MICROBIAL FRIENDS AND FOES: CHARACTERIZING THE CNIDARIAN RESPONSE TO PATHOGENIC AND MUTUALISITIC MICROORGANISMS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Morgan Elizabeth Mouchka

January 2014

MICROBIAL FRIENDS AND FOES: CHARACTERIZING THE CNIDARIAN RESPONSE TO PATHOGENIC AND MUTUALISITIC MICROORGANISMS

Morgan Elizabeth Mouchka, Ph. D.

Cornell University 2014

The ecology and evolution of cnidarians is driven by symbiotic and pathogenic host-microbe relationships. Research regarding these relationships is especially timely given the recent decline of coral reef ecosystems, in part due to disruptions in cnidarian-microbe interactions. My dissertation is an experimental analysis of how cnidarians respond to both harmful and beneficial microorganisms and explores the interplay between these two types of interactions.

Corals provide a multifaceted habitat that supports a rich bacterial assemblage, and in Chapter 1, I review our current knowledge regarding the diversity, specificity, development, and functions of these assemblages. With a meta-analysis of previous work, I quantitatively analyze what is known regarding the relationship between coral-associated microorganisms and disease. Finally, I examine evidence that these populations could be disrupted by climatic change.

One of the most well-known mutualistic relationships is that between cnidarians and unicellular dinoflagellates. To evaluate the molecular mechanisms that underlie the persistence of this relationship, in Chapter 2, I identify differentially expressed transcripts between symbiotic and aposymbiotic individuals of the model sea anemone, *Aiptasia pallida*. These transcripts include those with potential functions in several metabolic pathways, transport of nutrients between the partners, and host tolerance of the dinoflagellate.

To broaden our understanding of the cnidarian response to pathogenic microbes, in Chapter 3, I report the host transcriptome response of aposymbiotic *Aiptasia* to experimental

inoculation with *Serratia marcescens*. My results suggest that *Aiptasia* responds to bacterial challenge via the regulation of tumor necrosis factor receptor-associated factor-mediated signaling, apoptosis, and ubiquitination, thus suggesting that lower metazoans respond to immune challenge via highly conserved mechanisms.

To determine how the *Aiptasia* immune response is modulated via the presence of dinoflagellate symbionts, in Chapter 4, I compared gene expression and behavioral assays of *S. marcescens*-exposed anemones with and without their symbionts. The presence of dinoflagellates greatly alters the number and type of genes expressed in response to bacterial challenge. In addition, symbiotic anemones were less likely to recover from pathogen exposure and had lower survival rates than their aposymbiotic counterparts. These results are consistent with the hypothesis that symbiotic dinoflagellates suppress *Aiptasia* immunity, perhaps to promote symbiotic homeostasis.

BIOGRAPHICAL SKETCH

Morgan Elizabeth Mouchka was born on January 5, 1981 in Scottsbluff, Nebraska. Despite living more than 2000 miles from the nearest coast, Morgan became fascinated with the ocean at an early age via family trips to the California and Florida coasts and the British Virgin Islands. While in the 5th grade, Morgan participated in a multi-disciplinary education program entitled, "The Voyage of the Mimi." The program focused on the story of a young boy who spends the summer with his grandfather, the captain of a research vessel (the Mimi), and a group of marine biologists studying humpback whales. Morgan was surprised and elated that studying whales could be a career, and from then on was adamant about becoming a marine biologist. This despite her traumatic encounter with a Portuguese-man-of-war at age 16.

In 1999, Morgan began her undergraduate education at Oregon State University (OSU) in Corvallis, Oregon. One of her most defining educational experiences at OSU was the semester-long course in marine biology taught at the Hatfield Marine Science Center in Newport, Oregon. Here, Morgan got her boots wet, both literally and metaphorically, via classes and research relating to marine algae and invertebrates, intertidal ecology, physiology, and conservation biology.

Morgan's love of all animals sans backbone was fostered via her invertebrate biology instructor and honors thesis mentor, Dr. Virginia Weis. Through Dr. Weis, Morgan learned that invertebrates are diverse and kinky and she soon became far more interested in spineless species that charismatic megafauna. She was also introduced to symbiotic cnidarians via the Weis Laboratory and was instantly drawn to questions pertaining to how a foreign cell (the dinoflagellate) could be incorporated and remain viable in the cell of another organism (the cnidarian).

Following graduation, Morgan worked as a research technician in the laboratory of Dr. Bruce Menge and Dr. Jane Lubchenco at OSU. During her time in the LubMenge laboratory, Morgan was exposed to a wide variety of scientific approaches to address ecological questions in the rocky intertidal. She was also exposed to early morning tides, and it was years before she could see a full moon without associating it with the dread of having to get up at 3:00 am. As a research technician, Morgan had the opportunity to interact and collaborate with scientists at varying levels, including professors, post-docs, graduate students, fellow technicians, and undergraduate interns. These interactions, as well as her exposure to developing and conducting independent research, solidified her desire to pursue graduate study.

In 2006, two weeks after getting married and a cross-country move, Morgan began her Ph.D. program in the Department of Ecology and Evolutionary Biology at Cornell University. Suspecting that the cnidarian immune system played a role in the persistence of the symbiotic relationship, Morgan became interested in learning more about cnidarian immunity. Thus, she began her tenure in the laboratory of Dr. Drew Harvell, an expert in coral disease ecology and immunity. Via the guidance of Dr. Harvell and the other members of her doctoral committee, Dr. Jodi Schwarz, Dr. Brian Lazzaro, and Dr. Angela Douglas, Morgan developed a research program focusing on the interactions between cnidarians and both pathogenic and mutualistic microorganisms. Morgan also appreciated the teaching opportunities afforded to her via her Ph.D. program, as well as the stimulating intellectual and social environment of her department.

Morgan defended her Ph.D. two months after giving birth to her beautiful daughter, Elena Marie. The reader may thus notice a trend regarding concentrating major life events into short temporal spaces. For the moment, she is content with her new job as Dr. Mom, but is hoping to begin a post-doc on the West Coast where her and her family will be moving in early 2014.

To Greg, for his tremendous support, to Elena, for providing the ultimate motivation, and to my parents, Wil and Shari Packard, for their unwavering encouragement.

ACKNOWLEDGMENTS

They say it takes a village to raise a child, and likewise, it takes a village to earn a Ph.D. For that reason, I owe a great deal of thanks to many people that have supported me, logistically, intellectually and/or emotionally during this time.

First and foremost I would like to thank my family. My parents, Wil and Shari Packard, were a source of endless encouragement from my early education onward. I'm incredibly thankful for our family vacations, particularly to the British Virgin Islands, where I was first introduced to snorkeling and coral reefs. Although my parents understood that encouraging a career in marine biology would undoubtedly lead to my leaving Nebraska, they never wavered in their enthusiastically support. I would also like to thank my brothers, Jeff, Jeremy, and Will, and their families for their love and encouragement. In particular, Will and I shared a love of science and I'm thankful for his knowledge of the difficulties associated with research and his willingness to hear my complaints. I am forever indebted to my amazing husband, Greg, for taking a blind leap of faith and moving to Ithaca to support my graduate education. I don't think either of us completely understood what being a 'trailing spouse' entailed, and I'm so incredibly lucky that in the end, we both found our niches and thrived. Lastly, I would like to thank my daughter, Elena, for providing the motivation to finish and for teaching me about patience and priorities.

I am grateful to a long lineage of teachers that have fostered my love of science and research. My 5th grade teacher, Mrs. Bentley, introduced me to marine biology and shared her love of science. Virginia Weis was an amazing undergraduate thesis advisor that got me hooked on invertebrates and continues to be a source of knowledge regarding all things symbiotic. In that same vein, her graduate student, Laura Hauk, was the first person to teach me molecular

methods and I'm grateful for her patience and meticulous bench skills. I'm grateful to my committee members, Brian Lazzaro and Angela Douglas, for their plethora of knowledge concerning innate immunity and invertebrate physiology, respectively, as well as their insightful feedback over these last several years. A special thanks goes to my *ad hoc* committee member, Jodi Schwarz, who for all intents and purposes was really a co-advisor. Jodi and I formed our own 'symbiotic relationship' and she supported me financially, intellectually, and emotionally throughout this journey, especially during my time *in absentia* at Vassar College. Finally, I thank Drew Harvell and all that she did as my advisor. Through Drew, I was able to collaborate with an international network of wonderful scientists and to travel to diverse places. Drew also trusted me and gave me the independence to complete a molecular-based project knowing full well this would certainly not be the easiest and/or shortest route to obtaining a Ph.D. I would also like to thank Drew for sharing her love of the rocky intertidal, via both teaching and research, and for lovely kayaking trips in Friday Harbor.

Past and present members of the Harvell laboratory, affectionately known as sea fans, have always been a continual source of knowledge, support, assistance, and humor. Krystal Rypien, Jason Andras, David Baker, Courtney Couch, and Allison Tracy were amazing academic siblings and even better friends. Thanks also to post-docs Laura Mydlarz and Colleen Burge for their valuable feedback and assistance, particularly to Colleen who shared in the frustrations and elations of molecular work/bioinformatics. Nancy Douglas, our esteemed technician, kept the lab running smoothly and provided support and friendship. I also had the honor of working with many amazing undergraduates and/or recent graduates that kept my anemones and/or myself fed and happy, including Mary Rutz, Tiffany Walker, Catherine Kim, Jillian Lyles, Gabriel Ng, Reyn Yoshioka, Natalie Rivlin, Julie Geyer, and Ellie Bondra.

I would also like to thank a few people outside of Cornell that were instrumental to the completion of this thesis. John Pringle and his laboratory group at Stanford University were incredibly helpful by sharing data and knowhow and were incredibly hospitable during my visit. A special thanks to Erik Lehnert for helping me with bioinformatics and for his incredible insight regarding metabolism and transport in symbiotic cnidarians. I'm forever grateful for both John and Erik's willingness to collaborate and to make our stories that much stronger. I would also like to thank Christy Schniztler, a former Weis laboratory alum and current postdoc at NIH, who helped with bioinformatics and phylogenetics.

While sharing an office with up to seven other students was bit distracting at times, I wouldn't have had it any other way and I'm incredibly grateful to my Corson E231 dream team, including Paulo Llambias, David Cerasale, Joe Simonis, Mike Booth, Erica Larson, Findley Finseth, Chris Dalton, Ezra Lencer, and Ben Johnson. Their encouragement and insight kept me going and got me through graduate school. I'm so thankful that we shared the same twisted sense of humor and I'm excited to have a permanent network of friends all across the world. Special thanks to Eric and Findley, my officemates in isolation as we branched off in an attempt to finish. It was wonderful to start and end together and we'll always have Easter.

The graduate students in the Department of Ecology and Evolutionary Biology are an absolutely amazing group of people who are incredibly intelligent and thoughtful. There are too many to name here, but I appreciate the academic environment that you all helped to foster. I'm also grateful to so many fellow colleagues and mentors that provided friendship, support, and sane advice, including Willie Fetzer, Sarah Collins, April Melvin, Sara DeLeon, Sarah States, Danica Lombardazzi, Anna Forsman, Rayna Bell, Rulon Clark, Gretchen Gerrish, Kathy Morrow, Ian Hewson, Jed Sparks, Harry Greene, Nelson Hairston, Rick Harrison, and Rob

Raguso. I made some wonderful friends while here in Ithaca outside of the academic realm. James Dean, Scott Callan, Christopher Couch, Ben Zagorski, Ryan Finseth, Jon Larson, Terry Howell, Quinn Caldwell, Zug Thompson Ben and Joy Thomas, Devin and Michael Craven, and Whitney and Josh Hungerford, provided stimulating conversation and good company (and usually good food and drink).

The support of the departmental staff has been invaluable and I greatly appreciate all of the logistical and bureaucratic hoops they jump through to keep the place running. Thanks to LuAnne Kenjerska, DeeDee Albertsman, Alberta Jackson, Carol Damm, Janeen Orr, Patty Jordon, Christine Sancherico, Brian Mlodzinski, John Howell, and Gary Oltz. Lastly, my research was funded by many sources, including a Graduate Research Fellowship and a Doctoral Dissertation Improvement Grant from the National Science Foundation, the Cornell University Graduate School, Cornell Sigma Xi, the Andrew W. Mellon Foundation, the Orenstein Foundation, and the Cornell Department of Ecology and Evolutionary Biology.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

CORAL-ASSOCIATED BACTERIAL ASSEMBLAGES: CURRENT KNOWLEDGE AND THE POTENTIAL FOR CLIMATE DRIVEN IMPACTS[1]

Abstract

The importance of associations between microorganisms and their invertebrate hosts is becoming increasingly apparent. An emerging field, driven by the necessity to understand the microbial relationships that both maximize coral health and cause coral disease, is the study of coral-bacteria interactions. In this article, we review our current understanding of the diversity, specificity, development, and functions of coral-associated bacteria. We also summarize what is known regarding the role of coral microbiota in the health and disease of coral. We conduct a meta-analysis to determine whether the presence of unique taxa correlates with the state of coral health (i.e., healthy, diseased, or bleached), as well as whether coral reef habitats harbor clusters of distinct taxa. We find that healthy and bleached corals harbor similar dominant taxa, although bleached corals had higher proportions of *Vibrio* and *Acidobacteria*. Diseased corals generally had more *Rhodobacter*, *Clostridia*, and *Cyanobacteria* sequences, and fewer *Oceanospirillum* sequences. We caution, however, that while 16S rRNA is useful for microbial species identification, it is a poor predictor of habitat or lifestyle, and care should be taken in interpretation of 16S rRNA surveys to identify potential pathogens amongst complex coral-microbial assemblages. Finally, we highlight evidence that coral-bacterial assemblages could be sensitive to the effects of climate change. We suggest that the relationship between coral and their bacterial associates represents a valuable model that can be applied to the broader discipline of invertebrate-microbial interactions.

[1] Published as: Mouchka, M., I. Hewson, and C. Harvell. 2010. Coral-associated bacterial assemblages: current knowledge and the potential for climate-driven shifts. *Integrative and Comparative Biology* 50(4): 662-674.

Introduction

The close associations between animals and their microbiota have shaped the evolutionary paths of both host and symbiont alike. While interactions between microorganisms and vertebrates have been well studied, relatively little attention has been given to the examination of microbial-invertebrate associations. A frontier of invertebrate biology is the interaction between microorganisms and their hosts. Indeed, many of the biologically active compounds ascribed to marine invertebrates, like sponges (Flatt et al., 2005; Ridley et al., 2005) and bryozoans (Hildebrand et al., 2004) have been found to be produced by their bacterial associates. Increasingly, these associations show strong functional significance. For instance, the bacterial symbionts of sponges and bryozoans produce chemicals that protect their hosts from heterospecific settlement of larvae (Ridley et al., 2005) and from predation (Lopanik et al., 2004), respectively. Similar to other sessile invertebrates, research on coral-associated bacteria is revealing important symbiotic functions and this system is emerging as one of the best-studied examples of invertebrate-microbial interactions. While corals have been found to harbor a wide variety of microbes, including heterotrophic eukaryotes, bacteria, archaea and viruses, the majority of studies thus far have centered on bacteria associated with coral.

The coral organism is a complex host that forms associations with both external and internal microbiota. The coral animal, its intracellular algal symbionts, and the diverse microorganisms found in association with coral tissues and exudates have been termed the 'holobiont' (Rohwer et al., 2002; Reshef et al., 2006). While it has long been known that the algal symbiont is an obligate partner supplying up to 95% of the host's metabolic requirements for carbon and contributing to formation of the skeleton (Muscatine, 1973), the roles of coral-associated bacteria have not been well elucidated. The structure of the coral host provides a

multifaceted habitat, with distinct and diverse bacteria residing in the host skeleton, tissues, and surface mucus layer. As in terrestrial ecosystems, where bacterial assemblages play an essential role in ecosystem functioning (Balser et al., 2006), coral-associated bacteria are likely to drive biochemical and ecological processes within the reef environment. In this paper, we review the literature concerning coral-associated bacteria, summarizing the diversity, specificity, development, and functional roles of coral microbiota. We also consider the relationship between coral-associated microorganisms and disease, conducting a meta-analysis to determine whether diseased or bleached coral harbor unique taxa, as well as whether clusters of taxa are distinct to the coral reef habitat. Finally, we examine evidence that these populations have the potential to be disrupted by climate change. While this review focuses on coral-associated bacteria, we suggest it contains themes useful for a broader consideration of the importance of invertebrate-microbial interactions.

Diversity and specificity of coral-associated bacteria

Sequence-based assessments of microbial assemblages, which involve random sampling of bacterial rRNA genes amplified from nucleic acid (Olsen et al., 1986), provide high taxonomic resolution for environmental samples across large datasets based on nucleotide heterogeneity. While cultivation-based approaches provide important information on the metabolism of some microorganisms, the vast majority (> 99%) of marine microorganisms do not grow on enriched media (Azam, 1998). In coral, 16S rRNA surveys of bacteria have elucidated an astonishing diversity of bacterial ribotypes, many of which are not closely related to cultivated or uncultivated microorganisms identified in previous studies. For instance, Rohwer et al. (2002) characterized the bacterial assemblage of three Caribbean species and

estimated the presence of 6,000 ribotypes in libraries from 14 coral samples. Additional studies examining bacterial assemblages from multiple coral species and geographic regions have found similar results (Rohwer et al., 2001; Bourne and Munn, 2005; Klaus et al., 2005; Koren and Rosenberg, 2006; Sekar et al., 2006; Kapley et al., 2007; Wegley et al., 2007; Koren and Rosenberg, 2008; Lampert et al., 2008; Hong et al., 2009; Littman et al., 2009b; Reis et al., 2009). Like most microbial assemblages in marine ecosystems, coral-associated microbial assemblages contain microdiverse clusters (i.e., organisms varying by a handful of nucleotides across entire 16S rRNA genes) of closely related taxa, where rarely is exactly the same sequence retrieved twice in surveys. Microbial assemblages in corals, like plankton communities, are dominated by a few different taxonomic units with a long tail of the species-distribution curve (Rohwer et al., 2002), suggesting that much of the diversity within the coral microbiome exists within the 'rare' biosphere (Sogin et al., 2006).

A central question in microbial ecology is whether microorganisms fill defined niches within complex communities, or whether communities are comprised of functionally redundant, neutrally-selected taxa leading to random assemblages (Fuhrman et al., 2006). In marine plankton, microbial assemblages are heterogeneous between geochemical and productivity-defined habitats (Moeseneder et al., 2001; Hewson and Fuhrman, 2004), yet in richer habitats, like sediments, spatially distinct communities in the same habitat type are more similar to each other than to those in adjacent habitats (Hewson et al., 2007). It is therefore not surprising to see a similar pattern in studies of coral-associated bacteria, which presumably inhabit a productive environment, with similar bacterial ribotypes associated with the same coral species, but distinct from those in surrounding seawater and sediments (Frias-Lopez et al., 2002; Rohwer et al., 2002; Bourne and Munn, 2005; Pantos and Bythell, 2006; Littman et al., 2009b; Reis et al., 2009).

This is supported by the observation that some bacterial ribotypes form host-species-specific associations with coral (Rohwer et al., 2001; Frias-Lopez et al., 2002; Rohwer et al., 2002; Bourne, 2005; Klaus et al., 2005; Sekar et al., 2006; Lampert et al., 2008; Reis et al., 2009). It is hypothesized that this specificity is indicative of the importance of certain interactions to holobiont functioning, and that these interactions are structured in ways that maximize the health of the holobiont (Rohwer et al., 2002; Reshef et al., 2006).

While the existence of such coral-bacterial (and therefore microbial habitat) specificity is widely accepted, the spatial and temporal stability of these interactions is debated. In seawater, for example, bacterial assemblages can be heterogeneous within the same habitat at spatial scales ranging from μm to km (Long and Azam, 2001; Hewson et al., 2006a; Hewson et al., 2006b). In coral, some studies have shown that species-specific bacteria are geographically consistent (Rohwer et al., 2001; Rohwer et al., 2002). For instance, Rohwer et al. (2002) showed that bacteria associated with three coral species in Panama contained similar ribotypes to those of the same coral species in Bermuda. The opposite trend has also been observed, in which bacterial assemblages contained different ribotypes between geographic locations, but similar corals were inhabited by similar ribotypes (Klaus et al., 2005; Guppy and Bythell, 2006; Littman et al., 2009b). Trends observed by sequence library surveys of uncultivated communities are generally consistent with those using fingerprinting approaches, which have lower taxonomic resolution, but provide greater qualitative assessment of large numbers of samples or assemblages. These discrepancies could be explained, in part, by differences in methods (clone sequencing versus T-RFLP, DGGE), coral taxonomic resolution (comparing coral species within the same genus versus different genera) and the operator-defined taxonomic resolution of sequence analyses ('cutoffs' of sequence identity defining operational taxonomic units to permit comparisons

between communities) (Rohwer et al., 2001; Rohwer et al., 2002; Klaus et al., 2005; Guppy and Bythell, 2006; Littman et al., 2009b). Microbial taxonomic resolution influences similarity between assemblages based on sequencing; it is currently unclear which nucleotide identity cutoffs are appropriate for defining ecologically meaningful taxonomic levels. The varied trends over geographic scales and with host species may also reflect differential species responses (host and/or microbiota) to site-specific factors (Hong et al., 2009; Littman et al., 2009b). Taken together, differences between studies highlight the multifaceted and dynamic nature of coral-associated microbiota, and caution should be taken not to over-simplify or over-generalize the nature of these associations.

The onset of coral-bacterial associations

Determining when and how coral-microbial assemblages are established is fundamental to a better understanding of the coral holobiont. Apprill et al. (2009) examined the onset of microbial associations in the coral, *Pocillopora meandrina*, by comparing bacterial tRFLP profiles between pre-spawned oocyte bundles, spawned eggs, and week old planulae. They found that there were distinct ribotypes present within each stage, but that bacterial cells were not internally incorporated until the planulae are fully developed (Apprill et al., 2009). This suggests that, unlike the zooxanthellae, which are vertically transmitted in this system, bacteria that form associations with *P. meandrina* are acquired via horizontal uptake. As bacteria are internally incorporated during late development of the planulae, it is possible that bacteria play a role in processes specific to this life stage, such as benthic settlement (Apprill et al., 2009).

There is also evidence that coral-associated bacteria differ between adults and juveniles of coral. Nonmetric multidimensional scaling (nMDS) representations of bacterial profiles

assessed through random sequencing of clone libraries, DGGE, and T-RFLP, were all consistent in demonstrating that adult *Acropora tenuis* and *A. millepora* displayed tight grouping, whereas there was no apparent relationship between profiles of juveniles (Littman et al., 2009a). The bacterial complement of juvenile corals was also more diverse, and while there was some conservation in bacterial ribotypes between adult and juvenile corals, the vast majority of adult-associated bacterial ribotypes were not found in juveniles. This suggests a successional process whereby associates of adult corals gradually replace the diverse bacterial consortia of juveniles (Littman et al., 2009a). Future studies are required to examine this successional process throughout the ontogeny of the coral to determine when and how species-specificity is established and if these factors differ among coral species.

The role of coral-associated bacteria

While the presence of coral-associated bacteria has long been established (Di Salvo and Gundersen, 1971), little is known about how these microorganisms contribute to the functioning of the coral holobiont. There is increasing evidence that coral microbiota are crucial to at least two aspects of the host's physiology: biogeochemical cycling and pathogen resistance. The tight nutrient cycling that enables corals to thrive in oligotrophic waters was originally attributed to the mutualism between the coral host and its photosynthetic dinoflagellates. Recently, however, both culture-dependent and independent techniques have demonstrated that coral microbiota likely play a role in coral reef biogeochemistry (Williams et al., 1987; Szmant et al., 1990; Shashar et al., 1994; Ferrier-Pages et al., 2000; Lesser et al., 2007; Wegley et al., 2007; Chimetto et al., 2008; Olson et al., 2009; Raina et al., 2009; Kimes et al., 2010). For example, nitrogen fixation within the coral holobiont has been documented using acetylene reduction assays

(Williams et al., 1987; Shashar et al., 1994; Lesser et al., 2007; Chimetto et al., 2008;) and bacteria possessing genes for nitrogen-fixation have been identified within multiple coral species from varying geographic regions (Lesser et al., 2004; Wegley et al., 2007; Olson et al., 2009; Kimes et al., 2010). In addition, recent studies have found evidence that members of coral-associated microbiota may also be involved in additional nitrogen cycling processes, including nitrification, ammonium assimilation, ammonification, and denitrification (Wegley et al., 2007; Kimes et al., 2010). There is also evidence that coral-associate microbial assemblages function in carbon and sulfur cycling (Ferrier-Pages et al., 1998; Ferrier-Pages et al., 2000; Wegley et al., 2007; Raina et al., 2009; Kimes et al., 2010). Genes that regulate carbon fixation, carbon degradation, and methanogenesis have been detected in coral-associated bacteria (Wegley et al., 2007; Kimes et al., 2010), as have those that regulate assimilation of organic and inorganic sulfur sources (Wegley et al., 2007; Raina et al., 2009; Kimes et al., 2010). The ability of microbes to subsize the nutrient budgets of their coral host is likely a driver in the establishment of coral-associated microbial assemblages. Furthermore, niche partitioning of bacterial assemblages is likely to be controlled by availability of nutrients at the microscale of the coral host structure (van Duyl and Gast 2001; Scheffers et al., 2005; Raina et al., 2009; Ainsworth et al., 2010). However, it should be noted that the presence of a functional gene or gene fragment does not necessearily imply functionality, and that additional *in situ* or expression-based studies are required to elucidate the role that microbes play in driving nutrient cycling on coral reefs. The role of microbes in biogeochemical cycling and their distribution at the scale of the holobiont micro-niche are important areas of future research.

It has also been hypothesized that coral-associated bacteria play a role in resistance to disease (Ritchie and Smith, 2004; Rohwer and Kelley, 2004; Reshef et al., 2006) via competition

for nutrients and/or space, and/or production of antibiotics (Rohwer and Kelley, 2004). Several studies have demonstrated the antibacterial activity of isolates of coral mucus against indicator bacteria (e.g., *Escherichia coli, Staphylococcus aureus*), potentially invasive microbes (microbes from Florida Keys' canal water, African dust, and surrounding seawater) and putative pathogens of coral (*Vibrio shiloi*, *V. coralliilyticus*, and *Serratia marsecens*) (Ritchie 2006; Nissimov et al., 2009; Rypien et al., 2009; Shnit-Orland and Kushmaro, 2009). It has also been shown that the antibacterial properties of coral mucus select for a discrete set of commensal bacteria (Ritchie, 2006) and that antagonistic interactions are prevalent among co-occurring coral-associated microbes (Rypien et al., 2009). However, it should also be noted that coral mucus contains very high concentrations of organic and inorganic matter, leading to typically rare, r-selected (i.e. fast growing and nutrient sensitive) bacteria in seawater recruiting to the mucus matrix and increasing rapidly in abundance (Allers et al., 2008a). These findings suggest that the coral-associated microbiota is dynamic, self-regulating, and has the capacity to prevent settlement of exogenous bacteria, including pathogens. Future studies should focus on the factors that enable pathogens to become established, as well as the additional roles (e.g., competition and niche occupation) that symbiotic bacteria play in preventing colonization by pathogens.

Coral Disease and coral-associated bacteria

Over the past several decades, coral reef ecosystems have been degrading at an alarming rate (Hughes et al., 2003; Baker et al., 2008). This degradation, in part, is a consequence of coral disease (Harvell et al., 1999; Harvell, 2004), for which prevalence, severity, and host and geographic range have all been increasing (Harvell, 2004; Weil et al., 2006; Harvell et al., 2007). Like their terrestrial counterparts, marine epizootics cause marked declines in populations, alter

community structure, and therefore threaten biodiversity (Harvell et al., 2002). To date, there are more than 20 described coral diseases (Rosenberg et al., 2007). However, due to the difficulties of isolating and culturing putative pathogens and of aseptic cultivation of host tissues, there are only six diseases for which a causative agent has been identified (Rosenberg et al., 2007; Bourne et al., 2009). Knowledge of coral disease reservoirs, transmission, and pathogenesis is limited, as is the role that coral-associated microbial assemblages play in the health and disease of coral.

To gain a better understanding of how variation in microbial assemblages associated with corals may lead to the onset of disease, numerous studies have compared bacterial assemblages between healthy and diseased coral. These studies have shown that both the composition and function of microbiota associated with healthy and diseased corals are distinct (Ritchie and Smith, 1995; Cooney et al., 2002; Frias-Lopez et al., 2002; Pantos et al., 2003; Bourne, 2005; Gil-Agudelo et al., 2006; Pantos and Bythell, 2006; Sekar et al., 2006; Barneah et al., 2007; Gil-Agudelo et al., 2007; Voss et al., 2007; Sekar et al., 2008; Reis et al., 2009; Sunagawa et al., 2009). Furthermore, differences between the microbiota of healthy and diseased corals appear to be systemic in some cases (Pantos et al., 2003; Breitbart et al., 2005; Pantos and Bythell, 2006). That is, the bacterial assemblage of the entire diseased colony is the same and distinct from that of healthy colonies, even though only a small portion of the colony shows signs of disease. These results reinforce the idea that apparently healthy tissues of diseased colonies should not be used as control references, and that this systemic effect could be used as a diagnostic tool to identify stressed colonies susceptible to disease (Pantos et al., 2003).

There are several hypotheses for the observed variability in the structure of bacterial assemblages associated with disease: 1) changes in environmental conditions directly or indirectly alter the microbiota of healthy coral. For instance, increases in nutrients (e.g.,

nitrogen, dissolved organic carbon) may 'fertilize' nutrient-limited, r-selected, potentially pathogenic taxa, enabling them to dominate the community (Bruno et al., 2003; Kline et al., 2006; Smith et al., 2006; Voss and Richardson, 2006). Nutrient increases may also play a more indirect role by compromising normal function of beneficial coral residents, thereby leading to overgrowth of pathogenic taxa (Kline et al., 2006). 2) Changes in environmental conditions alter host physiology, subsequently leading to variable microbiota. For example, because coral mucus provides an important carbon source for coral-associated bacteria (Ferrier-Pages et al., 2000; Brown and Bythell, 2005; Wegley et al., 2007; Allers et al., 2008; Kimes et al., 2010), changes in production rates of mucus due to abiotic factors (e.g., temperature and/or irradiance)(Piggot et al., 2009) could also lead to variable structure of coral microbiota. 3) Colonization by pathogens directly or indirectly causes variation in the normal bacterial assemblage. For example, pathogens may directly alter community structure by outcompeting resident bacteria if they have a higher affinity for available substrates and/or are capable of producing antibiotics (Rypien et al., 2009). Pathogens may also indirectly cause variability in coral microbiota via degradation of host tissues, creating a nutrient-rich microenvironment that is readily colonized from surrounding waters by secondary r-selected invaders (Cooney et al., 2002; Frias-Lopez et al., 2002; Pantos et al., 2003; Bourne, 2005; Pantos and Bythell, 2006; Reis et al., 2009). In the tail of the species-distribution curve of bacterioplankton communities, there are many such r-selected taxa that maintain low abundances until prevailing conditions arise. A classical example of rare yet r-selected taxa that are present in bacterioplankton are marine *Vibrio*, which are easily cultivated on enriched solid media from seawater (Giovannoni and Stingl, 2005). Given that coral reef ecosystems are defined by complex multi-partner relationships and dynamic environmental conditions, it is likely that these hypotheses are not mutually exclusive and that a

combination of factors ultimately leads to the onset of coral disease. It is equally likely that coral disease is elicited by networks of interacting bacteria, and that interactions with physicochemical features of their habitat is complex and not easily disentangled by methods currently used in coral microbial ecology.

Meta-analysis of coral-associated bacterial assemblages

To gain insight into the extent to which the states of coral health are correlated with the composition of their microbial assemblages, we analyzed 16S rRNA sequence accessions to GenBank produced in 36 studies of coral microbial ecology. A summary of the studies, the methods employed, and the number and source of 16S rRNA sequences can be found in the appendix (Table S1.1). Our analysis is not a quantitative assessment of the composition of microbial assemblages, since different studies used distinctly different approaches (targeted 16S sequencing of DGGE amplicons to fully random sequencing of entire assemblages), with different approaches used in different compartments of the reef (e.g. coral, seawater). Note that the comparison between different habitat types is biased by variable numbers of sequences and approaches, making standardization difficult at this level of analysis. It is also important to point out that these data represent an inherent bias based on research interest. For instance, easily identifiable diseases (e.g., black band) and dominant reef species (e.g., *Montastrea spp*.) are sampled more often than are their less distinguishable, less dominant counterparts. This analysis does not include sequences derived from pyrosequencing technologies recently applied to coral microbiology (e.g. Sunagawa et al., 2010), which will likely provide extensive information on the 'tail' of the species distribution curve not sampled by Sanger-sequenced libraries.

Coral microbiota 16S rRNA sequences were mostly dominated by bacteria (~ 80 - 90%), with both healthy and bleached corals harboring similar dominant taxa (Figure 1.1). However, in diseased corals, there were generally more *Rhodobacter, Clostridia,* and *Cyanobacteria* sequences, and fewer *Oceanospirillum* sequences. Interestingly, the abundance of *Rhodobacter* sequences in diseased corals is not associated with a single disease or geographic locale. Rather, *Rhodobacter* seems to be abundant under many different conditions in many different locales, including assemblages associated with black band disease from the Caribbean, Red Sea, and Great Barrier Reef (Cooney et al., 2002; Frias-Lopez et al., 2002; Bourne, 2005; Sekar et al., 2006; Barneah et al., 2007; Sato et al., 2009), white plague and white band disease from the Caribbean (Pantos et al., 2003; Pantos and Bythell, 2006; Sunagawa et al., 2009), and two conditions, atramentous necrosis and cyanobacterial patches, from the Great Barrier Reef (Bourne, 2005; Sato et al., 2009). Surveys of bleached coral had a higher proportion of the r-selected opportunist genera, *Vibrio* and *Acidobacteria,* than did surveys of healthy coral. It is unclear whether bacterial sequence types detected in high abundance on the surface of diseased or bleached tissues represent pathogens driving the diseased state, or are merely opportunists taking advantage of shifts in the bacterial assemblage or in host physiology that are associated with bleaching and disease.

Despite these observations, it is important to note that 16S rRNA is a poor predictor of habitat or lifestyle, where closely related taxa can occupy disparate environments and carry out different functions within communities. For example, in our analysis, some sequences of α-Proteobacteria associated with disease were also closely related to those found in disparate environments, like deep-sea sediments and open-ocean plankton (Figure 1.2). While other

ribotypes were distinct to diseased tissues (i.e. their close relatives were only found in association with disease), our analysis emphasizes that care must be taken in interpretation of 16S rRNA surveys for identifying potential pathogens amongst complex microbial assemblages

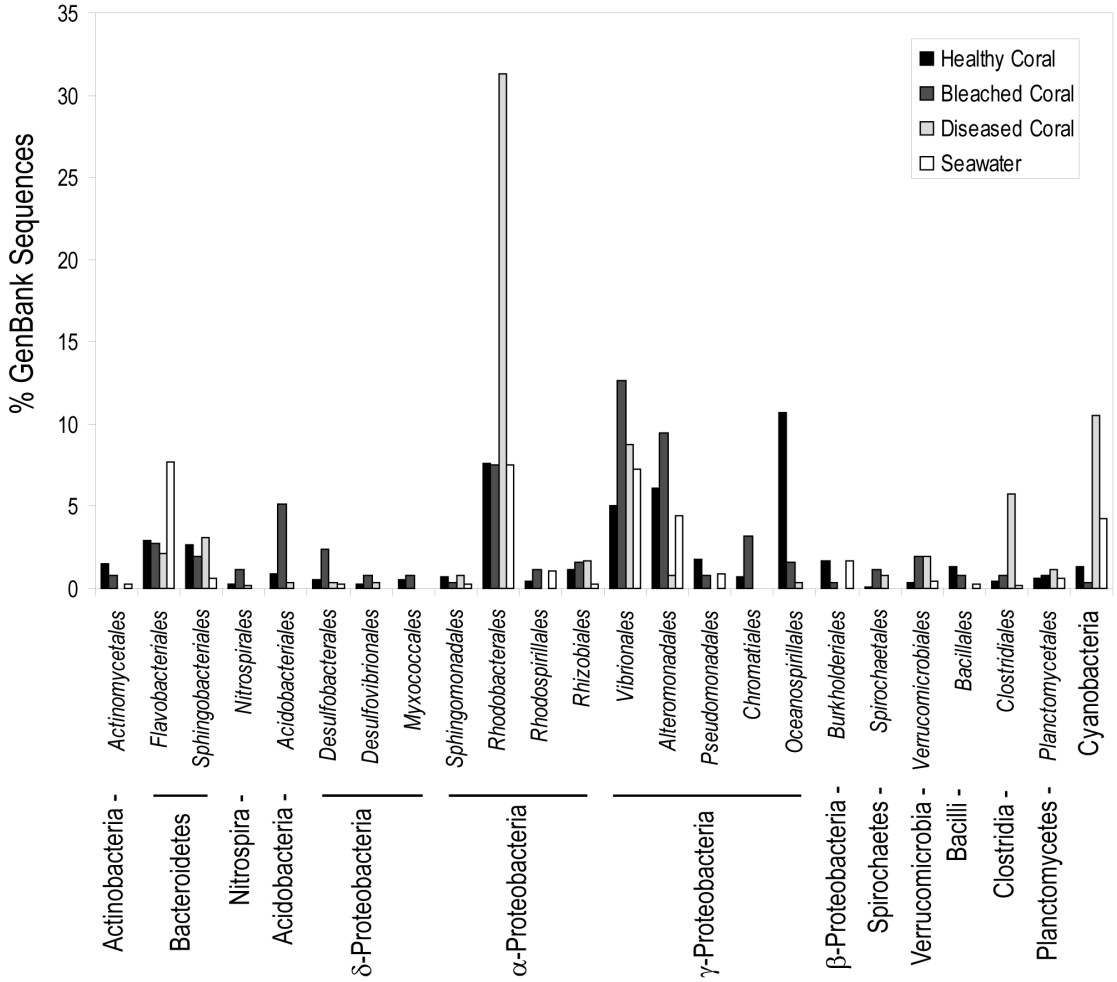


Figure 1.1 Analysis of coral reef-derived 16S rRNA sequence accessions to GenBank associated with healthy (n = 4,271 sequences), bleached (n = 254 sequences), and diseased (n = 524 sequences) coral and overlying seawater (n = 662 sequences). Sequence accessions were classified using the Bayesian classifier tool at the Ribosomal Database Project II. Unclassified sequences within each class are not included. Only orders representing > 1 % of all sequence accessions are shown.

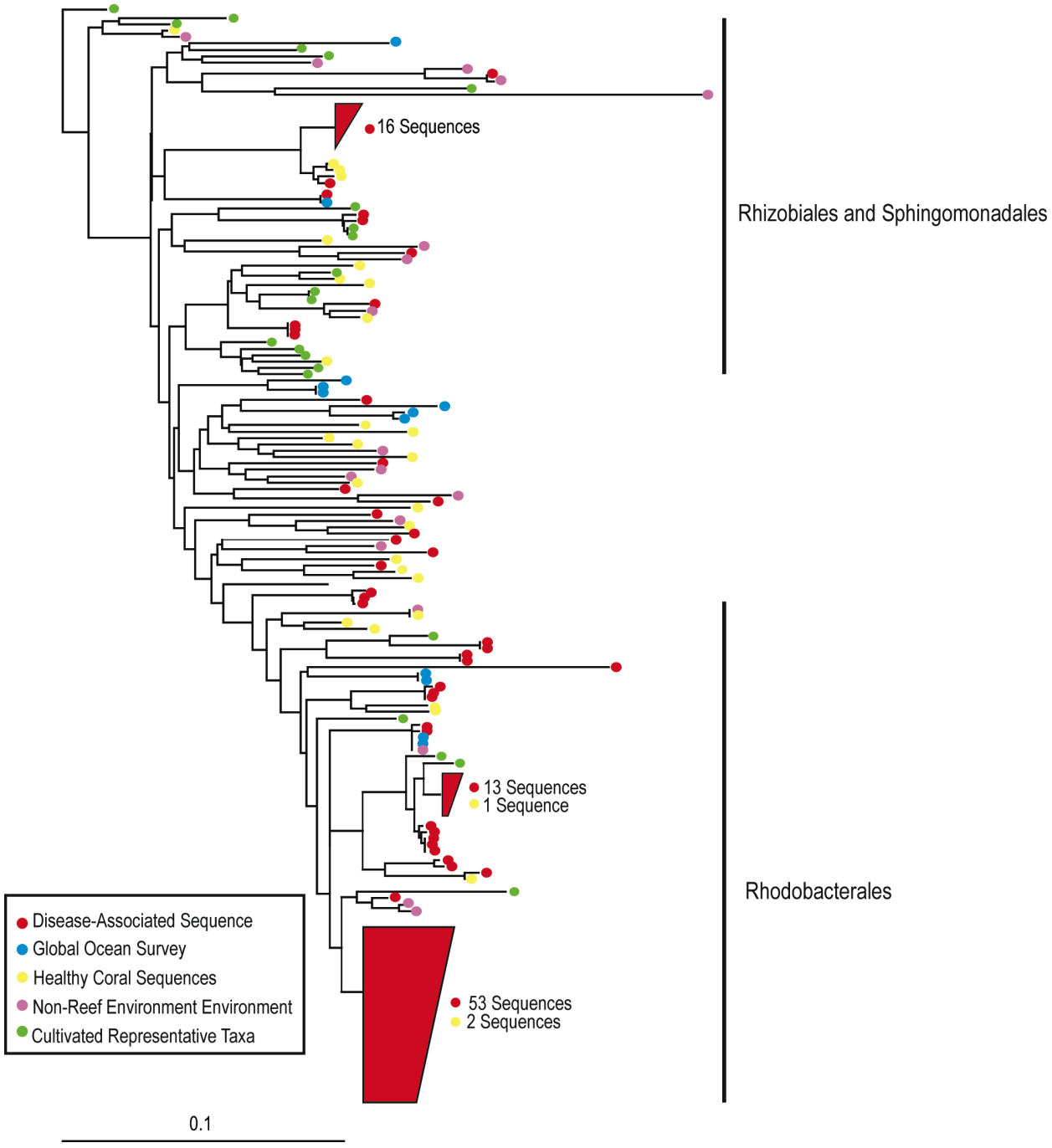


Figure 1.2 Phylogenetic analysis of 93 disease-associated α-Proteobacteria and their closest matches from genome sequences, the Global Ocean Survey of bacterioplankton, and the non-redundant database at NCBI. Branches have been collapsed where multiple sequences have been recovered. The tree was produced using neighbor-joining based on a 395 base pair alignment produced using the Ribosomal Database Project II. Scale bar = 0.1 substitutions per site. Non-reef sequences included, for example, those from deep-sea sediments, salt marsh sediments, and pelagic bacterioplankton.

in association with corals. While informative, comparative studies of coral-associated microbial assemblages are unable to answer a key question: are shifts in community structure the cause or the effect of the disease? Future comparative studies should focus on the temporal dynamics of bacterial replacement. In addition, the use of metagenomics, as opposed to 16S rRNA techniques, can provide concurrent information concerning both community function and structure, which could be useful in identifying potential pathogens through virulence genes and/or culturing conditions of putative pathogens based on physiological function (Wegley et al., 2007; Thurber et al., 2009; Ainsworth et al., 2010). Perhaps most importantly, however, these studies should follow up with active inoculations to determine the mechanisms underlying pathogen colonization and pathogenesis, and how the coral-associated microbial assemblage is altered via these mechanisms.

The potential impacts of climate change on coral-associated bacteria

Climate change is having measurable effects on marine and terrestrial ecosystems alike. In the ocean, anthropogenically-driven increases in atmospheric concentrations of carbon dioxide contribute to both ocean warming and acidification (Harvell et al., 2007; Doney et al., 2009; Feely et al., 2009). Warming and acidification alone, and synergistically, have the potential to not only alter coral physiology directly (Hoegh-Guldberg et al., 2007; De'ath et al., 2009; Kleypas and Yates 2009), but also indirectly through impacts on coral-associated microorganisms, thereby potentially disrupting the normal function of the coral holobiont. This loss of function, in turn, may impact coral reef ecosystems as a whole.

The hypotheses developed to explain variability in coral-bacteria assemblages as a result of disease, namely that environmental factors can directly or indirectly affect the microbiota

and/or host physiology, can also be used to predict the effects of ocean warming and acidification on the coral holobiont. While it is possible that increasing temperatures and decreasing pH of the sea surface will alter the biogeochemical role that coral microbiota potentially play, there is a paucity of research investigating this phenomena. However, there is considerably more data concerning how climate change, and more specifically, increasing temperatures, will affect the role that coral-associated microbiota play in disease.

Increases in seawater temperature can directly alter coral-associated bacterial structure and function, potentially leading to disease. Vega Thurber et al. (2009) demonstrated that elevated temperatures shifted the microbiome of *Porites compressa* to a more disease-associated state. That is, both the number of genes encoding virulence pathways and the abundance of ribosomal sequences associated with diseased organisms were greater in the microbial assemblage of corals exposed to elevated temperatures. Indeed, for a number of coral diseases, growth rates and/or virulence of pathogens are temperature-dependent (Alker et al., 2001; Ben-Haim et al., 2003; Cervino et al., 2004; Rosenberg and Falkovitz, 2004; Remily and Richardson, 2006; Ward et al., 2007). Therefore, increases in seawater temperature could potentially shift coral-associated microbial assemblages by selecting for more pathogenic taxa.

There is also evidence that increases in temperature can indirectly alter coral microbiota by compromising function of beneficial members that structure healthy communities. Several studies have demonstrated that antibacterial activity of mucus-associated bacteria is impaired under elevated temperatures (Ritchie, 2006; Rypien et al., 2009; Shnit-Orland and Kushmaro, 2009). For instance, Richie (2006) found that the antibacterial activity of apparently healthy *Acropora palmata* mucus was lost when corals were exposed to higher sea surface temperatures. Furthermore, culturable isolates from the mucus were dominated by *Vibrio*s, while this genus

was far less abundant in mucus sampled prior to the thermal event. These results suggest that increased temperatures can shift coral-associated microbial assemblages away from species that regulate unaffected communities toward dominance by potential pathogens.

Temperature-driven changes in host physiology could also affect coral-associated microbiota. Perhaps one of the most striking changes in the physiology of the host is bleaching. Coral bleaching is the breakdown of the symbiotic relationship between corals and their intracellular algae, leading to the loss of the algae and/or its photosynthetic pigments. Bleaching can be caused by a variety of factors (e.g., heavy metals, sediment, pathogens) (Coles and Brown 2003), but is most commonly caused by increases in sea surface temperatures that disrupt algal photosynthesis (Hoegh-Guldberg 1999; Hughes et al., 2003). Not surprisingly, variability in the structure of bacterial assemblages also occurs during bleaching (Ritchie, 2006; Bourne et al., 2008; Koren and Rosenberg, 2008). It is hypothesized that bleaching leaves the coral host more susceptible to disease, presumably due to alterations in both its physiology and its coral-associated microbial assemblages. Several studies have documented a link between bleaching events and subsequent outbreaks of disease (Guzman and Guevara, 1998; Harvell et al., 2001; Muller et al., 2008; Brandt and McManus, 2009; Croquer and Weil, 2009; McClanahan et al., 2009; Miller et al., 2009) further supporting this hypothesis.

To date, little work has been done to assess the role that ocean acidification will have on coral microbiota. The pH of the coral microenvironment is dynamic, changing in both space and in time. For instance, intracellular pH in the coral *Stylophora pistillata* ranges from 7.13 in the light to 7.41 in the dark (Venn et al., 2009), while the pH of the coral surface in *Favia* sp. varies from 7.3 in the dark to 8.5 in the light (Kuhl et al., 1995). Thus, bacteria that colonize the coral

microhabitat must be able to withstand diurnal fluctuations in pH associated with algal photosynthesis.

