her advice regarding the bacterial abundance counts.

REFERENCES:

1. **Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S.** 2013. Patterns and processes of microbial community assembly. Microbiology and Molecular Biology Reviews **77:**342-356.
2. **Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Ovreas L, Reysenbach A-L, Smith VH, Staley JT.** 2006. Microbial biogeography: putting microorganisms on the map. Nature reviews. Microbiology **4:**102-112.
3. **Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH.** 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature reviews. Microbiology **10:**497-506.
4. **Grossart HP, Dziallas C, Tang KW.** 2009. Bacterial diversity associated with freshwater zooplankton. Environ. Microbiol. Rep. **1:**50-55.
5. **Franzenburg S, Walter J, Künzel S, Wang J, Baines JF, Bosch TCG, Fraune S.** 2013. Distinct antimicrobial peptide expression determines host species-specific bacterial associations. Proceedings of the National Academy of Sciences **110:**E3730-E3738.
6. **Carlos C, Torres TT, Ottoboni LM.** 2013. Bacterial communities and species-specific associations with the mucus of Brazilian coral species. Sci. Rep. **3:**1624.
7. **Apprill A, Robbins J, Eren AM, Pack AA, Reveillaud J, Mattila D, Moore M, Niemeyer M, Moore KMT, Mincer TJ.** 2014. Humpback whale populations share a core skin bacterial ommunity: towards a health index for marine mammals? PLoS ONE **9:**e90785.
8. **Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C.** 2012. Microbial co-occurrence relationships in the human microbiome. PLoS Computational Biology **8:**e1002606.
9. **Heath-Heckman EAC, Peyer SM, Whistler CA, Apicella MA, Goldman WE, McFall-Ngai MJ.** 2013. Bacterial bioluminescence regulates expression of a host cryptochrome gene in the squid-*vibrio* symbiosis. mBio **4**(2)**:**10.1128/mBio.00167-00113.
10. **Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI.** 2009. A core gut microbiome in obese and lean twins. Nature **457:**480-484.
11. **Shade A, Handelsman J.** 2011. Beyond the Venn diagram: the hunt for a core microbiome. Environmental Microbiology **14:**4-12.
12. **Consortium HMP.** 2012. Structure, function and diversity of the healthy human microbiome. Nature **486:**207-214.
13. **Tschöp MH, Hugenholtz P, Karp CL.** 2009. Getting to the core of the gut microbiome. Nature Biotechnology **27:**344-346.
14. **Lazarevic V, Whiteson K, Hernandez D, Francois P, Schrenzel J.** 2011. Study of inter- and intra-individual variations in the salivary microbiota. BMC Genomics **11:**1471-2164.
15. **Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, Gordon JI, Knight R.** 2011. Moving pictures of the human microbiome. Genome Biology **12:**2011-2012.
16. **Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R.** 2009. Bacterial community variation in human body habitats across space and time. Science **326:**1694-1697.
17. **Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE.** 2011. Succession of microbial consortia in the developing infant gut microbiome. Proc. Natl. Acad. Sci. U.S.A. **108:**4578-4585.
18. **Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Program NCS, Bouffard GG, Blakesley RW, Murray PR, Green ED, Turner ML, Segre JA.** 2009. Topographical and temporal diversity of the human skin microbiome. Science **324:**1190-1192.

19. **Yan ML, Pamp SJ, Fukuyama J, Hwang PH, Cho DY, Holmes S, Relman DA.** 2013. Nasal microenvironments and interspecific interactions influence nasal microbiota complexity and *S. aureus* carriage. Cell Host & Microbe **14:**631-640.
20. **Libberton B, Coates RE, Brockhurst MA, Horsburgh MJ.** 2014. Evidence that intraspecific trait variation among nasal bacteria shapes the distribution of *Staphylococcus aureus*. Infect. Immun. **82:**3811-3815.
21. **Kluytmans JA, Wertheim HF.** 2005. Nasal carriage of *Staphylococcus aureus* and prevention of nosocomial infections. Infection **33:**3-8.
22. **Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA.** 2012. The application of ecological theory toward an understanding of the human microbiome. Science **336:**1255-1262.
23. **Davenport ER, Mizrahi-Man O, Michelini K, Barreiro LB, Ober C, Gilad Y.** 2014. Seasonal variation in human gut microbiome composition. PLoS ONE **9:**e90731.
24. **Wang J, Linnenbrink M, Kunzel S, Fernandes R, Nadeau MJ, Rosenstiel P, Baines JF.** 2014. Dietary history contributes to enterotype-like clustering and functional metagenomic content in the intestinal microbiome of wild mice. Proc. Natl. Acad. Sci. U.S.A. **111:**E2703-E2710.
25. **Spor A, Koren O, Ley R.** 2011. Unravelling the effects of the environment and host genotype on the gut microbiome. Nat. Rev. Microbiol. **9:**279-290.
26. **Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua K, Kachman SD, Moriyama EN, Walter J, Peterson DA, Pomp D.** 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc. Natl. Acad. Sci. U.S.A. **107:**18933-18938.
27. **Zhao L, Wang G, Siegel P, He C, Wang H, Zhao W, Zhai Z, Tian F, Zhao J, Zhang H, Sun Z, Chen W, Zhang Y, Meng H.** 2012. Quantitative genetic background of the host influences gut microbiomes in chickens. Scientific Reports **3**doi:10.1038/srep01163.
28. **Wegner KM, Volkenborn N, Peter H, Eiler A.** 2013. Disturbance induced decoupling between host genetics and composition of the associated microbiome. BMC Microbiology **13**(252)**:**doi:10.1186/1471-2180-1113-1252.
29. **Preheim SP, Boucher Y, Wildschutte H, David LA, Veneziano D, Alm EJ, Polz MF.** 2011. Metapopulation structure of *Vibrionaceae* among coastal marine invertebrates. Environ. Microbiol. **13:**265-275.
30. **Tang KW, Turk V, Grossart H-P.** 2010. Linkage between crustacean zooplankton and aquatic bacteria. Aquat. Microb. Ecol. **61:**261-277.
31. **Kirn TJ, Jude BA, Taylor RK.** 2005. A colonization factor links *Vibrio cholerae* environmental survival and human infection. Nature **438:**863-866.
32. **Hunt DE, Gevers D, Vahora NM, Polz MF.** 2008. Conservation of the chitin utilization pathway in the *Vibrionaceae*. Appl. Environ. Microbiol. **74:**44-51.
33. **Huq A, Colwell RR, Rahman R, Ali A, Chowdhury MAR, Parveen S, Sack DA, Russekcohen E.** 1990. Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and cell cultures Appl. Environ. Microbiol. **56:**2370-2373.
34. **Epstein SS, Colwell R.** 2009. Viable but Not Cultivable Bacteria, p. 121-129, Uncultivated Microorganisms, vol. 10. Springer Berlin Heidelberg.
35. **Møller EF, Riemann L, Sondergaard M.** 2007. Bacteria associated with copepods: abundance, activity and community composition. Aquatic Microbial Ecology **47:**99-106.
36. **De Corte D, Lekunberri I, Sintes E, Garcia J, Gonzales S, Herndl G.** 2014. Linkage between copepods and bacteria in the North Atlantic Ocean. Aquatic Microbial Ecology **72:**215-225.
37. **Hirche HJ.** 1996. Diapause in the marine copepod, *Calanus finmarchicus* - A review. Ophelia **44:**129-143.
38. **Kaartvedt S.** 1996. Habitat preference during overwintering and timing of seasonal vertical migration of *Calanus finmarchicus*. Ophelia **44:**145-156.

39. **Hirche HJ.** 1983. Overwintering of *Calanus finmarchicus* and *Calanus helgolandicus* Marine Ecology Progress Series **11:**281-290.
40. **Ingvarsdóttir A, Houlihan DF, Heath MR, Hay SJ.** 1999. Seasonal changes in respiration rates of copepodite stage V *Calanus finmarchicus* (Gunnerus). Fisheries Oceanography **8:**73-83.
41. **Miller CB, Crain JA, Morgan CA.** 2000. Oil storage variability in Calanus finmarchicus. ICES Journal of Marine Science: Journal du Conseil **57:**1786-1799.
42. **Tarrant A, Baumgartner M, Verslycke T, Johnson C.** 2008. Differential gene expression in diapausing and active *Calanus finmarchicus* (Copepoda). Marine Ecology Progress Series **355:**193-207.
43. **Mauchline J.** 1998. Chemical Composition, p. 220-252, Advances in Marine Biology: The Biology of Calanoid Copepods, vol. Volume 33. Academic Press.
44. **Bickel S, Tang K.** 2010. Microbial decomposition of proteins and lipids in copepod versus rotifer carcasses. Mar. Biol. **157:**1613-1624.
45. **Bickel SL, Tang KW, Grossart H-P.** 2012. Ciliate epibionts associated with crustacean zooplankton in german lakes: distribution, motility, and bacterivory, vol. 3.
46. **Tang K, Bickel S, Dziallas C, Grossart H.** 2009. Microbial activities accompanying decomposition of cladoceran and copepod carcasses under different environmental conditions. Aquatic Microbial Ecology **57:**89-100.
47. **Chen F** 2005, posting date. Protocols for Counting Viruses and Bacteria using SYBR Gold Stain. [Online.]
48. **Crump B.** 2007. DNA extraction from Sterivex filters. Oregon State University.
49. **Preheim SP, Perrotta AR, Martin-Platero AM, Gupta A, Alm EJ.** 2013. Distribution-based clustering: using ecology to refine the operational taxonomic unit. Applied and Environmental Microbiology **79:**6593-6603.
50. **Brandt P, Gerdts G, Boersma M, Wiltshire KH, Wichels A.** 2010. Comparison of different DNA-extraction techniques to investigate the bacterial community of marine copepods. Helgoland Marine Research **64:**331-342.
51. **Edgar R.** 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics **26:**2460-2461.
52. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology **75:**7537-7541.
53. **Friedman J, Alm EJ.** 2012. Inferring correlation networks from genomic survey data. PLoS Computational Biology **8:**e1002687.
54. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies Jr, Glöckner FO.** 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research **41:**D590-D596.
55. **Price MN, Dehal PS, Arkin AP.** 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular Biology and Evolution **26:**1641-1650.
56. **Tang KW, Turk V, Grossart HP.** 2010. Linkage between crustacean zooplankton and aquatic bacteria. Aquatic Microbial Ecology **61:**261-277.
57. **Gerdts G, Brandt P, Kreisel K, Boersma M, Schoo KL, Wichels A.** 2013. The microbiome of North Sea copepods. Helgoland Marine Research **67:**757-773.
58. **Holland RS, Hergenrader GL.** 1981. Bacterial epibionts of diaptomid copepods. Transactions of the American Microscopical Society **100:**56-65.
59. **Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR.** 1983. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Appl. Environ. Microbiol. **45:**275-283.