Despite being exposed to a large range of pH, there is some evidence that increasing acidity leads to variability in coral-associated microbiota. Similar to increasing temperatures, Vega Thurber (2009) found that decreasing the pH of seawater to 7.4 shifted the microbiome of *P. compressa* to a more disease-associated state. The mechanisms driving this shift are unknown, but like other environmental processes that drive changes in the structure of coral-associated bacterial assemblages, a complex interaction between direct and indirect effects on the coral holobiont is hypothesized. For instance, pH is an important factor regulating virulence pathways in other pathogens (Nakayama and Watanabe, 1995; Li et al., 2007; Fuentes et al., 2009; Gong et al., 2009; Werbrouck et al., 2009), and while this has not been investigated in corals, it has been shown that some pathogens of corals have an optimal growth rate within the range of the pH occurring in the coral microhabitat (Remily and Richardson, 2006; Rasoulouniriana et al., 2009). Furthermore, other than decreasing accretion rates, it is unknown how ocean acidification will alter the physiology and susceptibility to disease of the host. It is also possible that synergisms between increasing temperatures and decreasing pH could cause variation in coral-bacteria assemblages. For instance, Remily and Richardson (2006) found that increasing temperatures expanded the tolerance to pH of *Aurantimonas coralicida*, the causative agent of white plague II in the Caribbean. Therefore, the synergisms between the two environmental factors may enable niche expansion of potentially pathogenic bacteria.

Conclusions

Coral-microorganism interactions represent a useful model for the types of associations that are likely important for many marine invertebrates. Coral-bacteria assemblages have been relatively well studied because of the recognized role of bacteria in the biology of the coral holobiont, and the large climate-mediated stress coral reef ecosystems have suffered. From this emerging area we now know that 1) corals associate with a diverse array of bacteria, and some of these associations are species-specific, 2) bacteria can contribute both antibiotic resistance and some nutrient-cycling capabilities to their coral host, 3) there are differences in bacterial associations between healthy corals and those that are bleached and/or diseased, and 4) climate-driven temperature stress can alter coral-bacteria assemblages to become more characteristic of those found in diseased corals. However, huge gaps in knowledge remain regarding the function of coral-bacteria associations, the specificity of these associations, and the anticipated impact of climate change. The potential exists for very small modifications in temperature or pH associated with climate change to increase the variability of coral-bacterial populations and in turn affect the health, life history, and species composition of coral reefs.

REFERENCES

Ainsworth, T.D., Vega Thurber, R., and Gates, R.D. 2010. The future of coral reefs: a microbial perspective. Trends Ecol Evol 25:233-240.

Alker, A.P., Smith, G.W., and Kim, K. 2001. Characterization of *Aspergillus sydowii*, a fungal pathogen of Caribbean sea fan corals. Hydrobiologia 460:105-111.

Allers, E., Niesner, C., Wild, C., and Pernthaler, J. 2008b. Microbes enriched in seawater after addition of coral mucus. Appl Environ Microbiol 74:3274-3278.

Apprill, A., Marlow, H.Q., Martindale, M.Q., and Rappe, M.S. 2009. The onset of microbial associations in the coral *Pocillopora meandrina*. Isme J 3:685-699.

Arotsker, L., Siboni, N., Ben-Dov, E., Kramarsky-Winter, E., Loya, Y., and Kushmaro, A. 2009. *Vibrio* sp as a potentially important member of the Black Band Disease (BBD) consortium in *Favia* sp corals. FEMS Microbiol Ecol 70:515-524.

Azam, F. 1998. Microbial control of oceanic carbon flux: The plot thickens. Science 280:694-696.

Baker, A.C., Glynn, P.W., and Riegl, B. 2008. Climate change and coral reef bleaching: An ecological assessment of long-term impacts, recovery trends and future outlook. Estaur Coast Shelf S 80:435-471.

Balser, T.C., McMahon, K.D., Bart, D., Bronson, D., Coyle, D.R., Craig, N., Flores-Mangual, M.L., Forshay, K., Jones, S.E., Kent, A.E., and Shade, A.L. 2006. Bridging the gap between micro- and macro-scale perspectives on the role of microbial communities in global change ecology. Plant Soil 289:59-70.

Barneah, O., Ben-Dov, E., Kramarsky-Winter, E., and Kushmaro, A. 2007. Characterization of black band disease in Red Sea stony corals. Environ Microbiol 9:1995-2006.

Ben-Dov, E., Kramarsky-Winter, E., and Kushmaro, A. 2009. An in situ method for cultivating microorganisms using a double encapsulation technique. FEMS Microbiol Ecol 68:363-371.

Ben-Haim, Y., Zicherman-Keren, M., and Rosenberg, E. 2003. Temperature-regulated bleaching and lysis of the coral *Pocillopora damicornis* by the novel pathogen *Vibrio coralliilyticus*. Appl Environ Microbiol 69:4236-4242.

Bourne, D., Iida, Y., Uthicke, S., and Smith-Keune, C. 2008. Changes in coral-associated microbial communities during a bleaching event. Isme J 2:350-363.

Bourne, D.G. 2005. Microbiological assessment of a disease outbreak on corals from Magnetic Island (Great Barrier Reef, Australia). Coral Reefs 24:304-312.

Bourne, D.G., Garren, M., Work, T.M., Rosenberg, E., Smith, G.W., and Harvell, C.D. 2009. Microbial disease and the coral holobiont. Trends Microbiol 17:554-562.

Bourne, D.G., and Munn, C.B. 2005. Diversity of bacteria associated with the coral *Pocillopora damicornis* from the Great Barrier Reef. Environ Microbiol 7:1162-1174.

Brandt, M.E., and McManus, J.W. 2009. Disease incidence is related to bleaching extent in reef-building corals. Ecology 90:2859-2867.

Breitbart, M., Bhagooli, R., Griffin, S., Johnston, I., and Rohwer, F. 2005. Microbial communities associated with skeletal tumors on *Porites compressa*. FEMS Microbiol Lett 243:431-436.

Bruno, J.F., Harvell, C.D., Hettinger, A., and Petes, L.E. 2003. Nutrient enrichment can increase the severity of coral diseases. Ecol Lett 6:1056-1061.

Cervino, J.M., Hayes, R.L., Polson, S.W., Polson, S.C., Goreau, T.J., Martinez, R.J., and Smith, G.W. 2004. Relationship of *Vibrio* species infection and elevated temperatures to yellow blotch/band disease in Caribbean corals. Appl Environ Microbiol 70:6855-6864.

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

Coles, S.L., and Brown, B.E. 2003. Coral bleaching - Capacity for acclimatization and adaptation. Pp. 183-223. Advances in Marine Biology, Vol 46. Academic Press Ltd, London.

Cooney, R.P., Pantos, O., Le Tissier, M.D.A., Barer, M.R., O'Donnell, A.G., and Bythell, J.C. 2002. Characterization of the bacterial consortium associated with black band disease in coral using molecular microbiological techniques. Environ Microbiol 4:401-413.

Croquer, A., and Weil, E. 2009. Changes in Caribbean coral disease prevalence after the 2005 bleaching event. Dis Aquat Organ 87:33-43.

De'ath, G., Lough, J.M., and Fabricius, K.E. 2009. Declining Coral Calcification on the Great Barrier Reef. Science 323:116-119.

Di Salvo, L., and Gundersen, K. 1971. Regenerative functions and microbial ecology of coral reefs Part 1: Assay for microbial population. Can J Microbiol 17:1081-1089.

Doney, S.C., Balch, W.M., Fabry, V.J., and Feely, R.A. 2009. Ocean acidification: a critical emerging problem for the ocean sciences. Oceanography 22:16-26.

Feely, R.A., Doney, S.C., and Cooley, S.R. 2009. Ocean Acidification: Present Conditions and Future Changes in a High-CO_2 World. Oceanography 22:36-47.

Ferrier-Pages, C., Gattuso, J.P., Cauwet, G., Jaubert, J., and Allemand, D. 1998. Release of dissolved organic carbon and nitrogen by the zooxanthellate coral *Galaxea fascicularis*. Mar Ecol-Prog Ser 172:265-274.

Ferrier-Pages, C., Leclercq, N., Jaubert, J., and Pelegri, S.P. 2000. Enhancement of pico- and nanoplankton growth by coral exudates. Aquat Microb Ecol 21:203-209.

Flatt, P., Gautschi, J., Thacker, R., Musafija-Girt, M., Crews, P., and Gerwick, W. 2005. Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge *Dysidea (*Lamellodysidea*) herbacea* and symbiotic cyanobacterium *Oscillatoria spongeliae* by CARD-FISH analysis. Mar Biol 147:761-774.

Frias-Lopez, J., Klaus, J.S., Bonheyo, G.T., and Fouke, B.W. 2004. Bacterial community associated with black band disease in corals. Appl Environ Microbiol 70:5955-5962.

Frias-Lopez, J., Zerkle, A.L., Bonheyo, G.T., Fouke, and B.W. 2002. Partitioning of bacterial communities between seawater and healthy, black band diseased, and dead coral surfaces. Appl Environ Microbiol 68:2214-2228.

Fuentes, J.A., Jofre, M.R., Villagra, N.A., and Mora, G.C. 2009. RpoS- and Crp-dependent transcriptional control of *Salmonella typhi* taiA and hlyE genes: role of environmental conditions. Res Microbiol 160:800-808.

Fuhrman, J.A., Hewson, I., Schwalbach, M.S., Steele, J.A., Brown, M.V., and Naeem, S. 2006. Annually reoccurring bacterial communities are predictable from ocean conditions. Proc Natl Acad Sci USA 103:13104-13109.

Garren, M., Raymundo, L., Guest, J., Harvell, C.D., and Azam, F. 2009. Resilience of Coral-Associated Bacterial Communities Exposed to Fish Farm Effluent. PLoS One 4:9.

Garren, M., Smriga, S., and Azam, F. 2008. Gradients of coastal fish farm effluents and their effect on coral reef microbes. Environ Microbiol 10:2299-2312.

Gil-Agudelo, D.L., Fonseca, D.P., Weil, E., Garzon-Ferreira, J., and Smith, G.W. 2007. Bacterial communities associated with the mucopolysaccharide layers of three coral species affected and unaffected with dark spots disease. Can J Microbiol 53:465-471.

Gil-Agudelo, D.L., Myers, C., Smith, G.W., and Kim, K. 2006. Changes in the microbial communities associated with *Gorgonia ventalina* during aspergillosis infection. Dis Aquat Organ 69:89-94.

Giovannoni, S.J., and Stingl, U. 2005. Molecular diversity and ecology of microbial plankton. Nature 437:343-348.

Gong, Y.X., Tian, X.L., Sutherland, T., Sisson, G., Mai, J.N., Ling, J.Q., and Li, Y.H. 2009. Global transcriptional analysis of acid-inducible genes in *Streptococcus mutans*: multiple two-component systems involved in acid adaptation. Microbiology-Sgm 155:3322-3332.

Guppy, R., and Bythell, J.C. 2006. Environmental effects on bacterial diversity in the surface mucus layer of the reef coral *Montastraea faveolata*. Mar Ecol-Prog Ser 328:133-142.

Guzman, H.M., and Guevara, C. 1998. Massive mortality of zooxanthelate reef organisms during the 1995 bleaching in Cayos Cochinos, Honduras. Rev Biol Trop 46:165-173.

Harvell, C.D., Kim, K., Burkholder, J.M., Colwell, R.R., Epstein, P.R., Grimes, D.J., Hofmann, E.E., Lipp, E.K., Osterhaus, A.D.M.E., Overstreet, R.M., Porter, J.W., Smith, G.W., and Vasta, G.R. 1999. Emerging marine diseases-Climate links and anthropogenic factors. Science 285:1505-1510.

Harvell, C.D., Mitchell, C.E., Ward, J.R., Altizer, S., Dobson, A.P., Ostfeld, R.S., and Samuel, M.D. 2002. Ecology - Climate warming and disease risks for terrestrial and marine biota. Science 296:2158-2162.

Harvell, C.D. 2004. Ecology and evolution of host-pathogen interactions in nature. Am Nat 164:S1-S5.

Harvell, C.D., Jordan-Dahlgren, E., Merkel, S., Rosenberg, E., Raymundo, L., Smith, G., Weil, E., and Willis, B., Global Envrionm Facility C. 2007. Coral disease, environmental drivers, and the balance between coral and microbial associates. Oceanography 20:172-195.

Harvell, C.D., Kim, K., Quirolo, C., Weir, J., and Smith, G. 2001. Coral bleaching and disease: Contributors to 1998 mass mortality in *Briareum asbestinum* (*Octocorallia, Gorgonacea*). Hydrobiologia 460:97-104.

Hewson, I., and Fuhrman, J.A. 2004. Richness and diversity of bacterioplankton species along an estuarine gradient in Moreton Bay, Australia. Appl Environ Microbiol 70:3425-3433.

Hewson, I., Jacobson-Meyers, M., and Fuhrman, J. 2007. Diversity and biogeography of bacterial assemblages in surface sediments across the San Pedro Basin, Southern California Borderlands. Environmental Microbiology 9:923-933.

Hewson, I., Steele, J.A., Capone, D.G., and Fuhrman, J.A. 2006a. Remarkable heterogeneity in meso- and bathypelagic bacterioplankton assemblage composition. Limnol Oceanogr 51:1274-1283.

Hewson, I., Steele, J.A., Capone, D.G., and Fuhrman, J.A. 2006b. Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters. Mar Ecol Prog Ser 311:67-77.

Hildebrand, M., Waggoner, L.E., Lim, G.E., Sharp, K.H., Ridley, C.P., and Haygood, M.G. 2004. Approaches to identify, clone, and express symbiont bioactive metabolite genes. Nat Prod Rep 21:122-142.

Hoegh-Guldberg, O. 1999. Climate change, coral bleaching and the future of the world's coral reefs. Mar Freshw Res 50:839 - 866.

Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., Harvell, C.D., Sale, P.F., Edwards, A.J., Caldeira, K., Knowlton, N., Eakin, C.M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R.H., Dubi, A., and Hatziolos, M.E. 2007. Coral reefs under rapid climate change and ocean acidification. Science 318:1737-1742.

Hong MJ, Yu YT, Chen CA, Chiang PW, Tang SL. 2009. Influence of species Specificity and other factors on bacteria associated with the coral *Stylophora pistillata* in Taiwan. Appl Environ Microbiol 75:7797-7806.

Hughes, T.P., Baird, A.H., Bellwood, D.R., Card, M., Connolly, S.R., Folke, C., Grosberg, R., Hoegh-Guldberg, O., Jackson, J.B.C., Kleypas, J., Lough, J.M., Marshall, P., Nystrom, M., Palumbi, S.R., Pandolfi, J.M., Rosen, B., and Roughgarden, J. 2003. Climate change, human impacts, and the resilience of coral reefs. Science 301:929-933.

Ionescu, D., Penno, S., Haimovich, M., Rihtman, B., Goodwin, A., Schwartz, D., Hazanov, L., Chernihovsky, M., Post, A.F., and Oren, A. 2009. Archaea in the Gulf of Aqaba. FEMS Microbiol Ecol 69:425-438.

Jensen, S., Neufeld, J.D., Birkeland, N.K., Hovland, M., and Murrell, J.C. 2008. Insight into the microbial community structure of a Norwegian deep-water coral reef environment. Deep-Sea Res Part I-Oceanogr Res Pap 55:1554-1563.

Kapley, A., Siddiqui, S., Misra, K., Ahmad, S.M., and Purohit, H.J. 2007. Preliminary analysis of bacterial diversity associated with the *Porites* coral from the Arabian Sea. World J Microbiol Biotechnol 23:923-930.

Kellogg, C.A. 2004. Tropical Archaea: diversity associated with the surface microlayer of corals. Mar Ecol-Prog Ser 273:81-88.

Kellogg, C.A., Lisle, J.T., and Galkiewicz, J.P. 2009. Culture-independent characterization of bacterial communities associated with the cold-water coral *Lophelia pertusa* in the Northeastern Gulf of Mexico. Appl Environ Microbiol 75:2294-2303.

Kimes, N.E., Van Nostrand, J.D., Weil, E., Zhou, J.Z., and Morris, P.J. 2010. Microbial functional structure of *Montastraea faveolata*, an important Caribbean reef-building coral, differs between healthy and yellow-band diseased colonies. Environ Microbiol 12:541-556.

Klaus, J.S., Frias-Lopez, J., Bonheyo, G.T., Heikoop, J.M., and Fouke, B.W. 2005. Bacterial communities inhabiting the healthy tissues of two Caribbean reef corals: interspecific and spatial variation. Coral Reefs 24:129-137.

Kleypas, J.A., and Yates, K.K. 2009. Coral reefs and ocean Acidification. Oceanography 22:108-117.

Kline, D.I., Kuntz, N.M., Breitbart, M., Knowlton, N., and Rohwer, F. 2006. Role of elevated organic carbon levels and microbial activity in coral mortality. Mar Ecol Prog Ser 314:119-125.

Kooperman, N., Ben-Dov, E., Kramarsky-Winter, E., Barak, Z., and Kushmaro, A. 2007. Coral mucus-associated bacterial communities from natural and aquarium environments. FEMS Microbiol Lett 276:106-113.

Koren, O., and Rosenberg, E. 2006. Bacteria associated with mucus and tissues of the coral *Oculina patagonica* in summer and winter. Appl Environ Microbiol 72:5254-5259.

Koren, O., and Rosenberg, E. 2008. Bacteria associated with the bleached and cave coral *Oculina patagonica*. Microb Ecol 55:523-529.

Kuhl, M., Cohen, Y., Dalsgaard, T., Jorgensen, B.B., and Revsbech, N.P. 1995. Microenvironment and photosynthesis of zooxanthellae in scleractinian corals studied with microsensors for O_2, pH, and light. Mar Ecol-Prog Ser 117:159-172.

Lampert, Y., Kelman, D., Nitzan, Y., Dubinsky, Z., Behar, A., and Hill, R.T. 2008. Phylogenetic diversity of bacteria associated with the mucus of Red Sea corals. FEMS Microbiol Ecol 64:187-198.

Lesser, M.P., Falcon, L.I., Rodriguez-Roman, A., Enriquez, S., Hoegh-Guldberg, O., and Iglesias-Prieto, R. 2007. Nitrogen fixation by symbiotic cyanobacteria provides a source of nitrogen for the scleractinian coral *Montastraea cavernosa*. Mar Ecol Prog Ser 346:143-152.

Lesser, M.P., Mazel, C.H., Gorbunov, M.Y., and Falkowski, P.G. 2004. Discovery of symbiotic nitrogen-fixing cyanobacteria in corals. Science 305:997-1000.

Li, M., Liang, Y., Wei, D., and Xing, L. 2007. The response to environmental pH of RIM101 pathway in *Candida albicans*. Wei Sheng Wu Xue Bao 47:366-369.

Littman, R.A., Willis, B.L., and Bourne, D.G. 2009a. Bacterial communities of juvenile corals infected with different *Symbiodinium* (dinoflagellate) clades. Mar Ecol Prog Ser 389:45-59.

Littman, R.A., Willis, B.L., Pfeffer, C., and Bourne, D.G. 2009b. Diversities of coral-associated bacteria differ with location, but not species, for three acroporid corals on the Great Barrier Reef. FEMS Microbiol Ecol 68:152-163.

Long, R.A., and Azam, F. 2001. Microscale patchiness of bacterioplankton assemblage richness in seawater. Aquat Microb Ecol 26:103-113.

Lopanik, N., Lindquist, N., and Targett, N. 2004. Potent cytotoxins produced by a microbial symbiont protect host larvae from predation. Oecologia 139:131-139.

McClanahan, T.R., Weil, E., and Maina, J. 2009. Strong relationship between coral bleaching and growth anomalies in massive *Porites*. Glob Change Biol 15:1804-1816.

Miller, J., Muller, E., Rogers, C., Waara, R., Atkinson, A., Whelan, K.R.T., Patterson, M., and Witcher, B. 2009. Coral disease following massive bleaching in 2005 causes 60% decline in coral cover on reefs in the US Virgin Islands. Coral Reefs 28:925-937.

Moeseneder, M.M., Winter, C., and Herndl, G.J. 2001. Horizontal and vertical complexity of attached and free-living bacteria of the eastern Mediterranean Sea, determined by 16S rDNA and 16S rRNA fingerprints. Limnol Oceanogr 46:95-107.

Muller, E.M., Rogers, C.S., Spitzack, A.S., and van Woesik, R. 2008. Bleaching increases likelihood of disease on *Acropora palmata* (Lamarck) in Hawksnest Bay, St John, US Virgin Islands. Coral Reefs 27:191-195.

Muscatine, L. 1973. Nutrition of Corals. Biology and Geology of Coral Reefs. Acadmic Press.

Nakayama, S., and Watanabe, H. 1995. Involvement of cpxA, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei* virF gene. J Bacteriol 177:5062-5069.

Nissimov, J., Rosenberg, E., and Munn, C.B. 2009. Antimicrobial properties of resident coral mucus bacteria of *Oculina patagonica*. FEMS Microbiol Lett 292:210-215.

Olsen, G.J., Lane, D.J., Giovannoni, S.J., Pace, N.R., and Stahl, D.A. 1986. Microbial ecology and evolution-a ribosomal RNA approach. Annu Rev Microbiol 40:337-365.

Olson, N.D., Ainsworth, T.D., Gates, R.D., and Takabayashi, M. 2009. Diazotrophic bacteria associated with Hawaiian Montipora corals: Diversity and abundance in correlation with symbiotic dinoflagellates. J Exp Mar Biol Ecol 371:140-146.

Pantos, O., and Bythell, J.C. 2006. Bacterial community structure associated with white band disease in the elkhorn coral *Acropora palmata* determined using culture-independent 16S rRNA techniques. Dis Aquat Organ 69:79-88.

Pantos, O., Cooney, R.P., Le Tissier, M.D.A., Barer, M.R., O'Donnell, A.G., and Bythell, J.C. 2003. The bacterial ecology of a plague-like disease affecting the Caribbean coral *Montastrea annularis*. Environ Microbiol 5:370-382.

Piggot, A.M., Fouke, B.W., Sivaguru, M., Sanford, R.A., and Gaskins, H.R. 2009. Change in zooxanthellae and mucocyte tissue density as an adaptive response to environmental stress by the coral, *Montastraea annularis*. Mar Biol 156:2379-2389.

Raina, J., Tapiolas, D., Willis, B.L., and Bourne, D.G. 2009. Coral-associated bacteria and their role in the biogeochemical cycling of sulfur. Appl Environ Microbiol 75:3492-3501.

Rasoulouniriana, D., Siboni, N., Ben-Dov, E., Kramarsky-Winter, E., Loya, Y., and Kushmaro, A. 2009. *Pseudoscillatoria coralii* gen. nov., sp nov., a cyanobacterium associated with coral black band disease (BBD). Dis Aquat Organ 87:91-96.

Reis, A.M.M., Araujo, S.D., Moura, R.L., Francini, R.B., Pappas, G., Coelho, A.M.A., Kruger, R.H., and Thompson, F.L. 2009. Bacterial diversity associated with the Brazilian endemic reef coral *Mussismilia braziliensis*. J Appl Microbiol 106:1378-1387.

Remily, E.R., and Richardson, L.L. 2006. Ecological physiology of a coral pathogen and the coral reef environment. Microb Ecol 51:345-352.

Reshef, L., Koren, O., Loya, Y., Zilber-Rosenberg, I., and Rosenberg, E. 2006. The coral probiotic hypothesis. Environ Microbiol 8:2068-2073.

Ridley, C.P., Bergquist, P.R., Harper, M.K., Faulkner, D.J., Hooper, J.N.A., and Haygood, M.G. 2005. Speciation and biosynthetic variation in four dictyoceratid sponges and their cyanobacterial symbiont, *oscillatoria spongeliae*. Chem Biol 12:397-406.

Ritchie, K.B. 2006. Regulation of microbial populations by coral surface mucus and mucus-associated bacteria. Mar Ecol-Prog Ser 322:1-14.

Ritchie, K.B., and Smith, G.W. 1995. Preferential carbon utilization by surface bacterial communities from water mass, normal, and white-band diseased *Acropora cervicornis*. Mol Mar Biol Biotech 4:345-352.

Ritchie, K.B., and Smith, G.W. 2004. Microbial communities of coral surface mucopolysaccharide layers. Springer-Verlag.

Rohwer, F., Breitbart, M., Jara, J., Azam, F., and Knowlton, N. 2001. Diversity of bacteria associated with the Caribbean coral *Montastraea franksi*. Coral Reefs 20:85-91.

Rohwer, F., and Kelley, S. 2004. Culture-independent analyses of coral-associated microbes. Springer-Verlag.

Rohwer, F., Seguritan, V., Azam, F., and Knowlton, N. 2002. Diversity and distribution of coral-associated bacteria. Mar Ecol Prog Ser 243:1-10.

Rosenberg, E., and Falkovitz, L. 2004. The *Vibrio shiloi/Oculina patagonica* model system of coral bleaching. Annu Rev Microbiol 58:143-159.

Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. 2007. The role of microorganisms in coral health, disease and evolution. Nat Rev Microbiol 5:355-362.

Rypien, K.L., Ward, J.R., and Azam, F. 2009. Antagonistic interactions among coral-associated bacteria. Environ Microbiol 12:28-39.

Sato, Y., Willis, B.L., and Bourne, D.G. 2009. Successional changes in bacterial communities during the development of black band disease on the reef coral, *Montipora hispida*. Isme J 4:203-214.

Scheffers, S.R., Bak, R.P.M., and van Duyl, F.C. 2005. Why is bacterioplankton growth in coral reef framework cavities enhanced? Mar Ecol-Prog Ser 299:89-99.

Sekar, R., Kaczmarsky, L.T., and Richardson, L.L. 2008. Microbial community composition of black band disease on the coral host *Siderastrea siderea* from three regions of the wider Caribbean. Mar Ecol-Prog Ser 362:85-98.

Sekar, R., Mills, D.K., Remily, E.R., Voss, J.D., and Richardson, L.L. 2006. Microbial communities in the surface mucopolysaccharide layer and the black band microbial mat of black band-diseased *Siderastrea siderea*. Appl Environ Microbiol 72:5963-5973.

Shashar, N., Cohen, Y., Loya, Y., and Sar, N. 1994. Nitrogen fixation (acetylene reduction) in stony corals: Evidence for coral-bacteria interactions. Mar Ecol Prog Ser 111:259-264.

Shnit-Orland, M., and Kushmaro, A. 2009. Coral mucus-associated bacteria: a possible first line of defense. FEMS Microbiol Ecol 67:371-380.

Siboni, N., Ben-Dov, E., Sivan, A., and Kushmaro, A. 2008. Global distribution and diversity of coral-associated Archaea and their possible role in the coral holobiont nitrogen cycle. Environ Microbiol 10:2979-2990.

Smith, J.E., Shaw, M., Edwards, R.A., Obura, D., Pantos, O., Sala, E., Sandin, S.A., Smriga, S., Hatay, M., and Rohwer, F.L. 2006. Indirect effects of algae on coral: algae-mediated, microbe-induced coral mortality. Ecol Lett 9:835-845.

Sogin, M.L., Morrison, H.G., Huber, J.A., Mark, Welch, D., Huse, S.M., Neal, P.R., Arrieta, J.M., and Herndl, G.J. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc Nat Acad Sci USA 103:12115-12120.

Sunagawa, S., DeSantis, T.Z., Piceno, Y.M., Brodie, E.L., DeSalvo, M.K., Voolstra, C.R., Weil, E., Andersen, G.L., and Medina, M. 2009. Bacterial diversity and White Plague Disease-associated community changes in the Caribbean coral *Montastraea faveolata*. Isme J 3:512-521.

Sussman, M., Willis, B.L., Victor. S., Bourne, D.G. 2008. Coral pathogens identified for White Syndrome (WS) epizootics in the Indo-Pacific. PLoS One 3:e2393.

Szmant, A.M., Ferrer, L.M., and FitzGerald, L.M. 1990. Nitrogen excretion and O:N ratios in reef corals: evidence for conservation of nitrogen. Mar Biol 104:119-127.

Thurber, R.L.V., Barott, K.L., Hall, D., Liu, H., Rodriguez-Mueller, B., Desnues, C., Edwards, R.A., Haynes, M., Angly, F.E., Wegley, L., and Rohwer, F.L. 2008. Metagenomic analysis indicates that stressors induce production of herpes-like viruses in the coral *Porites compressa*. Proc Natl Acad Sci USA 105:18413-18418.

Thurber, R.V., Willner-Hall, D., Rodriguez-Mueller, B., Desnues, C., Edwards, R.A., Angly, F., Dinsdale, E., Kelly, L., and Rohwer, F. 2009. Metagenomic analysis of stressed coral holobionts. Environ Microbiol 11:2148-2163.

van Duyl, F.C., and Gast, G.J. 2001. Linkage of small-scale spatial variations in DOC, inorganic nutrients and bacterioplankton growth with different coral reef water types. Aquat Microb Ecol 24:17-26.

Venn, A.A., Tambutte, E., Lotto, S., Zoccola, D., Allemand, D., and Tambutte, S. 2009. Imaging intracellular pH in a reef coral and symbiotic anemone. Proc Natl Acad Sci USA 106:16574-16579.

Voss, J.D., Mills, D.K., Myers, J.L., Remily, E.R., and Richardson, L.L. 2007. Black band disease microbial community variation on corals in three regions of the wider Caribbean. Microb Ecol 54:730-739.

Voss, J.D., and Richardson, L.L. 2006. Nutrient enrichment enhances black band disease progression in corals. Coral Reefs 25:569-576.

Ward, J.R., Kim, K., and Harvell, C.D. 2007. Temperature affects coral disease resistance and pathogen growth. Mar Ecol Prog Ser 329:115-121.

Wegley, L., Edwards, R., Rodriguez-Brito, B., Liu, H., and Rohwer, F. 2007. Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. Environ Microbiol 9:2707-2719.

Weil, E., Smith, G., and Gil-Agudelo, D.L. 2006. Status and progress in coral reef disease research. Dis Aquat Organ 69:1-7.

Werbrouck, H., Vermeulen, A., Van Coillie, E., Messens, W., Herman, L., Devlieghere, F., and Uyttendaele, M. 2009. Influence of acid stress on survival, expression of virulence genes and invasion capacity into Caco-2 cells of *Listeria* monocytogenes strains of different origins. Int J Food Microbiol 134:140-146.

Williams, W.M., Viner, A.B., and Broughton, W.J. 1987. Nitrogen fixation (acetylene reduction) associated with the living coral *Acropora variabilis*. Mar Biol 94:531-535.

CHAPTER 2

EXTENSIVE DIFFERENCES IN GENE EXPRESSION BETWEEN SYMBIOTIC AND APOSYMBIOTIC ANEMONES[1]

Abstract

Coral reefs provide habitats for a disproportionate number of marine species relative to the small area of the oceans that they occupy. The mutualism between the cnidarian animal hosts and their intracellular dinoflagellate symbionts provides the nutritional foundation for coral growth and formation of large reef structures, as algal photosynthesis can provide >90% of the host's total energy. The large-scale disruption of this symbiosis, known as 'coral bleaching', is due largely to anthropogenic factors and poses a major threat to the future of coral reefs. Despite the importance of this symbiosis, the cellular mechanisms involved in its establishment, maintenance, and dissolution remain largely unknown. Here we report our continued development of genomic tools to study these mechanisms in *Aiptasia*, a small sea anemone that is emerging as a powerful model system for studies of cnidarian-dinoflagellate symbiosis. Specifically, we report a *de novo* assembly of the transcriptomes both of symbiotic anemones from a clonal line and of their endogenous dinoflagellate symbionts. We then demonstrate the utility of this resource by comparing transcript abundances in these anemones to those of animals from the same clonal line but lacking dinoflagellates (aposymbiotic). This analysis led to the identification of >900 differentially expressed transcripts and has allowed us to generate testable biological hypotheses about what cellular functions are affected by symbiosis establishment. The differentially regulated transcripts include >60 encoding distinct proteins that may play roles in transporting nutrients between the symbiotic partners; many more encoding proteins

[1] Co-authors include E. M. Lehnert, M. S. Burriesci, N. D. Gallo, J. A. Schwarz, and J. R. Pringle

functioning in several metabolic pathways, providing clues as to how the transported nutrients may be used by the partners; and several encoding proteins that may be involved in host tolerance of the dinoflagellate.

Introduction

Coral reefs comprise only a small part of the world's ocean environment but are habitats for a disproportionately large fraction of all marine species. Corals are able to produce the massive and biologically rich reef habitats despite growing in nutrient-poor waters because of the energy acquired through their mutualistic symbiosis with dinoflagellates of the genus *Symbiodinium*. These unicellular algae inhabit the symbiosomes (endosome-derived vacuoles) of gastrodermal cells in corals and other cnidarians (Figure 2.1) and transfer up to 95% of their photosynthetically fixed carbon to the host (Muscatine et al., 1984). Reef-building corals are declining worldwide due largely to anthropogenic causes, which include pollution, destructive fishing practices, and increasing sea-surface temperatures (De'ath et al., 2012). Such stresses can lead to coral "bleaching," in which the algae lose their photosynthetic capacity and/or are lost altogether by the host. In severe cases, bleaching can result in the death of the host. This is particularly alarming because many corals already live near the upper limits of their thermal tolerances, and most climate-change models predict that these tolerances will frequently be exceeded in the coming decades, potentially leading to widespread coral bleaching and death and a resulting loss of the reef habitats (Hughes et al., 2003).

Despite the great ecological importance of cnidarian-dinoflagellate symbioses, little is known about the cellular and molecular mechanisms by which these relationships are

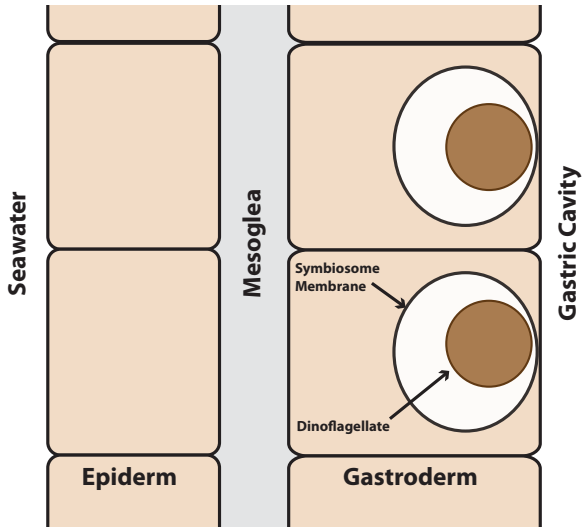


Figure 2.1 The spatial organization of cnidarian-dinoflagellate symbiosis. A simplified diagram of a section of cnidarian body wall is shown. The two major tissue layers are the epiderm, which faces the outside seawater and lacks both symbionts and direct access to food in the gastric cavity, and the gastroderm, which faces the gastric cavity and may contain dinoflagellate symbionts in some of its cells. These two cell layers are separated by the largely acellular mesoglea. After phagocytosis by a host gastrodermal cell, the dinoflagellate resides within a "symbiosome" (believed to be derived from a host endosome that does not fuse with lysosomes) and transfers fixed carbon to the host.

established, maintained, or disrupted. This situation has resulted in part from the difficulties inherent in studying corals directly (Weis, 2008). Thus, we and others have turned to the small sea anemone *Aiptasia*, which is normally symbiotic with dinoflagellates closely related to those found in corals but offers many experimental advantages (Lajeunesse et al., 2012). In particular, *Aiptasia* lacks the calcareous skeleton that renders biochemical and microscopic analyses of corals challenging, grows rapidly by asexual reproduction under standard aquarium conditions to form large clonal populations, can be induced to spawn and produce larvae throughout the year in the laboratory (Perez and Pringle, 2013), and (importantly for this study) can be maintained indefinitely in an aposymbiotic (bleached) state so long as it is fed regularly (Schoenberg and Trench, 1980).

The intracellular localization of the dinoflagellate (Figure 2.1) raises some key questions

about regulation of the symbiosis. First, how does the host recognize, take up, and maintain appropriate symbionts without generating a deleterious immune response that could result in a failure of algal uptake, digestion of the algae after uptake, or apoptosis of the host cells? Second, what metabolites do the two organisms exchange across the symbiosome membrane, and how? It seems likely that the symbiotic state involves both transporters and regulation of metabolic pathways that are distinct from those found in aposymbiotic animals. Third, what changes in transport occur at other membranes? For example, although both gastrodermal and epidermal cells in aposymbiotic anemones presumably excrete ammonium as a toxic waste product, as do other aquatic invertebrates (Wright, 1995), at least some of that ammonium must be redirected to the algae in symbiotic anemones. Particularly intriguing questions are how the epidermal tissue layer is nourished, as it lacks both dinoflagellate symbionts and direct access to food particles, and whether the nature and mechanisms of this nourishment change upon the establishment of symbiosis.

To begin to investigate these questions, we used RNA-Seq to generate an assembled and annotated transcriptome for symbiotic *Aiptasia*. This transcriptome was then used as a reference to compare global transcript abundances between symbiotic and aposymbiotic anemones. Previous studies using microarrays have identified few genes that were differentially expressed between the two states, possibly because of the low sensitivity of the technology, the low ratio of infected to uninfected cells, and/or an absence of probes for the relevant genes. In contrast, we identified nearly 1,000 genes in our study with significant and sometimes large expression differences between symbiotic and aposymbiotic states. Many of these suggest interesting and testable biological hypotheses.

Materials and Methods

Aiptasia strain and culture

All animals were from clonal population CC7 (Sunagawa et al., 2009), which in spawning experiments typically behaves as a male (Perez and Pringle, 2013). For experiments performed at Stanford, the stock cultures were grown in a circulating artificial seawater (ASW) system at ~25ºC with 20-40 µmol photons m^{-2} s^{-1} of photosynthetically active radiation (PAR) on an ~12 h light : 12 h dark (12L:12D) cycle and fed freshly hatched brine-shrimp nauplii approximately twice per week. To generate aposymbiotic anemones, animals were placed in a separate polycarbonate tub and subjected to several repetitions of the following process: cold-shocking by addition of 4ºC ASW and incubation at 4ºC for 4 h, followed by 1-2 days of treatment at ~25ºC in ASW containing the photosynthesis inhibitor diuron (Sigma-Aldrich D2425) at 50 µM (lighting approximately as above). After recovery for several weeks in ASW at ~25ºC in the light (as above) with feeding (as above, with a water change on the following day), putatively aposymbiotic anemones were inspected by fluorescence microscopy to confirm the complete absence of dinoflagellates (whose bright chlorophyll autofluorescence is conspicuous when they are present).

For experiments performed at Cornell, anemones were grown in incubators at 25°C in ASW in 1 L glass bowls and fed (as above) approximately three times per week. Symbiotic anemones were kept on a 12L:12D cycle at 18-22 µmol photons m^{-2} s^{-1} of PAR. Aposymbiotic animals were generated by exposing anemones under the same lighting and feeding regimen to 50 µM diuron in ASW, with daily water changes, for ~30 d or until the anemones were devoid of algae, as confirmed by fluorescence microscopy. Following bleaching, aposymbiotic anemones were maintained in the dark for ~2 years (with feeding as above) prior to experimentation.

Experimental design

Three separate experiments were performed using somewhat different conditions (Table 2.1). For Experiment 1 (RNA-Seq), both symbiotic and aposymbiotic anemones were held at 27°C on a 12L:12D cycle, with feeding and water changes as above, for one month before sampling to allow them to acclimate. The aposymbiotic anemones were checked immediately before sampling by fluorescence microscopy to ensure that they were still symbiont free. Anemones were collected ~2 d after the last feeding and ~5 h into the light period. Each of three biological replicates per condition consisted of two to five pooled anemones (for a total of ~35 mg wet weight); samples were stored in RNALater (Ambion AM7021) at -20ºC until processing.

For Experiment 2 (RNA-Seq), both symbiotic and aposymbiotic anemones were starved for 2 weeks before sampling. Symbiotic anemones were maintained at 25°C on a 12L:12D cycle, while aposymbiotic anemones were maintained at 25°C in constant dark. Anemones were collected 9 h into the symbiotic anemones' light period. Four symbiotic or eight aposymbiotic anemones (~50 mg wet weight in each case) were pooled in each of four biological replicates per treatment, flash frozen in liquid nitrogen, and held at -80ºC until processing.

For Experiment 3 (RT-qPCR), both symbiotic and aposymbiotic anemones were maintained at 25°C on a 12L:12D cycle with feeding every 2 d followed by water changes; samples were collected 2 d after the last feeding and 6 h into the light period. Four symbiotic or eight aposymbiotic anemones (~50 mg wet weight in each case), were pooled in each of four biological replicates per treatment, flash frozen in liquid nitrogen, and held at -80ºC until processing.

RNA isolation and sequencing

In Experiment 1, total RNA was extracted from whole anemones using the RNAqueous-4PCR Kit (Ambion AM1914) following the manufacturer's instructions. The RNA-integrity number (RIN) of each sample was determined using an Agilent 2100 Bioanalyzer, and only samples with a RIN ≥ 9 were used. ~3 µg of total RNA were processed (including a poly-A^+-selection step) using the TruSeq RNA Sample Prep Kit (Illumina FC-122-1001) following the manufacturer's instructions to produce indexed libraries. The resulting libraries were pooled based on their indices (as described in the kit instructions), and clustering and sequencing (both 101-bp paired-end reads and 36-bp single-end reads) were performed by the Stanford Center for Genomics and Personalized Medicine using an Illumina HiSeq 2000 sequencer.

In Experiments 2 and 3, total RNA was extracted using the ToTALLY RNA™ Total RNA Isolation Kit (Ambion AM1910) following the manufacturer's instructions, except that the RNA was precipitated using 0.1 volume of 3 M sodium acetate and 4 volumes of 100% ethanol. The resulting RNA was purified using the RNA Clean and Concentrator™-25 Kit (Zymo Research R1017). For RNA-Seq, the RIN of each sample was verified to be ≥9 using an Agilent 2100 Bioanalyzer, and ~4 µg of total RNA per sample were processed using the TruSeq Kit (as above) to produce indexed libraries. The resulting libraries were pooled into 8 samples per lane, and clustering and sequencing (101-bp paired-end reads) were performed by the Cornell Life Sciences Core Laboratory Center using an Illumina HiSeq 2000 sequencer. Processing of samples for Reverse Transcriptase quantitative PCR (RT-qPCR) is described below.

Read filtering and transcriptome assembly and annotation

Transcriptome assembly used all of the 101-bp paired-end reads obtained from symbiotic

anemones at both Stanford and Cornell (Table 2.1; reads available through NCBI Short-Read Archive, accession numbers SRR610288, SRR612165, and SRR696732). Prior to assembly, the reads were processed as follows: (1) reads of <60 bp or containing ≥1 N were discarded; (2) any read for which <25 of the first 35 bases had quality scores >30 was discarded; and (3) reads were trimmed to the first position for which a sliding 4-bp window had an average quality-score of <20. The remaining read-pairs were then processed using FLASH to join reads whose ends overlapped by ≥10 bp with no mismatches (Magoc and Salzberg, 2011). Finally, adapter sequencers were removed using cutadapt with default settings (Martin, 2011).

The processed reads were assembled in three sets due to memory constraints. Each set was assembled using an additive-multiple-*k*-mer approach (*k*-mers of 51, 59, 67, 75, 83, 91) with the Velvet/Oases assembler (Velvet version 1.1.07 and Oases version 0.2.02; Zerbino and Birney, 2008; Schultz et al., 2012) and merged using the Oases merge function with a *k*-mer of 27. The final outputs of each assembly were merged with one another using the Oases merge function again. Near-identical contigs (≥99% identical over their entire lengths) were merged using UCLUST v. 5.2.32 (Edgar, 2010). To cluster alternative transcripts from the same gene (and presumably also transcripts from highly similar paralogs), UCLUST was used again, as follows. Contigs were aligned locally in both directions and clustered together if the alignment consisted of ≥20% of the total length of each contig and the sequence was ≥99% identical over the alignment. These parameters were chosen because they produced valid clusters on a test dataset from zebrafish in 93% of cases (with the remaining cases being mostly the near-identical paralogs common in teleosts due to genome duplication) (EML and B. Benayoun, unpublished results).

To assign putative functional roles to the transcripts, we aligned them to the SwissProt

Table 2.1 Summary of experimental conditions.

Experiment	Site	Purpose	Light (µmol photons m^{-2} s^{-1})	Temperature (°C)	Feeding schedule
1, Apo	Stanford	Gene expression [a]	25 (12L:12D)	27	Every 2 d
1, Sym	Stanford	Transcriptome assembly and gene expression [b]	25 (12L:12D)	27	Every 2 d
2, Apo	Cornell	Gene expression[c]	0	25	Unfed 2 weeks
2, Sym	Cornell	Transcriptome assembly and gene expression [d]	18-22 (12L:12D)	25	Unfed 2 weeks
3, Apo	Cornell	RT-qPCR	18-22 (12L:12D)	25	Every 2 d
3, Sym	Cornell	RT-qPCR	18-22 (12L:12D)	25	Every 2 d

[a] ~49 million 36-bp single-end reads (Accession Number SRR612167).
[b] ~200 million 101-bp paired-end reads (Accession Number SRR610288) and 51 million 36-bp single end reads (Accession Number SRR612166).
[c] ~80 million 101-bp paired-end reads (Accession Number SRR612165).
[d] ~83 million 101-bp paired-end reads (Accession Number SRR696732).

protein database and the NCBI Non-Redundant Protein Database (nr) using the blastx program from the standalone BLAST 2.2.25+ software suite with an E-value cutoff of 1e-5 Camacho et al., 2009). The results of the alignment to SwissProt were imported using the Blast2GO software package and used to assign Enzyme Codes and Gene Ontology (GO) terms to the predicted proteins (Conesa et al., 2005; Ashburner et al., 2000).

Classification of contig origin using a transcript-sorting algorithm and alignment of genomic reads

To classify contigs into those derived from *Aiptasia,* those from the dinoflagellate symbionts, and those from other aquarium organisms that might have been associated with the mucous coats of the isolated anemones, we developed the machine-learning program TopSort, which uses support vector machines to classify transcripts as cnidarian, dinoflagellate, fungal, or bacterial (Burriesci, 2011). TopSort's basic principle is that if there are N features for each element in a dataset, each element can be represented as a point defined by these features in N-dimensional hyperspace. If classes of elements are distinguishable by the N features, then there should be a N-1-dimensional hyperplane that cuts the space such that one class can be separated from the others. If elements are clustered such that they are separable by another shape in N-space (e.g., a hypersphere), then an appropriate transform of the hyperspace to another space will make them separable by a hyperplane. The features used for TopSort were GC content; amino-acid and codon biases (where a strong BLAST hit allowed a reliable prediction of reading frame); phylogenetic classification of the top five best BLAST hits to the nr database (scoring each hit as cnidarian, non-cnidarian animal, dinoflagellate, non-dinoflagellate alveolate, plant, fungus, bacteria, or none-of-the-above); and best BLAST hit to a custom database composed of

the sequences of known origin that were not chosen for either the training or test set (see below and Supplementary Methods). BLAST hits to the species from which the training and test sequence sets were derived were discarded so as to avoid the development of a classifier that was highly accurate on the test and training sets but useless for a novel dataset.

To build the training and test sets and the custom database, we used publicly available sequences for the cnidarians *Nematostella vectensis* and *Hydra magnipapillata*; the dinoflagellates *Alexandrium tamarense*, *A. catenella*, *A. ostenteldii*, *A. mitum*, *Karlodinium micrum*, *Karenia brevis*, and *Symbiodinium* strain KB8 (clade A); the fungi *Saccharomyces cerevisiae, Schizosaccharomyces pombe*, *Aspergillus niger*, and *Neurospora crassa*); and the bacteria *Escherichia coli* and *Salmonella enterica* (see Supplementary Materials and Methods for Accession Numbers). We also included contigs from an earlier aposymbiotic *Aiptasia* transcriptome (Lehnert et al., 2012) that had ≥30 reads mapping to them from the aposymbiotic libraries produced during Experiment 1 of this study, as well as a large set of contigs from axenically cultured Clade B *Symbiodinium* strain SSB01 (Xiang et al., 2013; T. Xiang and A. Grossman, personal communication).

In addition, we tested the assembled contigs for alignment to *Aiptasia* genomic DNA sequences. We isolated genomic DNA from aposymbiotic *Aiptasia* and obtained about 101 Gb of untrimmed sequence reads from six separate libraries (Accession Numbers SRR646474 and SRR606428). We aligned the genomic reads to our contigs and obtained the mean read count for each contig from the six libraries. A previous test had shown that only 20 of ~60,000 contigs assembled from RNA isolated from cultured, axenic *Symbiodinium* strain SSB01 (Xiang et al., 2013; T. Xiang and A. Grossman, personal communication) had any *Aiptasia* genomic reads mapping to them. However, this clade B strain may have many sequence differences from the clade A strain found in CC7 anemones, so it seemed possible that low levels of *Symbiodinium* in

our putatively aposymbiotic anemones might lead to misclassification of dinoflagellate transcripts as cnidarian. We determined that 15,499 of the 23,794 contigs classified as dinoflagellate by TopSort had zero genomic reads mapping to them, whereas the median of the mean read counts for the contigs classified by TopSort as cnidarian was ~200. Thus, we chose a mean read count of 10 as the cut-off to classify a contig as cnidarian by genomic evidence.

Expression analysis by RNA-Seq

36-bp reads (Experiment 1) were trimmed as described above. With the 101-bp paired-end reads (Experiment 2), the forward reads were shortened to 36 bp for expression analysis and then trimmed as described above. Reads were aligned using bwa to the representative contigs (*i.e.,* the longest contig in each cluster produced by UCLUST; see above; Li and Durbin, 2009). The number of reads with a valid alignment to each contig was counted if it aligned with no errors or gaps to a unique region of the transcriptome. The R package DESeq was used to call contigs as differentially expressed if the false-discovery-rate (FDR)-adjusted *P*-value was ≤0.1 (Anders and Huber, 2010).

Expression analysis by RT-qPCR

RNA was extracted and purified as described above, treated with DNase using the TURBO DNA-free kit™ (Ambion AM1907) following the manufacturers' instructions, and diluted to a concentration of 200 ng per µl. cDNA was then synthesized using the GoScript™ Reverse Transcriptase System (Promega), following the manufacturers' instructions. Primers (Table S2.1B) were designed using Primer Quest (Integrated DNA Technologies) for 29 contigs with a variety of read counts and expression patterns; four of these contigs had previously been identified as appropriate internal reference standards as described below. The predicted product

sizes of 110-238 bp were confirmed by agarose-gel electrophoresis after standard PCR amplification. Primer efficiencies were determined using Real-time PCR Miner [23] and ranged from 90-100%. The RT-qPCR products were also sequenced (Cornell Life Sciences Core Laboratory Center), and all matched the expected product identities.

To quantify transcript levels, we used a ViiA™ 7 thermocycler (Applied Biosystems) with reaction conditions as follows: 12.5 µl of 2X Power SYBR® Green Master Mix (Applied Biosystems), 200 nmol of each primer, and 18 ng of cDNA in a total volume of 25 µl. Each sample and a no-template control was run in duplicate with thermocycler parameters of 95°C for 10 min, 40 cycles of 95°C for 15 s and 60°C for 60 s, and a subsequent dissociation curve to confirm the absence of non-specific products. To confirm the absence of genomic-DNA contamination, a pool of all eight RNA samples (see above) was used as template in a separate reaction as described above except omitting the reverse transcriptase. Real-time PCR Miner was used to calculate the critical threshold (C_T) of each gene from the raw fluorescence data.

To identify reliable reference standards to use for qPCR normalization, we evaluated six housekeeping genes that appeared to be plausible candidates and have indeed been used for this purpose in previous studies of cnidarian gene expression (see Supplemental Materials and Methods for details). Briefly, the expression levels of these genes were tested across a variety of experimental conditions (*e.g.,* heat shock and cold shock) in both aposymbiotic and symbiotic anemones and evaluated for stability of expression using the software geNorm (Vandesompele et al., 2002). Based on this analysis, the genes encoding 60*S* ribosomal protein L11, 40*S* ribosomal protein S7, NADH dehydrogenase subunit 5, and glyceraldehyde-3-phosphate dehydrogenase were selected as standards. The stability of these genes in Experiment 3 was confirmed using geNorm prior to calculating a normalization factor from the geometric mean of their expression

values (Vandesompele et al., 2002; see Table S2.1A). The expression levels of all 29 genes were then normalized via the normalization factor, and relative expression values were calculated using the equation 1/(1 + Primer Efficiency)^C_T. Log_2 fold-changes in expression in symbiotic relative to aposymbiotic anemones were then calculated as the quotients of the above equation. The R software package was used to perform correlations between log2 fold-change data from qPCR and RNA-Seq Experiment 1.

Bayesian phylogenetic analysis

Alignments of Npc2 proteins were generated using the MUSCLE software with its default parameters (Edgar, 2004). The alignments were inspected to identify regions conserved in all proteins and optimized manually over the conserved regions. We then generated a consensus phylogeny using MrBayes 3.1.2 with the following settings: prset aamodelpr = mixed and lset rates = invgamma (Ronquist and Huelsenbeck, 2003). Two separate runs were performed to ensure that identical consensus trees emerged regardless of starting conditions. The runs were terminated after 50 million generations with the average standard deviation of split frequencies ≤ 0.005.

Unbiased screening for functional groups among the differentially expressed genes

As one approach to identifying genes involved in the symbiosis, we used the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (Dennis et al., 2003; Huang et al., 2008). This program performs Fisher Exact tests to determine biological processes (based on GO terms) that are significantly overrepresented among differentially expressed transcripts relative to the background transcriptome. The Functional Annotation Clustering method was

employed, which clusters groups of similar biological processes and provides an enrichment score representative of the -log geometric mean of the *P*-values of the individual processes. Clusters were considered significantly enriched when the enrichment score was >1.3 (corresponding roughly to $P < 0.05$).

Results

Sequencing and assembly of the transcriptome of symbiotic Aiptasia

We isolated total RNA from a clonal population of symbiotic anemones raised under non-stressful culture conditions, enriched for poly-A^+ RNA, and used this RNA to synthesize paired-end Illumina libraries, from which we obtained a total of ~345 million pairs of reads containing ~70 Gb of sequence. The raw reads were trimmed and processed as described in Materials and Methods, leaving ~228 million pairs of reads and ~45 Gb of sequence. These reads were assembled in three batches using Velvet/Oases and a multiple-*k*-mer approach (see Materials and Methods). The resulting assemblies were merged using the Oases merge option, and redundant contigs (≥99% identical over their entire lengths) were collapsed using UCLUST, yielding an initial set of 140,945 contigs with lengths of 102 to 32,510 bp. To estimate the number of genes represented by these contigs and choose a representative contig for each gene, we clustered contigs with good alignments (≥99% identical over ≥20% of the length of the shorter contig). This resulted in 52,717 clusters, and the longest contig from each was taken as representative for further analysis. Although 31,014 clusters contained only a single contig, 19,380 contained two to nine contigs, and 2,323 clusters contained 10 or more contigs, with a largest cluster of 230 contigs (see Discussion).

Classification of contigs using TopSort and comparison to genomic sequence

It is difficult or impossible to obtain animal RNA without contamination by RNA from the intracellular algal symbionts and (although presumably in much smaller amounts) from other organisms in the non-sterile aquarium system. To address this issue, we developed the TopSort support-vector-machines algorithm to classify contigs as putatively of cnidarian, dinoflagellate, bacterial, or fungal origin. Sequences of reliably known origin were used to create training and test sets, and each contig was scored on several metrics (see Materials and Methods). After training on the training set, the accuracy of TopSort on the test set was ~95% for contigs of 150-300 bp and >99% for contigs of >300 bp, for an overall error rate of 2-3%.

We used TopSort to classify the 52,717 representative contigs in our dataset (Table 2.2, column B). As expected, most contigs were classified as cnidarian or dinoflagellate. However, the 2-3% error rate of TopSort with the test dataset suggested that some hundreds of the putative cnidarian contigs might actually be dinoflagellate contigs that had been misclassified, which would be a significant problem for subsequent analyses of gene-expression differences between symbiotic and aposymbiotic animals. Thus, we also aligned reads from *Aiptasia* genomic-DNA sequence libraries to the transcriptome (see Materials and Methods). ~94% of the contigs classified by TopSort as cnidarian had supporting genomic evidence, as compared to only ~5% of contigs classified by TopSort as non-cnidarian (Table 2.2, columns C-E). These results validated the performance of TopSort in initial classification and yielded a set of 26,219 high-confidence *Aiptasia* contigs (henceforth referred to as "cnidarian") on which we have focused for our further analyses. In the remainder of this paper, we also use the term "dinoflagellate" to refer to the 22,668 contigs classified as dinoflagellate by TopSort and lacking matches to *Aiptasia* genomic DNA, and we refer to contigs for which the classifications by TopSort and

genomic match conflicted as "ambiguous".

Table 2.2 Assignment of contigs to species of origin.

Type of Organism	No. by TopSort[a]	No. with genomic evidence[b]	No. without genomic evidence[b]	False-positive rate (%)[c]
Cnidarian	28,026	26,219	1,807	6.4
Dinoflagellate	23,794	1,126	22,668	4.7
Fungi	166	18	148	10.8 [d]
Bacteria	731	185	546	25.3 [d]

[a] See text.
[b] Genomic evidence was defined as ≥10 paired-end reads aligning from *Aiptasia* genomic-DNA libraries prepared from aposymbiotic anemones (see Materials and Methods).
[c] Classified as cnidarian by TopSort but lacking genomic evidence, or classified as non-cnidarian by TopSort but with apparent matches to *Aiptasia* genomic DNA.
[d] Many of these are presumably transcripts from contaminants that were present on the anemones from which the genomic DNA was prepared. However, the high rate of apparent false-positives among the putatively bacterial and fungal sequences probably also reflects Bayes's Rule, whereby the ratio of false-positives to true positives is high when the a priori probability of a true positive is low.

Characterization and annotation of transcriptome

The cnidarian contigs ranged in size up to >32 kb, with a median of 1,644 bp, whereas the dinoflagellate contigs had somewhat smaller maximum and median sizes (Table 2.3). The remaining contigs ("Other" in Table 2.3) had a size distribution similar to those of the cnidarian and dinoflagellate contigs. It is therefore unlikely that the failure to classify these contigs as cnidarian or dinoflagellate was due simply to their being shorter than average and thus more difficult to annotate by BLAST or align to genomic reads.

To assign putative functions to the representative cnidarian and dinoflagellate contigs, we used blastx to align them to SwissProt and the NCBI nr database, retaining only alignments with E-values $\leq 10^{-5}$. Of the 26,219 cnidarian contigs, 16,373 (62%) had such alignments to 9,386 unique accession numbers in SwissProt (Table 2.4). In contrast, of the 22,668 dinoflagellate

contigs, only 7,895 (~35%) had such alignments to 5,054 unique accession numbers (Table 2.4). Similar numbers were obtained by aligning sequences to nr (Table 2.4). Using Blast2GO with its default cut-off of 1e-3, we assigned GO terms based on the SwissProt annotations. We were able to assign 10,521 unique GO terms to cnidarian sequences and 5,747 unique GO terms to algal sequences.

Table 2.3 Size distribution of the representative contigs.

Parameter	Cnidarian	Dinoflagellate	Other [a]
Number of contigs	26,219	22,668	3,830
Median contig size (bp)	1,644	1,144	1,474
Mean contig size (bp)	2,227	1,355	1,789
Minimum contig size (bp)	106	108	102
Maximum contig size (bp)	32,510	20,508	18,089
Total length of contigs (Mb)	58	31	7

[a] Includes both the contigs classified as "ambiguous" (see text) and those classified as fungal or bacterial.

Table 2.4 Summary of alignments to SwissProt and nr.

Classification	No. Contigs	No. (%) of contigs aligned to SwissPro[a]	No. (%[b]) of unique accessions	No. (%) of contigs aligned to nr [a]	No. (%[b]) of unique accessions
Cnidarian	26,219	16,373 (62)	9,386 (57)	19,259 (74)	11,593 (60)
Dinoflagellate	22,668	7,895 (34)	5,054 (64)	11,184 (49)	7,789 (70)

[a] Alignments with E-value ≤ 10^{-5}.
[b] As % of all alignments.

To investigate why there were so few unique accession numbers relative to the numbers of representative contigs, we examined the distributions of contigs per accession number (Table 2.5). In both the cnidarian and dinoflagellate cases, ~76% of accession numbers annotated only one representative contig, and another ~14% annotated two representative contigs (as might occur with a duplicated gene or two sufficiently different alleles of the same gene). In contrast,

some accession numbers were hit by much larger numbers of representative contigs (Table 2.5). Although there are several possible explanations for such cases (including the existence of extended gene families, complex alternative splicing, and/or somatic differentiation), we suspect that most reflect a failure of contigs derived from the same gene to cluster with the algorithm used, perhaps because of repeat structures within the genes. In any case, if we assume (as a worst-case scenario) that all such cases result from such failures to cluster, and that the failure rate was identical between the successfully annotated and unannotated contigs, then we can infer that the numbers of 'unigenes' (sequences derived from unique genes) present in our dataset are ~14,500 for *Aiptasia* and ~14,000 for *Symbiodinium*, representing substantial fractions of the total gene numbers expected from information on other eukaryotes (see Discussion).

Table 2.5 Distribution of representative contigs among accession numbers. [a]

No. of contigs with best blast hit to a given accession number	No. of accession numbers (cnidarian)	No. of accession numbers (dinoflagellate)
1	7,108	3,874
2	1,332	703
3 to 5	643	361
6 to 10	190	85
11 to 25	89	25
26 to 50	14	3
>50 [b]	10	3
Total	**9,386**	**5,054**

[a] Analysis performed to investigate why there were so few unique BLAST hits relative to the numbers of representative contigs. See text for details.
[b] The largest numbers were 187 (cnidarian) and 71 (dinoflagellate).

Identification of differentially expressed transcripts

To compare gene expression in symbiotic relative to aposymbiotic anemones, we performed two RNA-Seq experiments using somewhat different conditions (see Materials and

Methods; Table 2.1). In each experiment, we identified many transcripts that appeared to be differentially expressed, including many in which the changes in abundance were ≥5-fold (Table 2.6, columns B and C). Although the two experiments identified many of the same genes, there were also differences that probably reflect both the noise inherent in such analyses and actual differences in expression due to the different experimental conditions. However, we hypothesized that any genes involved directly in the maintenance of symbiosis (*e.g.,* genes encoding proteins of the symbiosome) would show similar expression differences in both experiments. Therefore, we identified these contigs (Table 2.6, column D) and focused on them in subsequent analyses.

Although our further analyses to date have also focused on transcripts with convincing annotations by blastx, it is important to note that 101 of the cnidarian transcripts that appeared to be differentially expressed in both experiments, including 21 with ≥5-fold expression changes could not be annotated at this time (Table 2.6, column E). 52 of these transcripts (including four of the 21 with ≥5-fold expression changes) contained apparent open reading frames with ≥100 codons. Identifying the functions of these unknown proteins may be critical to understanding the structural and biochemical bases of the symbiosis.

To evaluate the reliability of the RNA-Seq data, we also performed an RT-qPCR experiment using culture conditions similar to those of Experiment 1 (Table 2.1). We tested 29 contigs that exhibited a range of fold-changes and read counts, including some that were of particular biological interest (Tables S2.1 A and B). In order to assess the overall agreement between the RNA-Seq and RT-qPCR experiments, we determined the Spearman's rank correlation coefficient of the Log_2 fold-change for all contigs, excluding those that had apparently infinite changes in expression (*i.e.* were only found in either symbiotic or

Table 2.6 Differential expression of cnidarian contigs.

	No. of Contigs		Shared (%[d])	
Contig Behavior [b]	**Experiment 1 [c]**	**Experiment 2 [c]**	**Total**	**Unannotated**
Upregulated	1,109	3,093	456 (41)	53 (5)
Upregulated ≥ 5-fold	138	631	79 (57)	17 (12) [e]
Downregulated	1,036	2,905	464 (45)	48 (5)
Downregulated ≥ 5-fold	23	388	6 (26)	4 (17)

[a] Classified as cnidarian by TopSort and confirmed by genomic match (see Table 2).
[b] Expression in symbiotic relative to aposymbiotic anemones. In all cases shown, the difference in expression was significant at a false-discovery-rate-adjusted $P \leq 0.1$.
[c] For experimental conditions, see Materials and Methods and Table 1.
[d] The percentage in each case is the number shared divided by the number from Experiment 1.
[e] Four of these 17 contigs had an ORF of >100 codons.

aposymbiotic anemones). The correlation coefficient of 0.96 (*P*-value = 3e-14) showed a strong correlation between the RNA-Seq and RT-qPCR datasets.

In what follows, we discuss several sets of cnidarian genes whose differential expression suggests testable biological hypotheses.

Genes involved in metabolite transport

Given the intimate relationship between the symbiotic partners, transporters involved in moving metabolites between compartments seem likely to be of special importance in maintaining the symbiosis. To identify such transporters, we screened the differentially expressed transcripts associated with the GO term "P:transport" for those encoding putative transporters of small molecules. Although the GO annotation of *Aiptasia* is incomplete, we were able to identify 48 up-regulated and 18 down-regulated transcripts encoding putative transporters and transport-related proteins (Tables S2.2). We focus in what follows on the 15 such proteins that were most highly up-regulated in symbiotic anemones.

Transport of photosynthetically fixed carbon and other organic metabolites

Among the transcripts strongly upregulated in symbiotic anemones were two (Table 2.7, lines 1 and 2) that encode proteins closely related (~39% identify in amino-acid sequence) to the mammalian facilitative glucose transporter GLUT8, which localizes to the endosome membrane (Augustin et al., 2005). This localization depends on an N-terminal dileucine motif, and indeed dileucines are present at amino acids 32-33 and 26-27 of the two *Aiptasia* GLUT8 proteins. One or both of the *Aiptasia* GLUT8 proteins are thus likely to be involved in the transport of photosynthetically produced glucose across the symbiosome membrane into the host cytoplasm (see Discussion). However, it should also be noted that the transcript encoding a predicted Na^+-glucose/*myo*-inositol co-transporter was detected only in symbiotic anemones (Table 2.7, line 3), while a transcript encoding a related protein was also upregulated 2.2-fold (Table S2.2, line 26). Interestingly, a transcript encoding a third member of this protein class was strongly downregulated in symbiotic anemones (Table S2.2, line 65).

Lipids may also be an important energy currency in symbiotic animals (see further discussion below), and in this regard it is interesting that the transcripts for a putative lipid-droplet surface-binding protein (potentially involved in the mobilization of stored fats for transport), a protein similar to scavenger receptor class B member 1 (related to CD36-type fatty-acid transport proteins), and a putative carnitine transporter (potentially involved in entry of fatty acids into mitochondria for degradation) were all strongly upregulated in symbiotic animals (Table 2.7, lines 4-6). In the last regard, it should also be noted that the transcripts for several putative acyl-carnitine transferases were also upregulated in symbiotic anemones (Table S2.2, lines 25, 43, and 44).

Also dramatically upregulated was the transcript for a member (Niemann-Pick disease

type, c2; Npc2D) of the Npc2 protein family (Table 2.7, line 7; Figure S2.1). In mammalian and *Drosophila* cells, Npc2 binds cholesterol in the lumen of the endosome and lysosome and transfers it to Npc1, a transmembrane protein that exports the cholesterol to other intracellular locations (Infante et al., 2008; Frolov et al., 2003; Sleat et al., 2004). Consistent with a previous study of the anemone *Anemonia viridis* (Ganot et al., 2011), we identified multiple transcripts encoding Npc2-like proteins in the *Aiptasia* transcriptome as well as in the transcriptomes of three other cnidarians. A multiple-sequence alignment and Bayesian phylogenetic analysis identified two subclades reflecting at least one duplication event in the Anthozoan lineage (Figure 2.2A). One subclade (including the Npc2A proteins of both *A. viridis* and *Aiptasia*) clustered with the canonical Npc2 proteins found in most animals (including mammals and *Drosophila*), while the second subclade contained both the *A. viridis* (Ganot et al., 2011; Sabourault et al., 2009) and *Aiptasia* Npc2D proteins that are upregulated during symbiosis. Strikingly, all of the proteins in this second subclade have sequence alterations at conserved positions in the sterol-binding site (Figure 2.2B). Mutations to alanine at these positions are known to disrupt cholesterol binding in human cells (Wang et al., 2010; Ko et al., 2003), raising interesting questions about the roles of these proteins in symbiotic cnidarians (see Discussion).

Other putative organic-metabolite transporters also showed large changes in expression. In particular, a transcript encoding a putative taurine transporter was detected only in symbiotic anemones (Table 2.7, line 8), while the transcripts for an aromatic-amino-acid transporter and a GABA/glycine transporter were upregulated 4.9- and 6.9-fold, respectively (Table 2.7, lines 9 and 10). Interestingly, taurine has been reported to comprise ~35% of the amino-acid pool in symbiotic *Aiptasia* (Swanson et al., 1998), although its specific functions are not well understood. Determining the intracellular localization of the transporter identified here could

Table 2.7 Transport-related proteins that were strongly up-regulated in symbiotic anemones. [a]

Line	Fold-change[b]	Fold-change[c]	Locus #/ Transcript #	Best Blast Hit to SwissPro	UniProt Accession No.	Blast-hit E-value
1	11	6.3	86800/1	Human facilitated glucose transporter (GLUT8)	Q9NY64	9.00E-89
2	3.7	n.d.	11708/1	Human facilitated glucose transporter (GLUT8)	Q9NY64	1.00E-88
3	∞	n.d.	36456/1	Rabbit Na^+/(glucose/*myo*-inositol) transporter 2	Q28728	3.00E-104
4	5.8	n.d.	45451/1	*Drosophila* lipid-droplet surface-binding protein 2	Q9VXY7	2.00E-08
5	28	3.7	77179/1	Human scavenger receptor class B member 1	Q8WTV0	9.00E-65
6	44	57	125065/1	*Drosophila* organic-cation (carnitine) transporter	Q9VCA2	6.00E-35
7	600	26	102514/1 [d]	Human Npc2 cholesterol transporter	P61916	2.00E-14
8	∞	29	58798/1	Bovine Na^+- and Cl^--dependent taurine transporter	Q9MZ34	1.00E-169
9	4.9	6.2	95114/1	Mouse aromatic-amino-acid transporter 1	Q3U9N9	3.00E-65
10	6.9	n.d.	12006/1	*Xenopus* GABA and glycine transporter	Q6PF45	8.00E-60
11	4.3	n.d.	84720/1	Fish (*Tribolodon*) carbonic anhydrase II	Q8UWA5	2.00E-36
12	13	2.2	65589/1	Sheep aquaporin-5	Q866S3	8.00E-37
13	4.3	n.d.	2130/2	Pig aquaporin-3	A9Y006	1.00E-68
14	130	n.d.	60777/1	Zebrafish NH_4^+ transporter rh type b	Q7T070	3.00E-98
15	5.9	7	70728/1	*C. elegans* NH_4^+ transporter 1 (AMT1-type)	P54145	6.00E-72

[a] Putative small-molecule transporters and some proteins of related function are arranged in the order of their discussion in the text.

[b] By RNA-Seq (see Table S2.2). The arithmetic mean of the values from Experiments 1 and 2 is shown except for transcript 77179/1 (line 5). ∞, expression was not detected in aposymbiotic animals. Transcript 77179/1 was detected in aposymbiotic anemones in Experiment 1 but not in Experiment 2, giving a nominal ∞-fold change in expression in that experiment. However, as the normalized read counts in both experiments were rather low, and the possible involvement of the 77179/1-encoded protein in lipid metabolism makes it likely to have been affected in its expression by the starvation conditions used in Experiment 2, we indicate here the more conservative value from Experiment 1 alone.

[c] By qPCR (Table S2.1). n.d., not determined.

[d] Encoding putative protein Npc2D.

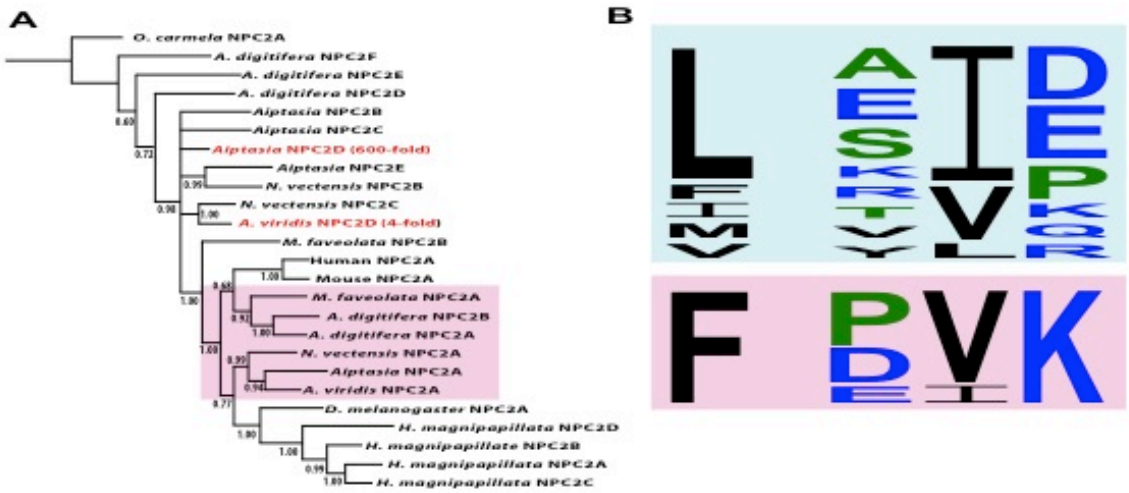


Figure 2.2 Npc2-like proteins that putatively do or do not have the ability to transport cholesterol. (A) Consensus phylogenetic tree constructed from alignments (Figure S2.1B) of 25 Npc2-like proteins (see Materials and Methods). *Oscarella carmela*, a sponge, served as the outgroup, and the single human Npc2 protein, the single mouse Npc2 protein, and one of the eight *Drosophila melanogaster* Npc2 proteins were included in the analysis. The cnidarian sequences included are from two corals (*Acropora digitifera* and *Montastraea faveolata*), three anemones (*Aiptasia* sp., *Nematostella vectensis*, and *Anemonia viridis*), and a hydrozoan (*Hydra magnipapillata*). The Npc2-encoding transcripts found to be upregulated in symbiotic anemones, which fall outside the clade containing the mammalian and *Drosophila* sequences, are shown in red with their fold-changes (Table 2.7, line 4; Ganot et al., 2011). Light blue and pink shading indicate the groups of anthozoan proteins in the cladogram to which the sequence displays in B correspond. Numbers indicated the bootstrap values for the branches indicated. (B, lower) Amino acids highly conserved in animal (including some cnidarian) Npc2 proteins and thought to be involved in cholesterol binding (see text). The mammalian proteins both have the sequences F…PVK, and the *Drosophila* Npc2A sequence is F…PVL. (B, upper) The variety of amino acids found at the corresponding positions in members of the other protein clade. The differentially regulated *Aiptasia* and *A. viridis* proteins both have the sequence L…SID.

provide insight into the possible function(s) of taurine. The transcripts for other putative amino-acid transporters also showed significant differences in expression between symbiotic and aposymbiotic anemones (Table S2.2, lines 16, 28, 30, 34, 39, 40, 45, 52, and 55), suggesting that the establishment of symbiosis produces profound changes in amino-acid transport and metabolism (see Discussion).

Transport of inorganic nutrients

CO_2 is an excreted waste product for animals such as aposymbiotic anemones, but it is required for photosynthesis when dinoflagellate symbionts are present. It may not require specific transporters if it can diffuse freely across cellular membranes. However, in order to maintain a high concentration of inorganic carbon in the symbiosome, the host may need to convert CO_2 to the less freely diffusing bicarbonate anion. We identified one carbonic-anhydrase gene that was upregulated 4.3-fold in symbiotic anemones (Table 2.7, line 11), while a second gene was downregulated 3-fold (Table S2.2, line 63). In addition, it is not clear that CO_2 diffuses sufficiently rapidly through the relevant membranes to support efficient photosynthesis, and some studies have suggested that aquaporins may play a role in facilitating this diffusion (Uehlein et al., 2012; Kaldenhoff, 2012). Thus, it is of interest that that we found two aquaporins to be up-regulated 13- and 4.3-fold in symbiotic anemones (Table 2.7, rows 12 and 13). Although aposymbiotic anemones, like other aquatic animals, excrete excess (and potentially toxic) ammonium produced by amino-acid breakdown (Pernice et al., 2012), symbiotic anemones need to supply nitrogen to their dinoflagellates. Thus, it was not surprising that we found differentially expressed genes encoding ammonium transporters. These genes were in both of the two major families found in animals: a "rhesus-like" gene and an "AMT-like" gene were upregulated 130- and 5.9-fold, respectively, in symbiotic anemones (Table 2.7, rows 14 and 15), suggesting that they might be involved with ammonium supply to the dinoflagellate, whereas another rhesus-like transporter was down-regulated 2.9-fold (Table S2.2, row 61), suggesting that it might be involved in ammonium excretion.

The host must also supply other inorganic nutrients to the algae. For example, phosphate and sulfate must be translocated across the symbiosome membrane either as the inorganic ions or

as part of some organic metabolite. In this regard, it is of interest that we found the genes for two putative inorganic-phosphate transporters to be up-regulated ~2-fold in symbiotic anemones, a gene for a putative UDP-sugar transporter to be up-regulated 2.7-fold, and a gene for a putative sulfate transporter to be up-regulated 3.1-fold (Table S2.2, lines 19, 22, 27 and 42). In addition, although it is not clear why, zinc is apparently absorbed to a greater extent by symbiotic than aposymbiotic anemones (Harland, 1990),with increased concentrations in both animal and dinoflagellate, and we found genes for three putative zinc transporters, in two different families, to be up-regulated 1.7- to 2.6-fold (Table S2.2, rows 23, 36, and 46).

Genes controlling certain metabolic pathways

To explore the integration of metabolite transport with the overall regulation of metabolic pathways, we looked for the presence and coordinated regulation of genes encoding the enzymes of particular pathways that we hypothesized might be involved in the animal's response to the presence of a symbiont. For these analyses, we used the full transcriptome but only the expression data from RNA-Seq Experiment 1, because the starvation of the aposymbiotic anemones in RNA-Seq Experiment 2 seemed likely to have had a strong effect on the expression of metabolic-pathway genes.

Lipid metabolism

There appear to be systematic changes in lipid metabolism between symbiotic and aposymbiotic anemones. Four genes encoding enzymes involved in fatty-acid synthesis (acetyl-CoA carboxylase, a fatty-acid elongase, and Δ^5- and Δ^6-fatty-acid desaturases) were upregulated 3.5- to 6.2-fold (Table S2.3, lines 1-4), and at least nine genes encoding proteins putatively

involved in lipid storage or its regulation were also differentially regulated (Table S2.3, lines 5-13). In addition, many genes involved in β-oxidation of fatty acids were upregulated in symbiotic anemones (Figure 2.3; Table S2.3, lines 17-21, 23, 24, and 30). Although some of the fold-changes were not large, the consistency is striking, and gastrodermal and epidermal cells may well differ in their expression patterns in ways that obscure the full extent of the changes in a particular cell population (see Discussion). Finally, although the glyoxylate cycle (which allows cells to achieve a net synthesis of longer carbon chains from two-carbon units such as those derived by β-oxidation) is not generally present in animal cells, we identified genes putatively encoding its two key enzymes, isocitrate lyase and malate synthase (Table S2.3, lines 31 and 32), consistent with a previous report of the presence of this cycle in cnidarians (Kondrashov et al., 2006). Although the malate-synthase transcript showed no statistically significant differential expression, the isocitrate-lyase transcript was upregulated 3.9-fold in symbiotic anemones. Interestingly, we did not see significant upregulation of the genes encoding the enzymes responsible for metabolizing medium- and short-chain fatty Acyl-CoA (MCAD, SCAD, crotonase, and M/SCHAD in Figure 2.3; Table S2.3, lines 25-28), suggesting that the metabolic change accompanying the establishment of symbiosis primarily involves long-chain and/or very-long-chain fatty acids.

Amino-acid metabolism and the SAM cycle

Consistent with previous observations (Wang and Douglas, 1998), we found that the transcript for a putative glutamine synthetase was upregulated in symbiotic anemones, as was the transcript for a putative NADPH-dependent glutamate synthase (Figure 2.4). These data suggest that the epidermal cells, the gastrodermal cells, or both synthesize glutamate *via* a complete GS-

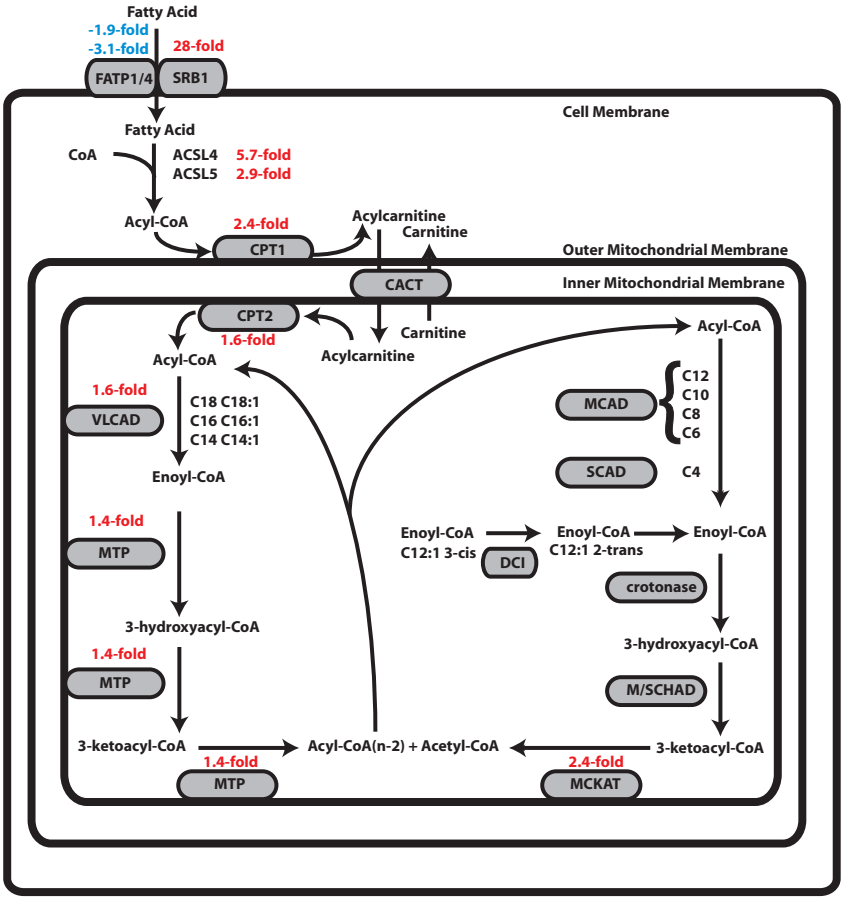


Figure 2.3 Expression changes of genes governing β-oxidation of fatty acids. The diagram (adapted from Houten and Wanders, 2010) shows the localization of proteins involved in fatty-acid transport and β-oxidation in relation to the membranes of the mitochondrion and cell (as known from other animal cells). Statistically significant expression changes from RNA-Seq Experiment 1 are shown where applicable; upregulation in symbiotic relative to aposymbiotic anemones is shown by positive/red numbers, and downregulation is shown by negative/blue numbers. Scavenger receptor class B member 1 (SRB1; CD36-related protein) and FATP1/4, putative fatty-acid transporters at the cell surface; ACSL4 and ACSL5, enzymes that convert free fatty acids to fatty acyl-CoA esters; CPT1, CPT2, and CACT, proteins involved in transporting fatty acyl-CoA esters across the mitochondrial membranes; VLCAD, MTP, MCAD, SCAD, M/SCHAD, and MCKAT, enzymes responsible for β-oxidation; DCI, converts fatty acids with double bonds starting at odd-numbered positions to fatty acids with double bonds starting at even-numbered positions; crotonase, hydrates double bonds that start at even-numbered positions. See Supplementary Table 3, lines 14-30, for full protein names, UniProt Accession Numbers, and transcript numbers.

GOGAT cycle (Miflin and Habash, 2002) rather than (or in addition to) simply obtaining it from the dinoflagellate. We also identified both upregulated and downregulated genes that encode putative glutamate dehydrogenases (Figure 2.4), which normally catabolize glutamate to α-ketoglutarate and ammonium in animal cells (where the concentrations of ammonium are typically too low to allow the reverse reaction to proceed effectively). The subcellular-localization program WoLF PSORT (Horton et al., 2007) predicts that the downregulated and upregulated enzymes should localize to the mitochondria and cytosol, respectively, consistent with a previous report that corals contain both mitochondrial and cytosolic glutamate dehydrogenases (Dudler et al., 1987). It seems likely that these initially rather puzzling observations (upregulation of one glutamate dehydrogenase and downregulation of another; upregulation of enzymes both of glutamate synthesis and of glutamate breakdown) reflect the differing metabolic needs of different cell types, and/or of different compartments within the same cells, in symbiotic anemones.

We also observed multiple changes in the expression of genes governing the metabolism of sulfur-containing amino-acids and the *S*-adenosylmethionine (SAM) cycle (Figure 2.5). Based on the failure to find a gene encoding cystathionine β-synthase (CBS) in the *A. digitifera* genome, it has been hypothesized that cysteine is an essential amino acid in cnidarians that must be obtained directly from either prey or the symbiont (Shinzato et al., 2011). However, we found an *Aiptasia* transcript encoding a CBS [best BLAST hit, rabbit CBS (Q9N0V7); E-value 9e-166], suggesting that anthozoans resemble other animals in their ability to synthesize cysteine from methionine. The CBS transcript was downregulated 2.1-fold in symbiotic anemones, perhaps reflecting a decreased need for cysteine synthesis in the host because it is being supplied directly by the dinoflagellate. Conceivably in the more obligately symbiotic corals, the enzyme

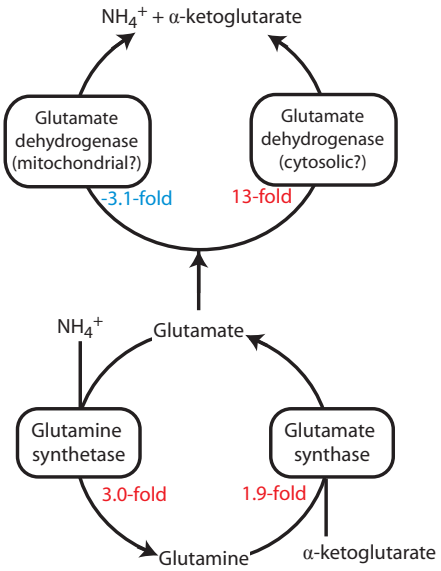


Figure 2.4 Expression changes of genes governing glutamine and glutamate metabolism. Upregulation in symbiotic relative to aposymbiotic anemones is shown by positive/red numbers, and downregulation is shown by negative/blue numbers. For UniProt and transcript numbers, see Supplementary Table 4, lines 1-4. The possible localizations of the glutamate dehydrogenases are discussed in the text.

is never needed and the gene has been lost altogether. It should also be noted that cysteine synthesis via the CBS pathway is a drain on the homocysteine pool, which otherwise remains available for the synthesis of methionine and the SAM cycle. The concordant upregulation of four genes encoding enzymes of the SAM cycle (Figure 2.5) suggests that it may assume an increased importance in symbiotic animals, although the very wide range of possible methylation targets makes it difficult to guess at the precise biological significance of this regulation. The apparent switch of pathways used for synthesis of methionine from homocysteine (Figure 2.5) may also be related, in that it could reflect an alteration in the kinetics and/or localization of the SAM cycle.

Interestingly, despite the presence of a CBS, we could not find a gene encoding aspartokinase or homoserine dehydrogenase in the transcriptomes of symbiotic or aposymbiotic *Aiptasia* or in the *A. digitifera* genome (Shinzato et al., 2011; Table S2.4, lines 15-17), implying

that anthozoans would be unable to achieve a net synthesis of homoserine (and hence of homocysteine and other sulfur-containing amino acids) from central metabolic intermediates. If confirmed, this would be consistent with the situation in other animals (where methionine is an amino acid essential in the diet) (Guedes et al., 2011) but surprisingly inconsistent with labeling results indicating synthesis of methionine by starved, aposymbiotic anemones (Wang and Douglas, 1999). A related puzzle is that the *Aiptasia* transcriptome and the *A. digitifera* genomeappear to contain genes encoding both a homoserine *O*-acetyltransferase and a cystathionine γ-synthase, which would allow the synthesis of cystathionine from homoserine, but not a cystathionine β-lyase, which in many microorganisms is responsible for the synthesis of

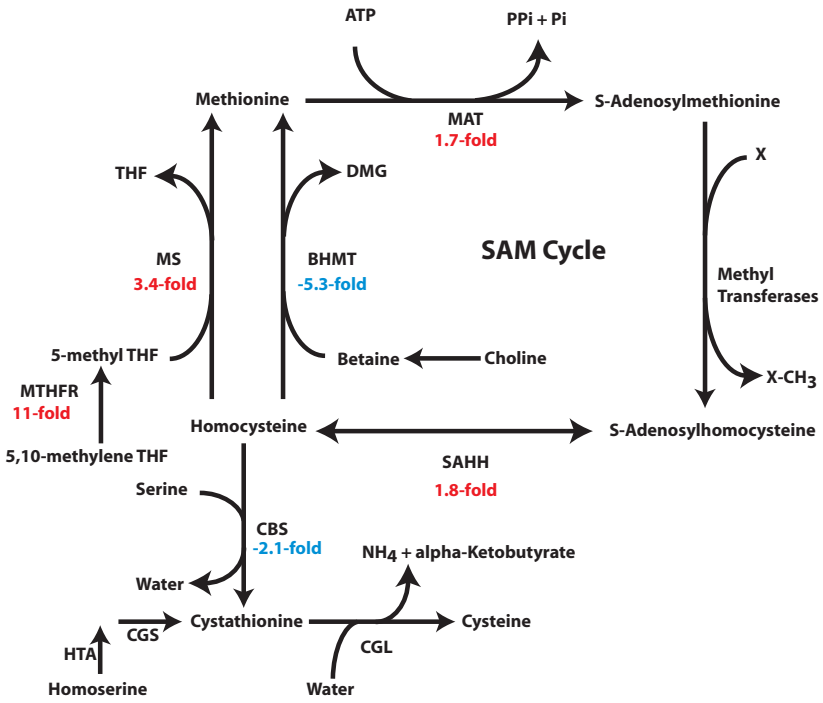


Figure 2.5 Expression changes of genes governing the metabolism of sulfur-containing amino acids and the *S*-adenosylmethionine (SAM) cycle. Upregulation in symbiotic relative to aposymbiotic anemones is shown by positive/red numbers, and downregulation is shown by negative/blue numbers. For full names of enzymes, UniProt Accession Numbers, and transcript numbers, see Supplementary Table 4, lines 7, 8, 10-14, 18, and 19. THF, tetrahydrofolate; DMG, dimethylglycine.

homocysteine from cystathionine (Table S2.4, lines 18-20). Further studies will be needed to resolve these issues.

These questions about methionine and cysteine metabolism raised a broader question about the degree to which the amino-acid-biosynthetic capabilities of anthozoans resemble those of better-characterized animals, in which 12 of the 20 amino acids needed for protein synthesis cannot be synthesized from central-pathway intermediates and so must be obtained (directly or indirectly) from the diet. To address this question, we asked if the elements of amino-acid-biosynthetic pathways were present in the *Aiptasia* transcriptome. As expected, it appears that *Aiptasia* should be able to synthesize the eight generally nonessential amino acids from intermediates in the central metabolic pathways (Table S2.4, lines 1, 2, 22-36). In addition, like other animals, they should be able to synthesize arginine from ornithine via the urea cycle (Table S2.4, lines 37-42), although a net synthesis of ornithine and arginine would not be possible because of the apparent lack of either an acetylglutamate kinase or an ornithine acetyltransferase (Table S2.4, lines 43 and 44). Similarly, although the *Aiptasia* transcriptome revealed genes encoding various enzymes involved in interconversions within other groups of amino acids, key enzymes needed to synthesize these groups of amino acids from central-pathway intermediates appear to be missing (Table S2.4, lines 15-17, 45-79). Thus, *Aiptasia,* like other animals, apparently must obtain 12 amino acids (or their amino-acid precursors) from their food, their dinoflagellate symbionts, or both. This conclusion is generally compatible with radiolabeling studies suggesting that leucine, isoleucine, valine, histidine, lysine, phenylalanine, and tyrosine are all translocated from the dinoflagellates to the host (Wang and Douglas, 1999).

Genes potentially involved in host tolerance of dinoflagellates

To take an unbiased approach to the identification of other genes that might be involved in maintenance of the symbiosis, we used the DAVID program to identify biological processes (based on GO terms) that were significantly overrepresented among the differentially expressed transcripts (see Materials and Methods). Among the groups of genes identified in this way were three that are potentially involved in the animal host's tolerance of the symbiotic dinoflagellates.

Response to oxidative stress

As the presence of an intracellular photosynthetic symbiont presumably imposes oxidative stress on an animal host (see Discussion), it was quite surprising that of the eight differentially expressed genes identified under this GO term, six (including a catalase gene) were actually downregulated in symbiotic animals (Figure 2.6A). Moreover, one of the two upregulated genes encodes a predicted guanylate cyclase, which may have many functions unrelated to oxidative stress. The other upregulated gene is one of a pair encoding distinct *Aiptasia* proteins (Figure S2.2A) that had a human peroxidasin as their top blastx hit (Table S2.5, section A). However, a function for these proteins in coping with oxidative stress is doubtful for two reasons. First, the second gene is downregulated in symbiotic animals (Figure 2.6A). Second, despite the blastx results, neither of the *Aiptasia* proteins contains the peroxidase domain found in canonical peroxidasins with peroxidase activity (Nelson et al., 1994; Figure S2.2 A and B).

Inflammation/tissue remodeling/response to wounding

Of the 15 differentially expressed genes associated with this cluster of GO terms, eight

were downregulated and seven were upregulated in symbiotic animals (Figure 2.6B). Despite this heterogeneity, a suggestive pattern was observed in which genes encoding proteins whose homologues are considered pro-inflammatory were mostly downregulated, whereas the three genes encoding proteins whose homologues are considered anti-inflammatory were all upregulated (Figure 2.6B; Table S2.5). The pattern appears even stronger when it is noted that one of the three upregulated genes with a putatively pro-inflammatory function encodes just one of at least three distinct *Aiptasia* plasma-kallikrein homologues (Figure S2.2C), the other two of which are downregulated (Figure 2.6B), and that the upregulated ficolin may function specifically in recognition of *Symbiodinium* rather than just as a general activator of the complement innate-immunity pathway (Logan et al., 2010). Thus, a downward modulation of the host's inflammatory response may contribute to allowing the persistence of dinoflagellate symbionts.

Apoptosis/cell death

Of the 13 differentially expressed genes associated with this pair of GO terms, five were downregulated in symbiotic animals but eight were upregulated, including several with large fold-changes in expression (Figure 2.6C). Thus, it seems possible that that an increased activity of cell-death pathways may be required for the animal to cope with the presence of the symbiotic dinoflagellate even under conditions considered to be nonstressful (see Discussion).