60. **Meibom KL, Li XB, Nielsen AT, Wu C-Y, Roseman S, Schoolnik GK.** 2004. The *Vibrio cholerae* chitin utilization program. Proc. Natl. Acad. Sci. U.S.A. **101:**2524-2529.
61. **Sato K, Kato Y, Fukamachi A, Nogawa M, Taguchi G, Shimosaka M.** 2010. Construction and analysis of a bacterial community exhibiting strong chitinolytic activity. Bioscience, Biotechnology, Biochemistry **74:**636-640.
62. **Shafquat A, Joice R, Simmons SL, Huttenhower C.** 2014. Functional and phylogenetic assembly of microbial communities in the human microbiome. Trends in Microbiology **22:**261-266.
63. **Möller EF.** 2005. Sloppy feeding in marine copepods: prey-size-dependent production of dissolved organic carbon. J. Plankton Res. **27:**27-35.
64. **Carman KR.** 1994. Stimulation of marine free-living and epibiotic bacterial activity by copepod excretions. FEMS Microbiol. Ecol. **14:**255-261.
65. **Caro A, Escalas A, Bouvier C, Grousset E, Lautredou-Audouy N, Roques C, Charmantier M, Gros O.** 2012. Epibiotic bacterial community of *Sphaeroma serratum* (Crustacea, Isopoda): relationship with molt status. Marine Ecology Progress Series **457:**11-27.
66. **Polz MF, Harbison C, Cavanaugh CM.** 1999. Diversity and heterogeneity of epibiotic bacterial communities on the marine nematode *Eubostrichus dianae*. Applied and Environmental Microbiology **65:**4271-4275.
67. **Rawlings TK, Ruiz GM, Colwell RR.** 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.
68. **Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J, Engelbrektson A, Kunin V, Rio TGd, Edgar RC, Eickhorst T, Ley RE, Hugenholtz P, Tringe SG, Dangl JL.** 2012. Defining the core *Arabidopsis thaliana* root microbiome. Nature **488:**86-90.
69. **Meriweather M, Matthews S, Rio R, Baucom RS.** 2013. A 454 survey reveals the community composition and core microbiome of the common bed bug (*Cimex lectularius*) across an urban landscape. PLoS ONE **8:**e61465.
70. **Dziallas C, Grossart H-P, Tang KW, Nielsen TG.** 2013. Distinct communities of free-living and copepod-associated microorganisms along a salinity gradient in Godthabsfjord, West Greenland. Arctic, Antarctic, and Alpine Research **45:**471-480.
71. **Soumana IH, Simo G, Njiokou F, Tchicaya B, Abd-Alla AMM, Cuny Gr, Geiger A.** 2013, p S89-S93. Journal of Invertebrate Pathology.
72. **Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, Towner K.** 2006. The Genus *Acinetobacter*, p. 746-758, The Prokaryotes. Springer New York.
73. **Gugliandolo C, Irrera GP, Lentini V, Maugeri TL.** 2008. Pathogenic *Vibrio*, *Aeromonas* and *Arcobacter spp*. associated with copepods in the Straits of Messina (Italy). Mar. Pollut. Bull. **56:**600-606.
74. **Bowman JP, Nichols DS.** 2005. Novel members of the family *Flavobacteriaceae* from Antarctic maritime habitats including *Subsaximicrobium wynnwilliamsii* gen. nov., sp. nov., *Subsaximicrobium saxinquilinus* sp. nov., *Subsaxibacter broadyi* gen. nov., sp. nov., *Lacinutrix copepodicola* gen. nov., sp. nov., and novel species of the genera *Bizionia*, *Gelidibacter* and *Gillisia*. International Journal of Systematic and Evolutionary Microbiology **55:**1471-1486.
75. **Suzuki M, Nakagawa Y, Harayama S, Yamamoto S.** 2001. Phylogenetic analysis and taxonomic study of marine *Cytophaga*-like bacteria: proposal for *Tenacibaculum* gen. nov. with *Tenacibaculum maritimum* comb. nov. and *Tenacibaculum ovolyticum* comb. nov., and description of *Tenacibaculum mesophilum* sp. nov. and *Tenacibaculum amylolyticum* sp. nov. International Journal of Systematic and Evolutionary Microbiology **51:**1639-1652.
76. **Nedashkovskaya OI, Kim SB, Han SK, Lysenko AM, Rohde M, Rhee M-S, Frolova GM, Falsen E, Mikhailov VV, Bae KS.** 2004. *Maribacter* gen. nov., a new member of the family Flavobacteriaceae, isolated from marine habitats, containing the species *Maribacter*

sedimenticola sp. nov., *Maribacter aquivivus* sp. nov., *Maribacter orientalis* sp. nov. and *Maribacter ulvicola* sp. nov. International Journal of Systematic and Evolutionary Microbiology **54:**1017-1023.

77. **Urbanczyk H, Ast JC, Higgins MJ, Carson J, Dunlap PV.** 2007. Reclassification of *Vibrio fischeri*, *Vibrio logei*, *Vibrio salmonicida* and *Vibrio wodanis* as *Aliivibrio fischeri* gen. nov., comb. nov., *Aliivibrio logei* comb. nov., *Aliivibrio salmonicida* comb. nov. and *Aliivibrio wodanis* comb. nov. International Journal of Systematic and Evolutionary Microbiology **57:**2823-2829.
78. **Pier GB, Madin SH.** 1976. *Streptococcus iniae* sp. nov., a Beta-Hemolytic *Streptococcus* Isolated from an Amazon Freshwater Dolphin, *Inia geoffrensis*. International Journal of Systematic Bacteriology **26:**545-553.
79. **Baumann P, Bowditch RD, Baumann L, Beamann B.** 1983. Taxonomy of Marine *Pseudomonas* Species: *P. stanieri* sp. nov.; *P. perfectomarina* sp. nov., nom. rev.; *P. nautica*: and *P. doudoroffii*. International Journal of Systematic Bacteriology **33:**857-865.
80. **Dang H, Li T, Chen M, Huang G.** 2008. Cross-ocean distribution of *Rhodobacterales* bacteria as primary surface colonizers in temperate coastal marine waters. Appl. Environ. Microbiol. **74:**52-60.
81. **Campbell RG, Wagner MM, Teegarden GJ, Boudreau CA, Durbin EG.** 2001. Growth and development rates of the copepod *Calanus finmarchicus* reared in the laboratory. Marine Ecology Progress Series **221:**161-183.
82. **Tang K, Dziallas C, Hutalle-Schmelzer K, Grossart HP.** 2009. Effects of food on bacterial community composition associated with the copepod Acartia tonsa Dana. Biology Letters **5:**549-553.
83. **Broderick NA, Buchon N, Lemaitre B.** 2014. Microbiota-induced changes in *Drosophila melanogaster* host gene expression and gut morphology. mBio **5**(3)**:**e01117-01410.01128/mBio.01117-01114.
84. **Cordero OX, Wildschutte H, Kirkup B, Proehl S, Ngo L, Hussain F, Le Roux F, Mincer T, Polz MF.** 2012. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science **337:**1228-1231.
85. **Tait K, Sutherland IW.** 2002. Antagonistic interactions amongst bacteriocin-producing enteric bacteria in dual species biofilms. J. Appl. Microbiol. **93:**345-352.
86. **Burmølle M, Webb JS, Rao D, Hansen LH, Sørensen SJ, Kjelleberg S.** 2006. Enhanced biofilm formation and increased resistance to antimicrobial agents and bacterial invasion are caused by synergistic interactions in multispecies biofilms. Appl. Environ. Microbiol. **72:**3916-3923.
87. **Holmstrom C, Kjelleberg S.** 1999. Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. Fems Microbiology Ecology **30:**285-293.
88. **Nair S, Simidu U.** 1987. Distribution and significance of heterotrophic marine bacteria with antibacterial activity. Appl. Environ. Microbiol. **53:**2957-2962.

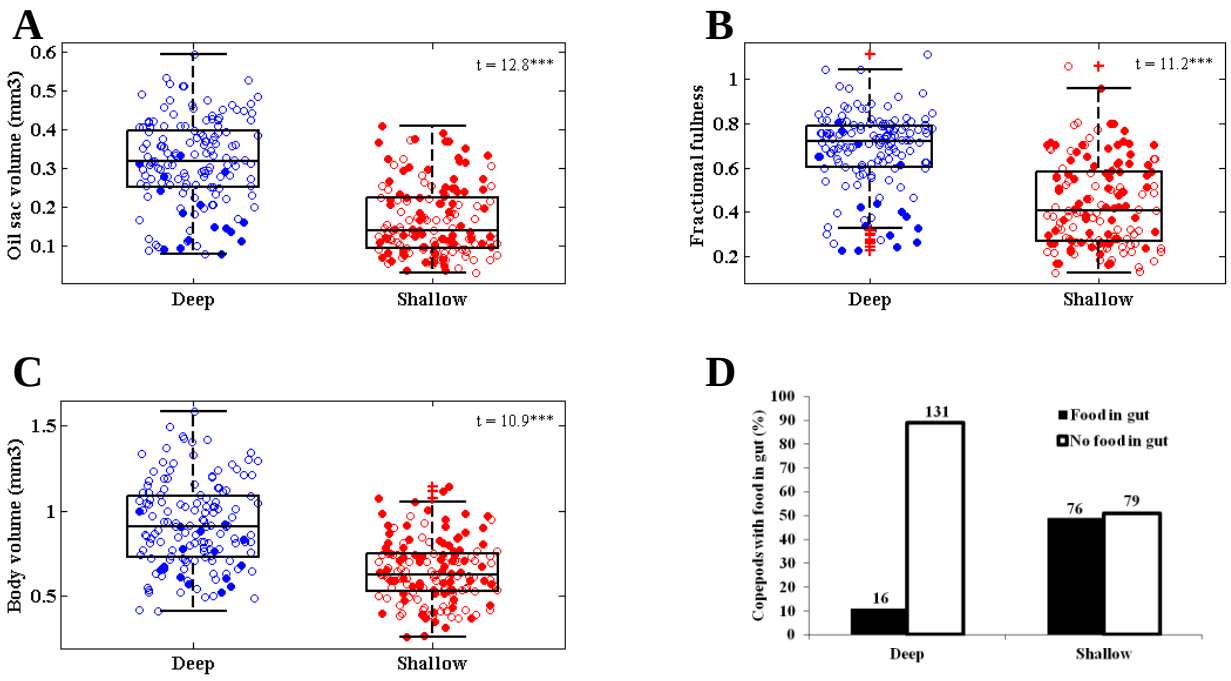


Figure 1: *C. finmarchicus* copepods demonstrate variability in physiology across and within deep and shallow populations. (A-C) The oil sac volumes (A), oil sac fractional fullness (B), and the body volumes (C) are higher in the deep population than in the shallow population. Boxplots are superimposed on jitter-plots of the morphometric values of all 16S rRNA and bacterial abundance samples from both sampling days (6/6/12, 6/11/12), with red crosses indicating outlier regions. Samples with food in guts and those without are represented by filled and empty circles, respectively. Two-sample, two-tailed t-test statistics from comparisons of deep and shallow populations for each morphometric parameter is indicated. *** indicates a p-value of significance < 0.0001. **(D)** Percentage of copepods with food in gut, with the total number of individuals in each category indicated above the bars. The higher values for all of the morphometric parameters, as well as the low percentage of individuals with food in their guts, suggest that the deep population is in diapause.

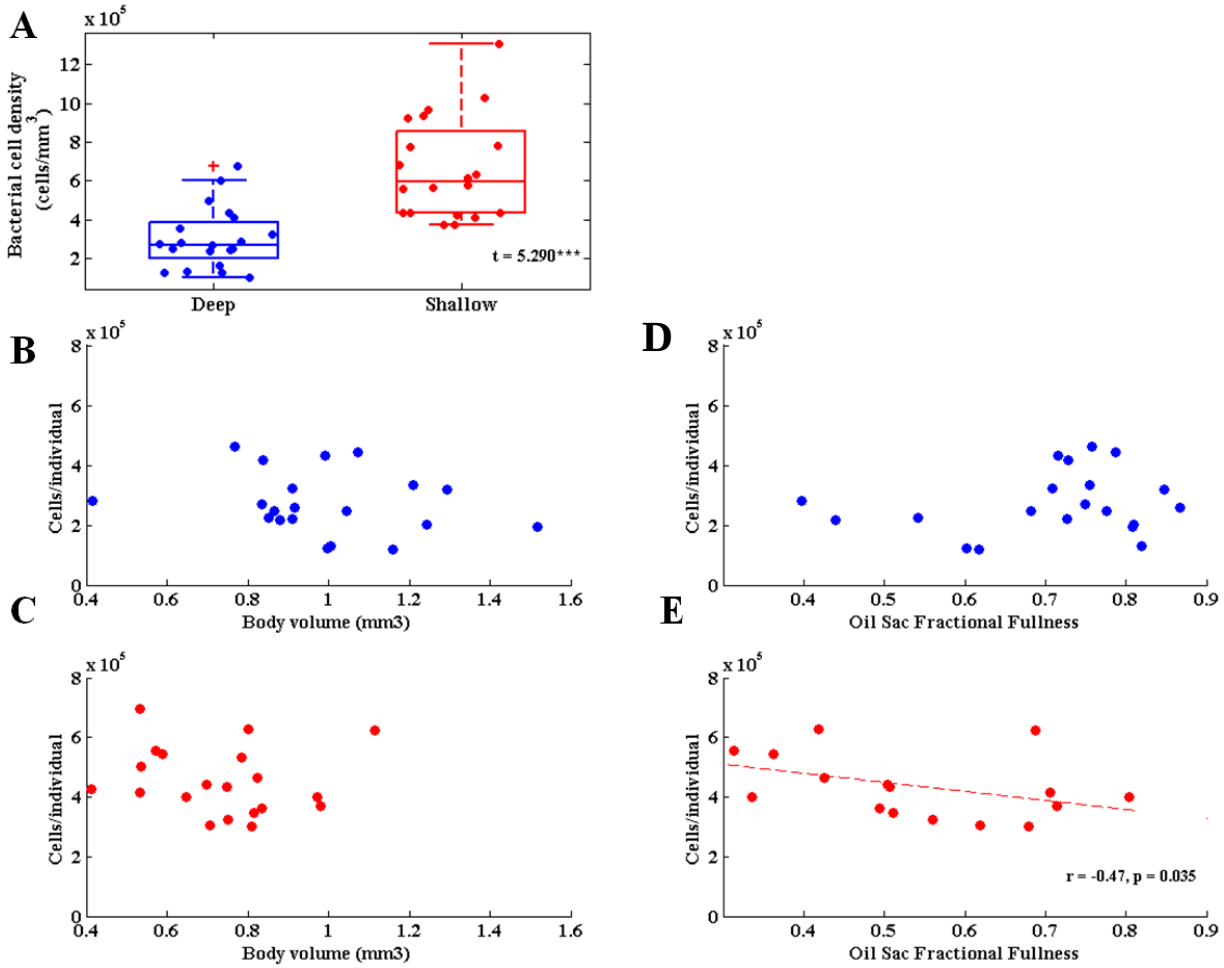


Fig. 2: *C. finmarchicus* bacterial load is significantly different between deep and shallow populations and is correlated with host physiology. (A) The mean bacterial cell density of the deep copepod population is 2-fold lower than that of the shallow copepod population ($t = 5.290$, $p < 0.0001$). Boxplots are superimposed on jitter-plots of the bacterial cell densities in each population, with red crosses indicating outlier regions. **(B-C)** Bacterial cell counts (counts/individual) are unrelated to copepod body volume in both deep (B) and shallow populations (C). **(D-E)** Bacterial cell count (counts/individual) is unrelated to copepod oil sac fractional fullness in the deep population (D) but is negatively correlated with oil sac fractional fullness in the shallow population (E). A regression line is shown only if the correlation coefficient is significantly different from 0. *** indicates a p-value of significance < 0.0001, *indicates a p-value < 0.05. Bacterial abundance counts were pooled across the two sampling days (6/6/12, 6/11/12).