Discussion

To explore the cellular and molecular basis of the cnidarian-dinoflagellate symbiosis, we undertook a global analysis of the transcriptomes of symbiotic and aposymbiotic *Aiptasia*. This

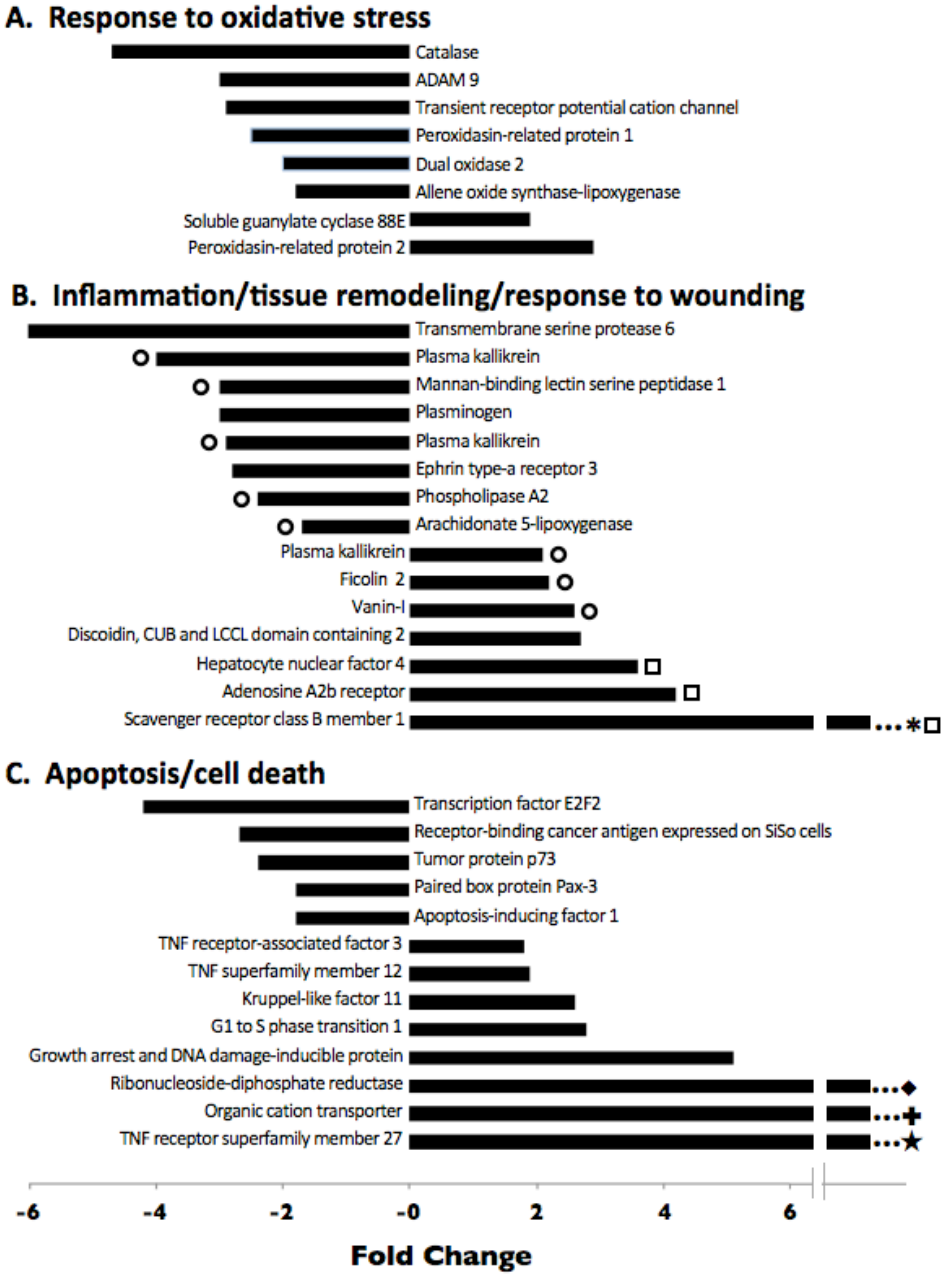


Figure 2.6. Expression changes of genes with functions that may relate to host tolerance of the symbiont. Functionally related groups of genes (by GO-term assignments) that were significantly enriched among the differentially expressed genes relative to the background transcriptome were identified as described in the text. Fold-changes are shown as expression in symbiotic anemones relative to that in aposymbiotic anemones (positive values are enriched in symbiotic anemones and negative values are depleted in symbiotic anemones). O, putatively pro-inflammatory; □, putatively anti inflammatory;*, highly up-regulated (28-fold in Experiment 1 and detected only in symbiotic animals in Experiment 2; see also Supplementary Table 2, footnote b); ◆, 12-fold-change; ✚, 44-fold-change; ★, 60-fold-change.

study has yielded (i) extensive transcriptome assemblies for both the anemone and its symbiont; (ii) novel hypotheses about changes in metabolism and metabolite transport that may occur in the host upon symbiosis establishment; and (iii) novel hypotheses about genes that may be involved in symbiont recognition and tolerance by the host. The transcriptomes also provide a reference for future studies of gene expression under other conditions such as exposure to various stresses.

Transcriptome assembly and annotation

We have sequenced, assembled, and partially characterized the transcriptomes of both a clonal stock of symbiotic *Aiptasia* and the endogenous clade A symbionts present in that stock. The animal and algal transcripts were separated bioinformatically using the TopSort algorithm that we developed for this purpose and comparisons to genomic sequence obtained from fully aposymbiotic animals. Although the assemblies appear to be of high quality overall, there remain some areas where improvements could be made. For example, it remains unclear both why reads from some genes assembled into multiple contigs that clustered based on regions of nucleotide identity (up to 230 contigs in a cluster) and why a small set of accession numbers have so many representative transcripts aligning to them. As it seems unlikely that alternative splicing and gene duplications alone could explain the magnitude of the effects observed, we presume that they result from some combination of these factors and the inherent complexities of *de novo* transcriptome assembly. For example, two of the contig clusters with the most members encode putative actins and olfactory C proteins. In most animals, actins are encoded by families of genes and expressed at high levels, whereas olfactory C proteins are members of very large gene families. Highly abundant transcripts may have some error-containing reads that recur with sufficient frequency that they are assembled into distinct contigs, and gene families with many

members could generate chimeric contigs if they have identical subsequences that are longer than the *k*-mers used for assembly. In addition, templates derived from more than one gene may arise during the reverse transcription or PCR steps of library preparation; the resulting fusion reads may lead to incorrect assembly of contigs, as well as increase the proportion of such contigs substantially when the Oases merge function is used (from 3.6 to 12.2% in one reported case: Chu et al., 2013). The question about accession numbers can perhaps be explained by similar mechanisms. One approach that might help with these problems would be to assemble the reads initially with a greater *k*-mer length and coverage cut-off to obtain fewer misassemblies for the most abundant transcripts and gene families, remove the reads that map to these transcripts, and then assemble the remaining lower-coverage reads with less stringent thresholds.

The assemblies could also be improved by investigating more closely the contigs for which the species of origin could not be determined with the methods used to date. To this end, the performance of TopSort could probably be improved in one or more of several ways. First, it could be retrained with a new dataset that includes transcriptome or genome sequence from an axenic Clade A *Symbiodinium* strain (to improve recognition of contigs from the Clade A strain resident in *Aiptasia* stock CC7). Second, it could be extended such that it assigns contigs not just to the four groups used to date (cnidarian, dinoflagellate, fungi, and bacteria) but also to other groups (such as diatoms and ciliates) that may be present in nontrivial amounts in the guts and/or mucous layers of the anemones; this would also require retraining with a dataset that included unequivocal sequences from those groups. Third, the classification metrics could be extended to include also other sequence features found to be specific to the phyla of interest (such as the spliced leader sequences thought to be present in many dinoflagellate transcripts; Zhang et al., 2007). In addition to improving the performance of TopSort, its assignments could also be tested

further using alignment to transcriptome or genome sequence from a clonal, axenic *Symbiodinium* strain (preferably of Clade A), essentially as we have already done using *Aiptasia* genome sequence. Ultimately, however, some contigs may remain ambiguous in assignment until assembled genomes of both partner organisms are available for alignment.

Several additional issues will require more investigation as the relevant resources become available. First, the numbers of unigenes found here for both symbiotic partners (~14,000 by conservative estimate) are significantly less than those expected for the full genomes. For *Symbiodinium,* this probably reflects, at least in part, the existence of many genes that are expressed at significant levels only in free-living and/or stressed organisms. For *Aiptasia,* it presumably reflects the absence in the current transcriptome of genes that are expressed at significant levels only in specialized and non-abundant cell types (*e.g.,* in nerve and muscle), at other stages in development (*e.g.,* in embryos and larvae), or under other environmental conditions. Second, only 62 and 74% of representative *Aiptasia* transcripts could be annotated using SwissProt and nr, respectively, and these numbers were even lower (35 and 49%) for *Symbiodinium*. This incomplete annotation presumably reflects the poor representation in the databases of genes and proteins unique to these relatively understudied organisms, as well as the great phylogenetic distance of the dinoflagellates from more intensively studied groups. Finally, the *Symbiodinium* transcriptome reported here has not yet been analyzed in depth, in part because of a lack of informative comparisons to be made at this time. However, we should soon be able to use this transcriptome to make interesting comparisons of gene expression in this *Symbiodinium* strain growing in culture vs. *in hospite,* in this strain after exposure to various stressors, and in different *Symbiodinium* strains grown in this same host.

Differential expression of animal genes

Based on their consistent behavior in two separate RNA-Seq experiments done under somewhat different conditions, at least 920 genes appear to have significantly different expression between symbiotic and aposymbiotic anemones, including ≥85 for which there is an ≥5-fold change in expression. These findings show the value of a comprehensive analysis of differential expression, as earlier studies using microarrays had found much smaller numbers of differentially expressed genes (Ganot et al., 2011; Weis and Levine, 1996; Yuyama et al., 2011; Barneah et al., 2006; Kuo et al., 2010; Kuo et al., 2004; Rodriguez-Lanetty et al., 2006). We have focused our more detailed analyses to date on several groups of genes whose differential expression suggests interesting hypotheses about the biology of the symbiosis.

Genes controlling metabolism and transport in gastrodermal and epidermal cells

Given the intimate relationship between the symbiotic partners (Figure 2.1), we were not surprised to observe substantial changes in the expression of many genes encoding proteins involved in small-molecule metabolism and transport. We present here some speculative but testable hypotheses about how these changes might reflect the establishment and maintenance of symbiosis.

Glucose transport within gastrodermal cells

As glucose appears to be the major form in which fixed carbon is transferred from the dinoflagellate to the host (Whitehead, 2003; Burriesci et al., 2012), it was not surprising to find the transcripts for three presumed glucose transporters among those highly upregulated in symbiotic animals (Table 2.7, lines 1-3). As mammalian GLUT8 is localized to the endosome

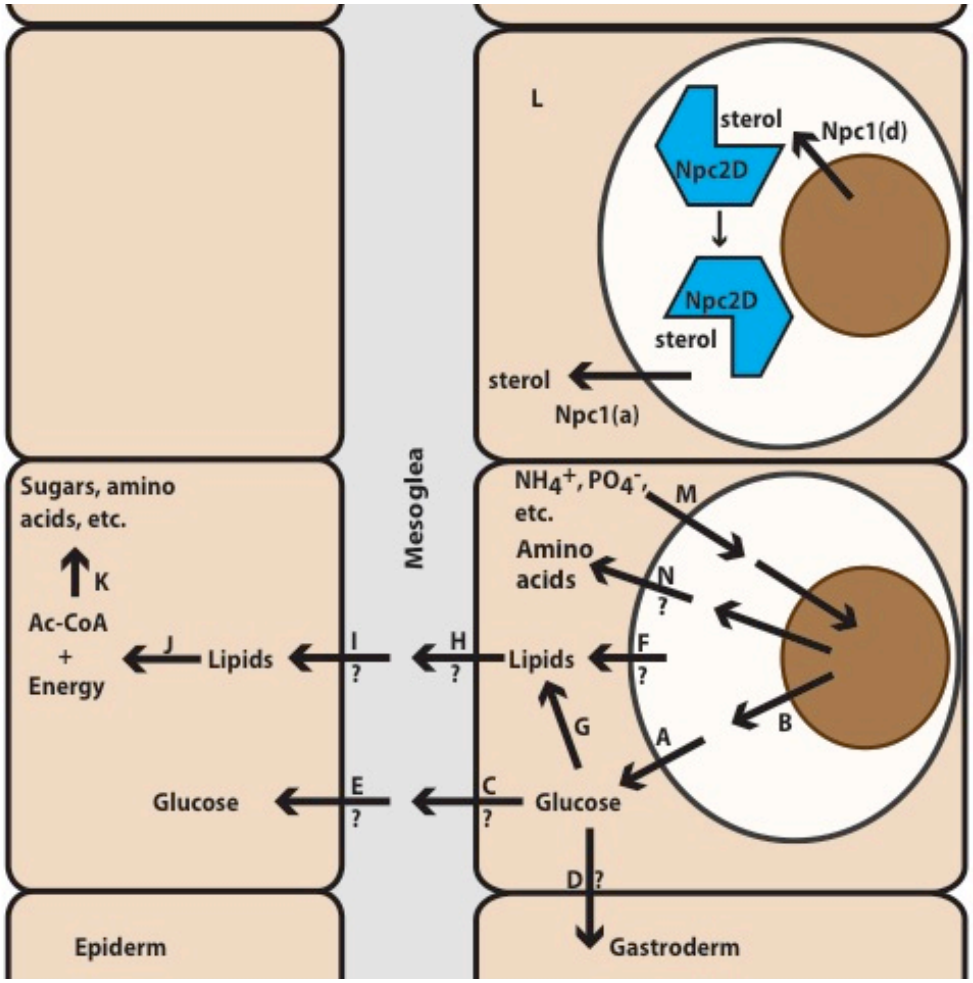


Figure 2.7 Summary of hypotheses about metabolism and metabolite transport as suggested by the gene-expression data and previously available information. Letters A through N are for reference in the text. Thick arrows across membranes, transporters hypothesized to be present in those membranes; thick arrows within cells, metabolic pathways hypothesized to be important in those cells; thin arrow, presumed diffusion of Npc2-sterol complexes; Npc1(a), the anemone-derived Npc1-like protein described in the text; Npc1(d), a presumed but as-yet-unidentified sterol transporter produced by the dinoflagellate and present in its plasma membrane; ?, hypotheses that we consider to be more problematic. Not shown because of their potential complexity are the other changes in inorganic-nutrient transport (e.g., a damping of NH_4^+ and CO_2 excretion across the apical plasma membranes of the gastrodermal cells) that are likely to occur upon the onset of symbiosis. All of these hypotheses should be testable through a combination of experiments including protein localization by immunofluorescence and/or cell fractionation, studies of separated gastrodermal and epidermal cell layers, sterol-binding experiments on Npc2 proteins, and others. See text for additional details.

membrane (Augustin et al., 2005), it is likely that one or both of the *Aiptasia* GLUT8 orthologs localize to the symbiosome membrane (Figure 2.7, A), a hypothesis that should be readily testable by immunolocalization studies once appropriate antibodies are available. The putative Na^+-glucose/*myo*-inositol co-transporters might also be involved in glucose transport across the symbiosome membrane. It should be noted that there must also be a *Symbiodinium* protein(s) that transports large amounts of glucose into the symbiosome lumen (Figure 2.7, B); it should be possible to identify the corresponding gene(s) among those expressed differentially in *Symbiodinium* cells growing *in hospite* relative to those growing in culture.

Possible glucose transport between cells

To our knowledge, there is currently no information as to whether and how the gastrodermal cells provide energy to the epidermal cells, cells in the mesoglea, and gastrodermal cells that lack dinoflagellates and/or access to nutrients from the gastric cavity, and as to whether these modes of nourishment change upon the establishment of symbiosis. Nourishment of the epidermal cells is a major issue because these cells neither contain algae nor have direct access to food, but they presumably require large amounts of energy for maintenance, reproduction, nematocyst replacement, and mucus production (which is extensive and has been reported to consume as much as 40% of the energy available to corals; Crossland et al., 1980). Thus, one or more of the upregulated glucose transporters might be found in the basolateral membranes of the gastrodermal cells (Figure 2.7, C and D), the basolateral membranes of the epidermal cells (Figure 2.7, E), or both. These questions should be resolvable by immunolocalization experiments and/or experiments in which gene expression is evaluated in separated tissue layers (Ganot el al, 2003).

Fatty-acid metabolism and transport

In a previous study, the transfer of fixed carbon from the algae to the host was more comprehensive in its evaluation of polar than of nonpolar compounds (Burriesci et al., 2012). Thus, it is conceivable that the upregulation of genes encoding proteins of fatty-acid metabolism and transport reflects a significant role of fatty acids in this transfer (Figure 2.7, F). Importantly, however (1) we saw upregulation of genes of both fatty-acid synthesis and fatty-acid breakdown (suggesting that gastrodermal and epidermal cells may be behaving differently) and (2) in most animals, high levels of glucose (such as those expected in the cytoplasm of gastrodermal cells harboring algae) inhibit β-oxidation and stimulate fatty-acid synthesis (Randle et al., 1998). Thus, we think it more likely that gastrodermal cells synthesize fatty acids from the glucose provided by the algae during the daytime (Figure 2.7, G), store them in neutral fats or wax esters, and subsequently mobilize them to serve as an energy supply at night and/or for transfer into the mesoglea (Figure 2.7, H) and then into the epidermal cells (Figure 2.7, I) (Crossland et al., 1980; Harland et al., 1993). The epidermal cells would metabolize the fatty acids by β-oxidation (Figure 2.7, J) to provide energy and acetyl-CoA building blocks. The apparent upregulation of the glyoxylate cycle (Figure 2.7, K; see Results) would be explained by the epidermal cells' need to synthesize carbohydrates, amino acids, etc., from acetyl-CoA. Testing of these hypotheses should be possible by immunolocalization of the relevant proteins, *in situ* mRNA hybridization, and/or examination of gene expression in separated gastrodermal and epidermal tissue layers.

Transport of sterols as building blocks or symbiosis signals

A previous study showed that the anemone *A. viridis* has at least two genes encoding Npc2-like proteins, one of which is in a sub-family distinct from that of the mammalian and

Drosophila Npc2 proteins and was upregulated in symbiotic gastrodermal tissue (Ganot et al., 2011). We have confirmed and extended these findings by showing that *Aiptasia* and several other cnidarians also contain multiple genes encoding Npc2-like proteins. In a phylogenetic analysis, one of the *Aiptasia* proteins clustered with the mammalian and *Drosophila* proteins and shares with them key residues implicated in cholesterol binding. Four other *Aiptasia* proteins, including the one (Npc2D) whose transcript is massively upregulated in symbiotic anemones, belong to a separate subfamily that also contains the upregulated *A. viridis* Npc2D; members of this subfamily do not share the residues implicated in cholesterol binding (Figure 2.2B). In contrast, we identified only one *Aiptasia* gene encoding an unequivocal Npc1-like protein; its expression did not change between aposymbiotic and symbiotic animals. These observations suggest the following speculative model (Figure 2.7, L). A *Symbiodinium*-encoded sterol transporter (*e.g.,* an Npc1-like protein) is present in the dinoflagellate plasma membrane and passes one or more dinoflagellate-synthesized sterols to the *Aiptasia*-encoded Npc2D, which is localized specifically to the symbiosome lumen. Npc2D in turn passes the sterol(s) to the *Aiptasia*-encoded Npc1 in the symbiosome membrane, which passes them to a sterol-carrier protein in the cytosol. If the same Npc1 protein functions in different membranes, in concert with all of the different Npc2 partners, and in other cell types as well as gastrodermal cells containing dinoflagellates, this could explain why its transcript is not significantly upregulated upon the onset of symbiosis.

The model to this point is agnostic about what sterols might be transferred, and to what end(s), but these are also important questions. The membranes of *Aiptasia,* like those of other animals, presumably contain cholesterol as an essential component. This cholesterol could be obtained from food, by *de novo* synthesis, or by modification of one or more of the distinctive,

non-cholesterol sterols (dinosterol, gorgosterols) produced in large quantities by many dinoflagellates (but not by other marine algae that have been investigated) (Giner et al 2001). Based on its sequence (Figure 2.2B), *Aiptasia* Npc2A seems the most likely to be involved in cholesterol traffic per se, whereas Npc2D (and Npc2B, C, and E: Figure 2.2) might all be involved in traffic of various other dinoflagellate-produced sterols. The latter might serve only as precursors of cholesterol, but the more intriguing possibility is that one or more of these molecules serves as the signal that a symbiosis-compatible dinoflagellate is present in the endosome/symbiosome.

Although the model of Figure 2.7L, is speculative, it is important to note that its major features should be testable by experiments that include (i) localization to the gastrodermal and/or epidermal cell layers of the expression of the several genes, (ii) protein localization by immunofluorescence and/or cell fractionation, (iii) using purified, labeled sterols to test the binding specificities of bacterially expressed Npc2 proteins, and (iv) tests of the abilities of exogenously added Npc2 proteins to complement the loss-of-cholesterol-transport phenotype in *NPC2*-knockout human cells (Ko et al., 2003). It is also worth noting that if this model is correct, Npc2D proteins would become an invaluable marker for the isolation of intact (i.e., non-ruptured) symbiosomes.

Transport and metabolism of inorganic nutrients and the coordination of nitrogen and carbon metabolism

It is clear that symbiotic cnidarians must transport CO_2, NH_4^+, and other inorganic nutrients across the symbiosome membrane – and indeed concentrate at least some of these materials within the symbiosome lumen – in order to provide their resident dinoflagellates with

these essential building blocks (Pernice et al., 2012). Thus, at least some of the many inorganic-nutrient transporters and transport-related proteins that are upregulated in symbiotic anemones (see Results) are presumably localized to the symbiosome membrane (Figure 2.7, M) or lumen. However, it seems virtually certain that the onset of symbiosis also induces changes in inorganic-nutrient transport across the various plasma-membrane domains. For example, aposymbiotic anemones presumably excrete CO_2 and NH_4^+ across the apical membranes of both epidermal and gastrodermal cells, whereas symbiotic anemones presumably reduce such excretion at least from the gastrodermal cells, and may even achieve a net uptake of both compounds from the environment. Although the possibilities are too many and complicated for ready depiction in Figure 2.7, determining the cellular (gastrodermal, epidermal, or both) and intracellular (symbiosome membrane, apical plasma membrane, basolateral plasma membrane, and/or other) localizations of the differentially regulated transporters and transport-related proteins should begin to answer many of these questions with a decisiveness that has not been possible before.

The changes in NH_4^+ movement upon symbiosis establishment also bear upon the probable linkage between carbon and nitrogen metabolism. In this regard, two non-mutually exclusive models have been put forward. The "nitrogen-recycling model" focuses on the possibility that continued host catabolism of amino acids produces NH_4^+ that is supplied to the dinoflagellate, which in turn releases amino acids for use by the host (Figure 2.7, N) (Wang and Douglas, 1998; Cates et al., 1976; Szmant-Froelich et al., 1977). In contrast, the "nitrogen-conservation model" focuses on the possibility that the fixed carbon provided by the dinoflagellate leads to a suppression of host amino-acid catabolism and therefore of the generation of NH_4^+ to be used by the dinoflagellate or excreted (Rees, 1986; Rees and Ellard, 1989; Wang and Douglas; 1998). Our data provide support for aspects of both models. In

particular, we observed upregulation of the genes of the GS-GOGAT cycle (the first dedicated step of NH_4^+ assimilation in animals) in symbiotic anemones, indicating that the release of carbon from the dinoflagellate promoted synthesis, rather than catabolism, of some amino acids. [The simultaneous upregulation of a presumably catabolic glutamate dehydrogenase might be taken as countervailing evidence, but this observation is difficult to interpret without knowing in which cell type(s) and cytoplasmic compartment this upregulation occurs.] Meanwhile, we also obtained strong support for previous evidence that cnidarians, like other animals, can only synthesize eight of the 20 amino acids found in proteins from intermediates of the central metabolic pathways. The remaining 12 amino acids must thus be obtained either from food or from the dinoflagellates, and our observation that the host genes for at least nine amino-acid transporters are upregulated in symbiotic anemones (see Results) suggests strongly that the dinoflagellates indeed contribute to the host's amino-acid supply. It should be informative to determine the localizations and amino-acid specificities of these host-encoded transporters as well as identify any amino-acid transporters expressed differentially by *Symbiodinium in hospite* (as it is unlikely that free-living dinoflagellates would export amino acids into the surrounding seawater).

Recognition and tolerance of dinoflagellate symbionts by the host

Establishment and maintenance of the mutualistic relationship also require that the host recognize and tolerate the dinoflagellate symbionts. Thus, it was not surprising that an unbiased screen for functional groups that were enriched among the differentially expressed genes revealed three groups that might be involved in these processes, as discussed below.

Response to oxidative stress

Both *a priori* logic and considerable experimental evidence support the view that possession of an intracellular photosynthetic symbiont imposes oxidative stress on the host, particularly under conditions in which chloroplast damage may result in enhanced production of reactive oxygen species (ROS) (D'Aoust et al., 1976; Lesser, 1997; Venn et al., 2008). If not detoxified, ROS can damage DNA, proteins, and lipids (Lesser, 2006), and it is widely believed that ROS production under stress is the major trigger of symbiosis breakdown during bleaching (Weis 2008; Venn et al., 2008; Jones et al., 1998; Downs et al., 2002; Lesser, 2011). Surprisingly, however, our results provide no support for this model: most of the differentially regulated genes in this GO category were actually downregulated in symbiotic animals, and the two that were upregulated do not seem likely to be involved in ROS detoxification (see Results). Previous studies of other species of anemones have also found host genes thought to be involved in ROS detoxification (copper/zinc superoxide dismutase and glutathione S-transferase) to be downregulated in symbiotic relative to aposymbiotic individuals (Ganot et al., 2011; Rodriguez-Lanetty et al., 2006). Although other studies have indicated that symbiotic cnidarians have higher superoxide-dismutase activities than their aposymbiotic counterparts (Dykens and Shick, 1982; Furla et al., 2005), it was not determined whether the enzyme was of host or dinoflagellate origin. Thus, it is possible that the hosts are protected from ROS by symbiont-generated antioxidants and can reduce the expression of their own enzymes (Rodriguez-Lanetty et al., 2006). These other studies, like our own, were conducted under conditions thought to be non-stressful, and it is possible that a different picture would emerge under stressful conditions. In that regard, however, we have also recently observed that bleaching under heat stress can occur rapidly in the dark, when photosynthetically produced ROS cannot be present (Tolleter et al.,

2013).

Inflammation/tissue remodeling/response to wounding

Inflammation is a protective tissue response to injury or pathogens that serves to destroy, dilute, and/or wall off both the injurious agent and the injured tissue (Sparks, 1985). In invertebrates, including anthozoans, the inflammation-like response involves both cellular and humoral aspects, including the infiltration of immune cells such as amoebocytes and granular cells (Patterson and Landolt, 1979; Palmer et al., 2008; Mydlarz et al., 2008), phagocytosis and/or encapsulation of foreign material (Patterson and Landolt, 1979; Mydlarz et al., 2008; Olano and Bigger, 2000; Petes et al., 2003;), and the production of cytotoxic molecules such as ROS, nitric oxide, lysozyme, antimicrobial peptides, and intermediates of the phenoloxidase cascade (Palmer et al., 2008; Mydlarz et al., 2008; Hutton and Smith, 1996; Perez and Weis, 2006). Our data suggest that the establishment of symbiosis is associated with an overall attenuation of the inflammatory response (see Results), presumably to allow the dinoflagellate to co-exist peacefully with the host rather than being attacked as a harmful invader. Other studies also support this conclusion and suggest that it may be a general feature of the means by which animal hosts accommodate symbiotic microbes. For example, when symbiotic and aposymbiotic *Aiptasia* were challenged with bacterial lipopolysaccharide (LPS), the former produced much less nitric oxide than did the latter (Detournay et al., 2012), and in two hard coral species with inflammatory-like responses, dinoflagellate densities were lower in the "inflamed" than in the adjacent healthy tissues (Palmer et al., 2008). Similarly, successful colonization of squid light organs by symbiotic bacteria is associated with an irreversible attenuation of host nitric oxide production (Davidson et al., 2004; Altura et al., 2011).

Of particular interest because of its massive upregulation in symbiotic anemones is the gene encoding scavenger receptor B class member 1 (SRB1); upregulation of SRB1 was also observed previously in symbiotic individuals of the anemone *Anthopleura* (Rodriguez-Lanetty et al., 2006). SRB1 is a member of the CD36 protein family and is a transmembrane cell-surface glycoprotein that has been implicated in multiple functions, including lipid transport (see Figure 2.2, Table 2.7, and associated text), cell adhesion, wound healing, apoptosis, and innate immunity (Areschoug and Gordon, 2009). Perhaps of most interest is its role in *Plasmodium* infection, as these apicomplexan parasites are a sister taxon to the dinoflagellates (Baldauf, 2003). SRB1 has been shown to boost host hepatocyte permissiveness to *Plasmodium* infection, promote parasite development by acting as major lipid provider, and enable adhesion between *Plasmodium*-infected and uninfected erythrocytes, thus allowing for movement of parasites between host cells (Adams et al., 2005; Yalaoui et al., 2008; Rodrigues et al., 2008). It is possible that SRB1 has similar functions in the cnidarian-dinoflagellate symbiosis, and further investigation of its function by protein localization and knockdown of function should be highly informative.

Apoptosis/cell death

The possible roles of apoptosis and necrotic cell death in the breakdown of symbiosis under stress have been investigated (Dunn et al., 2002; Dunn et al., 2004; Dunn et al., 2007; Richier et al., 2006; Pernice et al., 2011; Tchernov et al., 2011; Ainsworth et al., 2011; Kvitt et al., 2011), and a role for apoptosis in the post-phagocytic selection of compatible symbionts has also been suggested (Dunn et al., 2009). However, the possible role of apoptosis in maintenance of a stable symbiotic relationship has not been addressed experimentally. Others have suggested

that apoptosis might contribute to the dynamic equilibrium between host and symbiont cell growth and proliferation that is presumably necessary to ensure a stable relationship (Muscatine and Pool, 1979; Fitt, 2000; Davy et al., 2012), and our observation that 13 apoptosis/cell death-related genes were differentially expressed (some massively so) in symbiotic relative to aposymbiotic anemones is broadly consistent with this possibility. However, the complexity of the apoptotic pathways and the fact that a single protein can have either pro- or anti-apoptotic function depending on its localization and/or the presence/absence of other specific signals makes it impossible to draw firm conclusions from gene-expression data alone.

Nonetheless, it is worth noting the potential role of tumor-necrosis factor (TNF) family members and their associated proteins, which are prominent regulators of cell survival, proliferation, and differentiation in both vertebrates and invertebrates (reviewed by Branschädel et al., 2007). We found a TNF-family ligand, a TNF receptor, a receptor-associated factor, and the functionally related "growth-arrest and DNA-damage-inducible protein" all to be upregulated in symbiotic anemones (1.9-, 60-, 1.8-, and 5.1-fold, respectively). These proteins are capable of inducing caspase-dependent apoptosis via at least two different pathways (Zhang et al., 1999; Sinha and Chaudhary, 2004; Burkly et al., 2007; Sabour et al., 2012), as well as of activating the multi-functional NFκB and MAPK pathways (Sinha and Chaudhary, 2004; Burkly et al., 2007), so that they may coordinate multiple biological processes to regulate symbiotic stability. Interestingly, genes encoding TNF receptors and receptor-associated proteins were also prominent among the genes found to be upregulated in corals living under chronic, mild, heat stress (Barshis et al., 2013).

Symbiont tolerance likely requires the regulation of a number of biological processes. Response to oxidative stress, inflammation, and/or apoptosis could play important roles in the

symbiotic homeostasis via the regulation of cytotoxic symbiont products, enabling persistence of the symbiont in host cells/tissues, and coordinating host and symbiont populations. To test these hypotheses, target gene approaches, including knock-down or knock-out studies and in situ hybridization, should be conducted to determine if genes involved in these processes function similarly to their invertebrate or vertebrate counterparts and/or to determine if the onset and maintenance of the cnidarian-dinoflagellate mutualism is dependent on these candidate genes.

REFERENCES

Adams, Y., Smith, S.L., Schwartz-Albiez, R., and Andrews, K.T. (2005) Carrageenans inhibit the in vitro growth of *Plasmodium falciparum* and cytoadhesion to CD36. Parasitology Research 97: 290–294.

Ainsworth, T.D., Wasmund, K., Ukani, L., Seneca, F., Yellowlees, D., et al. (2011) Defining the tipping point. A complex cellular life/death balance in corals in response to stress. Scientific reports 1: 160.

Altura, M. A., Stabb, E., Goldman, W., Apicella, M., and McFall-Ngai, M.J. (2011) Attenuation of host NO production by MAMPs potentiates development of the host in the squid-*Vibrio* symbiosis. Cellular microbiology 13: 527–537.

Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. Genome biology 11: R106.

Areschoug, T., and Gordon, S. (2009) Scavenger receptors: role in innate immunity and microbial pathogenesis. Cellular Microbiology 11: 1160–1169.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics 25: 25–29.

Augustin, R., Riley, J., and Moley, K.H. (2005) GLUT8 contains a [DE]XXXL[LI] sorting motif and localizes to a late endosomal/lysosomal compartment. Traffic 6: 1196–1212.

Baldauf, S.L. (2003) The deep roots of eukaryotes. Science 300: 1703–1706.

Barneah, O., Benayahu, Y., and Weis, V.M. (2006) Comparative proteomics of symbiotic and aposymbiotic juvenile soft corals. Marine biotechnology 8: 11–16.

Barshis, D.J., Ladner, J.T., Oliver, T.A., Seneca, F.O., Traylor-Knowles N, et al. (2013) Genomic basis for coral resilience to climate change. Proceedings of the National Academy of Sciences of the United States of America 110: 1387–1392.

Branschädel, M., Boschert, V., and Krippner-Heidenreich, A. (2007) Tumour necrosis factors. eLS.

Burkly, L.C., Michaelson, J.S., Hahm, K., Jakubowski, A., and Zheng, T.S. (2007) TWEAKing tissue remodeling by a multifunctional cytokine: role of TWEAK/Fn14 pathway in health and disease. Cytokine 40: 1–16.

Burriesci, M.S., Raab, T.K., and Pringle, J.R. (2012) Evidence that glucose is the major transferred metabolite in dinoflagellate-cnidarian symbiosis. The Journal of experimental biology 215: 3467–3477.

Burriesci, M.S. (2011) Developing *Aiptaisa pallida* as a tractable model system for cnidarian-dinoflagellate symbiosis: identifying transferred metabolites and designing tools for analysis of ultra-high-thoughput-sequencing data, Stanford University.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., et al. (2009) BLAST+: architecture and applications. BMC Bioinformatics 10: 421.

Cates, N., and McLaughlin, J. (1976) Differences of ammonia metabolism in symbiotic and aposymbiotic *Condylactus* and *Cassiopea* spp. Journal of Experimental Marine Biology and Ecology 21: 1–5.

Chu, H.T., Hsiao, W.W.L., Chen, J.C., Yeh, T.J., Tsai, M.H., et al. (2013) EBARDenovo: Highly accurate de novo assembly of RNA-Seq with efficient chimera-detection. Bioinformatics: 1–7.

Conesa, A., Götz, S., García-Gómez, J., Terol, J., Talón, M., et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674.

Crossland, C.J., Barnes, D.J., and Borowitzka, M.A. (1980) Diurnal lipid and mucus production in the staghorn coral *Acropora acuminata*. Marine Biology 60: 81–90.

D'Aoust, B., White, R., Wells, J., and Olsen, D. (1976) Coral-algal associations: capacity for producing and sustaining elevated oxygen tensions in situ. Undersea Biomedical Research 3: 35–40.

Davidson, S.K., Koropatnick, T.A., Kossmehl, R., Sycuro, L., and McFall-Ngai, M.J. (2004) NO means "yes" in the squid-*Vibrio* symbiosis: nitric oxide (NO) during the initial stages of a beneficial association. Cellular Microbiology 6: 1139–1151.

Davy, S.K., Allemand ,D., and Weis, V.M. (2012) Cell biology of cnidarian-dinoflagellate symbiosis. Microbiology and molecular biology reviews : MMBR 76: 229–261.

De'ath, G., Fabricius, K.E., Sweatman, H., and Puotinen, M. (2012) The 27-year decline of coral cover on the Great Barrier Reef and its causes. Proceedings of the National Academy of Sciences of the United States of America 109: 17995–17999.

Dennis, G., Sherman, B.T., Hosack, D., Yang J., Gao, W., et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome biology 4: P3-P14.

Detournay, O., Schnitzler, C.E., Poole, A., and Weis, V.M. (2012) Regulation of cnidarian–dinoflagellate mutualisms: Evidence that activation of a host TGFβ innate immune

pathway promotes tolerance of the symbiont. Developmental & Comparative Immunology 38: 525–537.

Downs, C.A., Fauth, J..E, Halas, J.C., Dustan, P., Bemiss, J., et al. (2002) Oxidative stress and seasonal coral bleaching. Free Radical Biology and Medicine 33: 533–543.

Dudler, N., Yellowlees, D., and Miller, D.J. (1987) Localization of Two L-Glutamate Dehydrogenases in the Coral *Acropora latistella*. Archives of Biochemistry and Biophysics 254: 368–371.

Dunn, S.R., Bythell, J.C., Le Tissier, M.D.A., Burnett, W.J., and Thomason, J.C. (2002) Programmed cell death and cell necrosis activity during hyperthermic stress-induced bleaching of the symbiotic sea anemone *Aiptasia* sp. Journal of Experimental Marine Biology and Ecology 272: 29–53.

Dunn, S.R., Thomason, J.C., Le Tissier, M.D.A., and Bythell, J.C. (2004) Heat stress induces different forms of cell death in sea anemones and their endosymbiotic algae depending on temperature and duration. Cell death and differentiation 11: 1213–1222.

Dunn SR, Phillips WS, Green DR, Weis VM (2007) Knockdown of actin and caspase gene expression by RNA interference in the symbiotic anemone *Aiptasia pallida*. The Biological Bulletin 212: 250–258.

Dunn, S.R., and Weis, V.M. (2009) Apoptosis as a post-phagocytic winnowing mechanism in a coral-dinoflagellate mutualism. Environmental microbiology 11: 268–276.

Dykens, J.A., and Shick, J.M. (1982) Oxygen production by endosymbiotic algae controls superoxide dismutase activity in their animal host. Nature 297: 579–580.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32: 1792–1797.

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

Fitt, W.K. (2000) Cellular growth of host and symbiont in a cnidarian-zooxanthellar symbiosis. The Biological bulletin 198: 110–120.

Frolov, A., Zielinski, S.E., Crowley, J.R., Dudley-Rucker, N., Schaffer, J.E., et al. (2003) NPC1 and NPC2 regulate cellular cholesterol homeostasis through generation of low density lipoprotein cholesterol-derived oxysterols. The Journal of biological chemistry 278: 25517–25525.

Furla, P., Allemand, D., Shick, J.M., Ferrier-Pagès, C., Richier, S., et al. (2005) The Symbiotic Anthozoan: A Physiological Chimera between Alga and Animal. Integrative and comparative biology 45: 595–604.

Ganot, P., Moya, A., Magnone, V., Allemand, D., Furla, P., et al. (2011) Adaptations to endosymbiosis in a cnidarian-dinoflagellate association: differential gene expression and specific gene duplications. PLoS genetics 7: e1002187.

Giner, J.L., and Wikfors, G.H. (2011) "Dinoflagellate Sterols" in marine diatoms. Phytochemistry 72: 1896–1901.

Guedes, R.L.M., Prosdocimi, F., Fernandes, G.R., Moura, L.K., Ribeiro, H.L., et al. (2011) Amino acids biosynthesis and nitrogen assimilation pathways: a great genomic deletion during eukaryotes evolution. BMC genomics 12 Suppl 4: S2.

Harland, A. (1990) Zinc and cadmium absorption in the symbiotic anemone *Anemonia viridis* and the non-symbiotic anemone *Actinia equina*. Journal of the Marine Biological Association of the United Kingdom 70: 789–802.

Harland, A.D., Navarro, J.C., Spencer, Davies, P., and Fixter, L.M. (1993) Lipids of some Caribbean and Red Sea corals: total lipid, wax esters, triglycerides and fatty acids. Marine Biology 117: 113–117.

Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., et al. (2007) WoLF PSORT: protein localization predictor. Nucleic acids research 35: W585–7.

Huang, S., Yuan, S., Guo, L., Yu, Y., Li, J., et al. (2008) Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome research 18: 1112–1126.

Hughes, T.P., Baird, H., Bellwood, D.R., Card, M., Connolly, S.R., et al. (2003) Climate change, human impacts, and the resilience of coral reefs. Science 301: 929–933.

Hutton, D.M.C., and Smith, V.J. (1996) Antibacterial Properties of Isolated Amoebocytes from the Sea Anemone *Actinia equina*. Biological Bulletin 191: 441.

Infante, R.E., Wang, M.L., Radhakrishnan, A., Kwon, H.J., Brown, M.S., et al. (2008) NPC2 facilitates bidirectional transfer of cholesterol between NPC1 and lipid bilayers, a step in cholesterol egress from lysosomes. Proceedings of the National Academy of Sciences of the United States of America 105: 15287–15292.

Jones, R.J., Larkum, A.W.D., and Schreiber, U. (1998) Temperature-induced bleaching of corals begins with impairment of the CO_2 fixation mechanism in zooxanthellae. Plant, cell & environment 21: 1219–1230.

Kaldenhoff, R. (2012) Mechanisms underlying CO_2 diffusion in leaves. Current opinion in plant biology 15: 276–281.

Ko, D.C., Binkley, J., Sidow, A., and Scott, M.P. (2003) The integrity of a cholesterol-binding pocket in Niemann-Pick C2 protein is necessary to control lysosome cholesterol levels. Proceedings of the National Academy of Sciences of the United States of America 100: 2518–2525.

Kondrashov, F., Koonin, E. V., Morgunov, I.G., Finogenova, T. V., and Kondrashova, M.N. (2006) Evolution of glyoxylate cycle enzymes in Metazoa: evidence of multiple horizontal transfer events and pseudogene formation. Biology direct 1: 31.

Kuo, J., Liang, Z., and Lin, C. (2010) Suppression subtractive hybridization identifies genes correlated to symbiotic and aposymbiotic sea anemone associated with dinoflagellate. Journal of Experimental Marine Biology and Ecology 388: 11–19.

Kuo, J., Chen, M.C., Lin, C.H., and Fang, L.S. (2004) Comparative gene expression in the symbiotic and aposymbiotic *Aiptasia pulchella* by expressed sequence tag analysis. Biochemical and biophysical research communications 318: 176–186.

Kvitt, H., Rosenfeld, H., Zandbank, K., and Tchernov, D. (2011) Regulation of apoptotic pathways by *Stylophora pistillata* (Anthozoa, Pocilloporidae) to survive thermal stress and bleaching. PLoS ONE 6: e28665.

Lajeunesse, T., Parkinson, J., and Reimer, J. (2012) A genetics-based description of *Symbiodinium minutum* sp. nov. and *S. psygmophilum* sp. nov. (Dinophyceae), two dinoflagellates symbiotic with Cnidaria. Journal of phycology 48: 1380–1391.

Lehnert, E.M., Burriesci, M.S., and Pringle, J.R. (2012) Developing the anemone *Aiptasia* as a tractable model for cnidarian-dinoflagellate symbiosis: the transcriptome of aposymbiotic *A. pallida*. BMC genomics 13: 271.

Lesser, M.P. (1997) Oxidative stress causes coral bleaching during exposure to elevated temperatures. Planta 16: 187–192.

Lesser, M. (2006) Oxidative stress in marine environments: Biochemistry and Physiological Ecology. Annual Review of Physiology 68: 253–278.

Lesser, M.P. (2011) Coral bleaching: causes and mechanisms. Coral reefs: An ecosystem in transition. Springer. pp. 405–419.

Li, H., and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Magoc, T. and Salzberg, S.L. (2011) FLASH : Fast Length Adjustment of Short Reads to Improve Genome Assemblies Tanja Magoˇ. Bioinformatics: 1–8.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal: 10–12.

Miflin, B.J., and Habash, D.Z. (2002) The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. Journal of experimental botany 53: 979–987.

Muscatine, L., and Pool, R.R. (1979) Regulation of numbers of intracellular algae. Proceedings of the Royal Society B: Biological Sciences 204: 131–139.

Muscatine, L., Falkowski, P.G., Porter, J.W., Dubinsky, Z. (1984) Fate of Photosynthetic Fixed Carbon in Light- and Shade-Adapted Colonies of the Symbiotic Coral *Stylophora pistillata*. Proceedings of the Royal Society B: Biological Sciences 222: 181–202.

Mydlarz, L.D., Holthouse, S.F., Peters, E.C., Harvell, C.D. (2008) Cellular Responses in Sea Fan Corals: Granular Amoebocytes React to Pathogen and Climate Stressors. PLoS ONE 3: 9.

Nelson, R.E., Fessler, L.I., Takagi, Y., Blumberg, B., Keene, D.R., et al. (1994) Peroxidasin: a novel enzyme-matrix protein of *Drosophila* development. The EMBO journal 13: 3438–3447.

Olano, C.T,. and Bigger, C.H. (2000) Phagocytic activities of the gorgonian coral *Swiftia exserta*. Journal of Invertebrate Pathology 76: 176–184.

Palmer, C. V., Mydlarz, L.D., and Willis, B.L. (2008) Evidence of an inflammatory-like response in non-normally pigmented tissues of two scleractinian corals. Proceedings of the Royal Society B Biological Sciences 275: 2687–2693.

Patterson, M.J., and Landolt, M.L. (1979) Cellular reaction to injury in the anthozoan *Anthopleura elegantissima*. Journal of Invertebrate Pathology 33: 189–196.

Perez, S., and Pringle, J. (2013) Regular induction of spawning and larvae production by the anemone *Aiptasia pallida* under laboratory conditions. Submitted.

Perez, S., and Weis, V. (2006) Nitric oxide and cnidarian bleaching: an eviction notice mediates breakdown of a symbiosis. The Journal of experimental biology 209: 2804–2810.

Pernice, M., Dunn, S., Miard, T., and Dufour, S. (2011) Regulation of apoptotic mediators reveals dynamic responses to thermal stress in the reef building coral *Acropora millepora*. PloS ONE 6: e16095.

Pernice, M., Meibom, A., Van Den Heuvel, A., Kopp, C., Domart-Coulon, I., et al. (2012) A single-cell view of ammonium assimilation in coral-dinoflagellate symbiosis. The ISME journal: 1–11.

Petes, L.E., Harvell, C.D., Peters, E.C., Webb, M.A.H., and Mullen, K.M. (2003) Pathogens compromise reproduction and induce melanization in Caribbean sea fans. Marine Ecology Progress Series 264: 167–171.

Randle, P.J. (1998) Regulatory interactions between lipids and carbohydrates: the glucose fatty acid cycle after 35 years. Diabetes/metabolism reviews 14: 263–283.

Rees, T. A. V. (1986) The Green Hydra Symbiosis and Ammonium I. The Role of the Host in Ammonium Assimilation and its Possible Regulatory Significance. Proceedings of the Royal Society B: Biological Sciences 229: 299–314.

Rees T., and Ellard, F. (1989) Nitrogen conservation and the green hydra symbiosis. Society 236: 203–212.

Richier, S., Sabourault, C., Courtiade, J., Zucchini, N., Allemand, D., et al. (2006) Oxidative stress and apoptotic events during thermal stress in the symbiotic sea anemone, *Anemonia viridis*. The FEBS journal 273: 4186–4198.

Rodrigues, C.D., Hannus, M., Prudêncio, M., Martin, C., Gonçalves, L.A., et al. (2008) Host scavenger receptor SR-BI plays a dual role in the establishment of malaria parasite liver infection. Cell host microbe 4: 271–282.

Rodriguez-Lanetty, M., Phillips, W.S., and Weis, V.M. (2006) Transcriptome analysis of a cnidarian-dinoflagellate mutualism reveals complex modulation of host gene expression. BMC Genomics 7: 23.

Ronquist, F., and Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Sabour Alaoui, S., Dessirier, V., De Araujo, E., Alexaki, V.I., Pelekanou, V., et al. (2012) TWEAK Affects Keratinocyte G2/M Growth Arrest and Induces Apoptosis through the Translocation of the AIF Protein to the Nucleus. PLoS ONE 7: e33609.

Sabourault, C., Ganot, P., Deleury, E., Allemand, D., and Furla, P. (2009) Comprehensive EST analysis of the symbiotic sea anemone, *Anemonia viridis*. BMC Genomics 10: 333.

Schoenberg, D.A., and Trench, R.K. (1980) Genetic Variation in *Symbiodinium* (=Gymnodinium) *microadriaticum* Freudenthal, and Specificity in its Symbiosis with Marine Invertebrates. I. Isoenzyme and Soluble Protein Patterns of Axenic Cultures of *Symbiodinium microadriaticum*. Proceedings of the Royal Society B Biological Sciences 207: 405–427.

Schulz, M.H., Zerbino, D.R., Vingron, M., and Birney, E. (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 28: 1086–1092.

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., et al. (2011) Using the *Acropora digitifera* genome to understand coral responses to environmental change. Nature 476: 320–323.

Sinha. S.K., and Chaudhary, P.M. (2004) Induction of apoptosis by X-linked ectodermal dysplasia receptor via a caspase 8-dependent mechanism. The Journal of Biological Chemistry 279: 41873–41881.

Sleat, D.E., Wiseman, J., El-Banna, M., Price, S.M., Verot, L., et al. (2004) Genetic evidence for nonredundant functional cooperativity between NPC1 and NPC2 in lipid transport. Proceedings of the National Academy of Sciences of the United States of America 101: 5886–5891.

Sparks, A.K., (1985) Synopsis of invertebrate pathology-exclusive of insects. Elsevier Science Publishers BV (Biomedical Division).

Sunagawa, S., Wilson, E.C., Thaler, M., Smith, M.L., Caruso, C., et al. (2009) Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. BMC Genomics 10: 258.

Swanson, R., and Hoegh-Guldberg, O. (1998) Amino acid synthesis in the symbiotic sea anemone *Aiptasia pulchella*. Marine Biology 131: 83–93.

Szmant-Froelich, A., and Pilson, M. (1977) Nitrogen excretion by colonies of the temperate coral *Astrangia danae* with and without zooxanthellae. Proc. 3rd int. Coral Reef Symp. Vol. 1. pp. 417–424.

Tchernov, D., Kvitt, H., Haramaty, L., Bibby, T.S., Gorbunov, M.Y., et al. (2011) Apoptosis and the selective survival of host animals following thermal bleaching in zooxanthellate corals. Proceedings of the National Academy of Sciences of the United States of America 108: 9905–9909.

Tolleter, D., Seneca, F., DeNofrio, J., Palumbi, S., Pringle, J., et al. (2013) Coral bleaching independent of photosynthetic activity. Current Biology, in press.

Uehlein, N., Sperling, H., Heckwolf, M., and Kaldenhoff, R. (2012) The A*rabidopsis* aquaporin PIP1;2 rules cellular CO_2 uptake. Plant, cell & environment 35: 1077–1083.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 3: research0034.1-research0034.11.

Venn, A.A., Loram, J.E., and Douglas, A.E. (2008) Photosynthetic symbioses in animals. Journal of experimental botany 59: 1069–1080.

Wang, J., and Douglas, A. (1998) Nitrogen recycling or nitrogen conservation in an alga-invertebrate symbiosis? The Journal of Experimental Biology 201: 2445–2453.

Wang, J.T., and Douglas, A. E. (1999) Essential amino acid synthesis and nitrogen recycling in an alga-invertebrate symbiosis. Marine Biology 135: 219–222.

Wang, M.L., Motamed, M., Infante, R.E., Abi-Mosleh, L., Kwon, H.J., et al. (2010) Identification of surface residues on Niemann-Pick C2 essential for hydrophobic handoff of cholesterol to NPC1 in lysosomes. Cell metabolism 12: 166–173.

Weis V.M., and Levine, R.P. (1996) Differential protein profiles reflect the different lifestyles of symbiotic and aposymbiotic *Anthopleura Elegantissima*, a sea anemone from temperate waters. The Journal of experimental biology 199: 883–892.

Weis, V.M. (2008) Cellular mechanisms of Cnidarian bleaching: stress causes the collapse of symbiosis. The Journal of experimental biology 211: 3059–3066.

Whitehead, L.F. (2003) Metabolite comparisons and the identity of nutrients translocated from symbiotic algae to an animal host. Journal of Experimental Biology 206: 3149–3157.

Wright, P.A. (1995) Reciew of nitrogen excretion: three end products, many physiological roles. 281: 273–281.

Xiang, T., Hambleton, E., DeNofrio, J., Pringle, J.R., and Grossman, A. (2013) Isolation of clonal, axenic strains of the symbiotic dinoflagellate *Symbiodinium* and their growth and host specificity. Journal of Phycology 49: 447-458.

Yalaoui, S., Huby, T., Franetich, J.F., Gego, A., Rametti, A., et al. (2008) Scavenger receptor BI boosts hepatocyte permissiveness to *Plasmodium* infection. Cell host microbe 4: 283–292.

Yuyama, I., Watanabe, T., and Takei, Y. (2011) Profiling differential gene expression of symbiotic and aposymbiotic corals using a high coverage gene expression profiling (HiCEP) analysis. Marine Biotechnology 13: 32-40.

Zerbino, D.R., and Birney. E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome research 18: 821–829.

Zhang, W., Bae, I., Krishnaraju, K., Azam, N., Fan, W., et al. (1999) CR6: A third member in the MyD118 and Gadd45 gene family which functions in negative growth control. Oncogene 18: 4899–4907.

Zhang , H., Hou, Y., Miranda, L., Campbell, D. A., Sturm, N.R., et al. (2007) Spliced leader RNA trans-splicing in dinoflagellates. Proceedings of the National Academy of Sciences of the United States of America 104: 4618–4623.

Zhao, S., and Fernald, R.D. (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. Journal of Computational Biology 12: 1047–1064.

CHAPTER 3

CHARACTERIZATION OF THE *AIPTASIA PALLIDA* TRANSCRIPTIONAL RESPONSE TO PATHOGEN EXPOSURE PART 1: THE APOSYMBIOTIC HOST RESPONSE[1]

Abstract

The animal immune response interacts with both mutualistic and pathogenic microorganisms to enable resistance and/or tolerance. In cnidarians (*e.g.,* corals and sea anemones), we have a poor understanding of the host immune response to pathogenic microbes A better understanding of infection responses would not only provide information regarding the evolution of immunity in lower metazoans, but will also inform the study of how the host responds to mutualistic symbionts. Here we report the host transcriptome response of the aposymbiotic sea anemone, *Aiptasia pallida,* to exposure to *Serratia marcescens,* a pathogenic bacterium of cnidarians. We detected the differential expression of numerous tumor necrosis factor receptor-associated factor (TRAF) family members and their up- and down-stream mediators, which are ultimately responsible for the activation and/or repression of varied immune pathways. Apoptosis also appears to play a critical role in the host immune response, as numerous genes with apoptotic functions were differentially expressed and enrichment analysis revealed overrepresentation of apoptotic genes among those with altered expression. As apoptosis has also been found to play an important role in the host response to mutualists, further investigation into the role of apoptosis in microbial interactions may shed light on the similarities and differences between mutualists and pathogen infection. In addition, the differential expression of numerous transcripts with functions in ubiquitination suggests the importance of protein degradation and activation in the dynamic processes involved in the host response to

[1] Co-authors include J. A. Schwarz and C. D. Harvell

bacteria. Our experimental approach provides a holistic view of the pathways and processes elicited via pathogenic microbes and a better understanding of the evolution of immunity in lower metazoans. In addition, we have generated a list of candidate genes that can be investigated in further detail to better characterize the cnidarian immune repertoire and its role in both pathogenic and mutualistic interactions.

Introduction

Interactions between microbes and animal hosts are ubiquitous, and can range from pathogenic associations that can ultimately culminate in disease to mutualistic associations that promote partner fitness. Although the ultimate outcome of mutualistic and pathogenic relationships is by definition very different from the host's perspective, there are similarities between symbionts and pathogens. Both must overcome similar barriers to invasion, establish a niche, and multiply within the host. In addition, the two experience the similar coevolutionary pressure to adapt to changes in host morphology and/or physiology. Indeed, it has become increasing apparent that both host-pathogen and host-symbiont interactions rely on conserved features of the innate immune system that enable resistance and/or tolerance to microbial invaders (Chen et al., 2003; Chen et al., 2005; Dale et al., 2002; McFall-Ngai et al., 2012).

Cnidarian (*e.g.,* sea anemones and corals)-microbe interactions offer a tractable system to investigate the interplay between host immunity, pathogenesis, and symbiosis. Many cnidarians harbor intracellular photosynthetic dinoflagellates (*Symbiodinium* spp.) in a mutualistic relationship (Trench, 1987). This relationship underlies the trophic and structural foundation of reef ecosystems (Trench, 1997). Cnidarians are also regularly exposed to potential pathogens via their microbial-rich aqueous habitats. Research into the role of cnidarian immunity in host-

microbe interactions is especially important given the recent decline of coral reef ecosystems. Indeed, two of the most common causes of coral reef degradation, coral disease and coral bleaching (i.e. the stress-induced breakdown of the cnidarian-dinoflagellate symbiosis; Harvell et al., 2007; Hoegh-Guldberg, 1999; Hoegh-Guldberg et al., 2007), are characterized by alterations in host-microbe relationships that are likely mediated by the immune system. Unfortunately, we have a poor understanding of how cnidarians recognize and respond to pathogens, let alone the role of host immunity in regulating mutualistic interactions. A better characterization of the cnidarian response to pathogens is needed to provide a framework for understanding the host response to symbionts, as well as for probing questions relating to if/how symbionts modulate immunity.

The cnidarian immune response, like that of other animals, is characterized by recognition, signaling, and effector responses (Reviewed in Bosch, 2013; Dunn, 2009; and Palmer and Traylor-Knowles, 2012). Much of what we know about recognition and signaling comes from the identification of homologs of well-known immune mediators, including Toll-like receptors and MyD88, from several extensive cnidarian Expressed Sequence Tag (EST) and genomic sequencing projects (Chapman et al., 2010; Meyer et al., 2009; Miller et al., 2007; Polato et al., 2011; Putnam et al., 2007; Sunagawa et al., 2009; Voolstra et al., 2009). However, whether these genes function similarly to their structural homologs largely remains to be confirmed. The majority of functional studies in cnidarians have focused on systemic and cellular effector responses. Reactive oxygen species (ROS; *e.g.,* nitric oxide), antioxidant enzymes (*e.g.,* superoxide dismutase), and antimicrobial peptides are produced as a function of disease and/or immune elicitation in several cnidarian species (Bosch et al., 2009; Couch et al., 2008; Hawkridge et al., 2000; Hutton and Smith, 1996; Kim et al., 2000; Mydlarz et al., 2009;

Ovchinnikova et al., 2006; Perez and Weis, 2006). In addition, cnidarians possess motile phagocytic amoebocytes that aid in wound repair and histocompatibility (Bigger, 1982), and enable the formation of a melanized band to prevent pathogen progression (Mydlarz et al., 2008; Petes et al., 2003). To date, it is unclear what pathways and regulatory networks are ultimately responsible for the initiation of these effector responses. To elucidate these pathways and to gain a holistic understanding of the cnidarian immune response, the use of large-scale genomic analyses, which allow simultaneous profiling of thousands of genes, is needed.

The symbiotic sea anemone *Aiptasia pallida* provides a tractable model for elucidating the molecular biology of cnidarian immunity and symbiosis. Unlike coral, *A. pallida* can be easily reared in the laboratory and propagates quickly. In addition, anemones can be maintained as clonal lines, allowing for analysis of experimental conditions without the complication of genetic variability. Lacking a skeleton, their soft body makes them highly amenable to molecular and biochemical procedures and there is extensive genomic sequence data available *(e.g.,* ~80 GB of genomic sequence data and a published comprehensive transcriptome) for this species. Most importantly, *A. pallida* can be reared with or without symbiotic dinoflagellates and maintained in the aposymbiotic state indefinitely. The ability to rear aposymbiotic individuals enables the investigation of the anemone immune response to pathogens without the confounding influence of resident dinoflagellate endosymbionts.

In the current study, we used transcriptome sequencing to characterize the immune response of aposymbiotic *Aiptasia* to live pathogen exposure. The aim of this study was to identify genes that play a role in anemone immunity and to provide a better understanding of host-pathogen interactions in cnidarians. Our findings will also provide a framework for future studies to clarify the role of endosymbiotic dinoflagellates in anemone immunity.

Materials and Methods

Anemone maintenance

All anemones were from clonal population CC7 (Sunagawa et al., 2009). Anemones were maintained in incubators at 25°C in artificial seawater (ASW; Instant Ocean) in 1 L glass bowls and were fed freshly hatched brine-shrimp nauplii approximately three times per week with water changes days following feeding. To investigate the host immune response and to avoid the confounding influence of the role of dinoflagellates in host immunity, we used aposymbiotic anemones. Aposymbiotic animals were generated by exposing symbiotic anemones to 50 μM Diuron (DCMU), with daily water changes, for ~30 days, or until the anemones were devoid of algae, as confirmed with fluorescence microscopy. Following the bleaching process, aposymbiotic anemones were maintained in the dark and remained there for ~2 years to prevent algal repopulation. Two months prior to experimentation, aposymbiotic anemones were subjected to a 12 h light : 12 h dark photoperiod using 18-22 μmol photons m^{-2} s^{-1} of photosynthetically active radiation (PAR). Anemones were transferred to the light so that they could be directly compared to symbiotic anemones under these same conditions in future experiments. One week prior to the incubation experiment, feeding was terminated and anemones were maintained in 0.2 um-filtered ASW (FASW) with daily water changes.

Bacteria inoculate preparation and Incubation Experiment

Serratia marscescens strain PDL100, originally isolated from white-pox infected *Acropora palmata* in the Florida Keys (Patterson et al., 2002), was plated onto LB agar and grown overnight at 29°C. *S. marscescens* was chosen based on its ability to cause disease in both vertebrates and invertebrates, including coral. While not a known natural pathogen of

Aiptasia, it induces symptoms of disease (Alagely et al., 2011). A single colony was isolated and suspended in 7 ml sterile marine broth and grown overnight at 29°C on a shaker at 150 rpm. A subsequent overnight culture with identical growth conditions was prepared to amplify the amount of culture by adding 250 ul of culture to 250 ml of sterile marine broth. The bacteria were centrifuged for 10 minutes at 4000 rpm, the supernatant was discarded, and the bacterial pellet was resuspended in FASW and diluted to 1.5 X 10^7 cells per ml.

Prior to the incubation experiment, anemones were transferred to 12-well plates with 4 ml of FASW and were allowed to acclimate for two days. Treatment anemones were incubated in 4 ml of *S. marscescens* inoculate (1.5 X 10^7 cells per ml), while control anemones were incubated in 4 ml FASW. After 6 hours of exposure, both control and treatment anemones were rinsed 3 times with FASW, blotted dry, and pooled into groups of 6 anemones, representing ~50 mg wet tissue weight, for a total of 4 replicate pools per treatment. The tubes were flash frozen in liquid nitrogen and were stored at -80°C prior to processing.

RNA extraction, Library Preparation, and Sequencing

Total RNA was extracted using the ToTALLY RNA™ Total RNA Isolation Kit (Ambion, cat. no. AM1910) according to the manufacturer's instructions with the exception that the RNA was precipitated using 0.1 volume of 3 M sodium acetate and 4 volumes of 100% ethanol. The resulting RNA was purified using the RNA Clean and Concentrator™-25 Kit (Zymo Research, cat. no. R1017) and the RIN of each sample was verified to be ≥ 9 using Agilent 2100 Bioanalyzer.

Approximately 2 µg of total RNA per sample was processed using the TruSeq RNA Sample Prep Kit following the manufacturer's instructions to produce indexed libraries for

mutliplex sequencing. The resulting libraries were barcoded and pooled into 8 samples per lane and were single-end sequenced with a target read length of 100 bp. Clustering and sequencing were performed by the Cornell University Life Sciences Core Laboratory Center using an Illumina HiSeq 2000 sequencer.

Read Filtering, Alignment, and Expression Analysis

Initial quality filtering of reads and bar-code removal was performed by the Cornell University Life Sciences Core Laboratory Center. Fastq-mcf (Aronesty 2011; http://code.google.com/p/ea-utils) was used to remove Illumina adaptors, trim low-quality terminal ends, discard short sequences, and filter reads with phred scores less than 15. FastQC was used to assess the quality of the reads before and after filtering (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). To ensure that only host genes were being examined, reads were aligned to cnidarian-classified transcripts from the *Aiptasia* transcriptome (Chapter 2) using BWA (Li and Durbin, 2009). Transcripts were considered cnidarian if they were sorted as such via the machine learning program, TopSort, and they also mapped to *Aiptasia* genomic DNA sequences (Chapter 2). A python script was used to count the number of reads that aligned to transcript uniquely with no errors or gaps. The R package edgeR (Robinson et al., 2010), which uses empirical Bayes estimation and exact tests based on the negative binomial distribution, was to identify differentially expressed transcripts that were classified as such if the false discovery rate (FDR)-adjusted p-value was less than 0.05. Within edgeR, the data were normalized for differing sequencing depths between libraries and transcripts with low read counts were filtered out such that transcripts with at least one count per million in at least 4 samples were retained in the analysis.

Annotation and Enrichment Analysis

Transcript sequence homology was determined via blastx searches against the SwissProt and NCBI non-redundant (nr) databases with an e-value cut-off of less than 1e-6. All BLAST steps were performed in parallel via Cornell University's Computational Biology Application Suite for High Performance Computing (biohpc.org). For transcripts that did not have blastx hits, the program ORFPredictor (Min et al 2005) was used to predict open reading frames. SwissProt identifiers were aligned to Gene Ontology (GO) terms (Gene Ontology database: http://www.geneontology.org) and parental categories via the MGI GO Slim database (http://www.informatics.jax.org) using the table joining function in Galaxy (Blankenberg et al., 2001).

The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 was used (Dennis et al., 2003; Huang et al., 2007) to perform an enrichment analysis to determine biological processes that were significantly overrepresented in differentially expressed transcripts relative to the background transcriptome using a modified Fisher Exact test. The Functional Annotation Clustering method was employed, which clusters groups of similar biological processes and provides an enrichment score representative of the –log geometric mean of the p-values of the individual biological processes. Clusters were considered significantly enriched and highly significantly enriched when enrichment scores were greater than 1.3 and 2.0, respectively (which roughly corresponds to p-values less than 0.05 and .004, respectively)(Dennis et al., 2003; Huang et al., 2007).

To further investigate a subset of transcripts with homology to tumor necrosis factor receptor (TNFR)-associated factors (TRAFs), HMMER tools (Eddy, 2009; Finn et al., 2011) were used to determine the presence of meprin and TRAF-homology (MATH) domains,

conserved intracellular regions shared by all TRAFs, within the *Aiptasia* transcripts. We took full-length TRAF protein sequences from a variety of invertebrate and vertebrate species across the animal tree and added the predicted protein translation from the *Aiptasia* transcripts to this set. We included human meprins (closely related non-TRAF proteins with MATH domains) as outgroup sequences. We then used hmmsearch to automatically identify and align MATH domains within this set of sequences. For *Aiptasia* transcript translations with a MATH domain, duplicate sequences were removed from the resulting alignment.

To choose the best-fit model of protein evolution, we used the program ProtTest v2.4 (Abascal et al., 2005). The results indicated that the best fit model was LG+I+Γ, where 'LG' indicates the substitution matrix (Le and Gascuel, 2008), 'I' specifies a proportion of invariant sites, 'Γ' specifies gamma-distributed rates across sites. To construct a tree using MATH domain sequences, maximum likelihood analyses were performed with the MPI version of RaxML v7.2.8 (RAXML-HPC-MPI)(Stamatakis, 2006). Twenty parsimony starting trees were used in independent searches of tree space. 100 bootstrap searches were subsequently performed and the results were applied to the best tree from the initial search. Trees were rooted in FigTree (http://tree.bio.ed.ac.uk/software/figtree/), and then annotated manually using Adobe Illustrator. Nodes with support values between 50% and 100% were labeled on the trees.

Real-Time Quantitative PCR

To validate the RNA-Seq data, RT-qPCR was performed on the same RNA pools that were used for library preparation and sequencing for a total of 4 biological replicates per treatment. The RNA was treated with DNase using the TURBO DNA-free kit™ (Invitrogen AM1907) and cDNA was synthesized using the GoScript™ Reverse Transcriptase System, both

protocols performed according to manufacturers' instructions. Primers for 18 transcripts were designed using Integrated DNA Technologies' PrimerQuest. Primer sequences are listed in Table S3.1 and product sizes ranged from 99-272 bp. Expected length of the amplicons was checked by agarose gel electrophoresis following standard PCR amplification. Primer efficiencies were determined using Real-time PCR Miner (Zhao and Fernald, 2005) and ranged from 88-95%.

Transcript level quantification was performed using Power SYBR® Green Master Mix (Applied Biosystems) and the ViiA™ 7 (Applied Biosystems) thermocycler. The reactions conditions were as follows: 1X Power SYBR® Green Master Mix, 200 nM of each primer, and 18 ng of cDNA in a total volume of 25 ul. Each sample and no template control (NTC) were run in duplicate with the following thermocycler parameters: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 60 s, with a subsequent dissociation curve to confirm the absence of non-specific products. In addition, a pooled sample of RNA representing all 8 samples was used as template to confirm the absence of genomic DNA contamination. Real Time PCR miner was used to calculate the efficiencies and critical threshold (C_T) of the genes based on raw fluorescence data.

Two genes previously shown to be stable under a variety of conditions (cytochrome c oxidase and 40S ribosomal protein S7; Chapter 2), as well as two genes that showed similar expression patterns between control and pathogen-exposed anemones based on RNA-Seq data (60S ribosomal protein L10 and Polyadenylate-binding protein 1) were chosen as reference genes. A normalization factor was calculated from the geometric mean of the reference genes' expression values (Vandesompele et al., 2002) and was used to normalize target genes and relative expression values were calculated using the equation $1/(1 + \text{Efficiency})^{C_T}$. Fold

change in expression relative to control anemones was then calculated as the quotients of the above equation. The R software (Team, 2013) package was used to perform Pearson correlations between fold change expression from qPCR and RNA-Seq data.

Results

Sequencing, Alignment, and Identification of Differentially Expressed Genes

To gain insight into the *Aiptasia* immune response, we utilized RNA-Seq to analyze the transcriptome of control and pathogen-exposed aposymbiotic anemones. This approach resulted in ~175 million reads, with over 17.5 Gb of sequence information (Table 3.1). Following quality trimming, ~144 million reads and ~13.5 Gb of sequence remained. An average of 9.5 and 8.3 million reads of control and pathogen-exposed anemones, respectively, mapped uniquely to cnidarian-classified transcripts of the *Aiptasia* transcriptome.

Table 3.1 Summary statistics for *Aiptasia pallida* transcriptome sequencing (N=4 biological replicates per treatment).

	Control	**Pathogen-exposed**
Average number of reads before filtering	26,381,009	17,443,459
Average read length (bp)	100	100
Average number of reads after filtering	19,467,403	16,610,719
Average read legnth after filtering (bp)	95	95
Average number of reads uniquely mapped*	9,481,458	8,343,979

*mapped to cnidarian-classified transcripts of the *A. pallida* transcriptome (Chapter 2).

We identified 2,981 transcripts that are differentially expressed between control and bacteria-exposed anemones, which represents ~18% of the 16,788 transcripts included in statistical analysis following filtering of transcripts with low read counts. Approximately 58% of these transcripts are up-regulated in anemones exposed to *S. marcescens*. For 481 of these

transcripts (~16%), expression differences were greater than 2 fold, with ~69% being up-regulated in treatment anemones. Table 3.2 lists the most differentially expressed genes from our analysis, many of which are involved in ubiquitination processes, including NEDD8-conjugating enzyme ubc12, probable E3 ubiquitin-protein ligase ARI9l, and tripartite motif-containing protein 59.

To validate the results of the RNA-Seq experiment, we also performed RT-qPCR on 18 transcripts exhibiting a variety of expression patterns and read counts (Table S3.1). The expression values of the RNA-Seq and RT-qPCR experiments were highly correlated (Pearson coefficient of 0.99; p-value = 2.2e-16), indicating high similarity between the RNA-Seq and RT-qPCR datasets (Figure S3.1).

Annotation and Enrichment Analysis

Of the 2,981 identified differentially expressed transcripts, ~73% and 86% had significant blastx hits to the SwissProt and NCBI nr databases, respectively. Although our further analyses focus on transcripts with blastx annotations, it is important to note that 415 (13%) of the differentially expressed transcripts, including 116 with ≥2-fold expression could not be annotated. 218 of these transcripts (including 60 with ≥2-fold expression changes) contained apparent open reading frames with ≥100 codons. Identifying the functions of these unknown proteins may provide a better understanding of genes that potentially play critical roles in the cnidarian response to microbial exposure.

To examine functional categories of regulated genes in greater detail, we assigned biological process GO terms to differentially expressed transcripts and then conducted an enrichment analysis to identify biological processes that were overrepresented in our data set.

Approximately 63% of differentially expressed transcripts were assigned at least one GO biological process term, with 11,844 GO terms assigned in total. These GO annotations were then aligned with GO slim terms to provide a broad classification of transcript function by expression direction (Figure 3.1).

The enrichment analysis identified 19 clusters of biological processes representing 121 GO terms that were significantly enriched (Table S3.2), with the most highly enriched clusters shown in Table 3.3. Enrichment of clusters with functions related to transcription and protein transport suggests that immune-induction results in major changes to steady state transcription and protein localization. One of the mostly highly enriched clusters includes processes relating to apoptosis and the regulation of apoptosis. In addition, two clusters with functions in cell signaling were enriched and include GO terms relating to functions in cell surface receptor linked signal transduction and regulation of the protein kinase cascade. In particular, many genes involved in tumor TRAF signaling are characterized by these GO processes. Other clusters that were highly enriched included GO terms related to DNA replication and sterol metabolic processes.

Of the 14 *Aiptasia* transcripts with TRAF homology, 12 possessed MATH domains (Table S3.3). While transcripts 78485/1 and 32870/1, (top blastx hits to human TRAF3 and TRAF5, respectively), lacked MATH domains, we feel these sequences most likely represent true TRAF sequences based on the presence of additional domains characteristic to TRAF proteins (e.g., RING finger domains and a TRAF-type zinc fingers). It is possible these contigs represent partial transcripts from which the MATH domain is missing from the assembled sequence. For *Aiptasia* transcripts with a MATH domain, the following were considered duplicates based on identical MATH domain sequence: 46907/1 (TRAF2) and 46912/1

Table 3.2 Top-ten most differentially expressed up- and down-regulated genes in pathogen-exposed anemones relative to controls.

Locus/Contig	Fold Change	FDR p-value	Top blastx hit to NCBI NR or Swiss-Prot databases	Accession	e-value
106552/1	24.3	6.58E-236	NEDD8-conjugating enzyme ubc12	O74549	8.00E-19
32435/1	22.6	1.31E-299	Cytochrome P450 3A24	Q29496	1.00E-87
15721/1	19.4	1.74E-147	No Hit	NA	NA
128686/1	18.4	0	Predicted protein	XP_001623429	1.00E-26
12003/1	13.7	6.28E-75	No Hit	NA	NA
120889/1	13.6	1.50E-09	Cytochrome P450 3A41	Q9JMA7	8.00E-79
49438/1	10.4	2.84E-13	No Hit	NA	NA
21472/1	9.7	1.12E-129	Probable E3 ubiquitin-protein ligase ARI9l	Q9SKC3	7.00E-26
73392/1	8.3	1.94E-11	NAD-dependent protein deacetylase SRT1	Q9FE17	3.00E-53
23978/2	7.1	1.06E-127	Alternative oxidase, mitochondrial	O74180	9.00E-69
24947/1	-3.6	8.39E-07	Putative N-acetyltransferase 16	Q0P4Y1	7.00E-08
8584/1	-3.6	1.19E-09	Transmembrane protein 45B	Q3T130	5.00E-21
118088/1	-3.6	5.84E-24	Forkhead box protein D1	Q6F2E4	3.00E-36
122403/1	-3.7	2.25E-17	Tctex1 domain-containing protein 1	Q66IC8	2.00E-20
21855/1	-3.7	5.77E-05	No Hit	NA	NA
13121/1	-3.9	1.82E-39	DNA double-strand break repair Rad50 ATPase	O29230	6.00E-07
95866/1	-3.9	1.34E-05	Transcription activator BRG1	P51532	1.00E-59
64418/1	-4.1	2.99E-06	Tetratricopeptide repeat protein 28	Q96AY4	1.00E-34
119948/1	-4.5	1.57E-05	No Hit	NA	NA
95915/1	-5.8	2.29E-05	Tripartite motif-containing protein 59	Q922Y2	7.00E-14

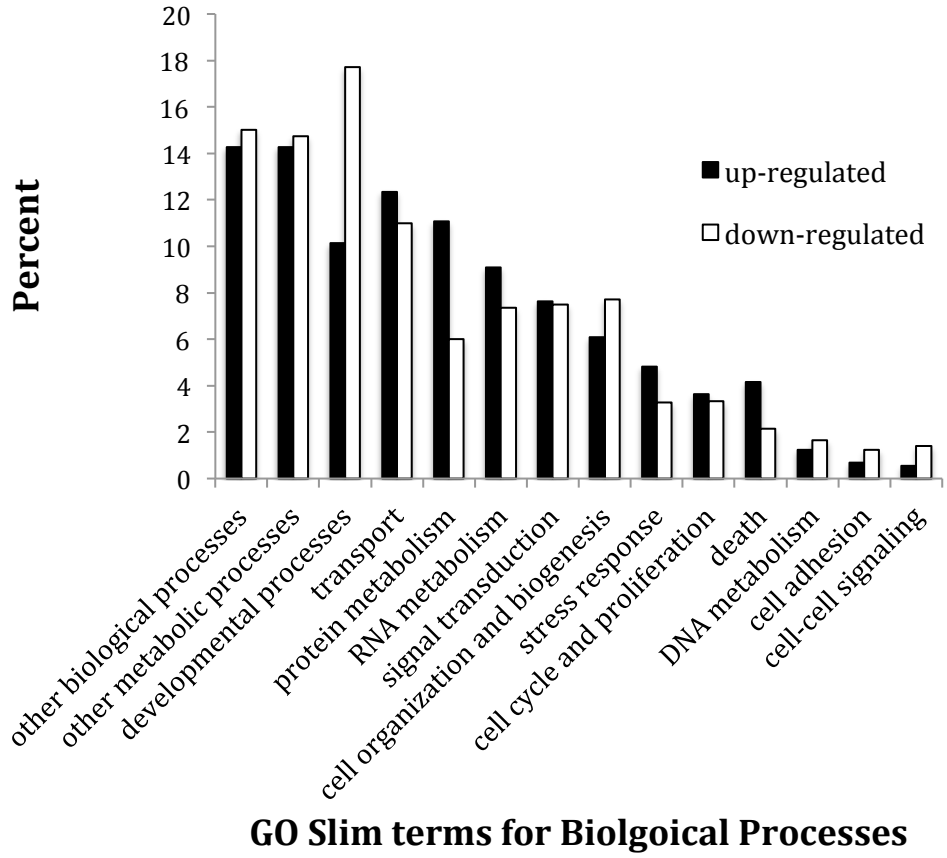


Figure 3.1 Percentage of differentially expressed genes classified by Biological Process Gene Ontology (GO) Slim terms for control and pathogen-exposed *A. pallida.* Up- and down-regulation are expressed in pathogen-exposed anemones relative to the controls.

(TRAF3), 46907/1 (TRAF2) and 46914/1 (TRAF3), and 67642/1(TRAF3) and 89050/1 (TRAF3). After removal of these sequences, nine unique *Aiptasia* MATH domain sequences remained. Figure S3.2 shows that there are at least 5 unique TRAF 'types' in this subset of differentially expressed transcripts, with the remaining sequences representing recent duplicates or alleles of the unique sequences. Transcript 30592/1 (TRAF4) falls into a group of cnidarian- and sponge-specific TRAFs while transcript 109431/1 (TRAF6) groups with TRAF6 sequences

Table 3.3 The most enriched clusters of Biological Process GO terms with similar functions via DAVID enrichment analysis. P-values are from a modified Fisher's Exact Test to identify overrepresented functions in differentially expressed genes as compared to the background transcriptome. The enrichment score (next to Cluster) is the –log geometric mean of the p-values in each cluster.

GO ID	GO term	p-value
Cluster 1: 4.64		
GO:0006278	RNA-dependent DNA replication	9.15E-06
GO:0015074	DNA integration	5.82E-05
Cluster 2: ***4:34***		
GO:0016265	death	2.91E-07
GO:0008219	cell death	4.69E-07
GO:0012501	programmed cell death	7.45E-07
GO:0006915	apoptosis	2.88E-06
GO:0042981	regulation of apoptosis	6.42E-04
GO:0043067	regulation of programmed cell death	8.21E-04
GO:0010941	regulation of cell death	9.65E-04
GO:0043065	positive regulation of apoptosis	4.40E-02
Cluster 3: 3.68		
GO:0045184	establishment of protein localization	2.04E-05
GO:0008104	protein localization	2.18E-05
GO:0015031	protein transport	3.67E-05
GO:0034613	cellular protein localization	1.72E-04
GO:0070727	cellular macromolecule localization	3.39E-04
GO:0006886	intracellular protein transport	7.57E-04
GO:0046907	intracellular transport	2.25E-02
Cluster 4: 3.66		
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	4.24E-05
GO:0007167	enzyme linked receptor protein signaling pathway	7.48E-05
GO:0007166	cell surface receptor linked signal transduction	3.36E-03
Cluster 5: 2.45		
GO:0045449	regulation of transcription	5.45E-04
GO:0006350	transcription	1.42E-03
GO:0006355	regulation of transcription, DNA-dependent	1.120E-02
GO:0051252	regulation of RNA metabolic process	1.68E-02
Cluster 6: 2.42		
GO:0010627	regulation of protein kinase cascade	3.28E-04
GO:0010740	positive regulation of protein kinase cascade	1.19E-03
GO:0043123	positive regulation of I-κB kinase/NF-κB cascade	1.29E-02
GO:0043122	regulation of I-κB kinase/NF-κB cascade	2.73E-02
Cluster 7: 2.3		
GO:0008203	cholesterol metabolic process	6.78E-04
GO:0008202	steroid metabolic process	1.17E-02
GO:0016125	sterol metabolic process	1.57E-02

from many different animals. The remaining *Aiptasia* TRAF transcripts fall into a large clade that includes human TRAFs 1, 2, 3, and 5.

Discussion

A better understanding of cnidarian immunity is necessary to clarify the host response to both pathogenic and mutualistic microbes, information that is crucial to inform how alterations in cnidarian-microbe associations lead to compromised host health. To this end, we used transcriptome sequencing to identify genes, pathways, and processes that are activated or repressed in response to a live pathogen in aposymbiotic *A. pallida*. We chose a live pathogen instead of an elicitor (e.g., LPS) to capture a more holistic view of immune expression in *A. pallida*. Our findings show that *S. marcescens* exposure results in a complex modulation of the host transcriptome, as demonstrated by the differential expression of genes involved in a wide variety of biological processes, as well as by the diverse processes that were enriched. Of particular interest are genes with known functions in response to microbial invaders. As such, our discussion highlights these functions and suggests their potential roles in cnidarian immunity.

TRAF-mediated signaling: convergence and transduction of multiple immune pathways

TRAFs are intracellular signaling molecules that function as critical transducers for members of TNFR, the Toll-like receptor (TLR), and the Interleukin-1 receptor (IL-1R) families (reviewed in Chung et al., 2002; Verstrepen et al., 2008). TRAF signaling ultimately results in the activation of transcription factors, such as NF-κB, and c-Jun (Verstrepen et al., 2008). In turn, these transcription factors regulate the expression of a large number of genes involved in

immunity, inflammation, and cell survival, proliferation, and differentiation (Verstrepen et al., 2008)(Figure 3.2).

While TRAFs share similar structures, their physiological functions differ, in part, due to their activation by different receptors and/or their varying downstream interactions (Verstrepen et al., 2008). For instance, in mammals, in which seven TRAFs have been identified and characterized (TRAF1-TRAF7; Bradley and Pober, 2001; Zotti et al., 2012), TRAF6 primarily functions in initiating pro-inflammatory cascades via TRL/IL-1R signaling and activation of NF-κB (Deng et al., 2000b), where as TRAF2 plays a central role in the regulation of cell survival and apoptosis via c-jun N-terminal kinase (JNK) via TNFR signaling (Chung et al., 2002). TRAFs have also been identified in many invertebrates, including cnidarians (Mali and Frank, 2004), nematodes (Wajant et al., 1998), molluscs (Goodson et al., 2005; Huang et al., 2012; Qiu et al., 2009), arthropods (Cha et al., 2003;Wang et al., 2011), and cephalochordates (Yuan et al., 2009), sharing both homologous and non-homologous functions with their vertebrate counterparts (Yuan et al., 2009; Cha et al., 2003). For instance, in amphioxus, like in mammals, TRAF6 activates the NF-κB pathway to initiate antimicrobial defense (Yuan et al., 2009). However, unlike in mammals, TRAF2 appears to negatively regulate the NF-κB pathway (Yuan et al., 2009).

Our findings suggest that in *Aiptasia*, TRAFs play an important role in host defense, potentially via the regulation of pro-inflammatory effectors and/or apoptotic pathways. We found that pathogen-exposure led to the differential expression of 14 transcripts with TRAF homology representing 5 'types' of TRAFs in *A. pallida* (Table 3.4, Table S3.3, Figure S3.2). These 'types' include one that is unique to lower metazoans (i.e., sponges and cnidarians) and others with homology to human TRAFs 1/2/ 3/5, and 6. Our results agree with Polato et al.,

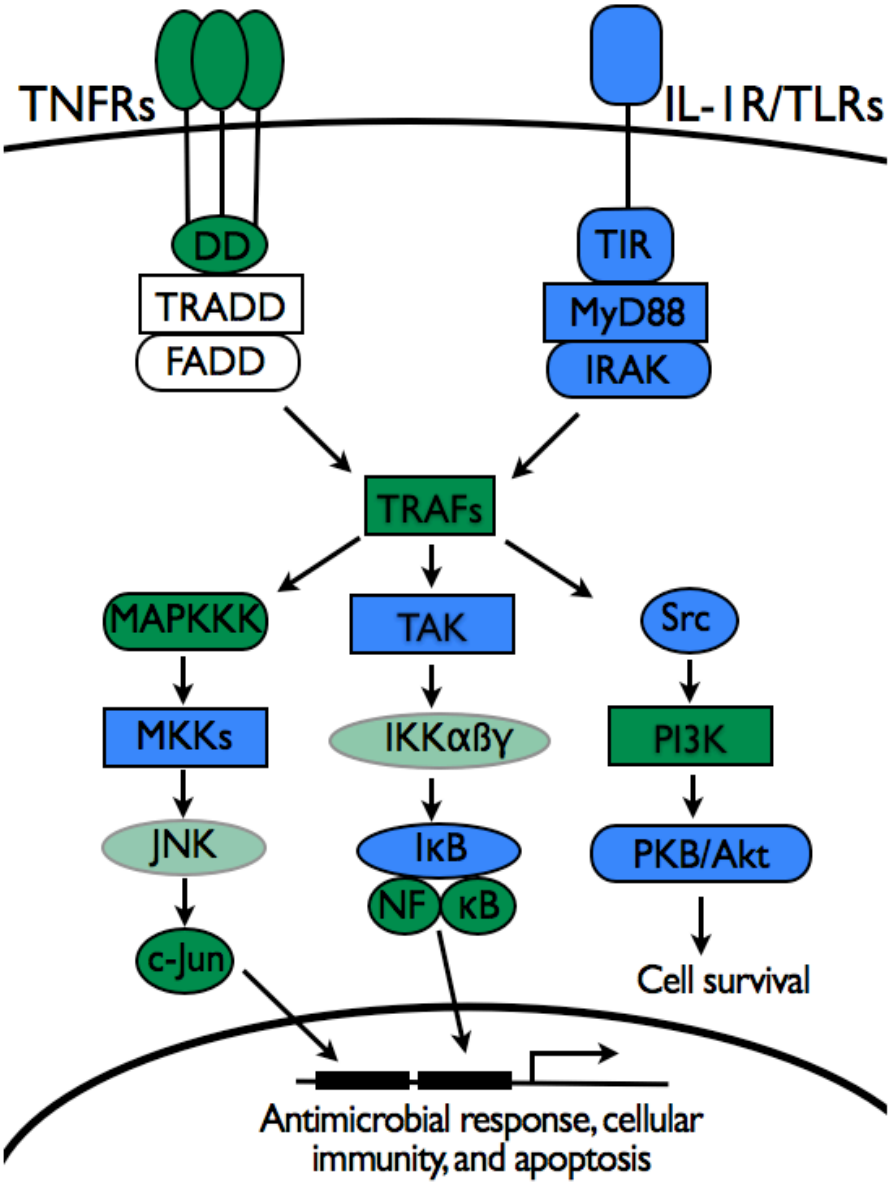


Figure 3.2 Simplified schematic of TRAF-mediated signaling pathways. Binding of ligands to TNFR and/or IL-1R/TLRs activates a series of pathway-specific adaptor and transducer proteins that ultimately rely on TRAFs for downstream signaling. Once activated, TRAFs signal via MAPK/JNK, NFkB, and PKB/Akt pathways to initiate cellular events. Proteins in green represent those that were differentially expressed in our data set, with those in lighter green being weakly differentially expressed (i.e., less than 1.5 fold change). Proteins in blue and white represent those that have been identified or not identified in the *Aiptasia* transcriptome, respectively. Table 3.4 lists annotations and expression values for transcripts with homology and their associated proteins depicted in this diagram.

(2011), who found the presence of multiple TRAF families in the transcriptome of the elkhorn coral, further providing evidence that the diversity of this gene family is more ancestral than previously acknowledged. Except for one transcript, all of the *A. pallida* TRAFs were up-regulated in response to bacterial challenge. In other invertebrates, TRAFs 2, 3, 4, 5, and 6 are differentially expressed upon immune-challenge (Huang et al., 2012; Wang et al., 2011; Yuan et al., 2009; Qiu et al., 2009, Cha et al., 2003). Our data especially mirror patterns found in amphioxus, where TRAFs 2 and 3 were up-regulated following bacterial challenge (Huang et al., 2008), and TRAFs 4 and 6 were up-regulated in LPS-exposed animals (Yuan et al., 2009).

Since TRAF signaling mechanisms have not been previously elucidated in cnidarians, and are not well described in other invertebrates, their differential expression in this study is a novel finding. A question raised by their activation in *Aiptasia* is whether cnidarian TRAFs signal via similar pathways and adaptor molecules as their vertebrate and higher invertebrate counterparts (Figure 3.2). Indeed, homologs of members of TLR/IL-1R and TNFR pathways have been identified in several cnidarians (Miller et al., 2007; Polato et al., 2011; Sullivan et al., 2007; Wolenski et al., 2011). In addition, enrichment analysis of our data set showed the overrepresentation of functions relating cell surface receptor linked signal transduction and regulation of the protein kinase cascade (Table 3.2). Furthermore, we found differential expression of other genes involved in TRAF-mediated signaling, including lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog and NF-κB p105 subunit (Figure 3.2, Table 3.4). Although cnidarian TRAFs have been detected in previous studies, our results provide the first evidence that TRAFs have critical functions in the response to pathogens. Given the large number of differentially expressed transcripts with TRAF homology in our study, coupled with

the expression patterns of TRAFs in other invertebrates, future studies should elucidate the specific functions that the various TRAF proteins play in cnidarian immunity.

Apoptosis: a dual role in host the response to pathogens and mutualists

Apoptosis, or programmed cell death, is a conserved and highly orchestrated process that plays a pivotal role in cellular and tissue homeostasis, embryonic development, and immunity of metazoans. Apoptosis can be initiated via the intrinsic pathway, which occurs in response to internal cellular damage, or the extrinsic pathway, which is triggered in response to extracellular 'death' signals (Hedrick et al., 2010). Both intrinsic and extrinsic pathways converge on the activation of proteases called caspases that initiate the proteolytic degradation of cellular structures (Hedrick et al., 2010). Apoptosis elicited during infection is thought to prevent pathogen spread, and because the intracellular components are sequestered and degraded without the induction of inflammation, to protect the integrity of surrounding tissues (Birge and Ucker, 2008; Hedrick et al., 2010; Sokolova, 2009). Exposure to a diverse array of pathogens, including viruses, protozoans, and bacteria, has been show to elicit apoptosis in both vertebrates and invertebrates (Sokolova, 2009; Al-Olayan et al., 2002; De Zoysa et al., 2012; Ishii et al., 2012; Jones et al., 2012; Li et al., 2011; Navarre and Zychlinsky, 2000; Zink et al., 2002), and indeed, many pathogens have evolved mechanisms that suppress or prevent host cell apoptosis to promote pathogen survival and proliferation (Behar et al., 2011; Heussler et al., 2001; Hughes et al., 2010).

Our experiment suggests that apoptosis plays an important role in the aposymbiotic host response to *S. marcescens*. A cluster of GO terms representing apoptotic-related processes was highly enriched (Table 3.3) and 204 transcripts with apoptotic functions were differentially

Table 3.4 Differentially expressed genes in pathogen-exposed anemones relative to controls with functions in TRAF-mediating signaling as pictorially represented in Figure 3.2.

Locus/Contig	Fold Change	FDR p-value	Top blastx hit to SwissProt database	Accession	e-value
TNF ligands and receptors					
60163/1	1.70	9.14E-11	Lipopolysaccharide-induced TNF-alpha factor	Q9JLJ0	2.00E-12
67712/1	-1.49	1.54E-05	Lipopolysaccharide-induced TNF-alpha factor	Q99732	1.00E-20
44804/1	-1.19	0.032837515	TNF ligand superfamily member 10	P50591	7.00E-15
95000/1	1.7	5.48E-10	TNF receptor superfamily member 27	Q8BX35	1.00E-10
95010/1	1.99	0.024871294	TNF receptor superfamily member 27	Q8BX35	1.00E-09
TRAFs					
46907/1	3.15	6.03E-41	TNF receptor-associated factor 2	P39429	2.00E-88
46888/1	2.07	0.00220459	TNF receptor-associated factor 2	P39429	1.00E-86
86236/1	1.2	0.037360178	TNF receptor-associated factor 2	Q12933	4.00E-69
67642/1	1.17	0.040216512	TNF receptor-associated factor 3	Q13114	1.00E-109
89049/1	2.34	1.22E-08	TNF receptor-associated factor 3	Q13114	6.00E-45
89050/1	1.66	5.32E-10	TNF receptor-associated factor 3	Q13114	4.00E-78
78485/1	4.16	2.19E-19	TNF receptor-associated factor 3	Q60803	1.00E-12
46912/1	2.25	1.45E-17	TNF receptor-associated factor 3	Q60803	7.00E-85
46914/1	1.51	2.47E-05	TNF receptor-associated factor 3	Q60803	3.00E-91
114598/1	-1.22	0.006704894	TNF receptor-associated factor 3	Q60803	2.00E-98
31324/1	4.90	4.73E-50	TNF receptor-associated factor 3	Q60803	5.00E-75
30592/1	1.88	0.012845507	TNF receptor-associated factor 4	Q9BUZ4	2.00E-28
32870/1	4.64	1.11E-19	TNF receptor-associated factor 5	P70191	6.00E-19
109431/1	1.21	0.032837515	TNF receptor-associated factor 6	B6CJY4	6.00E-62
Downstream signaling pathways					
32139/2	-1.78	0.007948327	MAPK kinase kinase	P28829	6.00E-16
13803/1	2.63	2.30E-07	Proto-oncogene c-Jun	P54864	9.00E-23
19796/1	1.21	0.00753963	Stress-activated protein kinase JNK	P92208	1.00E-152
114092/1	1.26	0.002641023	Inhibitor of NF-κB kinase subunit alpha	Q60680	1.00E-116
73699/1	1.58	7.94E-14	NFκB p105 subunit	P19838	3.00E-96
91253/1	1.51	0.012201961	Phosphatidylinositol 3-kinase	Q64143	1.00E-09

expressed (Figure 3.3, Table S3.4). Transcripts in this category include both pro- and anti-apoptotic members of the B-cell lymphoma 2 (Bcl-2) family, which initiate intrinsic apoptosis, as well as caspases involved in the extrinsic pathway. Numerous expression studies conducted in invertebrates have also detected differential expression of apoptosis-related genes upon exposure to a diverse array of pathogens (Aguilar et al., 2005; Araya et al., 2010; De Zoysa et al., 2012; Ding et al., 2005; Moreira et al., 2012; Wang et al., 2010). In addition, a recent study found that *S. marcescens* induced apoptosis in silkworm hemocytes via the activation of c-Jun NH2-terminal kinase (JNK) and caspases (Ishii et al., 2012). We also found the differential expression of genes involved in the MAPK/JNK signaling, including MAPKKK, MAP kinase-activated protein kinase 2, and c-Jun (Table 3.4), suggesting similar mechanisms may play a role in the *S. marcescens*-induction of apoptosis in *Aiptasia*.