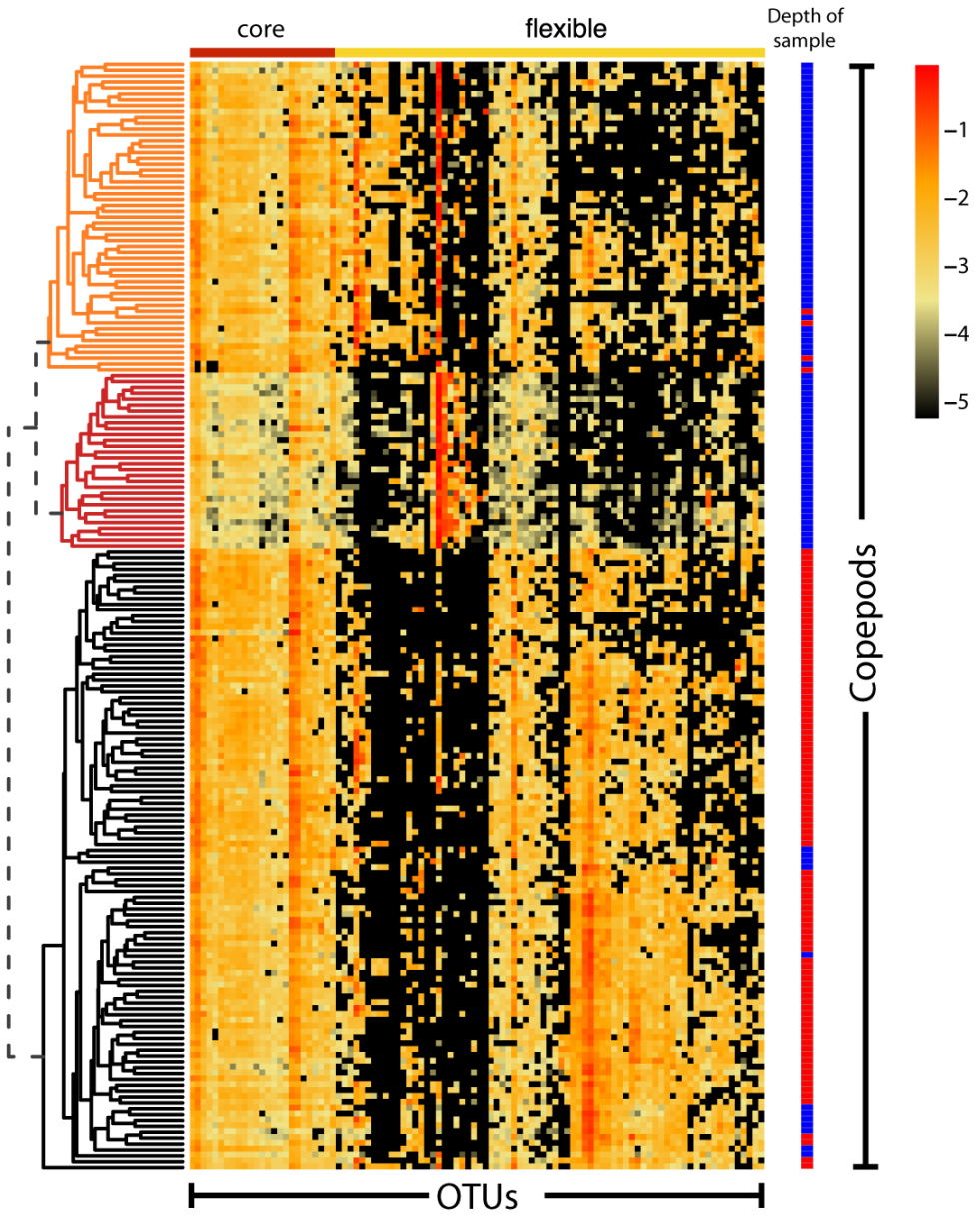


Figure 3: *C. finmarchicus* copepods have core microbial members that are persistent across deep and shallow populations and distinct flexible microbiomes between deep and shallow populations. The heat map represents the base-10 log-transformed relative abundance of the top OTUs (98 total) present at a mean relative abundance > 0.001 across all copepod individuals collected on both sampling dates (6/6/12, 6/11/12). The copepod samples are ordered on the matrix y-axis according to average-linkage clustering of the Spearman correlation values between all of the log-transformed copepod community relative abundance vectors. OTUs are ordered on the matrix x-axis by average-linkage clustering of the Spearman correlation values of the log-transformed abundances of the OTUs across all copepod individuals. The "core" microbial members are defined as those OTUs present at non-zero relative abundance in at least 95% of the copepod samples. Samples collected from shallow and deep populations are represented by blue and black bars, respectively, in the 'Depth of Sample' column.

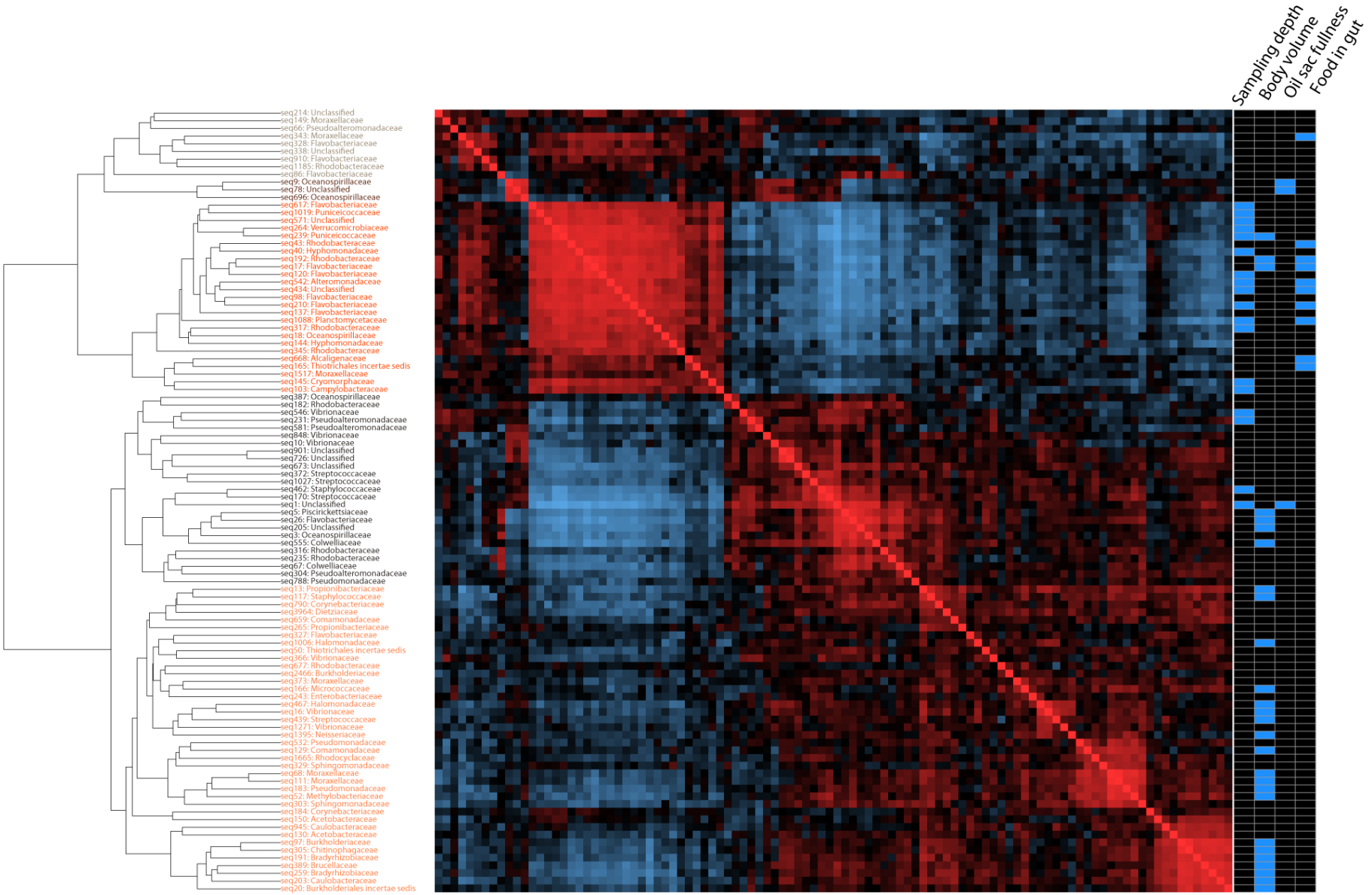

Sampling depth
Body volume
Oil sac fullness
Food in gut
seq214: Unclassified
seq149: Moraxellaceae
seq66: Pseudoalteromonadaceae
seq343: Moraxellaceae
seq328: Flavobacteriaceae
seq338: Unclassified
seq910: Flavobacteriaceae
seq1185: Rhodobacteraceae
seq86: Flavobacteriaceae
seq9: Oceanospirillaceae
seq78: Unclassified
seq696: Oceanospirillaceae
seq617: Flavobacteriaceae
seq1019: Puniceicoccaceae
seq571: Unclassified
seq264: Verrucomicrobiaceae
seq239: Puniceicoccaceae
seq43: Rhodobacteraceae
seq40: Hyphomonadaceae
seq192: Rhodobacteraceae
seq17: Flavobacteriaceae
seq120: Flavobacteriaceae
seq542: Alteromonadaceae
seq434: Unclassified
seq98: Flavobacteriaceae
seq210: Flavobacteriaceae
seq137: Flavobacteriaceae
seq1088: Planctomycetaceae
seq317: Rhodobacteraceae
seq18: Oceanospirillaceae
seq144: Hyphomonadaceae
seq345: Rhodobacteraceae
seq668: Alcaligenaceae
seq165: Thiotrichales incertae sedis
seq1517: Moraxellaceae
seq145: Cryomorphaceae
seq103: Campylobacteraceae
seq387: Oceanospirillaceae
seq182: Rhodobacteraceae
seq546: Vibrionaceae
seq231: Pseudoalteromonadaceae
seq581: Pseudoalteromonadaceae
seq848: Vibrionaceae
seq10: Vibrionaceae
seq901: Unclassified
seq726: Unclassified
seq673: Unclassified
seq372: Streptococcaceae
seq1027: Streptococcaceae
seq462: Staphylococcaceae
seq170: Streptococcaceae
seq1: Unclassified
seq5: Piscirickettsiaceae
seq26: Flavobacteriaceae
seq205: Unclassified
seq3: Oceanospirillaceae
seq555: Colwelliaceae
seq316: Rhodobacteraceae
seq235: Rhodobacteraceae
seq67: Colwelliaceae
seq304: Pseudoalteromonadaceae
seq788: Pseudomonadaceae
seq13: Propionibacteriaceae
seq117: Staphylococcaceae
seq790: Corynebacteriaceae
seq3964: Dietziaceae
seq659: Comamonadaceae
seq265: Propionibacteriaceae
seq327: Flavobacteriaceae
seq1006: Halomonadaceae
seq50: Thiotrichales incertae sedis
seq366: Vibrionaceae
seq677: Rhodobacteraceae
seq2466: Burkholderiaceae
seq373: Moraxellaceae
seq166: Micrococcaceae
seq243: Enterobacteriaceae
seq467: Halomonadaceae
seq16: Vibrionaceae
seq439: Streptococcaceae
seq1271: Vibrionaceae
seq1395: Neisseriaceae
seq532: Pseudomonadaceae
seq129: Comamonadaceae
seq1665: Rhodocyclaceae
seq329: Sphingomonadaceae
seq68: Moraxellaceae
seq111: Moraxellaceae
seq183: Pseudomonadaceae
seq52: Methylobacteriaceae
seq303: Sphingomonadaceae
seq184: Corynebacteriaceae
seq150: Acetobacteraceae
seq945: Caulobacteraceae
seq130: Acetobacteraceae
seq97: Burkholderiaceae
seq305: Chitinophagaceae
seq191: Bradyrhizobiaceae
seq389: Brucellaceae
seq259: Bradyrhizobiaceae
seq203: Caulobacteraceae
seq20: Burkholderiales incertae sedis

Figure 4: Co-occurrence analysis of copepod bacterial communities. SparCC was used to calculate correlations between the top 102 OTUs present across all copepods collected on 6/11/12. The values represented on the matrix are (1 – the absolute value of the correlation statistic) such that those OTUs that are strongly correlated have smaller values and are shown in bolder colors. OTU pairs that demonstrate positive correlations are represented in red hues while negative correlations are shown with blue hues. The OTUs are phylogenetically clustered according to generalized time reversible evolutionary model. Multivariate linear regression was performed to examine how predictive each morphometric parameter was in explaining the log-transformed relative abundance patterns of a given OTU across copepod individuals. Those morphometrics that were strongly predictive ($p < 0.001$) of OTU abundance patterns are indicated in blue.

SUPPORTING FIGURES AND TABLES:

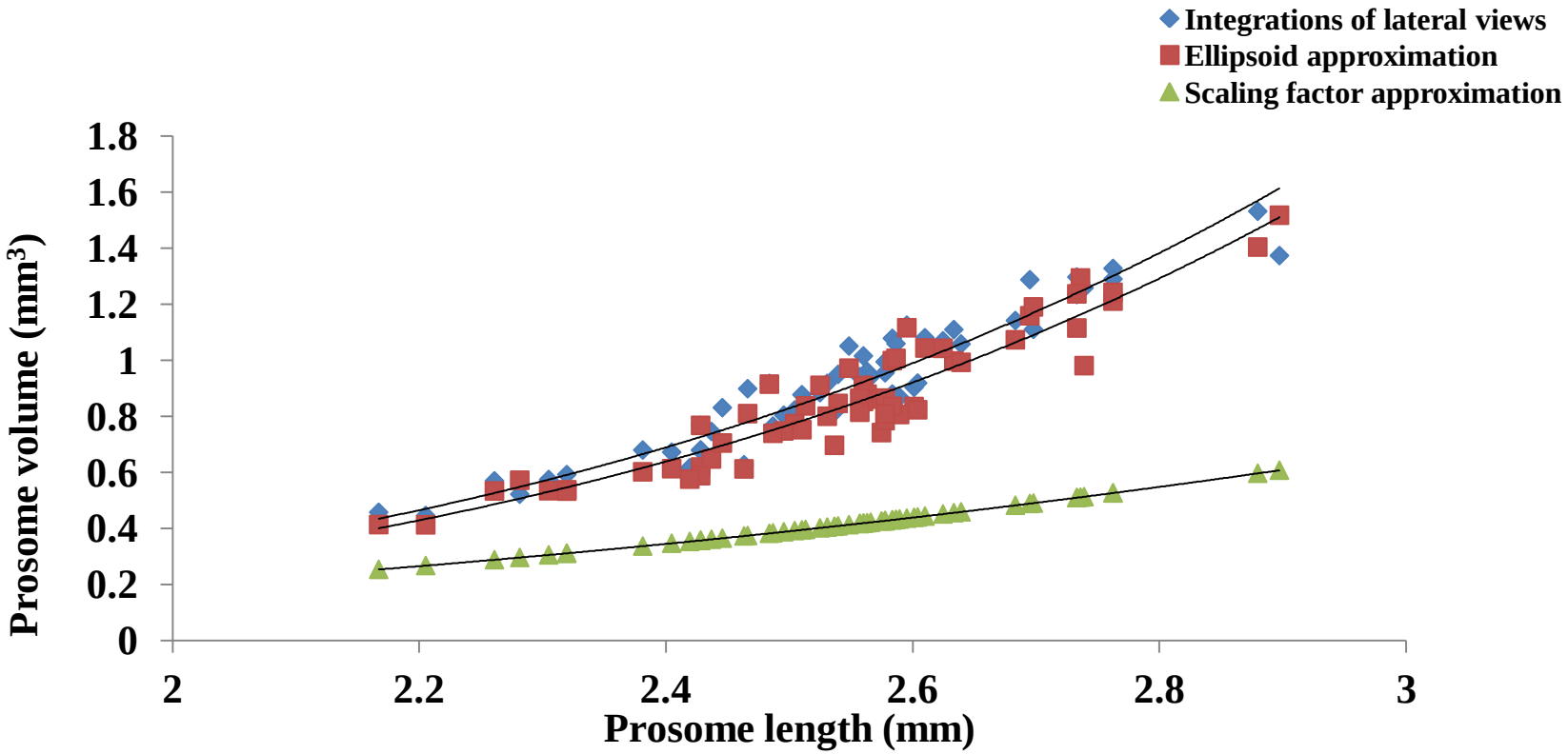


Figure S1: Comparison of standard methods for calculating copepod prosome volume. The scaling factor approximation estimates body volume as a cubic function of prosome length (volume = a*(length)^3 where 'a' is a scaling factor for the copepod *Acartia tonsa* equivalent to 2.5 x 10^{-5}). The ellipsoid approximation treats the prosome as an ellipse (volume = 1/6*(width)*(length)^2). See Miller et al., (2000) for a description of the approximation method that numerically integrates the areas of lateral slices of the prosome. The estimated body volumes are plotted as a function of prosome length to assess the accuracy of the various approximations as prosome length increases. The best-fit lines for each approximation are plotted and their equations are: Integrations of lateral views ($y = 0.0133x^{4.5113}$, $R^2 = 0.9204$), Ellipsoid approximation ($y = 0.0118x^{4.5649}$, $R^2 = 0.8972$), Scaling factor approximation ($y = 0.025x^3$, $R^2 = 1$).

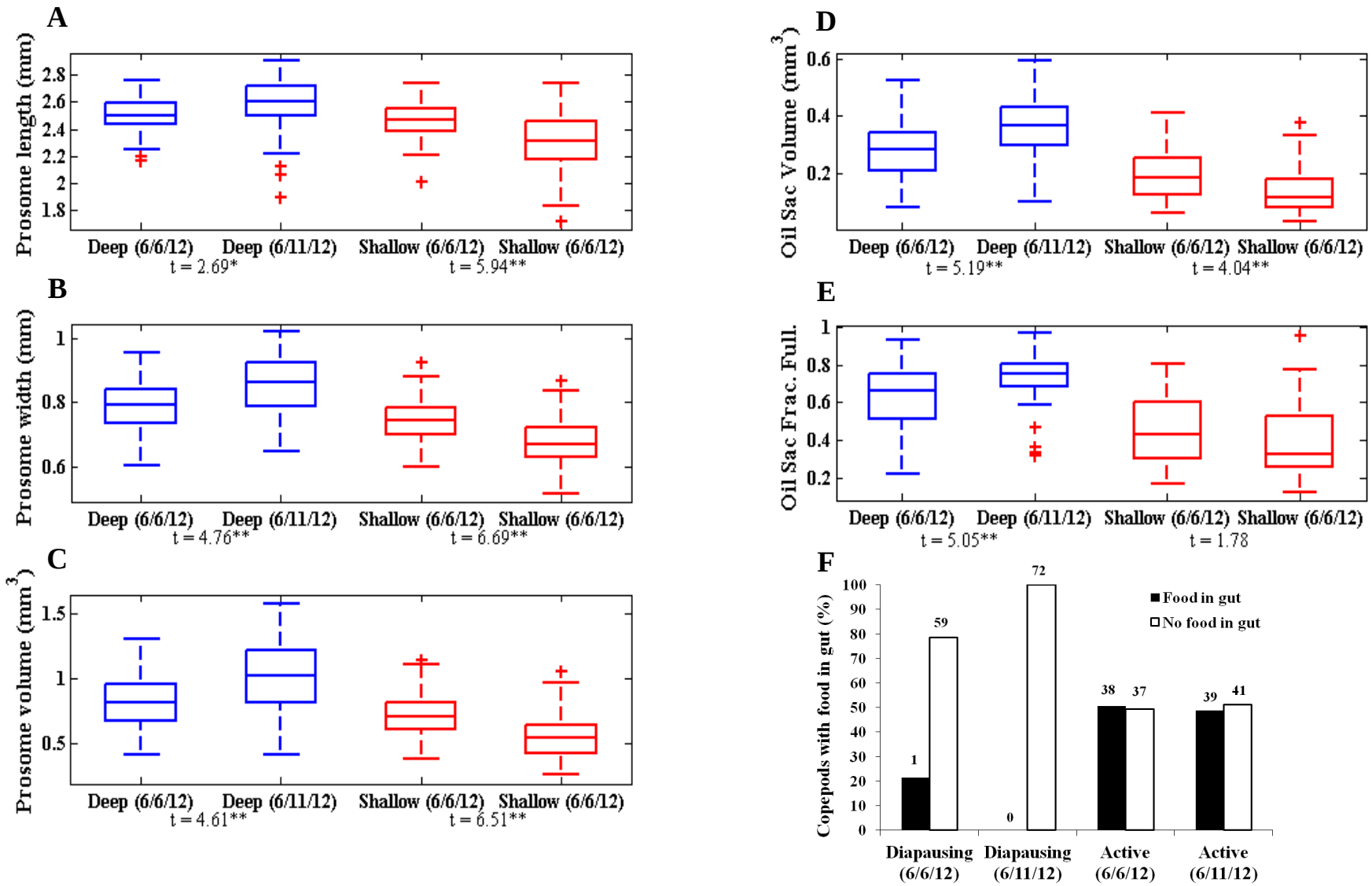


Figure S2: Comparison of the morphometric characteristics of the deep and shallow copepod populations across the two sampling days (6/6/12 and 6/11/12). Two-sample, two-tailed t-test statistics are shown for tests of significant differences between sampling dates within each population. * indicates a p-value < 0.05 ** indicates a p-value < 0.01, *** indicates a p-value < 0.0001

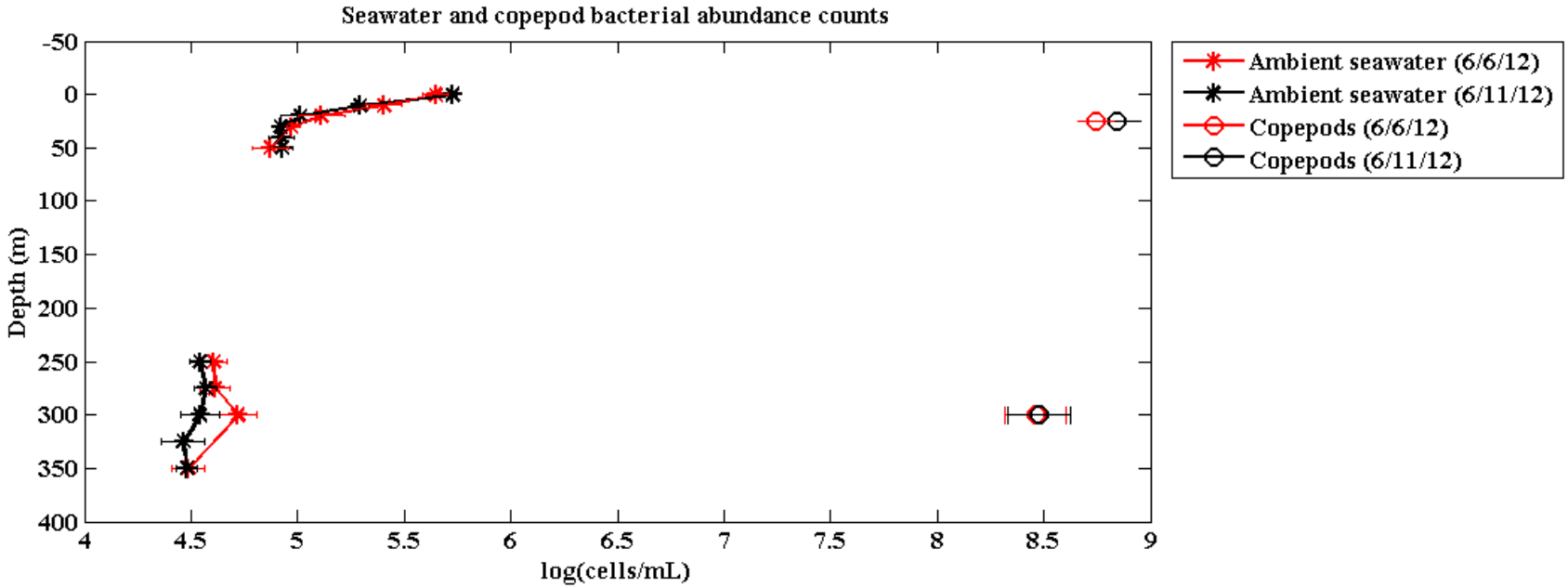


Figure S3: Bacterial cell density is enriched on copepods compared to ambient seawater. Copepod bacterial cell counts were converted to cells/mL by dividing the cell counts by body volume, which was estimated using an ellipsoid approximation. Copepod cell counts are plotted at the mid-range of their collection depth (Shallow: 0-50m; Deep: 250-350m).