Previous work on cnidarian apoptosis has identified and characterized apoptosis-related genes (e.g. caspases and members of the Bcl-2 family) (Ainsworth et al., 2008; Dunn et al., 2006; Lasi et al., 2010) and has confirmed that apoptosis occurs via similar mechanisms to vertebrates in this ancient animal lineage (Dunn et al., 2002; Dunn et al., 2007; Dunn and Weis, 2009; Pernice et al., 2011; Seipp et al., 2001; Tchernov et al., 2011). Cnidarian apoptosis has mainly been studied in the context of the dinoflagellate mutualism, including its potential role in the onset (Dunn and Weis, 2009) and breakdown (Dunn et al., 2002; Dunn et al., 2007; Pernice et al., 2011; Tchernov et al., 2011) of this relationship. Our recent study in *A. pallida* comparing gene expression between symbiotic and aposymbiotic anemones found enrichment of apoptosis-related genes (Chapter 2), suggesting that apoptosis may also play an important role in the maintenance of this relationship. Interestingly, there are several differentially expressed genes that overlap between that data set and our current data set, including TRAF3, krueppel-like factor

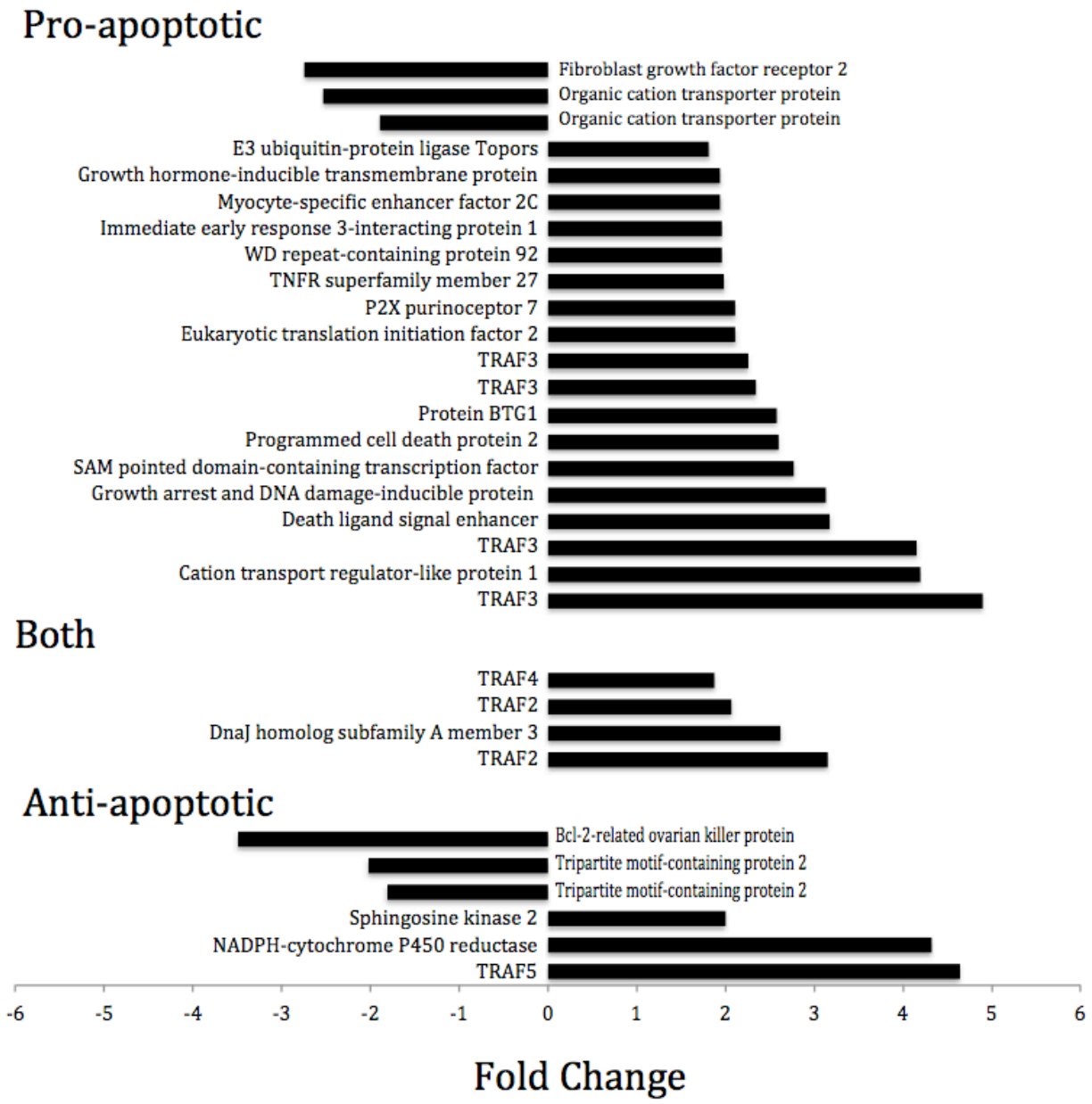


Figure 3.3 The most differentially expressed transcripts (>1.8 fold change) from pathogen-exposed anemones with functions in apoptosis as determined via Biological Process GO terms. Fold-changes are shown as expression in pathogen-exposed anemones relative to that in controls. Transcripts are categorized by pro- or anti-apoptotic or having both pro- and anti-apoptotic functions. See Table S3.4 for a list of all of the genes that were differentially expressed with apoptotic functions, as well as more detailed information regarding the above transcripts.

11, growth arrest and DNA damage-inducible protein, and paired box protein Pax-3. These findings provide further support that similar immune processes function in response to both mutualistic and pathogenic microorganisms, and comparing and contrasting the role of apoptosis during the onset of pathogen and mutualist infection would provide valuable insight into how pathogens are recognized and eliminated while mutualistic relationships are initiated and maintained.

Ubiquitination: protein activation and degradation to regulate dynamic cellular functions

Ubiquitination is a versatile post-translation modification that enables the precise regulation of proteins involved in diverse biological processes (reviewed in (Gao and Karin, 2005; Husnjak and Dikic, 2012; Komander, 2009). Ubiquitin is a small, highly conserved protein that is covalently attached to target proteins via an enzymatic cascade involving E1-activiating, E2-conjugating, and E3 ligase enzymes. This process can also be reversed through the action of deubiquitination enzymes (DUBs). Ubiquitination can marks proteins for degradation via the 26S proteasome or it can activate signal transduction (Komander, 2009). We found the differential expression of over a hundred transcripts with homology to genes involved in ubiquitination (Figure 3.4, Table S3.5), including E1, E2, and E3 enzymes, components of the 26S proteasome, DUBs, and ubiquitin-like modifiers (UBLs; i.e., proteins with similar structures and functions to ubiquitin). Indeed, some of the most highly up- and down-regulated transcripts in our dataset (Table 3.2) are homologous to ubiquitination proteins, including NEDD8-conjugating enzyme ubc12 (E2 enzyme for the UBL protein NEDD8), E3 ubiquitin-protein ligase ARI9l, and Tripartite motif-containing protein 59 (E3 ligase for ubiquitin and UBLs). The large number of differentially expressed transcripts with ubiquitination

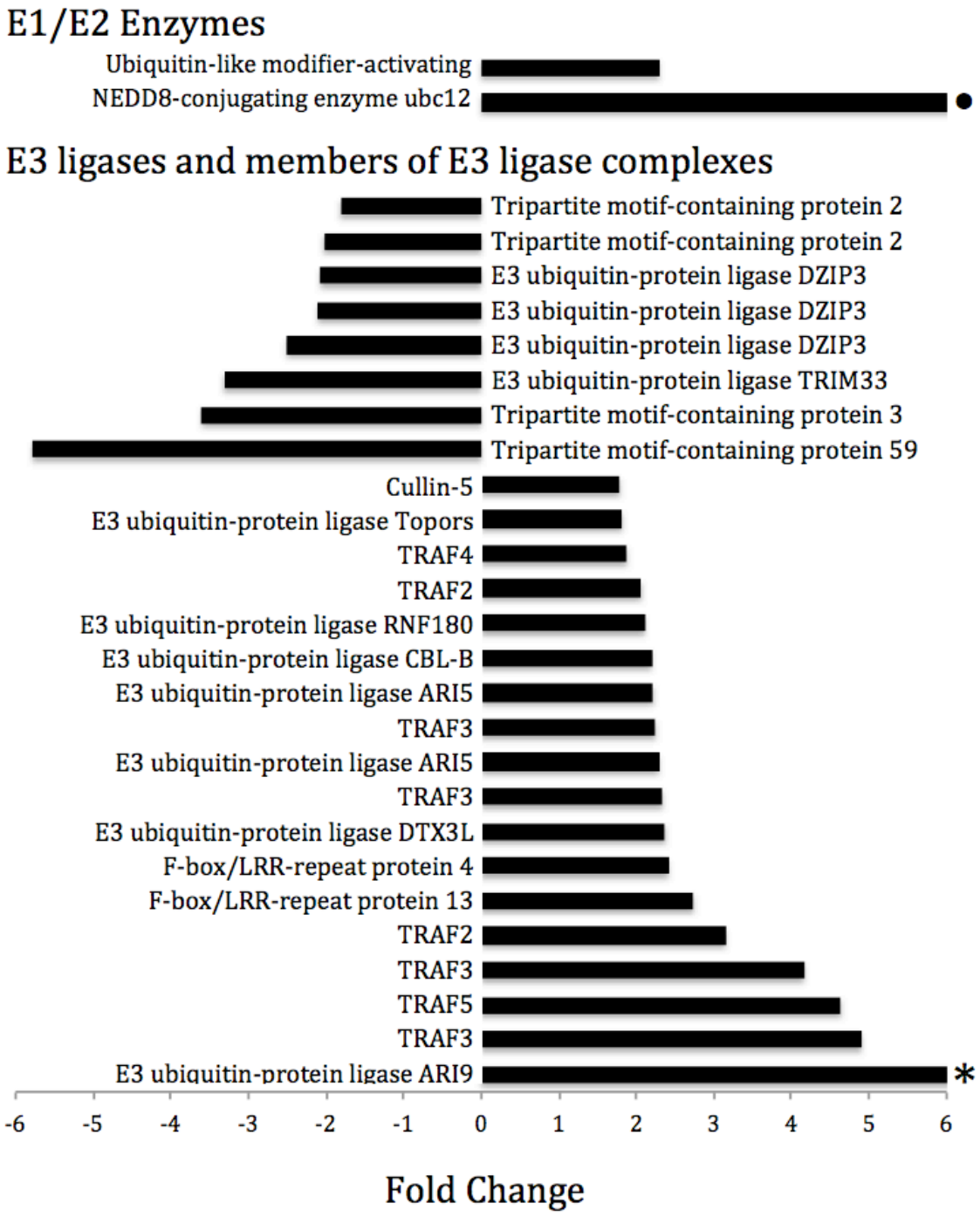


Figure 3.4. The most differentially expressed transcripts (>1.8 fold change) from pathogen-exposed anemones with functions in ubiquitination as determined via Biological Process GO terms. (*24.3 fold change; •9.7 fold change). See Supplementary Table 5 for a list of all of the genes that were differentially expressed with functions in ubiquitination, as well as more detailed information regarding the above transcripts.

functions most likely reflects the importance of the tight regulation of a large number of biological processes that are altered via pathogen exposure, including transcription, transport, and metabolism.

Not surprisingly, ubiquitination also plays an important role in immunity, specifically via regulation of MAPK and NF-κB pathways. In fact, apart from their role as adapter proteins, TRAFs activate downstream signaling events via their additional function as E3 ubiquitin ligases (Gao and Karin, 2005). In mammals, both TRAF2 and TRAF6 activate the TAK complex (Figure 3.2) via ubiquitination, thereby initiating MAPK and NF-κB signal transduction (Deng et al., 2000a). In addition to TRAFs, other ubiquitination enzymes that play roles in these signaling cascades were also differentially expressed. These include members of an E3 ligase complex (cullin-1 and F-box proteins) that mark IκB for degradation, subsequently leading to the activation and translocation of NF-κB into the nucleus (Karin and Ben-Neriah, 2000), as well as the DUB enzyme ubiquitin carboxyl-terminal hydrolase CYLD, which removes activating ubiquitin chains to negatively regulate NF-κB and prevent excessive immune stimulation (Bibeau-Poirier and Servant, 2008).

Our data suggest that, in response to bacterial challenge, *Aiptasia* regulates cellular processes not only via modulations in gene expression, but also via post-translation modification. While the role of protein ubiquitination in non-model invertebrates has not been extensively investigated, especially in the context of immunity, it appears that cnidarians depend on similar ubiquitin-mediated protein degradation and signaling to maintain host homeostasis.

Conclusions

This is one of the first studies to utilize a transcriptomic approach to characterize the genomic underpinnings of the pathways and processes elicited via pathogenic microbes in a model cnidarian. Our results suggest that, like their higher invertebrate and vertebrate counterparts, *A. pallida* responds to bacterial challenge via the regulation of TRAF-mediated signaling, apoptosis, and ubiquitination. In addition, the evolution of these processes for host defense appear to have ancestral origins and the results of this study will bolster a better understanding of cnidarian immunity and the subsequent changes that allowed for the development of more complex responses. We have also generated a list of candidate genes that can be investigated in further detail to better characterize the cnidarian immune repertoire and its role in both pathogenic and mutualistic interactions.

REFERENCES

Abascal, F., Zardoya, R., and Posada, D. (2005). ProTest: selection of best-fit models of protein evolution. Bioinformatics *21*, 2104-2105.

Aguilar, R., Jedlicka, A.E., Mintz, M., Mahairaki, V., Scott, A.L., and Dimopoulos, G. (2005). Global gene expression analysis of *Anopheles gambiae* responses to microbial challenge. Insect Biochemistry and Molecular Biology *35*, 709-719.

Ainsworth, T.D., Hoegh-Guldberg, O., Heron, S.F., Skirving, W.J., and Leggat, W. (2008). Early cellular changes are indicators of pre-bleaching thermal stress in the coral host. J Exp Mar Biol Ecol *364*, 63-71.

Al-Olayan, E.M., Williams, G.T., and Hurd, H. (2002). Apoptosis in the malaria protozoan, *Plasmodium berghei*: a possible mechanism for limiting intensity of infection in the mosquito. International Journal for Parasitology *32*, 1133-1143.

Alagely, A., Krediet, C.J., Ritchie, K.B., and Teplitski, M. (2011). Signaling-mediated cross-talk modulates swarming and biofilm formation in a coral pathogen *Serratia marcescens*. Isme J *5*, 1609-1620.

Araya, M.T., Markham, F., Mateo, D.R., McKenna, P., Johnson, G.R., Berthe, F.C.J., and Siah, A. (2010). Identification and expression of immune-related genes in hemocytes of soft-shell clams, *Mya arenaria*, challenged with *Vibrio splendidus*. Fish & Shellfish Immunology *29*, 557-564.

Behar, S.M., Martin, C.J., Booty, M.G., Nishimura, T., Zhao, X., Gan, H.X., Divangahi, M., and Remold, H.G. (2011). Apoptosis is an innate defense function of macrophages against *Mycobacterium tuberculosis*. Mucosal Immunol *4*, 279-287.

Bibeau-Poirier, A., and Servant, M.J. (2008). Roles of ubiquitination in pattern-recognition receptors and type I interferon receptor signaling. Cytokine *43*, 359-367.

Bigger, C., WH Hildenmann (1982). Cellular defense systems of the Coelenterata. (New York, Plenum).

Birge, R.B., and Ucker, D.S. (2008). Innate apoptotic immunity: the calming touch of death. Cell Death Differ *15*, 1096-1102.

Blankenberg, D., Kuster, G.V., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2001). Galaxy: A Web-Based Genome Analysis Tool for Experimentalists. In Current Protocols in Molecular Biology (John Wiley & Sons, Inc.).

Bosch, T.C.G. (2013). Cnidarian-Microbe Interactions and the Origin of Innate Immunity in Metazoans. Annu Rev Microbiol *67*, DOI: 10.1146/annurev-micro-092412-155626.

Bosch, T.C.G., Augustin, R., Anton-Erxleben, F., Fraune, S., Hemmrich, G., Zill, H., Rosenstiel, P., Jacobs, G., Schreiber, S., Leippe, M.*, et al.* (2009). Uncovering the evolutionary history of innate immunity: The simple metazoan Hydra uses epithelial cells for host defence. Dev Comp Immunol *33*, 559-569.

Bradley, J.R., and Pober, J.S. (2001). Tumor necrosis factor receptor-associated factors (TRAFs) Oncogene *20*, 6482-6491.

Cha, G.-H., Cho, K.S., Lee, J.H., Kim, M., Kim, E., Park, J., Lee, S.B., and Chung, J. (2003). Discrete Functions of TRAF1 and TRAF2 in *Drosophila melanogaster* Mediated by c-Jun N-Terminal Kinase and NF-kB-Dependent Signaling Pathways. Molecular and Cellular Biology *23*, 7982-7991.

Chapman, J.A., Kirkness, E.F., Simakov, O., Hampson, S.E., Mitros, T., Weinmaier, T., Rattei, T., Balasubramanian, P.G., Borman, J., Busam, D.*, et al.* (2010). The dynamic genome of Hydra. Nature *464*, 592-596.

Chen, M., Cheng, Y., Sung, P., Kuo, C., and Fang, L. (2003). Molecular identification of Rab7 (ApRab7) in *Aiptasia pulchella* and its exclusion from phagosomes harboring zooxanthellae. Biochem Biophys Res Commun *308*, 586 - 595.

Chen, M., Hong, M., Huang, Y., Liu, M., Cheng, Y., and Fang, L. (2005). ApRab11, a cnidarian homologue of the recycling regulatory protein Rab11, is involved in the establishment and maintenance of the *Aiptasia-Symbiodinium* endosymbiosis. Biochem Biophys Res Commun *338*, 1607 - 1616.

Chung, J., Park, Y., Ye, H., and Wu, H. (2002). All TRAFs are not created equal: common and distinct molecular mechanisms of TRAF-mediated signal transduction. J Cell Sci *115*, 679 - 688.

Couch, C.S., Mydlarz, L.D., Harvell, C.D., and Douglas, N.L. (2008). Variation in measures of immunocompetence of sea fan coral, *Gorgonia ventalina*, in the Florida Keys. Mar Biol *155*, 281-292.

Dale, C., Plague, G.R., Wang, B., Ochman, H., and Moran, N.A. (2002). Type III secretion systems and the evolution of mutualistic endosymbiosis. Proc Natl Acad Sci USA *99*, 12397-12402.

De Zoysa, M., Nikapitiya, C., Oh, C., Whang, I., Shin, H.-J., and Lee, J. (2012). cDNA microarray analysis of disk abalone genes in gills and hemocytes after viral hemorrhagic septicemia virus (VHSV) challenge. Fish & Shellfish Immunology *32*, 1205-1215.

Deng, L., Wang, C., Spencer, E., Yang, L., Braun, A., You, J., Slaughter, C., Pickart, C., and Chen, Z.J. (2000a). Activation of the IkB kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. Cell *103*, 351-361.

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol *4*, 11.

Ding, J.L., Tan, K.C., Thangamani, S., Kusuma, N., Seow, W.K., Bui, T.H.H., Wang, J., and Ho, B. (2005). Spatial and temporal coordination of expression of immune response genes during Pseudomonas infection of horseshoe crab, *Carcinoscorpius rotundicauda*. Genes Immun *6*, 557-574.

Dunn, S. (2009). Immunorecognition and immunoreceptors in the Cnidaria. Invertebrate Survival Journal *6*, 7-14.

Dunn, S., Phillips, W., Spatafora, J., Green, D., and Weis, V. (2006). Highly conserved caspase and Bcl-2 homologues from the sea anemone *Aiptasia pallida*: lower metazoans as models for the study of apoptosis evolution. Journal of Molecular Evolution *63*, 95-107.

Dunn, S.R., Bythell, J.C., Le Tissier, M.D.A., Burnett, W.J., and Thomason, J.C. (2002). Programmed cell death and cell necrosis activity during hyperthermic stress-induced bleaching of the symbiotic sea anemone *Aiptasia sp*. J Exp Mar Biol Ecol *272*, 29-53.

Dunn, S.R., Phillips, W.S., Green, D.R., and Weis, V.M. (2007). Knockdown of actin and caspase gene expression by RNA interference in the symbiotic anemone *Aiptasia pallida*. Biological Bulletin *212*, 250-258.

Dunn, S.R., and Weis, V.M. (2009). Apoptosis as a post-phagocytic winnowing mechanism in a coral-dinoflagellate mutualism. Environ Microbiol *11*, 268-276.

Eddy, S.R. (2009). A new generation of homology search tools based on probabilistic inference. Genome Informatics *23*, 205-211.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res *39*, W29-W37.

Gao, M., and Karin, M. (2005). Regulating the Regulators: Control of protein ubiquitination and ubiquitin-like modifications by extracellular stimuli. Molecular Cell *19*, 581-593.

Goodson, M.S., Kojadinovic, M., Troll, J.V., Scheetz, T.E., Casavant, T.L., Soares, M.B., and McFall-Ngai, M.J. (2005). Identifying components of the NF-kB pathway in the beneficial *Euprymna scolopes*-*Vibrio fischeri* light organ symbiosis. Appl Environ Microbiol *71*, 6934-6946.

Harvell, D., Jordan-Dahlgren, E., Merkel, S., Rosenberg, E., Raymundo, L., Smith, G., Weil, E., Willis, B., and Global Envrionm Facility, C. (2007). Coral disease, environmental drivers, and the balance between coral and microbial associates. Oceanography *20*, 172-195.

Hawkridge, J.M., Pipe, R.K., and Brown, B.E. (2000). Localisation of antioxidant enzymes in the cnidarians *Anemonia viridis* and *Goniopora stokesi*. Marine Biology (Berlin) *137*, 1-9.

Hedrick, S.M., Ch'en, I.L., and Alves, B.N. (2010). Intertwined pathways of programmed cell death in immunity. Immunol Rev *236*, 41-53.

Heussler, V.T., Kuenzi, P., and Rottenberg, S. (2001). Inhibition of apoptosis by intracellular protozoan parasites. International Journal for Parasitology *31*, 1166-1176.

Hoegh-Guldberg, O. (1999). Climate change, coral bleaching and the future of the world's coral reefs. Mar Freshw Res *50*, 839 - 866.

Hoegh-Guldberg, O., Mumby, P.J., Hooten, A.J., Steneck, R.S., Greenfield, P., Gomez, E., Harvell, C.D., Sale, P.F., Edwards, A.J., Caldeira, K.*, et al.* (2007). Coral reefs under rapid climate change and ocean acidification. Science *318*, 1737-1742.

Huang, D.W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C.*, et al.* (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res *35*, W169-W175.

Huang, S., Yuan, S., Guo, L., Yu, Y., Li, J., Wu, T., Liu, T., Yang, M., Wu, K., Liu, H.*, et al.* (2008). Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. Genome Research *18*, 1112-1126.

Huang, X.-D., Liu, W.-G., Guan, Y.-Y., Shi, Y., Wang, Q., Zhao, M., Wu, S.-Z., and He, M.-X. (2012). Molecular cloning, characterization and expression analysis of tumor necrosis factor receptor-associated factor 3 (TRAF3) from pearl oyster *Pinctada fucata*. Fish & Shellfish Immunology *33*, 652-658.

Hughes, F.M., Foster, B., Grewal, S., and Sokolova, I.M. (2010). Apoptosis as a host defense mechanism in *Crassostrea virginica* and its modulation by Perkinsus marinus. Fish & Shellfish Immunology *29*, 247-257.

Husnjak, K., and Dikic, I. (2012). Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. Annual Review of Biochemistry *81*, 291-322.

Hutton, D.M.C., and Smith, V.J. (1996). Antibacterial properties of isolated amoebocytes from the sea anemone *Actinia equina*. Biological Bulletin (Woods Hole) *191*, 441-451.

Ishii, K., Adachi, T., Imamura, K., Takano, S., Usui, K., Suzuki, K., Hamamoto, H., Watanabe, T., and Sekimizu, K. (2012). *Serratia marcescens* induces apoptotic cell death in host immune cells via a lipopolysaccharide- and flagella-dependent mechanism. Journal of Biological Chemistry *287*, 36582-36592.

Jones, R.M., Luo, L., and Moberg, K.H. (2012). *Aeromonas salmonicida*-secreted protein AopP is a potent inducer of apoptosis in a mammalian and a *Drosophila* model. Cell Microbiol *14*, 274-285.

Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: The control of NF-kB activity. Annual Review of Immunology *18*, 621-663.

Kim, K., Kim, P.D., Alker, A.P., and Harvell, C.D. (2000). Chemical resistance of gorgonian corals against fungal infections. Marine Biology (Berlin) *137*, 393-401.

Komander, D. (2009). The emerging complexity of protein ubiquitination. Biochemical Society Transactions *37*, 937-953.

Lasi, M., David, C., and B√∂ttger, A. (2010). Apoptosis in pre-Bilaterians: Hydra as a model. Apoptosis *15*, 269-278.

Le, S.Q., and Gascuel, O. (2008). An improved general amino acid replacement matrix. Molecular Biology and Evolution *25*, 1307-1320.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, M.-f., Wang, C.-L., and Sun, L. (2011). A pathogenic *Vibrio harveyi* lineage causes recurrent disease outbreaks in cultured Japanese flounder *(Paralichthys olivaceus)* and induces apoptosis in host cells. Aquaculture *319*, 30-36.

Mali, B., and Frank, U. (2004). Hydroid TNF-receptor-associated factor (TRAF) and its splice variant: a role in development. Molecular Immunology *41*, 377-384.

McFall-Ngai, M., Heath-Heckman, E.A.C., Gillette, A.A., Peyer, S.M., and Harvie, E.A. (2012). The secret languages of coevolved symbioses: Insights from the *Euprymna scolopes-Vibrio fischeri* symbiosis. Seminars in Immunology *24*, 3-8.

Meyer, E., Aglyamova, G., Wang, S., Buchanan-Carter, J., Abrego, D., Colbourne, J., Willis, B., and Matz, M. (2009). Sequencing and de novo analysis of a coral larval transcriptome using 454 GS-Flx. Bmc Genomics *10*, 219.

Miller, D.J., Hemmrich, G., Ball, E.E., Hayward, D.C., Khalturin, K., Funayama, N., Agata, K., and Bosch, T.C.G. (2007). The innate immune repertoire in Cnidaria - ancestral complexity and stochastic gene loss. Genome Biol *8*, 13.

Moreira, R., Balseiro, P., Planas, J.V., Fuste, B., Beltran, S., Novoa, B., and Figueras, A. (2012). Transcriptomics of *in vitro* immune-stimulated hemocytes from the manila clam *Ruditapes philippinarum* using high-throughput sequencing. PLoS One *7*, e35009.

Mydlarz, L.D., Couch, C.S., Weil, E., Smith, G., and Harvell, C.D. (2009). Immune defenses of healthy, bleached and diseased *Montastraea faveolata* during a natural bleaching event. Dis Aquat Organ *87*, 67-78.

Mydlarz, L.D., Holthouse, S.F., Peters, E.C., and Harvell, C.D. (2008). Cellular responses in sea fan corals: Granular amoebocytes react to pathogen and climate stressors. PLoS One *3*, Article No.: e1811.

Navarre, W.W., and Zychlinsky, A. (2000). Pathogen-induced apoptosis of macrophages: a common end for different pathogenic strategies. Cell Microbiol *2*, 265-273.

Ovchinnikova, T.V., Balandin, S.V., Aleshina, G.M., Tagaev, A.A., Leonova, Y.F., Krasnodembsky, E.D., Men'shenin, A.V., and Kokryakov, V.N. (2006). Aurelin, a novel antimicrobial peptide from jellyfish *Aurelia aurita* with structural features of defensins and channel-blocking toxins. Biochemical and Biophysical Research Communications *348*, 514-523.

Palmer, C.V., and Traylor-Knowles, N. (2012). Towards an integrated network of coral immune mechanisms. Proceedings of the Royal Society B: Biological Sciences *279*, 4106-4114.

Patterson, K.L., Porter, J.W., Ritchie, K.B., Polson, S.W., Mueller, E., Peters, E.C., Santavy, D.L., and Smith, G.W. (2002). The etiology of white pox, a lethal disease of the Caribbean elkhorn coral, *Acropora palmata*. Proceedings of the National Academy of Sciences *99*, 8725-8730.

Perez, S., and Weis, V. (2006). Nitric oxide and cnidarian bleaching: an eviction notice mediates breakdown of a symbiosis. J Exp Biol *209*, 2804 - 2810.

Pernice, M., Dunn, S.R., Miard, T., Dufour, S., Dove, S., and Hoegh-Guldberg, O. (2011). Regulation of apoptotic mediators reveals dynamic responses to thermal stress in the reef building coral *Acropora millepora.* PLoS One *6*, e16095.

Petes, L.E., Harvell, C.D., Peters, E.C., Webb, M.A.H., and Mullen, K.M. (2003). Pathogens compromise reproduction and induce melanization in Caribbean sea fans. Mar Ecol Prog Ser *264*, 167-171.

Polato, N.R., Vera, J.C., and Baums, I.B. (2011). Gene discovery in the threatened elkhorn coral: 454 sequencing of the *Acropora palmata* transcriptome. PLoS One *6*, e28634.

Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V.*, et al.* (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science *317*, 86-94.

Qiu, L., Song, L., Yu, Y., Zhao, J., Wang, L., and Zhang, Q. (2009). Identification and expression of TRAF6 (TNF receptor-associated factor 6) gene in Zhikong scallop *Chlamys farreri*. Fish & Shellfish Immunology *26*, 359-367.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Seipp, S., Schmich, J.r., and Leitz, T. (2001). Apoptosis: a death-inducing mechanism tightly linked with morphogenesis in *Hydractina echinata* (Cnidaria, Hydrozoa). Development *128*, 4891-4898.

Sokolova, I.M. (2009). Apoptosis in molluscan immune defense. Invertebrate Survival Journal *6*, 49-58.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688-2690.

Sullivan, J., Kalaitzidis, D., Gilmore, T., and Finnerty, J. (2007). Rel homology domain-containing transcription factors in the cnidarian *Nematostella vectensis*. Dev Genes Evol *217*, 63-72.

Sunagawa, S., Wilson, E., Thaler, M., Smith, M., Caruso, C., Pringle, J., Weis, V., Medina, M., and Schwarz, J. (2009). Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. Bmc Genomics *10*, 258.

Tchernov, D., Kvitt, H., Haramaty, L., Bibby, T.S., Gorbunov, M.Y., Rosenfeld, H., and Falkowski, P.G. (2011). Apoptosis and the selective survival of host animals following thermal bleaching in zooxanthellate corals. Proceedings of the National Academy of Sciences *108*, 9905-9909.

Team, R.D.C. (2013). R: A language and environment for statistical computing. (Vienna, Austria R Foundation for Statistical Computing).

Trench, R.K. (1997). Diversity of symbiotic dinoflagellates and the evolution of microalgal-invertebrate symbioses. Proceedings of the Eighth International Coral Reef Symposium *2*, 1275–1286.

Trench, R.K.a.R.J.B. (1987). *Symbiodinium microadriaticum Freudenthal, S. goreauii sp. nov., S. kawagutti sp. nov., and S. pilosumsp. nov.*: gymnodinioid dinoflagellate symbionts of marine invertebrates. Journal of Phycolology *23*, 469–481.479.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and 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.

Verstrepen, L., Bekaert, T., Chau, T.L., Tavernier, J., Chariot, A., and Beyaert, R. (2008). TLR-4, IL-1R and TNF-R signaling to NF-kB: variations on a common theme. Cellular and Molecular Life Sciences *65*, 2964-2978.

Voolstra, C.R., Schwarz, J.A., Schnetzer, J., Sunagawa, S., Desalvo, M.K., Szmant, A.M., Coffroth, M.A., and Medina, M. (2009). The host transcriptome remains unaltered during the establishment of coral-algal symbioses. Mol Ecol *18*, 1823-1833.

Wajant, H., Muhlenbeck, F., and Scheurich, P. (1998). Identification of a TRAF (TNF Receptor-Associated Factor) gene in *Caenorhabditis elegans*. Journal of Molecular Evolution *47*, 656-662.

Wang, P.-H., Wan, D.-H., Gu, Z.-H., Deng, X.-X., Weng, S.-P., Yu, X.-Q., and He, J.-G. (2011). *Litopenaeus vannamei* tumor necrosis factor receptor-associated factor 6 (TRAF6) responds to *Vibrio alginolyticus* and white spot syndrome virus (WSSV) infection and activates antimicrobial peptide genes. Developmental & Comparative Immunology *35*, 105-114.

Wang, S., Peatman, E., Liu, H., Bushek, D., Ford, S.E., Kucuktas, H., Quilang, J., Li, P., Wallace, R., Wang, Y*., et al.* (2010). Microarray analysis of gene expression in eastern oyster *(Crassostrea virginica)* reveals a novel combination of antimicrobial and oxidative stress host responses after dermo *(Perkinsus marinus)* challenge. Fish & Shellfish Immunology *29*, 921-929.

Wolenski, F.S., Garbati, M.R., Lubinski, T.J., Traylor-Knowles, N., Dresselhaus, E., Stefanik, D.J., Goucher, H., Finnerty, J.R., and Gilmore, T.D. (2011). Characterization of the Core Elements of the NF-kB Signaling Pathway of the Sea Anemone *Nematostella vectensis*. Molecular and Cellular Biology *31*, 1076-1087.

Yuan, S., Liu, T., Huang, S., Wu, T., Huang, L., Liu, H., Tao, X., Yang, M., Wu, K., Yu, Y*., et al.* (2009). Genomic and functional uniqueness of the TNF receptor-associated factor gene family in Amphioxus, the basal Chordate. The Journal of Immunology *183*, 4560-4568.

Zhao, S., and Fernald, R.D. (2005). Comprehensive algorithm for quantitative real-time polymerase chain reaction. Journal of Computational Biology *12*, 1047-1064.

Zink, S.D., Pedersen, L., Cianciotto, N.P., and Abu Kwaik, Y. (2002). The Dot/Icm type IV secretion system of *Legionella pneumophila* is essential for the induction of apoptosis in human macrophages. Infection and Immunity *70*, 1657-1663.

Zotti, T., Vito, P., and Stilo, R. (2012). The seventh ring: Exploring TRAF7 functions. Journal of Cellular Physiology *227*, 1280-1284.

CHAPTER 4

CHARACTERIZATION OF THE *AIPTASIA PALLIDA* TRANSCRIPTIONAL RESPONSE TO PATHOGEN EXPOSURE PART 2: MODULATION OF THE HOST RESPONSE BY THE PRESENE OF MUTUALISTIC DINOFLAGELLATES[1]

Abstract

Mutualisms between microorganisms and their animal hosts are ubiquitous and have played an important role in evolution. Growing evidence indicates that the establishment and maintenance of mutualistic relationships depends on interactions between microbes and the host immune system, however the influence of mutualists on host immunity is poorly understood. To this end, we investigated the modulation of the host response to the pathogenic bacterium, *Serratia marcescens*, by the presence of endosymbiotic dinoflagellates in the sea anemone, *Aiptasia pallida* using transcriptomics and behavioral and survival assays. Comparing host gene expression between symbiotic and aposymbiotic anemones, we found that the presence of dinoflagellates greatly alters host gene expression to pathogen challenge, both in terms of the number and type of genes expressed. Approximately 600 (50%) more genes were differentially expressed in aposymbiotic anemones, and for those genes that are shared between both symbiotic states, the majority of genes that are differentially expressed showed larger fold-changes in aposymbiotic anemones relative to their symbiotic counterparts. In addition, symbiotic anemones were less likely to recover from pathogen exposure and had lower survival rates (50%) than their aposymbiotic counterparts (100%). Based on gene expression data, two immune-related processes in particular, apoptosis and tumor necrosis factor receptor associated factor (TRAF)-mediated signaling, are modulated in symbiotic anemones. These results are consistent

[1] Co-authors include J. A. Schwarz and C. D. Harvell

with the hypothesis that symbiotic dinoflagellates suppress the *Aiptasia* response to *S. marcescens,* and thus support the growing body of evidence suggesting mutualistic dinoflagellates attenuate host immunity to achieve symbiotic homeostasis. In addition, we have provided candidate processes, pathways, and genes for further investigation to elucidate the mechanisms behind the altered host response to pathogens via symbiotic state.

Introduction

Mutualistic relationships between microorganisms and their hosts are widespread throughout the animal kingdom and have played an important role in evolution (Douglas, 2010; Margulis, 1993; Moran and Wernegreen, 2000). There is increasing evidence that the establishment and maintenance of mutualistic relationships relies on a complex interplay between the symbiont and the host innate immune system (Eberl, 2010; McFall-Ngai et al., 2012). However, the role of host immunity in mutualistic interactions is not well understood. Also lacking is an understanding of the influence of mutualists on host immunity, which has the potential to regulate host response to pathogenic invasion. A striking question concerns how the host tolerates mutualistic microbes while mounting a full immune response against pathogenic invaders.

The mutualistic relationship between cnidarians and intracellular dinoflagellates provides a tractable system to investigate the interactive dynamics of mutualism and parasitism. Symbiotic dinoflagellates are known to supplement host metabolism via the photosynthetic production of fixed carbon (Muscatine, 1973), but they may also contribute to host defense, either directly via the production of effector molecules (e.g. ROS and/or antimicrobial peptides), or indirectly by promoting increased constitutive levels of host immune effectors. In several

well-known mutualisms, symbionts confer protection from invading pathogens (Hedges et al., 2008; Oliver et al., 2003; Pais et al., 2008; Teixeira et al., 2008). For instance, carpenter ants with greater numbers of bacterial symbionts have improved pathogen encapsulation rates, fruit flies harboring *Wolbachia* populations have increased protection against viruses (Hedges et al., 2008; Teixeira et al., 2008), and symbiotic tsetse flies are less susceptible to infection by trypanosome parasites than their aposymbiotic (without symbiont) counterparts (Pais et al., 2008).

Alternatively, symbiotic dinoflagellates may actively suppress host immunity in order to successfully invade the host cell. This is a common strategy employed by both pathogenic and mutualistic bacteria, as well as protozoan parasites (Anselme et al., 2006; Buzoni-Gatel and Werts, 2006; Davidson et al., 2004; Lang et al., 2007; Plattner and Soldati-Favre, 2008; Pollard et al., 2009; Tato and Hunter, 2002). In the weevil, *Sitophilus zeamais*, the presence of bacterial symbionts leads to up-regulation of peptidoglycan recognition protein-LB (Anselme et al., 2006), a gene that has been found to suppress a key immune pathway in both fruit flies and mammals (Artis, 2008; Zaidman-Remy et al., 2006). Apicomplexan parasites, like *Toxoplasma* and *Plasmodium*, take advantage of host-cell machinery to become internalized, and then evade or suppress the host immune response via the release of proteins that interfere with host regulatory networks (Anselme et al., 2006; Buzoni-Gatel and Werts, 2006; Davidson et al., 2004; Lang et al., 2007; Plattner and Soldati-Favre, 2008; Pollard et al., 2009; Tato and Hunter, 2002). Dinoflagellates are a sister taxon to amplicomplexans (Cavalier-Smith, 1999), and it is therefore possible that because of a shared evolutionary history with apoicomplexan parasites, they utilize similar mechanisms to facilitate invasion and survival in cnidarian cells.

It is also possible that symbiont-mediated resistance and suppression are not mutually

exclusive. In a number of insect-bacterial symbioses, via processes that are not well understood, mutualistic symbionts contribute to host immunity while simultaneously suppressing particular immune pathways to enable successful host infection (Anselme et al., 2008; Anselme et al., 2006; Lhocine et al., 2008; Wang et al., 2009). For example, in *S. zeamais,* primary endosymbionts are recognized as microbial intruders by the host outside of the bacteriome, but are able to persist in the bacteriome via the upregulation of immune-modulating genes (Anselme et al., 2008).

To determine if the *Aiptasia pallida* immune response is modulated via the presence of mutualistic dinoflagellates, we used transcriptome sequencing to compare gene expression of *Serratia marcescens*-exposed anemones with and without their symbionts. We also monitored behavioral responses and survival of bacteria-exposed anemones to link changes in gene expression with alterations in organismal health. Via this multifaceted approach, we hope to gain a better understanding of how the mutualism persists and if there are tradeoffs associated with the symbiotic lifestyle.

Materials and Methods

Anemone maintenance

All anemones were from clonal population CC7 (Sunagawa et al. 2009). Anemones were maintained in incubators at 25°C in artificial seawater (ASW; Instant Ocean) in 1 L glass bowls and were fed freshly hatched brine-shrimp nauplii approximately three times per week with water changes following feeding. Symbiotic anemones were kept on a 12L:12D cycle at 18-22 μmol photons m^{-2} s^{-1} of PAR. Aposymbiotic animals were generated by exposing symbiotic anemones to 50 μM diuron, with daily water changes, for ~30 days, or until the anemones were

devoid of algae, as confirmed with fluorescence microscopy. Following the bleaching process, aposymbiotic anemones were maintained in the dark and remained there for ~2 years to prevent algal repopulation. Two months prior to experimentation, aposymbiotic anemones were subjected to light conditions identical to those for symbiotic anemones. One week prior to the incubation and survival experiments, feeding was terminated and anemones were maintained in 0.22-filtered ASW (FASW) with daily water changes.

Bacteria inoculate preparation and Incubation Experiment

Serratia marscescens strain PDL100, originally isolated from white-pox infected *Acropora palmata* in the Florida Keys (Patterson et al., 2002), was plated onto LB agar and grown overnight at 29°C. A single colony was isolated and suspended in 7 ml sterile marine broth and grown overnight at 29°C on a shaker at 150 rpm. A subsequent overnight culture with identical growth conditions was prepared to amplify the amount of culture by adding 250 ul of culture to 250 ml of sterile marine broth. The bacteria were centrifuged for 10 minutes at 4000 rpm, the supernatant was discarded, and the bacterial pellet was resuspended in FASW and diluted to 1.5 X 10^7 cells per ml as estimated via optical density at 600 nm and confirmed with plate counts.

Two days prior to the incubation experiment, anemones were transferred to 12-well plates with 4 ml of FASW to acclimate. To initiate an immune response, treatment symbiotic and aposymbiotic anemones were incubated in 4 ml of *S. marscescens* inoculate, while control symbiotic and aposymbiotic anemones were incubated in 4 ml FASW. After 6 hours, both control and treatment anemones were rinsed 3 times with FASW, blotted dry, and pooled into groups of 6 anemones, representing ~50 mg wet tissue weight, for a total of 4 replicate pools per

treatment. The tubes were flash frozen in liquid nitrogen and were stored at -80°C prior to processing.

RNA extraction, Library Preparation, and Sequencing

Total RNA was extracted using the ToTALLY RNA™ Total RNA Isolation Kit (Ambion, cat. no. AM1910) according to the manufacturer's instructions with the exception that the RNA was precipitated using 0.1 volume of 3 M sodium acetate and 4 volumes of 100% ethanol. The resulting RNA was purified using the RNA Clean and Concentrator™-25 Kit (Zymo Research, cat. no. R1017) and the RIN of each sample was verified to be ≥ 9 using Agilent 2100 Bioanalyzer.

Approximately 2 µg of total RNA per sample was processed using the TruSeq RNA Sample Prep Kit following the manufacturer's instructions to produce indexed libraries for mutliplex sequencing. The resulting libraries were pooled into 8 samples per lane and were single-end sequenced in two lanes with a target read length of 100 bp. Clustering and sequencing were performed by the Cornell University Life Sciences Core Laboratory Center using an Illumina HiSeq 2000 sequencer.

Read Filtering, Alignment, and Expression Analysis

Initial quality filtering of reads and bar-code removal was performed by the Cornell University Life Sciences Core Laboratory Center. Fastq-mcf was used to remove Illumina adaptors, trim low-quality terminal ends, discard short sequences, and filter reads with phred scores less than 15 (http://code.google.com/p/ea-utils). FastQC was used to assess the quality of the reads before and after filtering (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

To ensure that only host genes were being examined, reads were aligned to cnidarian-classified transcripts from the *Aiptasia* transcriptome (Chapter 2) using BWA (Li and Durbin, 2009). Transcripts were considered cnidarian if they were sorted as such via the machine learning program, TopSort, and they also mapped to *Aiptasia* genomic DNA sequences (Chapter 2). A python script was used to count the number of reads that aligned to transcript uniquely with no errors or gaps.

The R package edgeR (Robinson et al., 2010), which uses empirical Bayes estimation and exact tests based on the negative binomial distribution, was used to identify differentially expressed transcripts that were classified as such if the false discovery rate (FDR)-adjusted p-value was less than 0.05. Within edgeR, the data were normalized for differing sequencing depths between libraries and transcripts with low read counts were filtered out such that transcripts with at least one count per million in at least 4 samples were retained in the analysis. A generalized linear model (GLM) that incorporated state (e.g. symbiotic and aposymbiotic) and treatment (e.g. control and pathogen-exposed) was used to identify transcripts that were differentially expressed as a function of pathogen exposure for both aposymbiotic and symbiotic datasets (McCarthy et al., 2012). This inclusive model controls for differences in the variability between the aposymbiotic and symbiotic datasets so that the number of differentially expressed genes between the data sets can be directly compared. For genes that were differentially expressed as a function of pathogen exposure in both symbiotic and aposymbiotic datasets, an interaction term including both state and treatment was used to determine if these genes were significantly more differentially expressed in aposymbiotic anemones as compared to symbiotic anemones. By comparing gene expression of pathogen-exposed aposymbiotic and symbiotic anemones indirectly, we are able to identify genes that are differentially expressed solely as a

function of immune elicitation and not due to differences that occur due to symbiotic state (Figure 4.1).

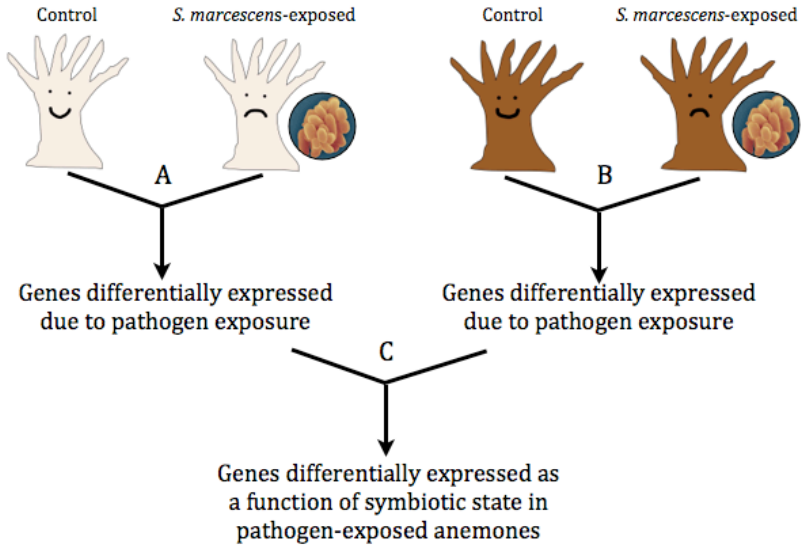


Figure 4.1 Design of pathogen-exposure experiment for transcriptome sequencing. RNA-Seq analysis of control and pathogen-exposed aposymbiotic (A) and symbiotic (B) anemones are conducted separately to identify genes that are expressed solely as a function of immune-challenge. The subsequent datasets are then compared to identify genes that are differentially expressed only in aposymbiotic or symbiotic anemones, or that are shared between the two datasets (C).

Annotation and Enrichment Analysis

Transcript sequence homology was determined via blastx searches against the SwissProt and NCBI non-redundant (nr) databases with an e-value cut-off of less than 1e-6. All blastx steps were performed in parallel via Cornell University's Computational Biology Application Suite for High Performance Computing (biohpc.org). For transcripts that did not have blastx hits, the program ORFPredictor (Min et al 2005) was used to predict open reading frames. SwissProt identifiers were aligned to Gene Ontology (GO) terms (Gene Ontology database: http://www.geneontology.org) and parental categories via the MGI GO Slim database

(http://www.informatics.jax.org) using the table joining function in Galaxy (Blankenberg et al., 2001).

The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 was used (Dennis et al., 2003; Huang et al., 2007) to perform an enrichment analysis to determine biological processes that were significantly overrepresented in differentially expressed transcripts relative to the background transcriptome using a modified Fisher Exact test. The Functional Annotation Clustering method was employed, which clusters groups of similar biological processes and provides an enrichment score representative of the –log geometric mean of the p-values of the individual biological processes. Clusters were considered highly significantly enriched when enrichment scores were greater than 2.0, which roughly corresponds to a p-value less than 0.004. Enrichment analysis was conducted on the subset of genes that were differentially expressed only in aposymbiotic or symbiotic anemones as a function of pathogen exposure, as well as genes that were differentially expressed in both datasets.

Real-Time Quantitative PCR

To validate the RNA-Seq data, RT-qPCR was performed on the same RNA pools that were used for library preparation and sequencing for a total of 4 biological replicates per treatment (e.g., symbiotic control and pathogen-exposed anemones and aposymbiotic control and pathogen-exposed anemones). The RNA was treated with DNase using the TURBO DNA-free kit™ (Invitrogen AM1907) and cDNA was synthesized using the GoScript™ Reverse Transcriptase System, both protocols performed according to manufacturers' instructions. Primers for 18 transcripts representing varying levels of expression (Table S4.1 A and B) were designed using Integrated DNA Technologies' PrimerQuest. Primer sequences are listed in

Supplementary Table 3 and product sizes ranged from 99-272 bp. Expected length of the amplicons was checked by agarose gel electrophoresis following standard PCR amplification. Primer efficiencies were determined using Real-time PCR Miner (Zhao and Fernald, 2005) and ranged from 88-95%.

Transcript level quantification was performed using Power SYBR® Green Master Mix (Applied Biosystems) and the ViiA™ 7 (Applied Biosystems) thermocycler. The reactions conditions are as follows: 1X Power SYBR® Green Master Mix, 200 nM of each primer, and 18 ng of cDNA in a total volume of 25 ul. Each sample and no template control were run in duplicate with the following thermocycler parameters: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 60 s, with a subsequent dissociation curve to confirm the absence of non-specific products. In addition, two pooled samples of RNA composed of 8 symbiotic and aposymbiotic samples were used as template to confirm the absence of genomic DNA contamination. Real Time PCR miner was used to calculate the efficiencies and critical threshold (C_T) of the genes based on raw fluorescence data.