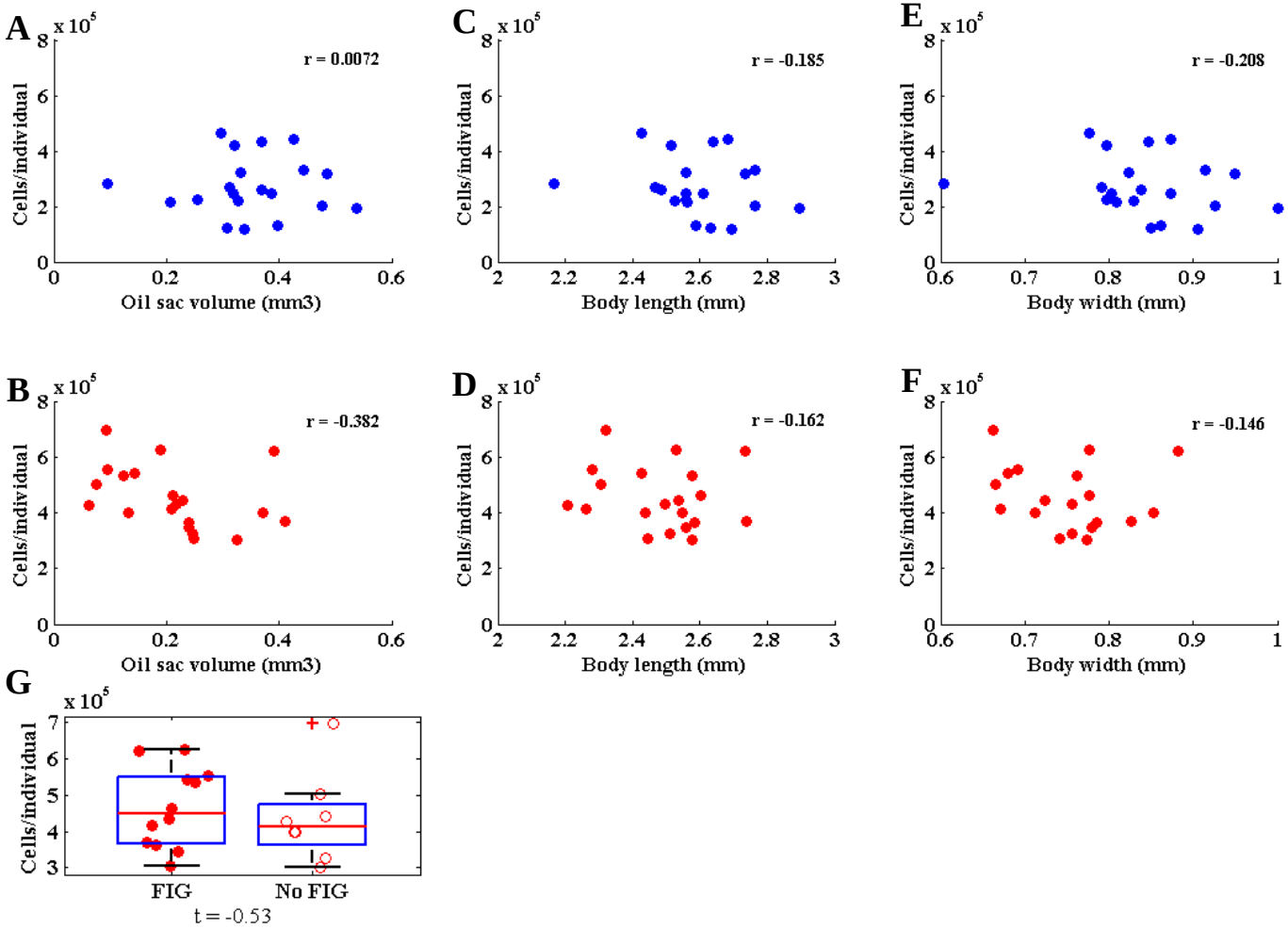


Figure S4: Bacterial load is not related to most *C. finmarchicus* morphometric variables. Bacterial cell counts (counts/individual) are unrelated to copepod oil sac volume **(A-B)**, body length **(C-D)**, and body width **(E-F)** within both deep and shallow copepod populations. All correlation coefficients are listed although none were statistically significant. Within the shallow population, a two-sample, two-tailed t-test did not demonstrate significant differences in cell abundance between those individuals with and without food in the gut. Abundance counts were pooled across the two sampling days (6/6/12, 6/11/12). Deep copepods are represented in blue and shallow copepods are represented in red.

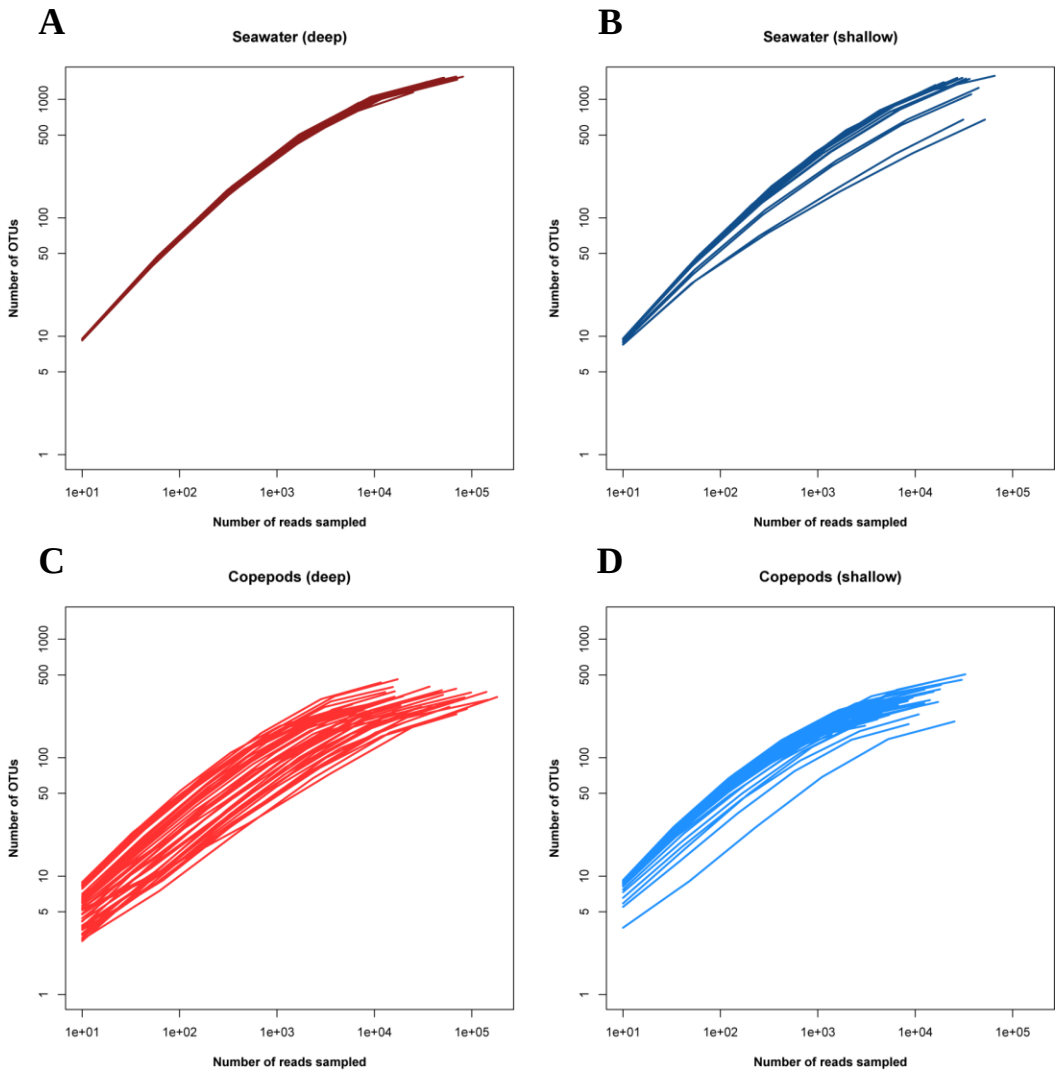


Figure S5: Rarefaction curves of observed operational taxonomic units (OTUs) in **(A-B)** seawater and **(C-D)** copepod samples determined by the R package vegan.

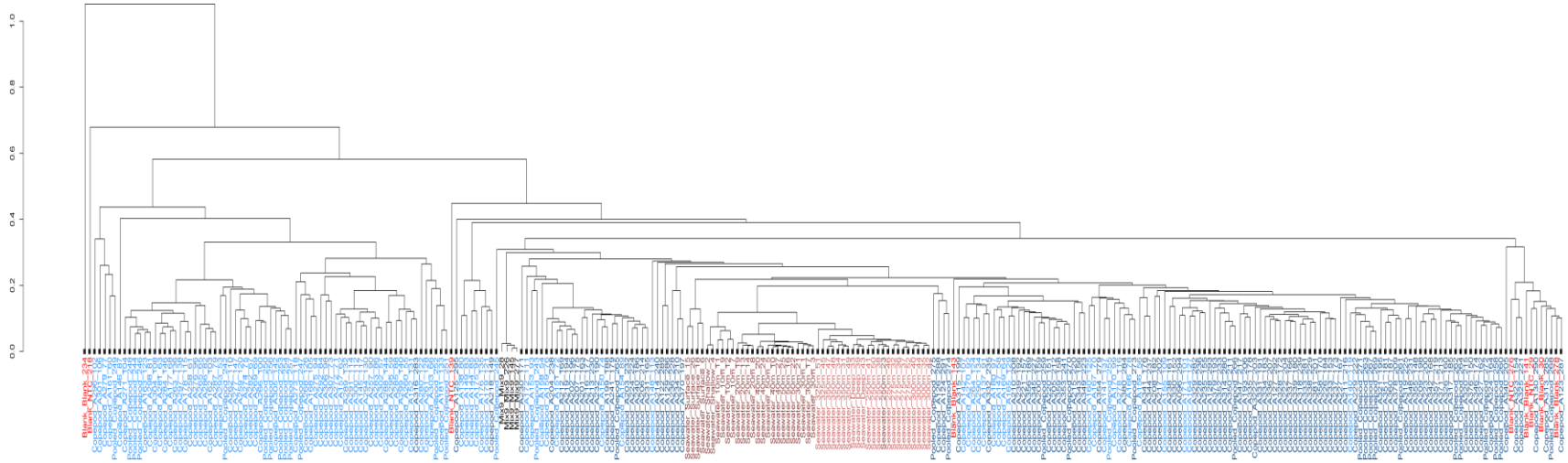


Figure S6: Hierarchical cluster analysis of the relative frequency of the most abundant OTUs across all sequenced samples using a Euclidean distance metric and average-linkage clustering. Deep copepod samples are represented in light blue, shallow copepod samples are represented in dark blue, the seawater samples are represented in maroon and light red, the Mix9 (positive control) samples are represented in black, and the blanks are represented in bright red.

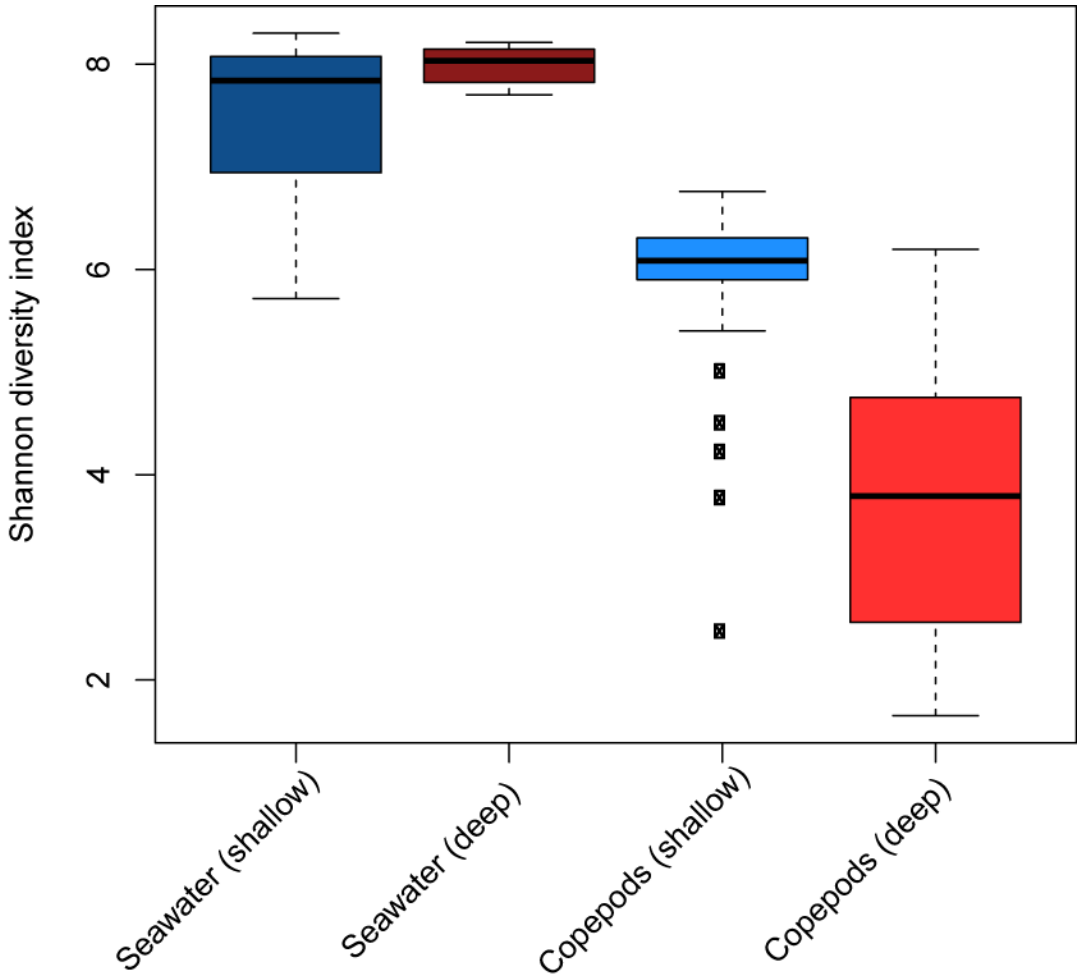


Figure S7: Boxplot of Shannon diversity index across seawater and copepod samples collected on 6/6/12 and 6/11/12.

Table S1: GPS locations, time, and sampling depth are given for all zooplankton tows and water samples across the two sampling dates.

Date of sampling	Sample type	Sampling Depth (meters)	Location of sampling	Time (GMT)	Depth of water at location (m)	Number of samples analyzed (Abundance counts)	Number of samples analyzed (16S sequencing)
6-Jun-12	Deep copepod tow	250-350	63° 28.98' N 010° 17.97' W	14:26 - 14:52	449	10	49
	CTD cast and deep water sampling	250, 275, 300, 350	63° 28.97' N 010° 18.28' W	14:58 - 15:24	390	19 (4-5 replicates/depth)	8 (2 replicates/depth)
	Shallow copepod tow	0 - 50	63° 28.96' N 010° 18.28' W	15:32 - 15:39	390	10	44
	CTD cast and shallow water sampling	0, 10, 20, 30, 40, 50	63° 28.94' N 010° 18.38' W	15:44 - 15:50	378	23 (4-5 replicates/depth; no data for 40 m)	12 (2 replicates/depth)
11-Jun-12	Deep copepod tow	250 - 350	63° 28.99' N 010° 18.00' W	07:00 - 07:30	448	10	43
	CTD cast and deep water sampling	200, 275, 300, 325, 350	63° 28.99 'N 010° 18.00' W	07:44 - 08:09	471	27 (3-5 replicates/depth)	10 (2 replicates/depth)
	Shallow copepod tow	0 - 50	63° 28.82' N 010° 17.67' W	08:14 - 08:21	471	10	54
	CTD cast and shallow water sampling	0, 10, 20, 30, 40, 50	63° 28.82' N 010° 17.68' W	08:32 - 08:36	475	24 (4-5 replicates/depth)	12 (2 replicates/depth)

Table S2: Correlation analysis of the different morphometric parameters. The correlation coefficients are listed unless otherwise noted (t-test statistics listed for food in gut comparisons). ***indicates a p-value <0.0001, ** indicates a p-value <0.001, * indicates a p-value <0.05

Across all samples

Morphometric	**Oil sac fractional fullness**	**Body volume**	**Body length**	**Body width**	**Food in gut**
Oil sac volume	0.851***	0.906***	0.726***	0.916***	t = 5.477***
Oil sac fractional fullness		0.614***	0.361***	0.658***	t = 4.757***
Body volume			0.894***	0.991***	t = 4.421***
Body length				0.872***	t = 2.170*
Body width					t = 4.493***

Within deep population

Morphometric	**Oil sac fractional fullness**	**Body volume**	**Body length**	**Body width**	**Food in gut**
Oil sac volume	0.771***	0.906***	0.906***	0.916***	t = 5.478***
Oil sac fractional fullness		0.496***	0.247**	0.559***	t = 6.303***
Body volume			0.913***	0.989***	t = 5.052***
Body length				0.876***	t = 1.882
Body width					t = 4.895***

Within shallow population

Morphometric	**Oil sac fractional fullness**	**Body volume**	**Body length**	**Body width**	**Food in gut**
Oil sac volume	0.805***	0.802***	0.802***	0.820***	t = 3.037**
Oil sac fractional fullness		0.427***	0.12	0.465***	t = 2.200*
Body volume			0.886***	0.990***	t = 2.875**
Body length				0.846***	t = 2.039*
Body width					t = 2.575**

Table S3: Evidence for diapause in deep copepod samples. Two-sample, two-tailed t-tests were conducted using reported expectations as the alternative hypothesis. Sample sizes were as follows: n = 75 for deep and shallow (6/6/12) and n = 72 for deep (6/11/12) and n = 80 for shallow (6/11/12) for the morphometric analyses. Sample sizes for the RNA:DNA were n = 5 for deep and shallow for each time point.

Indicator of diapause	**Expectation**	**Deep**	**Shallow**	***t***	**Deep (6/6/12)**	**Shallow (6/6/12)**	***t***	**Deep (6/11/12)**	**Shallow (6/11/12)**	***t***
Oil sac fractional fullness	Deep > Shallow	0.68 ± 0.030	0.44 ± 0.030	11.25* **	0.61 ± 0.041	0.46 ± 0.041	5.00* *	0.75 ± 0.033	0.41 ± 0.047	11.73* *
Oil sac volume (mm^3)	Deep > Shallow	0.32 ± 0.016	0.17 ± 0.016	12.86* **	0.27 ± 0.024	0.19 ± 0.021	4.91* *	0.36 ± 0.024	0.14 ± 0.018	14.83* *
Prosome size (mm^3)	Deep > Shallow	0.92 ± 0.035	0.64 ± 0.034	10.98* **	0.83 ± 0.047	0.73 ± 0.036	3.41*	1.01 ± 0.062	0.56 ± 0.036	12.70* *
Empty gut (%)	Deep > Shallow	89.1	50.9		78.7	50.7	N/A	100	51.3	N/A

95% confidence intervals are provided for the means

** Indicates significance of the *t*-test $p < 0.0001$

* Indicates significance of the t-test $p < 0.001$

Table S4: Abundance count results from our study are comparable with previous literature examining bacterial abundance on crustaceaous zooplankton. This is a modified table from Tang et al. (2010).

Zooplankton type	Zooplankton body volume (ml)	Bacterial abundance (cells ind. $^{-1}$)	Equivalent bacterial concentration (cells mm $^{-3}$)	Ambient bacterial concentration (cells mm^{-3})	Counting method	Source
Acartia tonsa	2.50E-05	2.00E+05	8.00E+06	1.00E+07	AODC direct counts	Hansen & Bech (1996)
Acartia tonsa	2.50E-05	2.0e3 - 4.5e5	8e4 - 1.8e7	NA	Plate counts	Tang (2005)
Artemia franciscana	3.10E-05	1.70E+04	5.60E+05	NA	Plate counts	Olsen et al. (2000)
Daphnia cucullata	1.20E-04	1.7e5-4.3e5	2.0e6 - 5.1e6	3.40E+06	SYBR Gold direct counts	Tang et al. (2009)
Calanus helgolandicus and Calanus finmarchicus	5.50E-04	1.90E+05	3.50E+05	1.4e5 - 5e5	DAPI direct counts (copepods); Flow cytometry (water sample)	Moller et al. (2007)
Calanus finmarchicus - Shallow	4.1e-4 ± 2.2e-5 (7.2e-4 ± 5.8e-5)	4.5e5 ± 5.0e4	1.2e6 ± 1.8e5 (6.6e5 ± 1.1e5)	9.1e4-6.1e5	SYBR Gold direct counts	This study
Calanus finmarchicus - Deep	4.4e-4 ± 2.3e-5 (9.7e-4 ± 7.9e-5)	2.8e5 ± 3.3e4	6.7e5 ± 1.3e5 (3.1e5 ± 7.2e4)	3.1-6.0e4	SYBR Gold direct counts	This study

Values calculated from ellipsoid approximation of copepod body volume in parentheses. All other calculations based on a scaling factor equation. Zooplankton body volume and equivalent bacterial concentration values from this study are given as an average ± 95% confidence interval values.

Table S5: Morphometrics of the distinct sub-group within the deep population as identified by Figure 3. The average ± 95% CI values for each morphometric variable as well as the t-test statistic derived from comparison of the distinct sub-group and the remainder of the deep population are listed. *** indicates a p-value < 0.0001, **p < 0.001. Sample sizes: n = 27 individuals in the distinct subgroup, n = 62 for the remainder of the deep population.

Morphometric	**Distinct deep subgroup**	**Remainder**	***t***
Prosome length	2.66 ± 0.056	2.51 ± 0.033	4.64***
Prosome width	0.89 ± 0.30	0.80 ± 0.20	5.17***
Prosome volume	1.13 ± 0.91	0.86 ± 0.053	5.39***
Oil sac volume	0.41 ± 0.37	0.29 ± 0.025	5.08***
Oil sac fractional fullness	0.76 ± 0.043	0.64 ± 0.043	3.27**

Table S6: Taxonomic classification of the *C. finmarchicus* "core" microbiome members identified in this study. The "core" microbiome was defined as those OTUs present at non-zero relative abundance in at least 95% of the copepod samples collected on 6/6/12 and 6/11/12 (25 total). The top four operational taxonomic units (OTUs) are shown in bold.

OTU	Phylum	Class	Order	Family	Genus
seq16	***Proteobacteria***	***Gammaproteobacteria***	***Vibrionales***	***Vibrionaceae***	***Vibrio***
seq13	***Actinobacteria***	***Actinobacteria***	***Actinomycetales***	***Propionibacteriaceae***	***Propionibacterium***
seq68	***Proteobacteria***	***Gammaproteobacteria***	***Pseudomonadales***	***Moraxellaceae***	***Acinetobacter***
seq111	***Proteobacteria***	***Gammaproteobacteria***	***Pseudomonadales***	***Moraxellaceae***	***Acinetobacter***
seq130	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	
seq117	*Firmicutes*	*Bacilli*	*Bacillales*	*Staphylococcaceae*	*Staphylococcus*
seq303	*Proteobacteria*	*Alphaproteobacteria*	*Sphingomonadales*	*Sphingomonadaceae*	*Sphingobium*
seq305	*Bacteroidetes*	*Sphingobacteria*	*Sphingobacteriales*	*Chitinophagaceae*	*Sediminibacterium*
seq170	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
seq790	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Corynebacteriaceae*	*Corynebacterium*
seq150	*Proteobacteria*	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	
seq265	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Propionibacteriaceae*	*Propionibacterium*
seq389	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Brucellaceae*	*Brucella*
seq97	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Burkholderiaceae*	*Burkholderia*
seq203	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Phenylobacterium*
seq52	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Methylobacteriaceae*	*Methylobacterium*
seq439	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
seq183	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
seq659	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Pelomonas*
seq129	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Comamonadaceae*	*Delftia*
seq945	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Caulobacteraceae*	*Phenylobacterium*
seq166	*Actinobacteria*	*Actinobacteria*	*Actinomycetales*	*Micrococcaceae*	*Micrococcus*
seq259	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Bradyrhizobiaceae*	*Bradyrhizobium*
seq467	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Halomonadaceae*	*Halomonas*
seq191	*Proteobacteria*	*Alphaproteobacteria*	*Rhizobiales*	*Bradyrhizobiaceae*	*Bradyrhizobium*

Table S7: Taxonomic classification of the "flexible" microbiome members in the deep (diapausing) *C. finmarchicus* population identified in this study. The "flexible" deep microbiome was defined as those remaining members of the top 98 OTUs that were not in the core microbiome and that had higher relative abundances in the deep population.