The software package geNorm (Vandesompele et al., 2002) was used to evaluate relative expression stability of potential reference genes, including those that have been previously used in *Aiptasia* (cytochrome c oxidase, glyceraldehyde-3-phosphate dehydrogenase, 40S ribosomal protein S7, and 60S ribosomal protein L11; Chapter 2) and others that showed similar expression patterns between control and pathogen-exposed anemones based on RNA-Seq data (60S ribosomal protein L10 and Polyadenylate-binding protein 1). The genes with the highest stability (lowest M-values; Vandesompele et al. 2002) between control and pathogen-exposure for each symbiotic state were subsequently used as reference genes. For symbiotic anemones, cytochrome c oxidase, glyceraldehyde-3-phosphate dehydrogenase, 40S ribosomal protein S7,

and 60S ribosomal protein L10, were chosen, whereas cytochrome c oxidase, 40S ribosomal protein S7, 60S ribosomal protein L10 and Polyadenylate-binding protein 1 were chosen for aposymbiotic anemones. A normalization factor that was calculated from the geometric mean of the reference genes' expression values (Vandesompele et al., 2002) was used to normalize target genes and relative expression values were calculated using the equation $1/(1 + \text{Efficiency})^{\wedge}C_T$. Fold change in expression relative to control anemones was then calculated as the quotients of the above equation. The R software package (R Core Team, 2013) was used to perform Pearson correlations between fold change expression from qPCR and RNASeq data.

Behavioral/survival experiment

S. marscescens strain PDL100 was cultured as previously described and diluted to high, medium, and low concentrations as represented by 2.43×10^9, 4.2×10^8, and 5.38×10^7 cells per ml, respectively. As before, two days prior to the experiment, anemones were transferred to 12-well plates with 4 ml of FASW for acclimation. During the entirely of the experiment, anemones were maintained in incubators at 25°C and were kept on a 12L:12D cycle at 18-22 µmol photons m^{-2} s^{-1} of PAR. Twelve symbiotic and aposymbiotic anemones were incubated in 4 ml of *S. marscescens* inoculate at high, medium, and low concentrations, while 12 control symbiotic and aposymbiotic anemones were incubated in 4 ml FASW. Each plate had 3 anemones exposed to control and low, medium, and high concentrations of bacteria for a total of 8 plates. The anemones were exposed to the inoculate for 24 hours, during which they were placed on a shaker table at 90 rpm. After 24 hrs, the inoculate was removed and replaced with FASW, following three rinses with FASW. and water changes were made daily.

Anemones were sampled every 6 hrs for the first 48 hrs, 12 hours for the next 48 hrs, and then every 24 hrs until the termination of the experiment at 9 days. At each sampling point, anemones were photographed under a dissecting scope (Wild Heerbrugg, 6X magnification) and the behavioral response was monitored on a scale from 1 to 4, with 1 and 2 representing anemones for which the tentacles and column were extended and 3 and 4 representing anemones for which the tentacles and column were retracted (Figure 4.2). Anemones under ambient condition spend the majority of time with their tentacles and/or column extended, while anemones under various stressors (e.g., pathogen-exposure, heat-stress) retract their tentacles and/or column (Mouchka, unpuplished data). Survival was also monitored at each sampling point with the following criteria for death: 1) the anemone had detached from the well walls, 2) the anemone was producing excess mucus, and 3) the anemone was nonresponsive to light and/or touch.

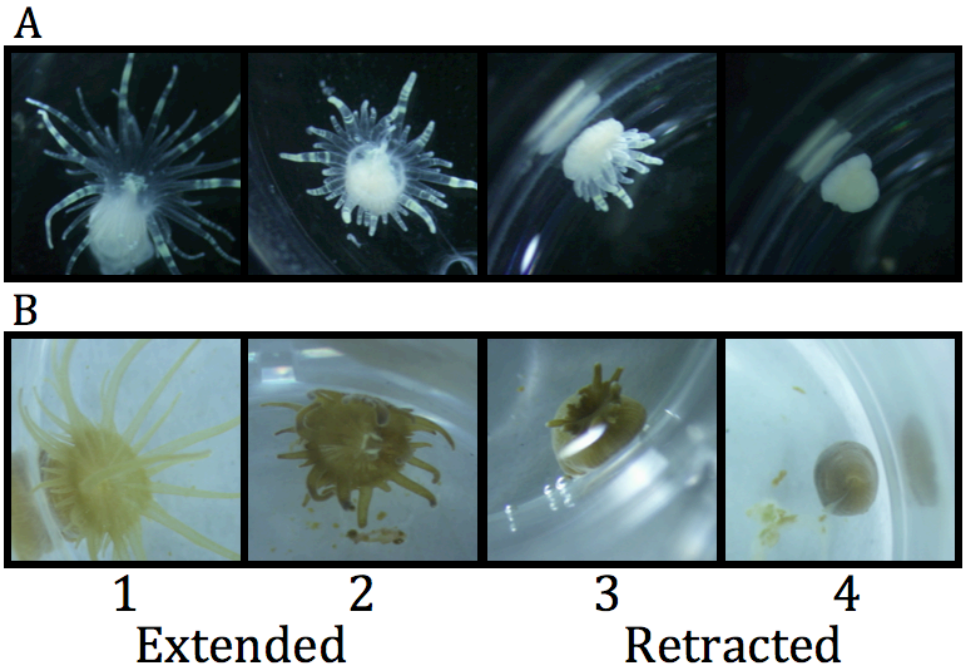


Figure 4. 2 Score of behavioral responses of aposymbiotic (A) and symbiotic (B) anemones that were monitored at each sampling point post-incubation with *S. marcescens*.

All statistics for the survival experiment were performed in the R software package (Team, 2013). To determine if there was a difference in the proportion of anemones that survived to the end of the experiment between aposymbiotic and symbiotic individuals, a two-sided test of equal proportions and a Kaplan-Meier survival distribution was performed using a parametric model with an exponential distribution and a constant hazard with censoring (Crawley, 2007). The proportion of anemones that had recovered following bacterial incubation was also investigated. Recovery was defined as returning to a behavioral response of 1 or 2 indefinitely following bacterial inoculation and a two-sided test of equal proportions was used to determine if there were differences in recovery between aposymbiotic and symbiotic anemones for each treatment. Lastly, the proportion of sampling periods that each anemone exhibited a particular behavioral response was documented. A GLM with binomial sampling was used to determine if the proportion of time spent in the extended state differed by treatment (con, low, med, high) and/or state (apo, sym).

Results

The effect of the presence of symbionts on the host response to pathogen-exposure

To determine if differences in gene expression between symbiotic and aposymbiotic anemones resulted in differences in organismal health, we monitored behavioral responses and survival of anemones following pathogen-exposure. For behavioral responses, because anemones frequently oscillate between being completely and partially extended or completely and partially retracted (states 1 and 2 or 3 and 4; data not shown), these behavioral states were combined into two distinct states, extended or retracted. At six hours post-incubation, all symbiotic and aposymbiotic anemones in all three bacterial concentrations were retracted (data

not shown). Figure 4.3A shows the proportion of sampling times the anemones spent in the extended state under control and pathogen-exposure treatments. For control and medium bacterial concentrations, there was no difference between aposymbiotic and symbiotic anemones in the proportion of sampling times spent in the extended state (Supplementary Table 5). However, at low and high bacterial concentrations, symbiotic anemones spent less time in the extended state, although only significantly so for the low bacterial concentrations (Table S4.3). Treatment did have a significant effect on the percent of time spent extended, with anemones in medium and high concentrations being extended less frequently than those in control and low concentrations (Table S4.3).

Under control conditions and low and medium concentrations of *S. marcescens*, both symbiotic and aposymbiotic anemones exhibited 100% survival at the conclusion of the experiment, or 9 days post-incubation (Figure 4.3B). Under high bacterial concentrations, 50% of symbiotic anemones died, whereas 0% of aposymbiotic anemones died by 6 days post-incubation (proportion test: χ -squared = 8, df = 1, p-value = 0.004678; Kaplan-Meier survival distribution: χ-squared = 10.7, df =1, p-value = 0.0011), and these percentages remained consistent until the end of the experiment at nine days post-incubation (Figure 4.3B and C). Recovery post-incubation was also monitored until experiment termination. For low and medium bacterial concentrations, 100% and 92% of both symbiotic and aposymbiotic anemones recovered, respectively. For high bacterial concentrations, 67% and 25% of aposymbiotic and symbiotic anemones, respectively, recovered by day 9 post-incubation (Figure 4.3D; χ-squared = 4.1958, df = 1, p-value = 0.02026).

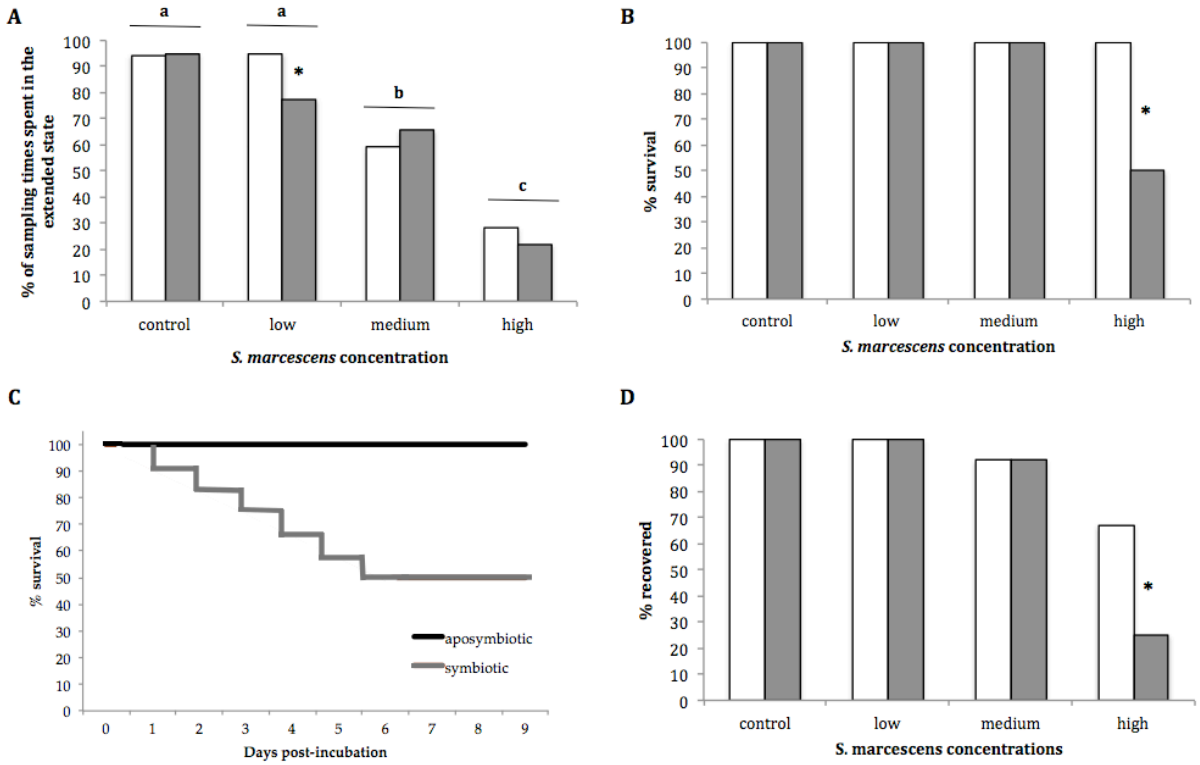


Figure 4.3 Behavioral, survival, and recovery data for aposymbiotic and symbiotic anemones exposed to low, medium, and high concentrations of *S. marcescens*. Unless otherwise stated, white bars =aposymbiotic and grey bars = symbiotic anemones. Letters and asterisks denote statistical significance between treatments and symbiotic states, respectively. **A.** Percent of sampling times anemones spent in the extended behavioral state. See Supplementary Table x for results of the GLM. **B.** Percent of anemones that survived until the termination of the experiment at nine days post-incubation (χ-squared = 8, df = 1, p-value = 0.004678). **C.** Survival curve for anemones incubated in high *S. marcescens* concentrations (χ-squared = 10.7, df =1, p-value = 0.0011). **D.** Percent of anemones that had recovered, as defined as returning to the extended state indefinitely following bacterial innoculation, by the termination of the experiment (χ -squared = 4.1958, df = 1, p-value = 0.02026).

Sequencing, Alignment, and Identification of Differentially Expressed Genes

To gain insight into the *Aiptasia* immune response and its potential modulation via symbiont dinoflagellates, we utilized RNASeq to analyze the transcriptome of control and pathogen-exposed aposymbiotic and symbiotic anemones. This approach resulted in greater than 394 million reads (~39 Gb) of sequence information. Following quality trimming, ~352 million

reads (~33 Gb) remained, for an average of ~36 and 52 million reads per aposymbiotic and symbiotic anemones, respectively (Table 4.1). An average of ~17.8 and ~17.5 million reads of aposymbiotic and symbiotic anemones, respectively, mapped uniquely to cnidarian-classified transcripts of the *Aiptasia* transcriptome (Chapter 2).

Table 4.1 Summary statistics for *Aiptasia pallida* transcriptome sequencing (N=4 biological replicates per treatment).

	Aposymbiotic		**Symbiotic**	
	Control	**Pathogen-exposed**	**Control**	**Pathogen-exposed**
Avg. # of reads before filtering	26,381,009	17,443,459	22,244,465	43,228,003
Avg. read length (bp)	100	100	100	100
Avg. # of reads after filtering	19,467,403	16,610,719	18,212,061	33,699,212
Avg. read length after filtering (bp)	95	95	94	95
Avg. # of reads uniquely mapped*	9,481,458	8,343,979	5,782,145	11,673,127

Figure 4.4A shows a multi-dimensional scaling plot in which the distances between samples is the biological coefficient of variation, a measure of the variation with which the abundance of the gene varies between replicate samples (EdgeR manual). The figure confirms the paired nature of the samples and suggests there are large differences in gene abundances between both control (C) and *S. marcescens*-exposed (T) anemones and aposymbiotic (A) and symbiotic (S) anemones. There is more variability between samples for symbiotic anemones, as the individual samples within the treatment groups do not cluster as tightly, and there is less separation across Dimension 2 between control and pathogen-exposed samples. However, the GLM applied to the data (see methods) accounts for this variability when calculating differentially expressed genes.

In the aposymbiotic dataset, we identified 1648 transcripts that were differentially expressed between control and pathogen-exposed anemones, whereas 1029 transcripts were differentially expressed in the symbiotic dataset (Figure 4.4B). 481 (~29%) and 252 (~23%) of the transcripts from the aposymbiotic and symbiotic datasets, respectively, exhibited expression differences greater than 2 fold. 485 of these transcripts were differentially expressed in both datasets, with 347 (72%) of these transcripts being more differentially expressed in aposymbiotic anemones, 60 (12%) of the 485 being significantly more differentially expressed. Approximately 66% and 52% of the transcripts are up-regulated in bacteria-challenged aposymbiotic and symbiotic anemones, respectively. Interestingly, all but seven of the 485 genes that overlap between the two datasets are expressed in the same direction, with six of the seven being up-regulated in aposymbiotic anemones and down regulated in symbiotic anemones.

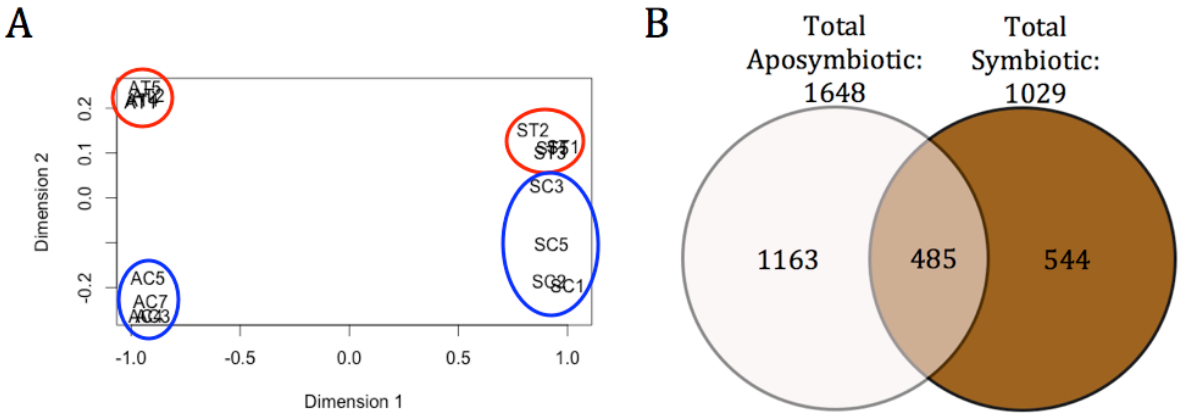


Figure 4.4 Comparison of gene expression between control and *S. marcescens*-exposed aposymbiotic and symbiotic anemones**. A.** Multi-dimensional scaling plot illustrating biological coefficient of variation between samples generated in EdgeR**.** (AT=aposymbiotic treatment; AC =aposymbiotic control; ST=symbiotic treatment; SC=symbiotic control; blue circles = control samples; red circles = treatment samples). **B.** Venn Diagram illustrating numbers of differentially expressed genes identified between control and *S. marcescens*-exposed aposymbiotic and symbiotic anemones.

To validate the results of the RNAseq experiment, we also performed RT-qPCR on 18 transcripts exhibiting a variety of expression patterns and read counts (Table S4.1). The expression values of the RNAseq and RT-qPCR experiments were highly correlated (Pearson coefficient of 0.99 and 0.87 for aposymbiotic and symbiotic datasets, respectively; both p-values = 2.2e-16), indicating high similarity between the RNAseq and RT-qPCR datasets (Figure 4.5).

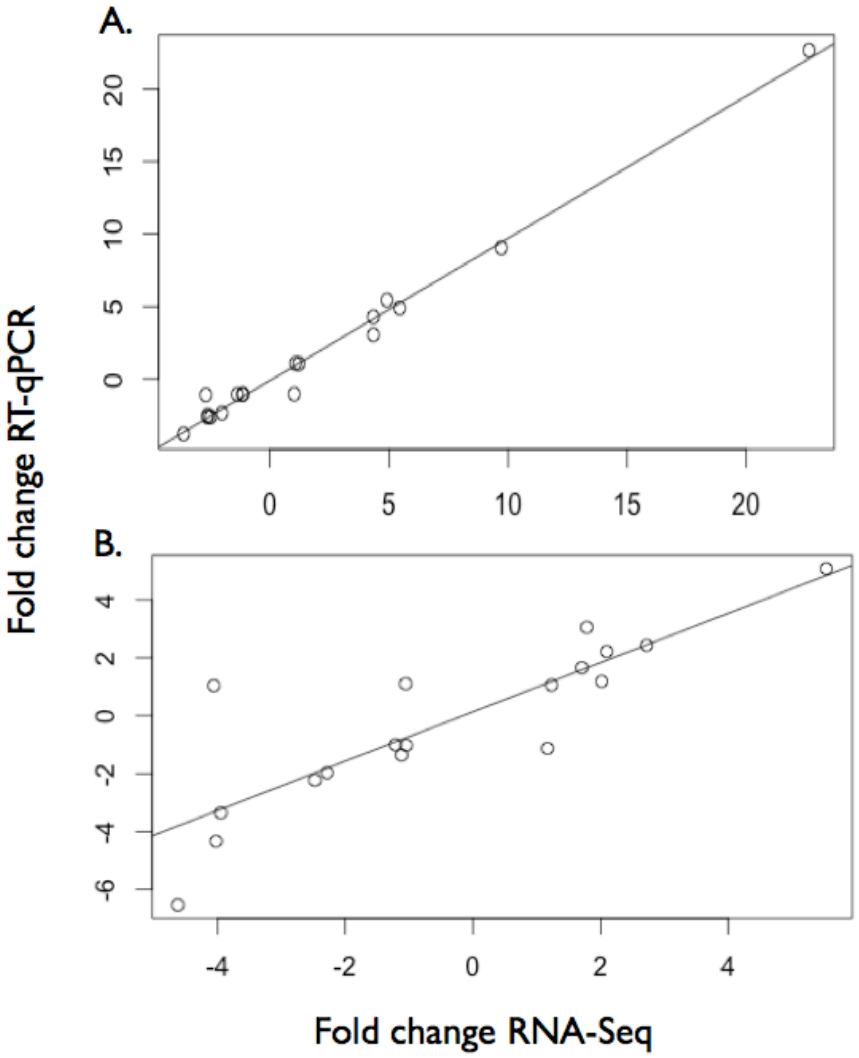


Figure 4.5 Validation of the RNA-Seq results with RT-qPCR. Fold change expression data of 18 genes obtained via RNA-Seq plotted against the corresponding expression data obtained with RT-qPCR (see also Supplementary Table 3) for the aposymbiotic (**A**) and symbiotic (**B**) data sets. A Pearson correlation coefficient of 0.99 ($p<2.22e-16$) and 0.87 ($p<2.2e-16$) for aposymbitoic and symbitoic datsets, respectively, demonstrates the high level of consistency between the RNA-Seq and RT-qPCR.

Annotation and Enrichment Analysis

Of the 1648 transcripts identified as differentially expressed in the aposymbiotic dataset, ~73% and 84% had significant blastx hits to the SwissProt and NCBI nr databases, respectively. For the symbiotic dataset, ~70% and 84% of differentially expressed transcripts had significant blastx hits to the SwissProt and NCBI nr databases, respectively. Although our further analyses focus on transcripts with blastx annotations, it is important to note that 256 and 168 of the differentially expressed transcripts in the aposymbiotic and symbiotic datasets, respectively, did not have significant blastx hits. However, 193 of these transcripts (including 59 with ≥2-fold expression changes) contained apparent open reading frames with ≥100 codons.

To examine functional categories of regulated genes in greater detail, we assigned biological process GO terms to differentially expressed transcripts and then conducted an enrichment analysis to identify biological processes that were overrepresented in the two datasets. Approximately 61% and 57% of differentially expressed transcripts from the aposymbiotic and symbiotic datasets, respectively, were assigned at least one GO biological process term, with 7,231 unique GO biological process terms assigned in total. These GO annotations were then aligned with GO slim terms to provide a broad classification of transcript function by dataset (Figure 4.6).

To highlight the differences between the two datasets, we conducted enrichment analysis on the subset of genes that were differentially expressed in aposymbiotic or symbiotic anemones only, and the genes that were differentially expressed between the two datasets (i.e., the overlapping dataset). For the aposymbiotic dataset, three clusters were highly enriched (i.e. enrichment score >2) and these clusters include biological process related to protein transport, apoptosis, and regulation of protein kinase cascade (more specifically, regulation of the I-κB

kinase/NF-κB cascade;Table 4.2). For the symbiotic data set, one cluster was highly enriched and includes procsses with developmental functions. Lastly, for the overlapping dataset, two clusters were highly enriched and include processes related to translation and transcription. These results are consistent with the GO analysis in Figure 4.6.

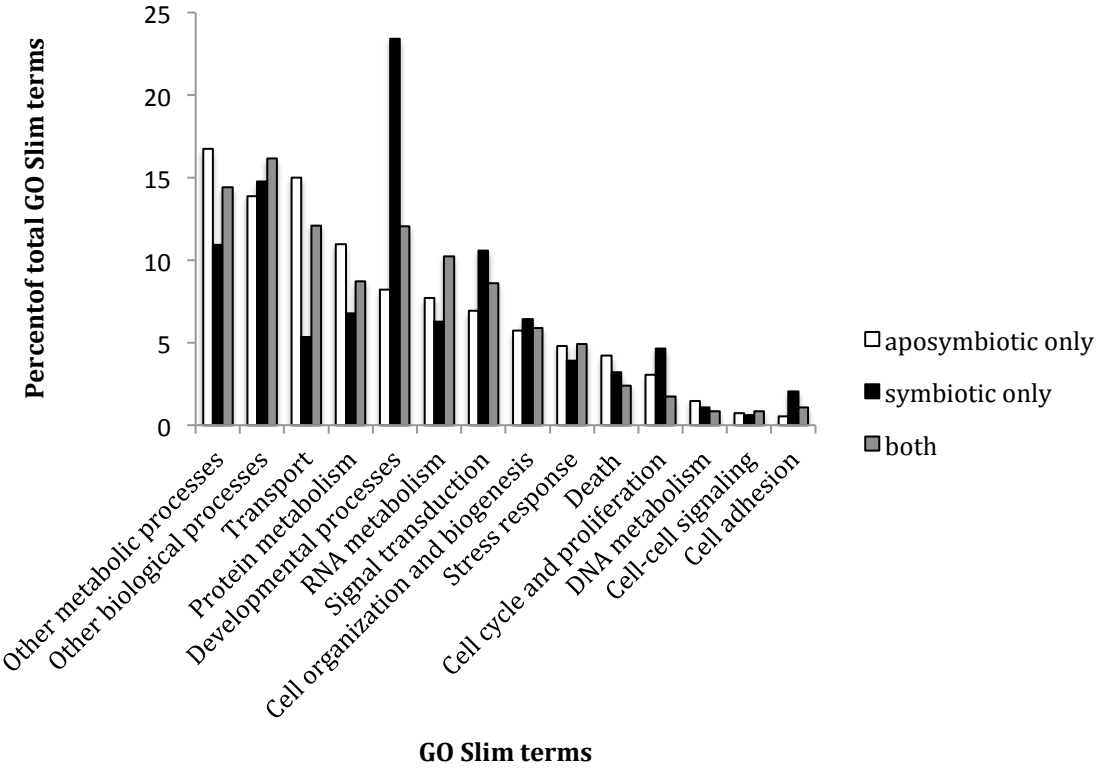


Figure 4.6 GO Slim biological process terms for genes that were differentially expressed between control and *S. marcescens*-exposed anemones. GO slim terms are parsed according to whether they were expressed uniquely in aposymbiotic or symbiotic anemones or in both.

A closer look at both datasets reveals there are 64 and 24 differentially expressed transcripts with apoptotic functions that are unique to the aposymbiotic and symbiotic datasets, respectively (Figure 4.7, Table S4.2). Twenty-one transcripts are differentially expressed in both

Table 4.2 The most enriched clusters of Biological Process GO terms with similar functions via DAVID enrichment analysis. The analysis was conducted on the subset of differentially expressed genes unique to aposymbiotic or symbiotic anemones, as well as the genes that were expressed in both datasets. P-values are from a modified Fisher's Exact Test to identify overrepresented functions in differentially expressed genes as compared to the background transcriptome. The enrichment score (next to cluster) is the –log geometric mean of the p-values in each cluster.

GO ID	GO term	p-value
Aposymbiotic Cluster 1: 3.71		
GO:0015031	protein transport	4.75E-05
GO:0045184	establishment of protein localization	7.63E-05
GO:0008104	protein localization	2.08E-04
GO:0046907	intracellular transport	2.38E-04
GO:0034613	cellular protein localization	3.40E-04
GO:0006886	intracellular protein transport	3.50E-04
GO:0070727	cellular macromolecule localization	5.09E-04
Aposymbiotic Cluster 2: 3.47		
GO:0012501	programmed cell death	1.48E-04
GO:0016265	death	2.71E-04
GO:0008219	cell death	5.27E-04
GO:0006915	apoptosis	5.99E-04
Aposymbiotic Cluster 3: 2.18		
GO:0010740	positive regulation of protein kinase cascade	4.74E-04
GO:0010627	regulation of protein kinase cascade	2.66E-03
GO:0043123	positive regulation of I-κB kinase/NF-κB cascade	3.13E-02
GO:0043122	regulation of I-κB kinase/NF-κB cascade	4.71E-02
Symbiotic Cluster 1: 2.36		
GO:0001501	skeletal system development	5.12E-04
GO:0030199	collagen fibril organization	6.39E-04
GO:0009792	embryonic development ending in birth or egg hatching	6.79E-04
GO:0048706	embryonic skeletal system development	7.35E-04
GO:0043062	extracellular structure organization	2.05E-03
GO:0043009	chordate embryonic development	4.06E-03
GO:0051216	cartilage development	4.61E-03
GO:0060351	cartilage development involved in endochondral bone morphogenesis	5.00E-03
GO:0030198	extracellular matrix organization	6.03E-03
GO:0048568	embryonic organ development	7.57E-03
GO:0044236	multicellular organismal metabolic process	8.78E-03
GO:0048705	skeletal system morphogenesis	9.63E-03
GO:0001958	endochondral ossification	1.02E-02

Table 4.2 (Continued)

GO:0048562	embryonic organ morphogenesis	1.64E-02
GO:0060350	endochondral bone morphogenesis	1.69E-02
GO:0060349	bone morphogenesis	4.51E-02
Overlapping Cluster 1: 6.50		
GO:0043038	amino acid activation	4.82E-10
GO:0006418	tRNA aminoacylation for protein translation	4.82E-10
GO:0043039	tRNA aminoacylation	4.82E-10
GO:0006399	tRNA metabolic process	1.88E-06
GO:0034660	ncRNA metabolic process	9.52E-04
GO:0006412	translation	5.76E-03
Overlapping Cluster 2: 3.07		
GO:0006350	transcription	4.78E-04
GO:0045449	regulation of transcription	7.63E-04
GO:0006355	regulation of transcription, DNA-dependent	1.06E-03
GO:0051252	regulation of RNA metabolic process	1.30E-03

datasets, with five transcripts being significantly more up-regulated in aposymbiotic anemones. For transcripts that are only differentially expressed in the aposymbiotic dataset, 66% of the up-regulated genes have pro-apoptotic functions, including the majority of the most up-regulated transcripts, while 66% of down-regulated genes have pro-apoptotic functions. For transcripts that are only differentially expressed in the symbiotic dataset, 63% of up-regulated genes have pro-apoptotic functions, while 93% of down-regulated genes have pro-apoptotic functions, including the most down-regulated genes.

Tumor necrosis factor receptor-associated factors (TRAF)-mediated signaling is an additional process for which there appears to be a modulated response via the symbiotic state. In aposymbiotic anemones, 25 genes with functions in TRAF-mediated signaling were differentially expressed as a function of bacteria exposure, whereas only 8 genes were

differentially expressed in symbiotic counterparts (Table 4.3). Furthermore, there were no genes involved in up-stream signaling (e.g., TNF ligands and receptors) and there were fewer TRAFs and down-stream signaling molecules that were differentially expressed in *S. marcescens*-exposed symbiotic anemones.

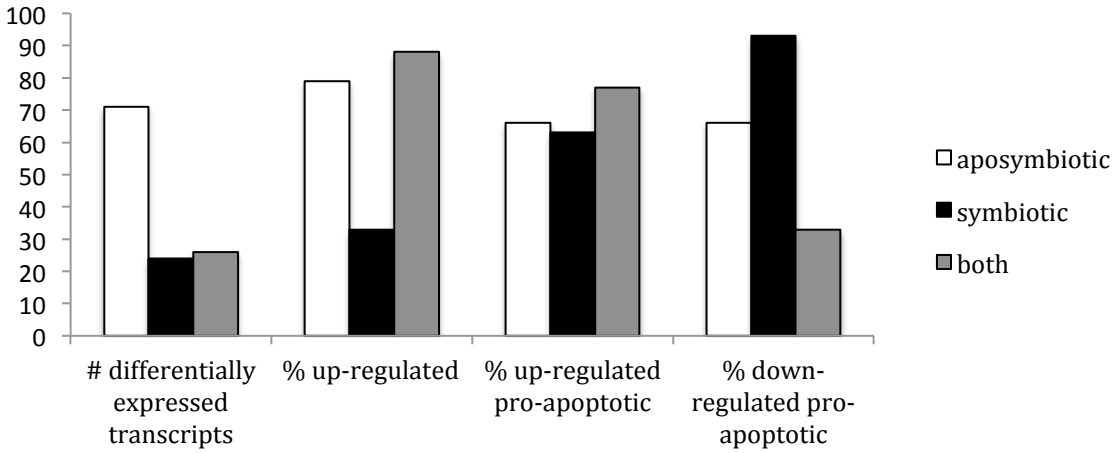


Figure 4.7 Summary of transcripts with GO biological processes relating to apoptosis uniquely expressed in aposymbiotic or symbiotic anemones or in both datasets.

Discussion

The characterization of the molecular and cellular mechanisms that underlie interactions between invertebrates and microorganisms has emerged as an important field of study. A better understanding of these interactions will broaden our understanding of the evolution of immunity and the ecological trade-offs that enable complex, multi-partner relationships. While much focus has been placed on host-pathogen interactions, more recent emphasis has centered on the role of host immunity in the establishment and maintenance of mutualistic symbioses. In addition, the impact of symbionts on host immunocompetence has also garnered attention. To date, the modulation of the host immune response to pathogens by the presence of mutualistic symbionts

Table 4.3 Transcripts with functions in TRAF-mediated signaling in *S. marcescens*-exposed aposymbiotic (Apo) and Symbiotic (Sym) anemones. P-values have been adjusted for a false discovery rate. NDE = not differentially expressed in symbiotic anemones.

Locus/Contig	Fold Change (Apo)	p-value	Top blastx hit to SwissProt database	Fold Change (Sym)	p-value
TNF ligands and receptors					
60163/1	1.7	9.14E-11	Lipopolysaccharide-induced TNF-α factor	NDE	
67712/1	-1.49	1.54E-05	Lipopolysaccharide-induced TNF-α factor	NDE	
44804/1	-1.19	3.00E-02	TNF ligand superfamily member 10	NDE	
95000/1	1.7	5.48E-10	TNF receptor superfamily member 27	NDE	
95010/1	1.99	2.00E-02	TNF receptor superfamily member 27	NDE	
TRAFs					
46907/1	3.15	6.03E-41	TRAF 2	1.90	1.67E-03
46888/1	2.07	2.00E-03	TRAF 2	NDE	
86236/1	1.2	3.00E-02	TRAF 2	NDE	
67642/1	1.17	4.00E-02	TRAF 3	NDE	
89049/1	2.34	1.22E-08	TRAF 3	1.64	5.00E-02
89050/1	1.66	5.32E-10	TRAF 3	1.43	1.18E-02
78485/1	4.16	2.19E-19	TRAF 3	NDE	
46912/1	2.25	1.45E-17	TRAF 3	1.92	1.15E-03
Table 4.3 (Continued)					
46914/1	1.51	2.47E-05	TRAF 3	NDE	
114598/1	-1.22	7.00E-03	TRAF 3	NDE	
31324/1	4.9	4.73E-50	TRAF 3	2.13[a]	1.43E-05
30592/1	1.88	1.00E-02	TRAF 4	NDE[b]	
32870/1	4.64	1.11E-19	TRAF 5	3.03	9.07E-05
109431/1	1.21	3.00E-02	TRAF 6	NDE	
Downstream signaling pathways					
32139/2	-1.78	8.00E-03	MAPK kinase kinase	NDE	
13803/1	2.63	2.30E-07	Proto-oncogene c-Jun	3.42	6.85E-05
19796/1	1.21	8.00E-03	Stress-activated protein kinase JNK	NDE	
114092/1	1.26	3.00E-03	Inhibitor of NF-κB kinase subunit alpha	NDE	
73699/1	1.58	7.94E-14	NFκB p105 subunit	NDE	
91253/1	1.51	1.00E-02	Phosphatidylinositol 3-kinase	NDE	

[a] Significantly more differentially expressed in the aposymbiotic anemones.
[b] An alternative transcript with homology to TRAF 4 was found in the symbiotic data set with a fold change of -1.5 and a p-value of 0.04.

has not been addressed in cnidarian-dinoflagellate symbiosis, nor has it been addressed in an animal-protozoan mutualism. To this end, we took a transcriptomic approach to determine how the response of *A. pallida* to a live pathogen, *S. marcescens,* differs as a consequence of the presence of symbionts. To directly assess functional consequences of symbiotic state, we also monitored behavioral responses and survival of symbiotic and apo-symbiotic anemones following pathogen-exposure.

One of the major findings to emerge from our study is the very different nature of the aposymbiotic and symbiotic response to *S. marcescens*, both at the levels of organismal health and at the level of gene expression. Under high concentrations of *S. marcescens*, symbiotic anemones had lower survival rates and were less likely to recover from pathogen exposure than their aposymbiotic counterparts. In addition, pathogen-exposed aposymbiotic and symbiotic anemones showed dramatically different gene expression responses, with few of the differentially expressed genes shared between them. Furthermore, enriched biological processes were distinct for the subset of genes that were uniquely expressed in each type of anemone: processes relating to protein transport, apoptosis, and protein kinase cascade regulation were enriched in aposymbiotic anemones, while processes relating to multicellular organismal development were enriched in symbiotic anemones. There is also evidence that the magnitude of changes in gene expression were dampened in symbiotic anemones relative to aposymbiotic anemones. Approximately 600 (50%) more genes were differentially expressed in aposymbiotic anemones, and for those genes that are shared between both symbiotic states, the majority of genes that are differentially expressed showed larger fold-changes in aposymbiotic relative to symbiotic anemones. Together these major findings all suggest that aposymbiotic anemones display a more robust response to potential pathogens.

Another major trend is the modulation of two pathways with central functions in innate immunity. In a previous study solely focusing the response of aposymbiotic anemones to *S. marcescens* (Chapter 3), we found that exposure to bacteria provoked changes in processes relating to apoptosis and TRAF-mediated signaling. Interestingly, these processes appear to be a response that is unique to aposymbiotic anemones, as we did not find these processes to be altered substantively in symbiotic anemones exposed to bacteria. The failure of symbiotic anemones to mount a robust apoptotic and TRAF-mediated signaling response, which are known to be standard responses to pathogens (Chung et al., 2002; Hedrick et al., 2010; Sokolova, 2009), suggests that the symbiosis itself either dampens or eliminates some avenues for responding to bacterial infection.

As intracellular symbionts, dinoflagellates rely on the sustained viability of their host cells. Therefore, during pathogen challenge, which can often lead to the induction of apoptosis as a mechanism to prevent pathogen spread and/or inflammation (Hedrick et al., 2010), modulation of apoptosis may be necessary to maintain symbiotic homeostasis. For instance, *T. gondii* suppresses apoptosis induced by a variety of stimuli (e.g., growth factor deprivation, toxins, TNF-alpha, and death ligands) at both the transcriptional and post-transcriptional levels in several types of mammalian cell lines (Reviewed in Lang et al., 2007; Molestina et al., 2003).

Interestingly, when I compared gene expression between symbiotic and aposymbiotic anemones under ambient conditions in Chapter 2, I found the opposite pattern. Specifically, I saw enrichment of apoptotic processes in symbiotic anemones and hypothesized that apoptosis plays a role in the maintenance of the mutualism by regulating host-symbiont biomass. Taken together, these studies suggest that apoptosis is context dependent and that pathogen exposure may shift conditions such that the suppression of apoptosis is necessary to preserve the symbiotic

state. Because these studies examined gene expression patterns in the entire organism, we have been unable to distinguish tissue-specific responses, and it may be that apoptotic processes occurring in healthy symbiotic anemones represent tissue remodeling/repair processes in some tissues, whereas apoptotic processes occurring in aposymbiotic anemones exposed to bacteria represent an immune-related response serving to remove bacteria. To tease apart the role of apoptosis in the intact symbiosis vs. in host responses to potential pathogens will require a more fine-scaled study.

TRAFs are intracellular signaling molecules that function as critical transducers for members of TNFR, Toll-like receptor, and Interleukin-1 receptor families (reviewed in Chung et al., 2002 and Verstrepen et al., 2008). TRAF signaling ultimately results in the activation of transcription factors, such as NF-κB, (Verstrepen et al., 2008) that regulate the expression of a large number of genes involved in immunity, inflammation, and cell survival and proliferation (Verstrepen et al., 2008). Modulation of TRAF-mediated signaling pathways, particularly suppression of NF-κB, appears to be a common theme in maintaining mutualisms between gut microbes and their human or *Drosophila* hosts (Kumar et al., 2007; Neish et al.; Ryu et al.). This strategy has also been adopted by *T. gondii*, which is able to invade macrophages without activating NF-κB and can inhibit LPS-induced activation of the NF-κB pathway once established (Butcher et al., 2001; Shapira et al., 2005).

The differences in host gene expression and recovery and survival between aposymbiotic and symbiotic anemones, as well as the modulation of key immune pathways via the symbiotic state, are consistent with the hypothesis that attenuation of host immunity by dinoflagellates is necessary to promote symbiotic homeostasis. This is a common strategy employed to establish and maintain microbial relationships by both pathogenic and mutualistic microorganisms.

Indeed, several members of *Apicomplexa*, the sister taxa to dinoflagellates, successfully infect and persist within host cells via the suppression of host immune mechanisms (reviewed in Plattner and Soldati-Favre, 2008). This pattern has also been documented in several insect-bacterial mutualisms. For instance, *Drosophila simulans* harboring beneficial *Wolbachia* show a reduced ability to encapsulate parasitoid eggs (Fytrou et al., 2006) aphids infected with the secondary symbionts, *Hamiltonella defensa* or *Regiella insecticola*, have lower adherent hemocyte numbers and granulocyte proportions (Schmitz et al., 2012), and pill bugs infected with beneficial *Wolbachia* display lower hemocytes densities and prophenoloxidase activity, more severe septicemia, and have reduced expression of immune-related genes (Braquart-Varnier et al., 2008; Chevalier et al., 2011; Chevalier et al., 2012).

There is also molecular and cellular evidence for modulation of the cnidarian host response via symbionts (Chen et al., 2003; Chen et al., 2004; Detournay et al., 2012; Detournay and Weis, 2011; Voolstra et al., 2009; Chapter 2). Detournay et al (2013) found that aposymbiotic anemones have a higher capacity to respond to a microbial elicitor than symbiotic anemones and this response is attenuated in aposymbiotic anemones by the addition of the anti-inflammatory protein transforming growth factor-β. Aposymbiotic coral larvae exposed to compatible strains of dinoflagellates show little alteration in host gene expression, while exposure to incompatible strains results in dramatic changes (Voolstra et al., 2009). Dinoflagellates actively exclude host endocytosis regulatory proteins from the host phagosome in which they reside to prevent phagosome maturation, and thus symbiont digestion via lysosome fusion (Chen et al., 2003). Our results support these studies and add to the growing body of evidence suggesting successful cnidarian-dinoflagellate mutualisms rely on reduced host immunocompetence.

It should be noted, however, that based on our approach, we are unable to rule out alternative hypotheses as to the mechanisms underlying the altered host response to *S. marcescens* in symbiotic anemones. For instance, it's possible that symbionts do not directly suppress host immunity. Rather, the symbiotic state may alter host physiology such that the magnitude of gene expression is dampened and different pathways are invoked. Indeed, the presence of dinoflagellates leads to large shifts in host physiology (Ganot et al., 2011; Reynolds et al., 2000; Rodriguez-Lanetty et al., 2006; Chapter 2), and these shifts could also include host-mediated changes in immunity. Alternatively, gene expression patterns in symbiotic anemones could be explained via the production of symbiont-derived antimicrobial compounds. At low bacterial concentrations, the symbionts may contribute to defense such that the host does not upregulate a response. At high concentrations, however, the bacteria may overwhelm both the symbiont and host response, resulting in anemone death. This scenario implies that the presence of symbionts may hinder immunity at high bacterial concentrations, and it is therefore possible that the symbiotic state can both enhance an suppress host immunity in a context dependent manner.

Our results could also be explained via differences in surface mucus layers between aposymbiotic and symbiotic anemones. Because they exude photosynthate into their mucus (Brown and Bythell, 2005), it's possible that symbiotic anemones experience a more lethal bacterial infection via fertilization of bacterial growth. Moreover, differences in surface microbial assemblages between aposymbiotic and symbiotic anemones may facilitate differential infection success. However, it seems less likely that different bacterial concentrations would alter the outcome under these scenarios. Indeed, future functional studies are required to investigate the molecular and cellular underpinnings that explain the major findings of our study.

Regardless of the mechanism, the decreased resistance of symbiotic anemones to bacterial infection that we observed in this study implies that the host incurs a cost of mutualism. Indeed, tradeoffs are an inherent facet of mutualistic relationships (Bronstein, 2001; Herre and West, 1997). For instance, the presence of secondary bacterial symbionts in the pea aphid, *Acyrthosiphon pisum,* provides resistance to parasitoid wasp attack, but at the cost of reduced fecundity (Gwynn et al., 2005). It stands to reason that a tradeoff may exist between nutritional supplementation and immunity in cnidarian-dinoflagellate mutualisms. From the host's perspective, under stable conditions, the benefits of metabolic provisions supplied by their symbionts outweigh the costs of hosting the symbionts. However, under stressful conditions, like those experienced during pathogen-exposure, the presence of dinoflagellates might tip the cost/benefit balance, thereby undermining the mutualistic nature of the interaction. This scenario has also been proposed to explain coral bleaching (Weis, 2008), or the stress-induced breakdown of the cnidarian-dinoflagellate relationship. Under this model, dinoflagellates effectively become parasitic under stressful conditions due to the production of excessive reactive oxygen species, and are thus eliminated to prevent host damage (Weis, 2008). This cost/benefit balance could have important implications for coral, as being symbiotic may compromise health under some scenarios. With increasing environmental pressures on coral reefs, it may be that the benefit offered by being symbiotic becomes outweighed when the symbiotic state acts to intensify stressors (e.g, thermal anomalies and/or increased pathogen load), rather than reliving them.

Our results, which suggest that aposymbiotic anemones are better able to resist bacterial infection, at first may seem to be in direct conflict with previous studies that have shown that corals that have lost their symbionts from coral bleaching show increased disease susceptibility

(Harvell et al., 2001; Miller et al., 2009; Muller et al., 2008). However, the phenomenon of coral bleaching is not equivalent to the state of aposymbiosis in the lab. Coral bleaching occurs after protracted stressful conditions (e.g., elevated sea surface temperatures), causing the host to become physiologically stressed and potentially more susceptible to infection. Furthermore, bleached corals are also effectively starved, as up to 95% of their metabolic needs are derived from symbiont photosynthesis (Muscatine et al., 1984). The anemones used in this experiment, in contrast, were not subjected to stressful conditions prior to the experiment, and were fed regularly until one week before the initiation of the experiment. We hypothesize that the advantage incurred by aposymbiotic anemones in our study would not apply to bleached corals due to their altered physiological state.

While we were able to annotate ~70% of differentially expressed genes from both symbiotic and aposymbiotic datasets, many of the most differentially expressed transcripts in our study did not share sequence homology with proteins in the SwissProt or NCBI nr databases. Of particular interest are two transcripts (15721/1 and 2003/1) that were expressed as a function of bacterial challenge in both aposymbiotic and symbiotic datasets, but were ~4 and 3 fold more differentially expressed in aposymbiotic anemones. The expression of these genes appears to be modulated via the symbiotic state and their functions may clarify some of the physiological mechanisms that underlie this modulation. A future bioinformatic characterization of these genes to determine whether they are present in other cnidarians, and whether their expression appears to be connected to stress or disease will provide first steps towards understanding the role of these genes.

Conclusions

In summary, we found that the *Aiptasia* response to *S. marcescens* is greatly altered via the presence of dinoflagellates. This result is consistent with previous studies in cnidarians that suggest modulation of host immunity is necessary for the establishment and maintenance of the cnidarian-dinoflagellate relationship. Our approach, pairing gene expression with behavioral/survival assays, provides new insights into the host response to bacteria-challenge and allows for an improved interpretation of gene expression data. Moreover, we have generated a list of candidate pathways, processes, and genes that can be investigated in further detail to elucidate the mechanisms behind how the symbiotic state alters the host response to bacteria.

REFERENCES

Anselme, C., Perez-Brocal, V., Vallier, A., Vincent-Monegat, C., Charif, D., Latorre, A., Moya, A., and Heddi, A. (2008). Identification of the weevil immune genes and their expression in the bacteriome tissue. BMC Biology *6*, 1-13.

Anselme, C., Vallier, A., Balmand, S., Fauvarque, M.-O., and Heddi, A. (2006). Host PGRP Gene Expression and Bacterial Release in Endosymbiosis of the Weevil *Sitophilus zeamais*. Appl Environ Microbiol *72*, 6766-6772.

Artis, D. (2008). Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. Nature Reviews Immunology *8*, 411-420.

Blankenberg, D., Kuster, G.V., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2001). Galaxy: A Web-Based Genome Analysis Tool for Experimentalists. In Current Protocols in Molecular Biology (John Wiley & Sons, Inc.).