OTU	Phylum	Class	Order	Family	Genus
seq696	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Oceanospirillaceae*	
seq1149	*Proteobacteria*	*Gammaproteobacteria*			
seq901	*Proteobacteria*	*Gammaproteobacteria*			
seq726	*Unknown bacteria*				
seq673	*Proteobacteria*				
seq506	*Proteobacteria*				
seq231	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Pseudoalteromonadaceae*	*Pseudoalteromonas*
seq581	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Pseudoalteromonadaceae*	*Pseudoalteromonas*
seq366	*Proteobacteria*	*Gammaproteobacteria*	*Vibrionales*	*Vibrionaceae*	*Aliivibrio*
seq546	*Proteobacteria*	*Gammaproteobacteria*	*Vibrionales*	*Vibrionaceae*	*Aliivibrio*
seq848	*Proteobacteria*	*Gammaproteobacteria*	*Vibrionales*	*Vibrionaceae*	*Vibrio*
seq205	*Proteobacteria*	*Gammaproteobacteria*			
seq1	*Proteobacteria*	*Gammaproteobacteria*			
seq26	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Tenacibaculum*
seq5	*Proteobacteria*	*Gammaproteobacteria*	*Thiotrichales*	*Piscirickettsiaceae*	*Methylophaga*
seq3	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Oceanospirillaceae*	
seq555	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Colwelliaceae*	*Colwellia*
seq67	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Colwelliaceae*	*Colwellia*
seq235	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	
seq316	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	
seq304	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Pseudoalteromonadaceae*	*Pseudoalteromonas*
seq462	*Firmicutes*	*Bacilli*	*Bacillales*	*Staphylococcaceae*	*Gemella*
seq372	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
seq1027	*Firmicutes*	*Bacilli*	*Lactobacillales*	*Streptococcaceae*	*Streptococcus*
seq9	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Oceanospirillaceae*	
seq78	*Proteobacteria*	*Gammaproteobacteria*			

Table S8: Taxonomic classification of the "flexible" microbiome members in the shallow (active) *C. finmarchicus* population identified in this study. The "flexible" shallow microbiome was defined as those remaining members of the top 98 OTUs that were not in the core microbiome and that had higher relative abundances in the shallow population.

OTU	Phylum	Class	Order	Family	Genus
seq243	*Proteobacteria*	*Gammaproteobacteria*	*Enterobacteriales*	*Enterobacteriaceae*	*Escherichia/Shigella*
seq788	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Pseudomonadaceae*	*Pseudomonas*
seq165	*Proteobacteria*	*Gammaproteobacteria*	*Thiotrichales*	*Thiotrichales_incertae_sedis*	*Fangia*
seq668	*Proteobacteria*	*Betaproteobacteria*	*Burkholderiales*	*Alcaligenaceae*	*Tetrathiobacter*
seq40	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Hyphomonadaceae*	*Hyphomonas*
seq120	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Lacinutrix*
seq192	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Marivita*
seq17	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Maribacter*
seq43	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	
seq210	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	
seq98	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	
seq571	*Proteobacteria*	*Gammaproteobacteria*			
seq144	*Proteobacteria*	*Alphaproteobacteria*	*Caulobacterales*	*Hyphomonadaceae*	*Hyphomonas*
seq317	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Roseibium*
seq239	*Verrucomicrobia*	*Opitutae*	*Puniceicoccales*	*Puniceicoccaceae*	*Coraliomargarita*
seq264	*Verrucomicrobia*	*Verrucomicrobiae*	*Verrucomicrobiales*	*Verrucomicrobiaceae*	
seq1019	*Verrucomicrobia*	*Opitutae*	*Puniceicoccales*	*Puniceicoccaceae*	*Coraliomargarita*
seq434	*Unknown bacteria*				
seq542	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Alteromonadaceae*	*Haliea*
seq1088	*Planctomycetes*	*Planctomycetacia*	*Planctomycetales*	*Planctomycetaceae*	
seq18	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Oceanospirillaceae*	
seq137	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Tenacibaculum*
seq214	*Proteobacteria*	*Gammaproteobacteria*			
seq345	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Loktanella*
seq864	*Unknown bacteria*				
seq1185	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Sulfitobacter*
seq338	*Unknown bacteria*				
seq86	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Formosa*
seq328	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Bizionia*
seq343	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Psychrobacter*
seq910	*Bacteroidetes*	*Flavobacteria*	*Flavobacteriales*	*Flavobacteriaceae*	*Krokinobacter*
seq149	*Proteobacteria*	*Gammaproteobacteria*	*Pseudomonadales*	*Moraxellaceae*	*Psychrobacter*
seq677	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Paracoccus*
seq66	*Proteobacteria*	*Gammaproteobacteria*	*Alteromonadales*	*Pseudoalteromonadaceae*	*Pseudoalteromonas*
seq188	*Bacteroidetes*	*Sphingobacteria*	*Sphingobacteriales*	*Saprospiraceae*	*Aureispira*
seq182	*Proteobacteria*	*Alphaproteobacteria*	*Rhodobacterales*	*Rhodobacteraceae*	*Loktanella*
seq387	*Proteobacteria*	*Gammaproteobacteria*	*Oceanospirillales*	*Oceanospirillaceae*	

Chapter Five

Conclusions and Future Work

CONCLUSIONS:

Microbial ecologists seek to understand how microorganisms interact with one another and their environment. In marine ecosystems, microbial ecology has traditionally focused on those bacteria that are "free-living" in the environment, but it is increasingly apparent that particles and aggregates, including living organisms, are hotspots of microbial activity. Such "particle-associated" lifestyles have profound influences on microbial ecology and evolution, as well as on the fitness of the living "particles" themselves. Zooplankton, such as copepods, are highly abundant environmental reservoirs of diverse microbial communities. Association with copepods is known to have dramatic impacts on the persistence and physiology of many *Vibrio* species, yet whether copepods respond to or even shape these interactions is unknown. This thesis work investigated whether copepods are dynamic vectors of bacterial communities by exploring whether copepods elicit molecular responses to colonizing *Vibrio* bacteria and whether copepod host physiology influences the abundance and composition of their bacterial communities.

While this thesis focused primarily on copepod-microbial interactions, there are also many knowledge gaps in how copepods themselves detect and respond to their physical environment. Copepods are an integral part of marine food webs due to their role in transferring energy from phytoplankton to higher trophic levels (e.g. commercially important fish, sea birds). The copepod *Calanus finmarchicus* is a particularly well-studied species in light of its predominance in the temperate North Atlantic, which is facilitated by its ability to avoid adverse seasonal conditions and high predation risk by vertically migrating to depth and entering a facultative diapause during the last juvenile stages (typically stage C5). Despite the ecological implications of seasonal dormancy in *C. finmarchicus*, very little is understood about the factors that regulate this diapause response. To this end, we characterized the expression patterns of several heat shock proteins (Hsps), a highly conserved superfamily of molecular chaperones, known to be important regulators of the stress response and diapause in other invertebrates. Unlike the classic stress response which is characterized by transient and universal up-regulation of a wide range of *Hsp*s, *Hsp* expression patterns during diapause may be prolonged and highly variable among species and *Hsp* types (1). We found that several Hsps (i.e., Hsp21, Hsp22, and Hsp70) induced in response to stressors such as high temperatures and exposure to toxins in *C. finmarchicus,* were also

induced by handling stress in our study (2-5). We identified a small *Hsp (Hsp22)* that is also up-regulated during diapause, suggesting that it may play a role both in short-term stress responses and in protecting proteins from degradation during diapause. Our findings also suggest that the inducibility of invertebrate Hsps during diapause is dependent on the cellular localization (e.g., cytosolic vs. ER subfamilies) and on the particular isoforms examined. Overall, it seems that Hsps play different roles in mediating diapause in different invertebrate species, which may be attributed to the intensity of the environmental conditions experienced during the diapause period. We noted limited changes in *Hsp* expression during diapause in *C. finmarchicus*, suggesting that the environmental conditions experienced during this diapause period may be relatively minor (i.e., small variation in temperature, not subject to freezing or desiccation) compared to those experienced by overwintering insects or *Artemia* cysts which strongly regulate Hsp expression (Chapter 2).

A growing number of studies have highlighted the complexity of the innate immune response and the potential for invertebrate hosts to specifically regulate their associated microbiota (6-10). Therefore, we characterized the transcriptional response of the copepod *Eurytemora affinis* to distinct *Vibrio* species in order to investigate whether copepods can actively respond to different colonizing vibrios (Chapter 3). Transcriptional profiling (RNA-Seq) and quantitative PCR were used to examine the transcriptomic response of adult female *E. affinis* exposed to a putative copepod symbiont (*Vibrio sp. F10 9ZB36*) or a free-living *Vibrio* (*V. ordalii 12B09*) for 24 hours. Although we originally intended to perform these *Vibrio* exposure experiments with *V. sp. F10* and a *Vibrio* species pathogenic to *E. affinis,* we found that *E. affinis* was highly resilient to the many *Vibrio* species tested (8 overall). *V. ordalii 12B09* demonstrated some weak pathogenicity towards *E. affinis* at high inoculation densities and incubation times and is also known to be classically free-living in the environment, so we felt this would be an appropriate *Vibrio* strain to use as a contrast to *V. sp. F10* in our exposure assays. Our findings provide evidence that the copepod *E. affinis* does distinctly recognize and respond to colonizing vibrios via transcriptional regulation of innate immune response elements and transcripts involved in maintaining cuticle integrity. The expression of six genes that were further profiled by qPCR were consistently up-regulated upon *V. sp. F10* association and not differentially expressed upon *V. ordalii* exposure, suggesting that the expression of these genes (3 C-type lectin-like, 2 chitin-binding, 1 saposin-like transcript) may represent a mechanism by which *E.*

affinis recognizes and maintains symbiotic *Vibrio* bacteria. The two *Vibrio* species also demonstrated differences in their colonization densities of *E. affinis*, which suggests that *Vibrio* species may exhibit distinct affinities for copepods. Interestingly, the culturability of *V. sp. F10* was diminished after colonization of copepods, which is reminiscent of previous observations of viable but non-culturable (VBNC) *Vibrio* species associated with zooplankton (11-14). Together these findings suggest that rather than solely being passive vectors, copepods may actively regulate their *Vibrio* colonizers to ultimately influence the composition and activity of their associated *Vibrio* communities.

The ecological factors that control the assembly of microbial communities on zooplankton are not well understood. In Chapter 4, the inter-individual variability of the *C. finmarchicus* microbiome was examined to identify whether host physiology influences the bacterial community structure and how specifically and predictably bacterial communities assemble on copepods. The dramatic physiological changes coupled to the vertical migration in the water column associated with diapause in *C. finmarchicus* provides a unique system to study how a sustained physiological change of a copepod host may influence its bacterial associates. By comparing the abundance and composition of the microbial communities on individual active and diapausing *Calanus finmarchicus*, our findings suggest that these copepods have a predictable "core microbiome" that persists throughout the host's entry into diapause. Furthermore, the differences in the structure of the 'flexible microbiome' between diapausing and active individuals may be driven by host factors including feeding history and body size as well as microbial interactions. These findings suggest that the bacterial communities associated with copepods may be rather consistent and stable, with some influence of host physiology on the variability in community composition between active and diapausing copepod populations.

Overall, this thesis work highlights the role of copepods as dynamic hosts of diverse microbial communities and implicates copepod host physiology and its innate immune response as important contributors to the prevalence and activity of its associated microbiota.

FUTURE DIRECTIONS:

Immune response of diapausing Calanus finmarchicus

In many invertebrates, entry into diapause is characterized by enhanced resistance to environmental stressors (15). There is growing evidence that elements of the innate immune response are also induced by diapause, particularly those elements involved in anti-microbial responses. For example, the transcription of hemolin, a bacterial-induced cell-adhesion protein, is enhanced in the gut during the embryonic diapause of the gypsy moth *Lymantria dispar* (16). In addition, diapausing larvae of the blowfly *Calliphora vicina* have higher antimicrobial activity in the hemolymph in response to septic injury than do non-diapausing bacteria-challenged larvae (17). However, it seems that there is not just an overall change in the intrinsic activity and inducibility of the innate immune response of diapausing insects, but particularly a shift in the functioning elements of the immune response. Chernysh *et al.* (2000) noted that the relative composition of antimicrobial peptides present in diapausing blowfly larvae was markedly different than that of non-diapausing individuals, in terms of which peptides were induced and which peptides were no longer detectable. Although many diapausing insects do seem to have enhanced immune potential and inducibility by stressors, the influence of diapause on the immune response of other arthropods is not well understood.

Diapausing *C. finmarchicus* exhibit gene expression profiles distinct from those of active copepods, including up-regulation of stress-related genes including ferritin (18) and heat shock proteins (*Hsp22*) (19). In light of previously observed differences in the infectivity of active and diapausing freshwater copepodids (20), it is possible that differential expression of immune responsive genes is also a characteristic that distinguishes active and diapausing *C. finmarchicus.* A search for homologs of the *E. affinis Vibrio*-responsive genes (Chapter 3) in the *C. finmarchicus* transcriptome followed by qPCR experiments in active and diapausing *C. finmarchicus* individuals would be an ideal way to begin to examine the question of whether diapausing *C. finmarchicus* have an enhanced anti-microbial response. Further field experiments could perform *Vibrio* colonization experiments of active and diapausing *C. finmarchicus* to compare the extent of bacterial colonization and the transcriptomic response of the active and diapausing animals.

One of our hypotheses for why the bacterial load of active *C. finmarchicus* individuals was negatively correlated to oil sac fractional fullness (Chapter 4) is that an up-regulation of the immune response as the copepods approach diapause caused a decrease in the abundance of the associated bacterial communities. To test this, the basal expression of immune elements in laboratory cultures of *C. finmarchicus* could be monitored as the individuals progress through the fifth copepodite stage. An alternative hypothesis for the results we observed in Chapter 4 is that the decrease in bacterial load is a by-product of bacterial community succession on the surface of the copepod. One way to test this would be to dissect the jaws of active *C. finmarchicus* individuals in order to estimate their age since the last molt (21) before performing bacterial abundance counts. This would allow us to determine if older animals do indeed have lower bacterial loads.

Characterization of the copepod-Vibrio molecular dialogue

Our study is the first to characterize the copepod response to *Vibrio* colonists, with the ultimate determination that copepods are not merely passive vectors of *Vibrio*. We hypothesize that our observation of up-regulation of a targeted set of genes known to be involved in invertebrate maintenance of bacterial symbionts reflects the *E. affinis* response to maintain a commensal relationship with the putative copepod symbiont *V. sp. F10*. Further work is needed to confirm this hypothesis, which could be tested by methods such as gene knockdown of the targets identified in Chapter 3 followed by colonization experiments with *V. sp. F10*. If with gene knockdown the localization of *V. sp. F10* on *E. affinis* dramatically shifts, the colonization densities of *V. sp. F10* increase, and/or the health of *E. affinis* is compromised, as measured by molecular or physiological markers, this would suggest that the up-regulated candidate genes we identified are in fact keeping the *V. sp. F10* symbiont in check. Further study is also needed to identify the potential role that *V. sp. F10* may play in its association with *E. affinis*. If the copepod is also benefiting from an association with *V. sp. F10*, this could be potentially detected by monitoring the survival rates and fecundity of *V. sp. F10*-colonized individuals. *V. sp. F10* could also benefit the copepod host through production of allelopathic chemicals that prevent colonization by other pathogenic bacteria or fungi, which could be tested by competitive colonization assays with *V. sp. F10* and other additional microbial colonists.

Although we were unable to identify a *Vibrio* strain highly pathogenic to *E. affinis* in our experiments, it would warrant further study to test the pathogenicity of other known crustacean bacterial pathogens that persist at lower temperatures (18 °C) and salinities (15 PSU) as many of the bacterial pathogens we tested were originally isolated from tropical, marine systems. Our method of infection, specifically immersion, may also have contributed to the low mortality rates of the *E. affinis* exposed to the various vibrios. Injecting the copepods with the *Vibrio* pathogens or encouraging ingestion of the bacteria by allowing the vibrios to colonize their algal food sources may also be important alterations to the exposure methodology that could enhance the pathogenic potential of the *Vibrio* species. Furthermore, fungi are also known to be important pathogens of many crustaceans and could serve as an appropriate substitute for bacterial pathogens if a copepod bacterial pathogen is not easily found.

In addition, it would be interesting to extend the RNA-Seq experiments to explore the *E. affinis* transcriptomic response to other ecologically relevant *Vibrio* species (e.g., *Vibrio cholerae*, *V. parahaemolyticus*) to see whether the expression of these identified *Vibrio*-response elements are well-conserved in the *E. affinis* response to other vibrios known to naturally associate with copepods. Furthermore, performing these same exposure experiments with *Vibrio* or other bacteria highly pathogenic to *E. affinis*, could help elucidate the range of transcriptional elements *E. affinis* utilizes to respond to harmful versus commensal bacterial associates.

To gain a more comprehensive picture of the molecular elements utilized by both copepod and vibrios to interact with one another, future colonization studies could utilize transposon mutant libraries of the *V. sp. F10* and strains to identify the genes required by *Vibrio* to colonize their copepod hosts. Transposon-insertion mutagenesis of *Vibrio* strains would provide libraries of randomly-generated mutants that will be evaluated for their ability to colonize copepods. By sequencing the entire library of mutants both before and after colonization of the copepods, it would be possible to determine which mutants were unable to compete with other members of the population based on a decrease in their relative frequency. Those mutants unable to compete will have lost key genetic elements during the insertion mutagenesis that prevents them from successfully colonizing their copepod host.

Implications of copepod host association on Vibrio persistence

Association with copepods has been demonstrated to have profound implications on the proliferation, virulence, and physiology of many *Vibrio* species (14, 22-25). Given our observation of a change in culturability of *V. sp. F10* upon association with the copepod *E. affinis*, further work should be done to identify whether this shift in culturability is also observed when other ecologically relevant *Vibrio* species (e.g., *Vibrio cholerae*, *V. parahaemolyticus*) colonize *E. affinis* for 24 hours or longer. Further comparisons of the colonization densities of these different *Vibrio* species on *E. affinis* could provide further insight into how association with copepods may help to regulate the abundance and activity of *Vibrio* species in the natural environment. Colonization experiments with different *Vibrio* strains on live and dead copepods would also help determine whether it is the chitinous structure of the copepods or selection by host physiology that influences the abundance and culturability patterns of the host-associated vibrios.

Conservation and function of the copepod microbiome

Host-associated microbial communities are known to vary across individuals, space and time (26-31). Our work demonstrates that *C. finmarchicus* individuals from Trondheimfjord can exhibit a 'core' microbiome that is consistent across individuals in both active and early diapausing populations. However, further work is needed to identify whether the core microbiome we have identified is conserved at other temporal (e.g., beginning and end of diapause, across years) and spatial (e.g., North Atlantic basin) scales. Examination of the microbiomes of individuals from other diapausing copepod species in similar environments (e.g., *Calanus glacialis*) would also uncover whether *C. finmarchicus* exhibits a species-specific microbiome as has been observed in other invertebrates (10) or if many copepod species are colonized by similar bacteria.

Our study may also be missing important host-microbial interactions by not distinguishing between surface and gut-associated microbial members of *C. finmarchicus*. Future studies could explore the localization of the microbial community as the diapause period progresses to determine whether it is primarily the surface or gut-associated community that

shifts between the active and diapausing copepod populations. Further examination of the functional changes in the flexible microbiome during the transition to the diapausing period through metatranscriptomics would provide further information into how the associated microbiota are shifting with changes in the host environment.

REFERENCES:

1. **Denlinger DL, Rinehart JP, Yocum GD, Denlinger DL, Giebultowicz JM, Saunders DS.** 2001. Stress proteins: A role in insect diapause?, p. 155-171, Insect Timing: Circadian Rhythmicity to Seasonality. Elsevier Science B.V., Amsterdam.
2. **Qiu Z, Macrae TH.** 2008. ArHsp21, a developmentally regulated small heat-shock protein synthesized in diapausing embryos of *Artemia franciscana*. Biochem. J. **411:**605-611.
3. **Hansen BH, Altin D, Vang SH, Nordtug T, Olsen AJ.** 2008. Effects of naphthalene on gene transcription in *Calanus finmarchicus* (Crustacea: Copepoda). Aquat. Toxicol. **86:**157-165.
4. **Qiu Z, MacRae TH.** 2008. ArHsp22, a developmentally regulated small heat shock protein produced in diapause-destined Artemia embryos, is stress inducible in adults. Febs. J. **275:**3556-3566.
5. **Voznesensky M, Lenz PH, Spanings-Pierrot Cl, Towle DW.** 2004. Genomic approaches to detecting thermal stress in *Calanus finmarchicus* (Copepoda: Calanoida). J. Exp. Mar. Biol. Ecol. **311:**37-46.
6. **Douglas A.** 2011. Lessons from studying insect symbioses. Cell Host Microbe **10:**359-367.
7. **Fraune S, Augustin R, Anton-Erxleben F, Wittlieb J, Gelhaus C, Klimovich VB, Samoilovich MP, Bosch TCG.** 2010. In an early branching metazoan, bacterial colonization of the embryo is controlled by maternal antimicrobial peptides. Proc. Natl. Acad. Sci. U.S.A. **107:**18067-18072.
8. **Heath-Heckman EAC, Gillette AA, Augustin R, Gillette MX, Goldman WE, McFall-Ngai MJ.** 2014. Shaping the microenvironment: evidence for the influence of a host galaxin on symbiont acquisition and maintenance in the squid-vibrio symbiosis. Environ. Microbiol.**:**Epub.
9. **Login FH, Balmand S, Vallier A, Vincent-Monégat C, Vigneron A, Weiss-Gayet M, Rochat D, Heddi A.** 2011. Antimicrobial peptides keep insect endosymbionts under control. Science **334:**362-365.
10. **Franzenburg S, Walter J, Künzel S, Wang J, Baines JF, Bosch TCG, Fraune S.** 2013. Distinct antimicrobial peptide expression determines host species-specific bacterial associations. Proc. Natl. Acad. Sci. U.S.A. **110:**E3730-E3738.
11. **Signoretto C, Burlacchini G, Pruzzo C, Canepari P.** 2005. Persistence of *Enterococcus faecalis* in aquatic environments via surface interactions with copepods. Appl. Environ. Microbiol. **71:**2756-2761.
12. **Huq A, Small EB, West PA, Huq MI, Rahman R, Colwell RR.** 1983. Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. Appl. Environ. Microbiol. **45:**275-283.
13. **Thomas KU, Joseph N, Raveendran O, Nair S.** 2006. Salinity-induced survival strategy of *Vibrio cholerae* associated with copepods in Cochin backwaters. Mar. Pollut. Bull. **52:**1425-1430.
14. **Epstein SS, Colwell R.** 2009. Viable but Not Cultivable Bacteria, p. 121-129, Uncultivated Microorganisms, vol. 10. Springer Berlin Heidelberg.
15. **MacRae TH.** 2010. Gene expression, metabolic regulation and stress tolerance during diapause. Cell. Mol. Life Sci. **67:**2405-2424.
16. **Lee KY, Horodyski FM, Valaitis AP, Denlinger DL.** 2002. Molecular characterization of the insect immune protein hemolin and its high induction during embryonic diapause in the gypsy moth, *Lymantria dispar*. Insect Biochem. Mol. Biol. **32:**1457-1467.
17. **Chernysh SI, Gordja NA, Simonenko NP.** 2000. Diapause and immune response : Induction of antimicrobial peptides synthesis in the blowfly, *Calliphora vicina* R.-D. (Diptera : Calliphoridae). Entomol. Sci. **3:**139-144.
18. **Tarrant A, Baumgartner M, Verslycke T, Johnson C.** 2008. Differential gene expression in diapausing and active *Calanus finmarchicus* (Copepoda). Mar. Ecol. Prog. Ser. **355:**193-207.
19. **Aruda AM, Baumgartner MF, Reitzel AM, Tarrant AM.** 2011. Heat shock protein expression during stress and diapause in the marine copepod *Calanus finmarchicus*. J Insect Physiol **57:**665-675.
20. **Evseeva N.** 1996. Diapause of copepods as an element for stabilizing the parasite system of some fish helminths. Hydrobiologia **320:**229-233.
21. **Miller CB, Nelson DM, Weiss C, Soeldner AH.** 1990. Morphogenesis of opal teeth in calanoid copepods. Mar. Biol. **106:**91-101.
22. **Tang KW, Turk V, Grossart H-P.** 2010. Linkage between crustacean zooplankton and aquatic bacteria. Aquat. Microb. Ecol. **61:**261-277.
23. **Kirn TJ, Jude BA, Taylor RK.** 2005. A colonization factor links *Vibrio cholerae* environmental survival and human infection. Nature **438:**863-866.
24. **Hunt DE, Gevers D, Vahora NM, Polz MF.** 2008. Conservation of the chitin utilization pathway in the *Vibrionaceae*. Appl. Environ. Microbiol. **74:**44-51.

25. **Huq A, Colwell RR, Rahman R, Ali A, Chowdhury MAR, Parveen S, Sack DA, Russekcohen E.** 1990. Detection of *Vibrio cholerae* O1 in the aquatic environment by fluorescent-monoclonal antibody and cell cultures Appl. Environ. Microbiol. **56:**2370-2373.
26. **Lazarevic V, Whiteson K, Hernandez D, Francois P, Schrenzel J.** 2011. Study of inter- and intra-individual variations in the salivary microbiota. BMC Genomics **11:**1471-2164.
27. **Consortium HMP.** 2012. Structure, function and diversity of the healthy human microbiome. Nature **486:**207-214.
28. **Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, Gordon JI, Knight R.** 2011. Moving pictures of the human microbiome. Genome Biol. **12:**2011-2012.
29. **Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R.** 2009. Bacterial community variation in human body habitats across space and time. Science **326:**1694-1697.
30. **Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE.** 2011. Succession of microbial consortia in the developing infant gut microbiome. Proc. Natl. Acad. Sci. U.S.A. **108:**4578-4585.
31. **Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Program NCS, Bouffard GG, Blakesley RW, Murray PR, Green ED, Turner ML, Segre JA.** 2009. Topographical and temporal diversity of the human skin microbiome. Science **324:**1190-1192.