Braquart-Varnier, C., Lachat, M., Herbiniere, J., Johnson, M., Caubet, Y., Bouchon, D., and Sicard, M. (2008). *Wolbachia* Mediate Variation of Host Immunocompetence. PLoS One *3*, e3286.

Bronstein, J.L. (2001). The Costs of Mutualism. American Zoologist *41*, 825-839.

Brown, B.E., and Bythell, J.C. (2005). Perspectives on mucus secretion in reef corals. Mar Ecol-Prog Ser *296*, 291-309.

Butcher, B.A., Kim, L., Johnson, P.F., and Denkers, E.Y. (2001). *Toxoplasma gondii* Tachyzoites Inhibit Proinflammatory Cytokine Induction in Infected Macrophages by Preventing Nuclear Translocation of the Transcription Factor NF-kB. The Journal of Immunology *167*, 2193-2201.

Buzoni-Gatel, D., and Werts, C. (2006). *Toxoplasma gondii* and subversion of the immune system. Trends Parasitol *22*, 448-452.

Cavalier-Smith, T. (1999). Principles of protein and lipid targting in secondary symbiogenesis: Euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. . Journal of Eukaryote Microbiology *46*, 347-366.

Chen, M., Cheng, Y., Sung, P., Kuo, C., and Fang, L. (2003). Molecular identification of Rab7 (ApRab7) in *Aiptasia pulchella* and its exclusion from phagosomes harboring zooxanthellae. Biochem Biophys Res Commun *308*, 586 - 595.

Chen, M.-C., Cheng, Y.-M., Hong, M.-C., and Fang, L.-S. (2004). Molecular cloning of Rab5 (ApRab5) in *Aiptasia pulchella* and its retention in phagosomes harboring live zooxanthellae. Biochemical and Biophysical Research Communications *324*, 1024-1033.

Chevalier, F.d.r., Herbiniere-Gaboreau, J., Bertaux, J., Raimond, M., Morel, F., Bouchon, D., Greve, P., and Braquart-Varnier, C. (2011). The Immune Cellular Effectors of Terrestrial Isopod *Armadillidium vulgare*: Meeting with Their Invaders, *Wolbachia*. PLoS One *6*, e18531.

Chevalier, F.d.r., Herbiniere-Gaboreau, J., Charif, D., Mitta, G., Gavory, F.d.r., Wincker, P., Greve, P., Braquart-Varnier, C., and Bouchon, D. (2012). Feminizing *Wolbachia*: a transcriptomics approach with insights on the immune response genes in *Armadillidium vulgare*. BMC Microbiology *12*, 1-18.

Chung, J., Park, Y., Ye, H., and Wu, H. (2002). All TRAFs are not created equal: common and distinct molecular mechanisms of TRAF-mediated signal transduction. J Cell Sci *115*, 679 - 688.

Crawley, M.J. (2007). The R Book (Wiley Publishing).

Davidson, S.K., Koropatnick, T.A., Kossmehl, R., Sycuro, L., and McFall-Ngai, M.J. (2004). NO means 'yes' in the squid-vibrio symbiosis: nitric oxide (NO) during the initial stages of a beneficial association. Cell Microbiol *6*, 1139-1151.

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol *4*, 11.

Detournay, O., Schnitzler, C.E., Poole, A., and Weis, V.M. (2012). Regulation of cnidarian-dinoflagellate mutualisms: Evidence that activation of a host TGFβ innate immune pathway promotes tolerance of the symbiont. Developmental & Comparative Immunology *38*, 525-537.

Detournay, O., and Weis, V.M. (2011). Role of the Sphingosine Rheostat in the Regulation of Cnidarian-Dinoflagellate Symbioses. The Biological Bulletin *221*, 261-269.

Douglas, A.E. (2010). The Symbiotic Habit (Princeton Princeton University Press).

Eberl, G. (2010). A new vision of immunity: homeostasis of the superorganism. Mucosal Immunol *3*, 450-460.

Fytrou, A., Schofield, P.G., Kraaijeveld, A.R., and Hubbard, S.F. (2006). *Wolbachia* infection suppresses both host defence and parasitoid counter-defence. Proceedings of the Royal Society B: Biological Sciences *273*, 791-796.

Ganot, P., Moya, A.l., Magnone, V., Allemand, D., Furla, P., and Sabourault, C.c. (2011). Adaptations to Endosymbiosis in a Cnidarian-Dinoflagellate Association: Differential Gene Expression and Specific Gene Duplications. PLoS Genet *7*, e1002187.

Gwynn, D.M., Callaghan, A., Gorham, J., Walters, K.F.A., and Fellowes, M.D.E. (2005). Resistance is costly: trade-offs between immunity, fecundity and survival in the pea aphid. Proceedings of the Royal Society B: Biological Sciences *272*, 1803-1808.

Harvell, D., Kim, K., Quirolo, C., Weir, J., and Smith, G. (2001). Coral bleaching and disease: Contributors to 1998 mass mortality in *Briareum asbestinum* (Octocorallia, Gorgonacea). Hydrobiologia *460*, 97-104.

Hedges, L.M., Brownlie, J.C., O'Neill, S.L., and Johnson, K.N. (2008). *Wolbachia* and Virus Protection in Insects. Science *322*, 702-702.

Hedrick, S.M., Ch'en, I.L., and Alves, B.N. (2010). Intertwined pathways of programmed cell death in immunity. Immunol Rev *236*, 41-53.

Herre, E.A., and West, S.A. (1997). Conflict of interest in a mutualism: documenting the elusive fig wasp-seed tradeoff. Proceedings of the Royal Society of London Series B: Biological Sciences *264*, 1501-1507.

Huang, D.W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C.*, et al.* (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res *35*, W169-W175.

Kumar, A., Wu, H., Collier-Hyams, L.S., Hansen, J.M., Li, T., Yamoah, K., Pan, Z.-Q., Jones, D.P., and Neish, A.S. (2007). Commensal bacteria modulate cullin-dependent signaling via generation of reactive oxygen species. EMBO J *26*, 4457-4466.

Lang, C., Gross, U., and Luder, C.G.K. (2007). Subversion of innate and adaptive immune responses by *Toxoplasma Gondii*. Parasitol Res *100*, 191-203.

Lhocine, N., Ribeiro, P.S., Buchon, N., Wepf, A., Wilson, R., Tenev, T., Lemaitre, B., Gstaiger, M., Meier, P., and Leulier, F. (2008). PIMS Modulates Immune Tolerance by Negatively Regulating *Drosophila* Innate Immune Signaling. Cell Host & Microbe *4*, 147-158.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Margulis, L. (1993). Symbiosis in cell evolution (New York, W. H. Freeman).

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

McFall-Ngai, M., Heath-Heckman, E.A.C., Gillette, A.A., Peyer, S.M., and Harvie, E.A. (2012). The secret languages of coevolved symbioses: Insights from the *Euprymna scolopes*-*Vibrio fischeri* symbiosis. Seminars in Immunology *24*, 3-8.

Miller, J., Muller, E., Rogers, C., Waara, R., Atkinson, A., Whelan, K.R.T., Patterson, M., and Witcher, B. (2009). Coral disease following massive bleaching in 2005 causes 60% decline in coral cover on reefs in the US Virgin Islands. Coral Reefs *28*, 925-937.

Molestina, R.E., Payne, T.M., Coppens, I., and Sinai, A.P. (2003). Activation of NF-kB by *Toxoplasma gondii* correlates with increased expression of antiapoptotic genes and localization of phosphorylated Ik-B to the parasitophorous vacuole membrane. Journal of Cell Science *116*, 4359-4371.

Moran, N.A., and Wernegreen, J.J. (2000). Lifestyle evolution in symbiotic bacteria: insights from genomics. Trends Ecol Evol *15*, 321-326.

Muller, E.M., Rogers, C.S., Spitzack, A.S., and van Woesik, R. (2008). Bleaching increases likelihood of disease on *Acropora palmata* (Lamarck) in Hawksnest Bay, St John, US Virgin Islands. Coral Reefs *27*, 191-195.

Muscatine, L. (1973). Nutrition of Corals (New York).

Muscatine, L., Falkowski, P.G., Porter, J.W., and Dubinsky, Z. (1984). Fate of Photosynthetic Fixed Carbon in Light- and Shade-Adapted Colonies of the Symbiotic Coral *Stylophora pistillata*. Proceedings of the Royal Society of London Series B Biological Sciences *222*, 181-202.

Neish, A. S., Gewirtz, A.T., Zeng, H., Young, A.N., Hobert, M.E., Karmali, V., Rao, A.S., and Madara, J.L. Prokaryotic regulation of epithelial responses by inhibition of IkappaB-alpha ubiquitination.

Oliver, K.M., Russell, J.A., Moran, N.A., and Hunter, M.S. (2003). Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. Proceedings of the National Academy of Sciences *100*, 1803-1807.

Pais, R., Lohs, C., Wu, Y.N., Wang, J.W., and 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.

Patterson, K.L., Porter, J.W., Ritchie, K.B., Polson, S.W., Mueller, E., Peters, E.C., Santavy, D.L., and Smith, G.W. (2002). The etiology of white pox, a lethal disease of the Caribbean elkhorn coral, *Acropora palmata*. Proceedings of the National Academy of Sciences *99*, 8725-8730.

Plattner, F., and Soldati-Favre, D. (2008). Hijacking of Host Cellular Functions by the Apicomplexa. Annu Rev Microbiol *62*, 471-487.

Pollard, A.M., Knoll, L.J., and Mordue, D.G. (2009). The role of specific *Toxoplasma gondi*i molecules in manipulation of innate immunity. Trends Parasitol *25*, 491-494.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rodriguez-Lanetty, M., Phillips, W.S., and Weis, V.M. (2006). Transcriptome analysis of a cnidarian-dinoflagellate mutualism reveals complex modulation of host gene expression. Bmc Genomics *7*, 11.

Ryu J H, Kim, S.H., Lee, H.Y., Bai, J.Y., Nam Y.D. , Bae J.W. , Lee, D.G., Shin, S.C., Ha, E. M. and Lee, W.J. Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in *Drosophila*.

Schmitz, A., Anselme, C., Ravallec, M., Rebuf, C., Simon, J.-C., Gatti, J.-L., and Poirie, M.n. (2012). The Cellular Immune Response of the Pea Aphid to Foreign Intrusion and Symbiotic Challenge. PLoS One *7*, e42114.

Shapira, S., Harb, O.S., Margarit, J., Matrajt, M., Han, J., Hoffmann, A., Freedman, B., May, M.J., Roos, D.S., and Hunter, C.A. (2005). Initiation and termination of NF-kB signaling by the intracellular protozoan parasite *Toxoplasma gondii*. Journal of Cell Science *118*, 3501-3508.

Sokolova, I.M. (2009). Apoptosis in molluscan immune defense. Invertebrate Survival Journal *6*, 49-58.

Tato, C.M., and Hunter, C.A. (2002). Host-pathogen interactions: Subversion and utilization of the NF-kappaB pathway during infection. Infection and Immunity *70*, 3311-3317.

Team, R.D.C. (2013). R: A language and environment for statistical computing. (Vienna, Austria R Foundation for Statistical Computing).

Teixeira, L., Ferreira, A., and Ashburner, M. (2008). The Bacterial Symbiont *Wolbachia* Induces Resistance to RNA Viral Infections in *Drosophila melanogaster*. PLoS Biol *6*, 2753-2763.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and 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.

Verstrepen, L., Bekaert, T., Chau, T.L., Tavernier, J., Chariot, A., and Beyaert, R. (2008). TLR-4, IL-1R and TNF-R signaling to NF-kB: variations on a common theme. Cellular and Molecular Life Sciences *65*, 2964-2978.

Voolstra, C.R., Schwarz, J.A., Schnetzer, J., Sunagawa, S., Desalvo, M.K., Szmant, A.M., Coffroth, M.A., and Medina, M. (2009). The host transcriptome remains unaltered during the establishment of coral-algal symbioses. Mol Ecol *18*, 1823-1833.

Wang, J., Wu, Y., Yang, G., and Aksoy, S. (2009). Interactions between mutualist *Wigglesworthia* and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. Proc Natl Acad Sci USA *106*, 12133-12138.

Weis, V.M. (2008). Cellular mechanisms of Cnidarian bleaching: stress causes the collapse of symbiosis. J Exp Biol *211*, 3059-3066.

Zaidman-Remy, A., Herve, M., Poidevin, M., Pili-Floury, S., Kim, M.-S., Blanot, D., Oh, B.-H., Ueda, R., Mengin-Lecreulx, D., and Lemaitre, B. (2006). The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. Immunity *24*, 463-473.

Zhao, S., and Fernald, R.D. (2005). Comprehensive algorithm for quantitative real-time polymerase chain reaction. Journal of Computational Biology *12*, 1047-1064.

APPENDIX

SUPPLEMENTARY MATERIAL

CHAPTER 1

Table S1.1 List of studies investigating ecology of coral microbiota from which 16S rRNA sequences were analyzed in the meta-analysis.

Study	Domain	Method(s)	# of sequences per reef compartment			
			Coral			Seawater
			Healthy	Bleached	Diseased	
Allers et al. 2008	bacteria	16S rRNA surveys				22
Apprill et al. 2009	archaea		13			
	bacteria	16S rRNA surveys	151			
Arotsker et al. 2009	bacteria	16S rRNA surveys	131			
		Cultivation-based survey	36			
Ben-Dov et al. 2009	archaea	Selected 16S rRNA survey of RFLP-selected OTUs*	2			
	bacteria		31			
Bourne et al. 2005	bacteria	16S rRNA surveys	43			
		Selected 16S rRNA survey of DGGE-selected OTUs	27			7
Cervino et al. 2004	bacteria	Cultivation-based survey			4	
Chimetto et al. 2008	bacteria	Cultivation-based survey	95			
Cooney et al. 2002	bacteria	Selected 16S rRNA survey of ARDRA-selected OTUs	23		27	
		Selected 16S rRNA survey of DGGE-selected OTUs			10	
Frias-Lopez et al. 2004	bacteria	Selected 16S rRNA survey of T-RFLP-selected OTUs			8	
Garren et al. 2008	bacteria	16S rRNA surveys	48			89
Garren et al. 2009	bacteria	16S rRNA surveys	602			89
Ionescu et al. 2009	archaea	16S rRNA surveys				37
Jensen et al. 2008	bacteria	16S rRNA surveys				41
Kapely et al. 2007	bacteria	16S rRNA surveys	82			
Kellogg et al. 2004	archaea	16S rRNA surveys	160			
Kellogg et al. 2009	bacteria	16S rRNA surveys	508			
		Selected 16S rRNA survey of DGGE-selected OTUs	10			
Klaus et al. 2007	bacteria	16S rRNA surveys	248			
Kooperman et al. 2007	bacteria	16S rRNA surveys	125			22
Koren et al. 2006	bacteria	16S rRNA surveys	475			
Koren et al. 2008	bacteria	16S rRNA surveys	100	232		

Table S1.1 (Continued)

Lampert et al. 2008	bacteria	16S rRNA surveys	79			
Littman et al. 2009a	bacteria	16S rRNA surveys	329			
Littman et al. 2009b	bacteria	16S rRNA surveys	155			
Pantos et al. 2003	bacteria	Selected 16S rRNA survey of ARDRA-selected OTUs	32		44	
		Selected 16S rRNA survey of DGGE-selected OTUs	16		18	
Pantos et al. 2006	bacteria	Selected 16S rRNA survey of ARDRA-selected OTUs	23		43	
Raina et al. 2009	bacteria	Cultivation-based survey	27			
		16S rRNA surveys	589			81
Reis et al. 2009	bacteria	16S rRNA surveys	232		143	221
Ritchie et al. 2006	bacteria	Cultivation-based survey	34	21		33
Sekar et al. 2006	bacteria	16S rRNA surveys	89		94	
Siboni et al 2008	archaea	16S rRNA surveys	212			69
Sunagawa et al. 2009	bacteria	16S rRNA surveys	317		340	
Sussman et al. 2008	bacteria	Cultivation-based survey			23	
Yakimov et al. 2006	bacteria	Selected 16S rRNA survey of RFLP-selected OTUs	12			
Total # of sequences			**5056**	**253**	**754**	**711**

*Operation taxonomic unit

Table S2.1A Correlation between RNA-Seq and RT-qPCR measurements of differential gene expression in symbiotic relative to aposymbiotic anemones. [a]

Locus #/ transcript #	Top Blast Hit	UniProt accession number	Read count [b]	Fold-change (RNA-Seq)	Fold-change (RT-qPCR)
58798/1	Bovine Na^+- and Cl^--dependent taurine transporter	Q9MZ34	61	∞	29
102514/1	Human Npc2 cholesterol transporter	P61916	269	1197	26
95010/1	Mouse tumor necrosis factor receptor superfamily member 27	Q8BX35	202	240	33
125065/1	*Drosophila* organic-cation (carnitine) transporter	Q9VCA2	255	131	57
77179/1	Human scavenger receptor class B member 1 (SRB1; CD36-related)	Q8WTV0	11	28	3.7
95925/1	*Bacteroides thetaiotaomicron* glutamate dehydrogenase	P94598	852	13	2.9
86800/1	Human facilitated glucose transporter (GLUT8)	Q9NY64	57	12	6.3
65589/1	Sheep aquaporin-5	Q866S3	71	11	2.2
70728/1	*C. elegans* NH_4^+ transporter 1 (AMT1-type)	P54145	1382	6.4	7.0
95114/1	Mouse aromatic-amino-acid transporter 1	Q3U9N9	54	5.9	6.2
101012/1	*Bacillus halodurans* isocitrate lyase	Q9K9H0	79	3.9	4.6
66644/1	Human carnitine O-palmitoyltransferase 1	P50416	1237	2.4	2.8
101000/1	*S. cerevisiae* delta(24(24(1)))-sterol reductase	P25340	40	2.0	∞
105631/1	Rat Na^+- and Cl^-dependent GABA transporter 1	P23978	1302	2.0	1.9
125822/1	*Cerberus rynchops* ficolin (collagen/fibrinogen domain containing lectin) 2	D8VNS9	187	1.7	1.8
27493/1	*Salmo salar* Golgi pH regulator	B5X1G3	61	1.2	1.3
12296/1	60*S* ribosomal protein L11	P46222	280	1.1	0.9
119098/1	Rat 40*S* ribosomal protein s7	Q9ZNS1	94	1.0	1.1
12335/1	*Dictyostelium* F-box/WD repeat-containing protein A-like protein	Q54N86	239	-1.0	-1.3
84201/1	*Metridium senile* cytochrome c oxidase	Q35101	1784	-1.4	-1.4
58671/1	*Coturnix japonica* glyceraldehyde-3-phosphate dehydrogenase	Q05025	237	-1.4	-1.1
77428/1	Superoxide dismutase	P81926	987	-1.6	-1.7
21845/2	Rat apoptosis-inducing factor mitochondrial	Q9JM53	769	-1.6	-1.1
59465/1	Rat calmodulin-like protein 3	Q5U206	679	-1.6	-1.5

Table S2.1A (Continued)

13527/1	Rat monocarboxylate transporter 10	Q91Y77	47	-1.7	-1.8
12461/1	Rat mannan-binding lectin serine protease 1	Q8CHN8	223	-3.1	-2.8
431/2	Human Na^+/glucose cotransporter 4	Q2M3M2	67	-3.2	-2.0
1568/1	Mouse E2F transcription factor 2	P56931	136	-3.5	-2.1
20440/1	Zebrafish delta-like protein c	Q9IAT6	2	$-\infty$	-1.8

[a] Transcripts are arranged (top to bottom) in order of their degree of expression in symbiotic relative to aposymbiotic anemones as determined by RNA-Seq. Only the data from RNA-Seq Experiment 1 are used, because its conditions matched more closely those of the RT-qPCR experiment (see Materials and Methods and Table 2.1).

[b] The baseMean expression value as calculated by DESeq (Anders and Huber, 2010).

Table S2.1B Primer sequences and product sizes for RT-qPCR data. [a]

Locus #/ transcript #	Forward Primer	Reverse Primer	Product Size
58798/1	AAAGATCTGCTGGCTGACCCTGA	AACACCAACCAATTGCCTCACCC	134
102514/1	AAGTGACCCGTGCGTTCTCAAA	TGCGTTTGGGTTGGGAATGTGT	148
95010/1	TTTGACATGCTGCGCGAACTGCT	AATGGCCACGACGTGTTTGAAGG	225
125065/1	TGTCAGTGGCGTTGCACAGTCTT	ACATTGCCAATTCTTGCGCGGT	159
77179/1	GAAATGGCGGAAAAAGCATA	GGTGGAAATTGTGTCCCATC	225
95925/1	CAAAGCCTGGACATCGACGCAAA	CAATGACACAGGCCCGCAGAAA	194
86800/1	AGCTGGAGGGAAGGCACCAATAA	TGGGAGCTGTCAATCAACTTGGGA	110
65589/1	TTTGCCGGGAACACGTGCATT	TGAGCGCCGAGTGATGTAGGA	177
70728/1	ACCAACGGATTCCCATTCTCGTCA	TTTGCGGGCAGCAGTGTTGTT	110
95114/1	TGTCGCGCTGTTGCCTTTGTT	TGGCCAAAGCAAGGCGTTTGTGA	187
101012/1	GGTCAGCACGCATGAAAGCATTGT	AAGCAATCCAGATGGCAAAGGCAG	171
66644/1	TCCAAGACCAAGTGTTGGTGGACT	TGATCCAAGTCAGGGACAGGCAAA	110
101000/1	TCTGTCGTGGACACTGCTGTTGA	ATCCAACCGAACTTCTCCGTGGT	189
105631/1	ACCGTGAACACTTCTTGAGAGCCA	GCCTCGGTTGAATGCTTTGTTCGT	210
125822/1	ACCTCGCGCCTTGTCCTTATCAAA	AATGGGACTGTTAAGGCGGTTCGT	225
27493/1	GGTTTGCTGCATCTTCACAGGTCA	AGAAACAGCTGGCGACTAAGCTCT	133
12296/1	AGCCAAGGTCTTGGAGCAGCTTA	TTGGGCCTCTGACAGTACAGTGAACA	125
119098/1	ACTGCAGTCCACGATGCTATCCTT	GTCTGTTGTGCTTTGTCGAGATGC	125
12335/1	TGAAACCTCCTTTCAGCCTCCCA	TCACTTCACTCATCTCGGCAGCA	172
84201/1	AGCAGTTGGTAAGTCTGCACAA	GTAACCATGGTAGCAGCATGAA	105
58671/1	AACAGCTTTGGCAGCACCTGTAGA	TGCTTTCACAGCAACCCAGAAGAC	114
77428/1	AAGGCAAGCGGTAACGAGGTTT	TGCTTTCCTTCTGTCAGCCCAGT	177
21845/2	TCATGGCAAGGACGACGAGTGAA	TCACCCATGGCAGTAAAGAGCGA	156
59465/1	TCGGCAGGATTGTGTCCAAGTGA	AAACGAGCGACACAACGTCAGCA	197
13527/1	AGACACCCAACTGTTCCTTCCCA	ACACGCCGTAAGTAAACGCCAA	212
12461/1	AGCAAAGGGCACGAACAACCAAC	TTGACTCGCTATGGCCGCTAACA	125
431/2	TGGCCTTCAACAAACCTTCACGCT	ACGTTTGTAGTCCCAGCCAGTCA	238
1568/1	AAGTTCGTTGGAGGGTACTGCGA	CCACCAAAGACTTCACACAGCCA	110
20440/1	AATGGCGGAGTTTGTCAAGACGG	TGCCGATGCATTTGCCTGAGTT	118

[a] Transcripts are listed in the same order as in Table S2.1A.

Table S2.2 Transport-related genes showing differential expression in symbiotic relative to aposymbiotic anemones.[a]

Line	Fold-change [b]	Read count [c]	Locus#/ transcript#	Best BLAST hit	UniProt accession number	BLAST-hit E-value
1	∞	78	58798/1	Bovine Na^+- and Cl^--dependent taurine transporter	Q9MZ34	1.00E-169
2	∞	81	36456/1	Rabbit Na^+/(glucose/*myo*-inositol) transporter 2	Q28728	3.00E-104
3	600	659	102514/1	Human Npc2 cholesterol transporter	P61916	2.00E-14
4	131	437	60777/1	Zebrafish NH_4^+ transporter rh type b	Q7T070	3.00E-98
5	44	150	125065/1	*Drosophila* organic-cation (carnitine) transporter	Q9VCA2	6.00E-35
6	28	11	77179/1	Human scavenger receptor class B member 1 (SRB1; CD36-related)	Q8WTV0	9.00E-65
7	13	70	65589/1	Sheep aquaporin-5	Q866S3	8.00E-37
8	11	52	86800/1	Human facilitated glucose transporter (GLUT8)	Q9NY64	9.00E-89
9	6.9	94	12006/1	*Xenopus* GABA and glycine transporter	Q6PF45	8.00E-60
10	5.9	881	70728/1	*C. elegans* NH_4^+ transporter 1 (AMT1-type)	P54145	6.00E-72
11	5.8	1667	45451/1	*Drosophila* lipid-droplet surface-binding protein 2	Q9VXY7	2.00E-08
12	4.9	45	95114/1	Mouse aromatic-amino-acid transporter 1	Q3U9N9	3.00E-65
13	4.3	198	84722/1	Fish (*Tribolodon*) carbonic anhydrase II	Q8UWA5	2.00E-36
14	4.3	288	2130/2	Pig aquaporin-3	A9Y006	1.00E-68
15	3.7	111	11708/1	Human facilitated glucose transporter (GLUT8)	Q9NY64	1.00E-88
16	3.6	2547	101327/1	Rat neutral- and basic-amino-acid transporter 1	P82252	2.00E-117
17	3.5	52	103419/1	Chicken monocarboxylate transporter 4 (slc16a3)	P57788	1.00E-28
18	3.1	71	37788/1	Rabbit hyperpolarization-activated cation channel 4	Q9TV66	9.00E-129
19	3.1	707	56440/1	Mouse Na^+-independent SO_4^- transporter	Q80ZD3	1.00E-126
20	2.9	241	97639/1	Bovine ABC subfamily f member 2	Q2KJA2	0
21	2.9	264	11677/1	*Arabidopsis* ABC transporter g family member 27	Q9FT51	7.00E-67
22	2.7	35	26261/3	*Dictyostelium* UDP-sugar transporter	Q54YK1	1.00E-32
23	2.6	108	49092/1	Mouse zinc transporter 1 (znt-type)	Q60738	4.00E-70
24	2.4	61	15916/1	Human major-facilitator-superfamily-domain-containing protein 12	Q6NUT3	9.00E-37
25	2.4	991	66644/1	Human carnitine *O*-palmitoyltransferase 1	P50416	0

Table S2.2 (Continued)

26	2.2	35	119860/1	Mouse Na^+/(glucose/*myo*-inositol) cotransporter 2	Q8K0E3	4.00E-125
27	2.1	3879	76979/1	Human Na^+-dependent phosphate-transport protein 2b	O95436	2.00E-112
28	2.1	4804	109479/1	Human neutral- and basic-amino-acid transport protein	Q07837	1.00E-49
29	2.1	554	12947/1	Rat very-low-density-lipoprotein receptor	P98166	4.00E-162
30	2.1	53	37499/1	Human aromatic-amino-acid transporter 1	Q8TF71	6.00E-49
31	2.1	714	14877/3	Zebrafish pyrimidine-nucleotide carrier	Q6DG32	1.00E-69
32	2.1	243	16360/1	*Xenopus* monocarboxylate transporter 12 (slc16a12)	Q6P2X9	3.00E-35
33	2	88	22123/1	Rat chloride channel clic-like protein	Q9WU61	1.00E-16
34	2	61	120787/1	Mouse aromatic amino acid transporter 1	Q3U9N9	7.00E-34
35	2	989	105631/1	Rat Na^+- and Cl^--dependent GABA transporter 1	P23978	9.00E-124
36	2	231	2338/1	Bovine zinc transporter (zip-type)	A5D7L5	1.00E-43
37	2	582	49156/1	Human ABC subfamily b member 1	P08183	3.00E-108
38	1.9	1076	86906/1	Human lipid-transfer protein	Q9NQZ5	1.00E-52
39	1.8	218	120269/1	Mouse Na^+-dependent neutral-amino-acid transporter	O88576	2.00E-117
40	1.8	1272	36717/5	Rat neutral- and basic-amino-acid transporter 1	P82252	2.00E-116
41	1.8	381	19286/1	Rat v-ATPase subunit f	P50408	3.00E-41
42	1.8	211	81279/1	Rat Na^+-dependent phosphate transporter 1	Q9JJP0	2.00E-71
43	1.7	105	82158/1	*Columba livia* carnitine *O*-acetyltransferase	P52826	5.00E-151
44	1.7	176	12043/1	Mouse carnitine *O*-palmitoyltransferase 2	P52825	0
45	1.7	1020	70022/1	Rat neutral- and basic-amino-acid transporter 1	P82252	2.00E-128
46	1.7	193	126133/1	Zebrafish zinc transporter (znt-type)	Q5PQZ3	2.00E-112
47	1.6	1058	33916/1	Human transitional-ER ATPase	P55072	0
48	1.6	479	86475/1	*Xenopus* peptide transporter 4	Q68F72	1.00E-98
49	-1.6	164	22834/1	Chicken monocarboxylate transporter 3	Q90632	2.00E-30
50	-1.6	684	108875/1	Chicken low-density-lipoprotein receptor-related protein 1	P98157	0
51	-1.6	404	28605/1	Rat TRP cation-channel subfamily a, member 1	Q6RI86	1.00E-115

Table S2.2 (Continued)

52	-1.8	1323	33284/1	Mouse aromatic amino acid transporter 1	Q3U9N9	3.00E-34
53	-1.8	5989	71915/1	Rabbit non-specific lipid-transfer protein	O62742	0
54	-1.8	817	16745/1	Rat plasma membrane Ca^{2+}-transporting ATPase	P11505	3.00E-98
55	-2	39	13527/1	Rat monocarboxylate transporter 10 (aromatic amino acid transporter 1)	Q91Y77	6.00E-19
56	-2.1	33	122320/1	Human long-chain fatty acid transport protein 1	Q6PCB7	3.00E-96
57	-2.2	848	43841/1	Chicken ovotransferrin	P02789	3.00E-47
58	-2.3	4512	98994/1 [d]	Human Npc2 cholesterol transporter	P61916	5.00E-09
59	-2.4	5556	44110/1	Human low-density lipoprotein receptor-related protein 4	O75096	8.00E-61
60	-2.8	40	76106/1	*C. elegans* TRP-like cation channel protein 1	P34586	3.00E-55
61	-2.9	190	93152/1	Chimpanzee NH_4^+ transporter rh type c	Q3BCQ7	5.00E-111
62	-2.9	629	104248/1	Rat serotransferrin	P12346	2.00E-47
63	-3	506	76019/1	Carbonic anhydrase	P83299	1.00E-37
64	-3.5	1893	56973/1	Mouse organic cation carnitine transporter 3	Q9WTN6	5.00E-31
65	-3.5	301	432/1	Human Na^+/glucose cotransporter 4	Q2M3M2	2.00E-130
66	-3.8	325	129624/1	*E. coli* high-affinity choline-transport protein	P0ABD0	2.00E-78

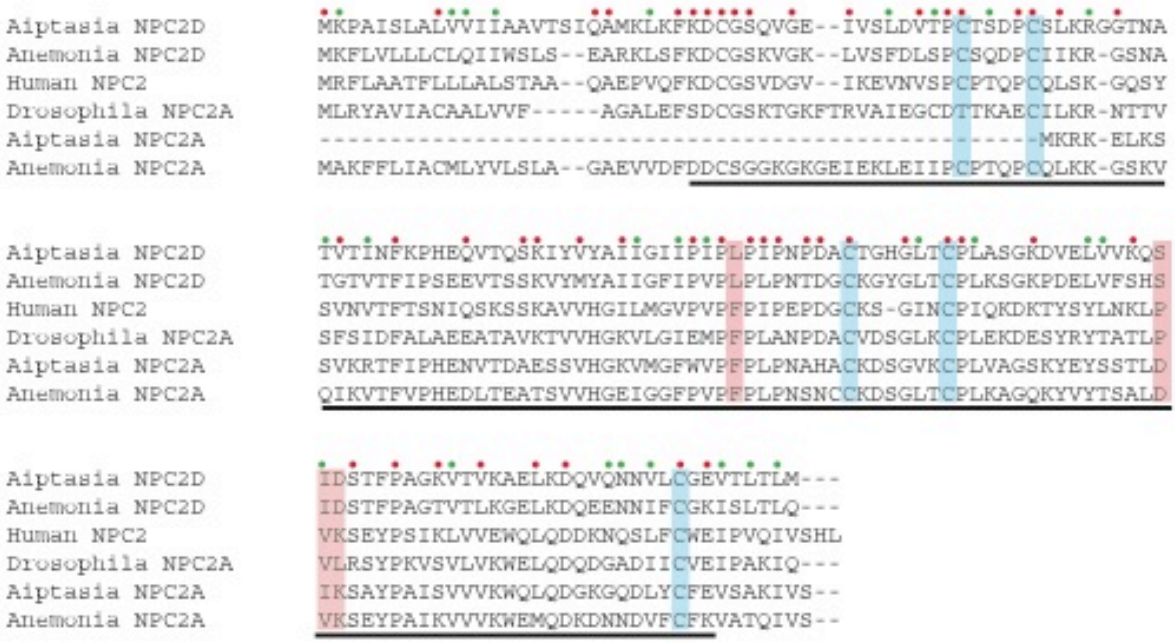
[a] Putative small-molecule transporters and some proteins of related function (see text) are arranged in order of their degree of differential expression in symbiotic anemones relative to aposymbiotic anemones. Positive fold-changes, expression higher in symbiotic anemones; negative fold-changes, expression higher in aposymbiotic anemones.

[b] The arithmetic mean of the values from the two RNA-Seq experiments, except in line 6 (transcript 77179/1). ∞, expression was not detected in aposymbiotic animals. Transcript 77179/1 was not detected in aposymbiotic anemones in Experiment 2, giving a nominal ∞-fold change in expression. However, as the normalized read counts in both experiments were rather low, and the possible involvement of the 77179/1-encoded protein in lipid metabolism makes it likely to have been affected in its expression by the starvation conditions used in Experiment 2, we report in line 6 the more conservative value from Experiment 1 alone.

[c] Except for line 6, the average of the baseMean expression values (as calculated by DESeq (Anders and Huber, 2010) for Experiment 1 and Experiment 2. As explained in footnote b, for transcript 77179/1 (line 6), we show the value for Experiment 1 alone.

[d] Appears to represent a truncated version of transcript 98999/1, whose predicted protein product was used for the phylogenetic analysis of Figure 2.2A.

A


Aiptasia NPC2D MKPAISLALVVIIAAVTSIQAMKLKFKDCGSQVGE--IVSLDVTPCTSDPCSLKRGGTNA
Anemonia NPC2D MKFLVLLLCLQIIWSLS--EARKLSFKDCGSKVGK--LVSFDLSPCSQDPCIIKR-GSNA
Human NPC2 MRFLAATFLLLALSTAA--QAEPVQFKDCGSVDGV--IKEVNVSPCPTQPCQLSK-GQSY
Drosophila NPC2A MLRYAVIACAALVVF-----AGALEFSDCGSKTGKFTRVAIEGCDTTKAECILKR-NTTV
Aiptasia NPC2A ----------MKRK-ELKS
Anemonia NPC2A MAKFFLIACMLYVLSLA--GAEVVDFDDCSGGKGKGEIEKLEIIPCPTQPCQLKK-GSKV
Aiptasia NPC2D TVTINFKPHEQVTQSKIYVYAIIGIIPIPLPIPNPDACTGHGLTCPLASGKDVELVVKQS
Anemonia NPC2D TGTVTFIPSEEVTSSKVYMYAIIGFIPVPLPLPNTDGCKGYGLTCPLKSGKPDELVFSHS
Human NPC2 SVNVTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPDGCKS-GINCPIQKDKTYSYLNKLP
Drosophila NPC2A SFSIDFALAEEATAVKTVVHGKVLGIEMPFPLANPDACVDSGLKCPLEKDESYRYTATLP
Aiptasia NPC2A SVKRTFIPHENVTDAESSVHGKVMGFWVPFPLPNAHACKDSGVKCPLVAGSKYEYSSTLD
Anemonia NPC2A QIKVTFVPHEDLTEATSVVHGEIGGFPVPFPLPNSNCCKDSGLTCPLKAGQKYVYTSALD
Aiptasia NPC2D IDSTFPAGKVTVKAELKDQVQNNVLCGEVTLTLM---
Anemonia NPC2D IDSTFPAGTVTLKGELKDQEENNIFCGKISLTLQ---
Human NPC2 VKSEYPSIKLVVEWQLQDDKNQSLFCWEIPVQIVSHL
Drosophila NPC2A VLRSYPKVSVLVKWELQDQDGADIICVEIPAKIQ---
Aiptasia NPC2A IKSAYPAISVVVKWQLQDGKGQDLYCFEVSAKIVS--
Anemonia NPC2A VKSEYPAIKVVVKWEMQDKDNNDVFCFKVATQIVS--

B

```
A_digitifera_NPC2E          -KNCT--KNDDVTVESLDIN---PCSE-EP---CIFHK-GSTVSVTVAF-TPLEEVKSGE
A_digitifera_NPC2F          -KNCA--SRKYALPLKVAIN---PCTK-QP---CTLHP-GKKASIAVVV-KPLVTIRRGT
O_carmela_NPC2a             -SNCTSNPGPSTLGKTVNVTAVPPCDT-AP---CVVHQ-GESLNVTVTF-VPNVAIENFT
A_digitifera_NPC2D          -QTC---DKPSGRLNSVDVT---PCNG-NP---CVFKR-GTNETITVTF-TPNEVVSKGK
Aiptasia_NPC2E              -KDC---GSKGATIVRLDIS---PCEE-EP---CNFKT-GTTVTGTLTF-VAKEYFTSGR
N_vectensis_NPC2B           -RDC---GSQEGEIVGMDIS---PCDS-EP---CVLKR-GTSVDGSLTF-IPHEDLKRAK
Aiptasia_NPC2B              VVVL---VVVVGIVVVVDVD---QCTSDDP---CSLKR-GTNVTSTATM-IPLEEVTQAT
Aiptasia_NPC2C              -TDC---GSYLGEIHSLEVN---PCTS-DP---CVLKR-GDNMTSVISF-TPHEQVSAAK
Aiptasia_NPC2D              -KDC---GSQVGEIVSLDVT---PCTS-DP---CSLKRGGTNATVTINF-KPHEQVTQSK
N_vectensis_NPC2C           -QDC---GSKKGELISVDLT---PCSS-DP---CVIKR-GANASGVITF-IPHEVVTSSK
A_viridis_NPC2D             -KDC---GSKVGKLVSFDLS---PCSQ-DP---CIIKR-GSNATGTVTF-IPSEEVTSSK
M_faveolata_NPC2B           -ANCSDVTALEGKLISVDLT---PCPS-QP---CVFHK-GTNVTATIKF-SPEEMVTDGT
D_melanogaster_NPC2A        -SDC---GSKTGKFTRVAIE---GCDT-TK-AECILKR-NTTVSFSIDF-ALAEEATAVK
M_faveolata_NPC2A           -ADC---GSL-AKINFVDVS---PCVM-EP---CELKK-GTNESIEIQF-IPNSNITEGK
A_digitifera_NPC2B          -RDC---GNKELSPAQVIIT---PCPA-EP---CQLKK-GVNESVEVIF-KPTEVVTSAK
A_digitifera_NPC2A          -SYI---GSKESSISQVIVT---PCPA-EP---CQLKK-GVNESIEVIF-KPGEVVTSSK
Human_NPC2a                 -KDC---GSVDGVIKEVNVS---PCPT-QP---CQLSK-GQSYSVNVTF-TSNIQSKSSK
Mouse_NPC2A                 -KDC---GSKVGVIKEVNVS---PCPT-DP---CQLHK-GQSYSVNITF-TSGTQSQNST
Aiptasia_NPC2A              ------------------------------------K-ELKSSVKRTF-IPHENVTDAE
N_vectensis_NPC2A           -KDCSG-GKGEGEIVELDIS---PCPT-QP---CTLHK-GTTVSVNITF-VPHVTLDSGK
A_viridis_NPC2A             -DDCSG-GKGKGEIEKLEII---PCPT-QP---CQLKK-GSKVQIKVTF-VPHEDLTEAT
H_magnipapillata_NPC2D      -QNC---GHLDSNTI-VSIT---PCEK-EP---CTLVR-GSNATLEIQF-KAKHFSKQLK
H_magnipapillata_NPC2B      -KPC----DMSSTVGDVAIS---PCDK-QP---CAFQR-GGSANIEISF-TAAKDADKLT
H_magnipapillata_NPC2A      -KKCTS-PASSAVIGDVIIT---PCDS-LP---CSFKR-GGSGNIKINF-QATKNNSELT
H_magnipapillata_NPC2C      -KKCSS-PASSAVVGDVVIS---PCDN-QP---CQFIR-GGNANIQIHF-QAKKDNSNIT

A_digitifera_NPC2E          LSVDAI-AFGHRLP-M--VRKE--NICEG--HGVT-----CPLEKGKKQTFTINQKVERY
A_digitifera_NPC2F          LELYGIHWLGIKFP-LS-VPNP--DICHG--YGTR-----CPMIANSRVVLSISQTLPSF
O_carmela_NPC2A             VVVHAS-VGIIHVP-YP-VTDP--NGCDTAVTGVT-----CPLKANVAVEWHHSFSVPSI
A_digitifera_NPC2D          ILLYAK-LVLGWIE-LS-LRNP--NICEG--YGLK-----CPLAKGVREELSVTERVPQV
Aiptasia_NPC2E              VKAYAV-IEGVDLP-LP-IPT---DACQG--YGLT-----CPINNGQTANFVIKQEIQAD
N_vectensis_NPC2B           LSAHAI-IDKLPLP-LP-IPS---DACQG--YGLS-----CPVDSGVKSMFKIHQAIESE
Aiptasia_NPC2B              IYMHAT-VSGITIP-ID-IPNP--NACSG--HGLS-----CPLKSGETVELSMVLEVEAK
Aiptasia_NPC2C              IDINAI-IAGSPIH-VH-IPNP--NACDG--HGLK-----CPLEKGKKVELVVSQVIRRS
Aiptasia_NPC2D              IYVYAI-IGIIPIP-LP-IPNP--DACTG--HGLT-----CPLASGKDVELVVKQSIDST
N_vectensis_NPC2C           VLAYAI-PGLIPVP-LP-LPNS--DGCKG--YGLT-----CPLKSGKQVELVFEHYIDQT
A_viridis_NPC2D             VYMYAI-IGFIPVP-LP-LPNT--DGCKG--YGLT-----CPLKSGKPDELVFSHSIDST
M_faveolata_NPC2B           LQVYGF-IEGIKTP-FP-LEQP--DACKE--HGLE-----CPLKSGVTYSLEITLAIKPA
D_melanogaster_2A           TVVHGK-VLGIEMP-FP-LANP--DACVD--SGLK-----CPLEKDESYRYTATLPVLRS
M_faveolata_NPC2A           TVVYGI-IEGVQVP-FP-VDNP--EVCKE--HGIT-----CPMPAEKTQTFKATLPVKSE
A_digitifera_NPC2B          VVIHGI-IEGVRFP-FP-FPHP--NGCKE--HGLE-----CPLKPNKEYTFKATLPVKRT
A_digitifera_NPC2A          VVVHGI-IAGVPVP-FP-ISQP--NGCED--HGLD-----CPLQPNKEYTFKATLPVKSA
Human_NPC2a                 AVVHGI-LMGVPVP-FP-IPEP--DGCKS---GIN-----CPIQKDKTYSYLNKLPVKSE
Mouse_NPC2A                 ALVHGI-LEGIRVP-FP-IPEP--DGCKS---GIN-----CPIQKDKVYSYLNKLPVKNE
Aiptasia_NPC2A              SSVHGK-VMGFWVP-FP-LPNA--HACKD--SGVK-----CPLVAGSKYEYSSTLDIKSA
N_vectensis_NPC2A           AIVHGV-IAGIPVP-FP-LPNA--DVCKN--SGLK-----CPLEPGTKYVYQSSLEVKTM
A_viridis_NPC2A             SVVHGE-IGGFPVP-FP-LPNS--NCCKD--SGLT-----CPLKAGQKYVYTSALDVKSE
H_magnipapillata_NPC2D      TKVYGK--LLFWVPYYN-FGKE--DSCLD--NGIT-----CPVIEDEEYSYSQSLHISKL
H_magnipapillata_NPC2B      TVVKGK-IGPIWVP-FP-LSQP--DACNN--EGLT-----CPIKSSQKYTYQYSLPISES
H_magnipapillata_NPC2A      SVVKGK-IGPLWVP-FP-LSQP--DACQN--EGIT-----CPIKDGQSYLFSYDLPISTT
H_magnipapillata_NPC2C      TIVKGK-IGPLWVP-FP-LSQP--DGCLN--DGII-----CPVKTDQQYVYSYDLPISKS

A_digitifera_NPC2E          YPPLPI-DVEAYVENDNRK----ILC
A_digitifera_NPC2F          VPMGSY-QLQAVMKDQLGR--M-VLC
O_carmela_NPC2a             APKGPVEIITWELQAPSKE--D-VAC
A_digitifera_NPC2D          LPSSTR-EVKAKLVDQNGG--T-VVC
Aiptasia_NPC2E              FPKVKL-QLKGEVMDPQGN--M-LFC
N_vectensis_NPC2B           FPVGNL-TLKAAVTDSDTS--QVVFC
Aiptasia_NPC2B              FPRGKV-ILKTELKDQAKN--D-IFC
Aiptasia_NPC2C              APPGRY-RIRTELKEQYGI--D-VFC
Aiptasia_NPC2D              FPAGKV-TVKAELKDQVQN--N-VLC
N_vectensis_NPC2C           FPTGHL-TLKAELKDQDSD--V-VIC
A_viridis_NPC2D             FPAGTV-TLKGELKDQEEN--N-IFC
M_faveolata_NPC2B           YPSIQL-VAQMDFKLPDDG--Y-LFC
D_melanogaster_2A           YPKVSV-LVKWELQDQDGA--D-IIC
M_faveolata_NPC2A           YPALQL-DVKWELHDQDAK--V-VYC
A_digitifera_NPC2B          YQDVCM-I---RLL---------CSC
A_digitifera_NPC2A          YPDIKL-VVKWQLLDQNAN--S-VFC
Human_NPC2a                 YPSIKL-VVEWQLQDDKNQ--S-LFC
Mouse_NPC2A                 YPSIKL-VVEWKLEDDKKN--N-LFC
Aiptasia_NPC2A              YPAISV-VVKWQLQDGKGQ--D-LYC
N_vectensis_NPC2A           YPSLKL-VVRWEIQDNKNK--D-VLC
A_viridis_NPC2A             YPAIKV-VVKWEMQDKDNN--D-VFC
H_magnipapillata_NPC2D      NPKISI-PVKWLIQNEAEK--D-LVC
H_magnipapillata_NPC2B      YPKINL-PVSWELKDEKGE--S-LVC
H_magnipapillata_NPC2A      YPAISL-VVSWEIQDENGN--D-VVC
H_magnipapillata_NPC2C      YPAISV-VVSWELQDENGN--D-LVC
```

Figure S2.1 Alignments of Npc2 sequences from *Aiptasia* and other organisms. (A) Full-length alignments of selected Npc2-like proteins from *Aiptasia* sp. (this study), *A. viridis* human, and *D. melanogaster*. Red and green dots, amino acids that are identical (red) or similar (green: I,V,L; S,T; D,E; K,R; Q,N) between *Aiptasia* NpcD and human Npc2; red shading, amino acids whose mutation to alanine ablates the cholesterol-binding function of Npc2 in mammalian cells [33]; blue shading, conserved cysteines used to identify conserved regions of the proteins for phylogenetic analysis; dark underline, the conserved region used for phylogenetic analysis. (B) The multiple-sequence alignment of the conserved regions used to produce the phylogenetic tree in Figure 2.2.

Table S2.3 Lipid-metabolism genes showing differential expression in symbiotic relative to aposymbiotic anemones. [a]

Line	Metabolic Process	Putative protein function (best BLAST hit)	UniProt accession number	BLAST -hit E-value	Locus #/ transcript #	Fold-change [b]
1	FA synthesis	ACC1: Acetyl-CoA carboxylase (chicken)	P11029	0	26166/1	3.9
2	FA synthesis	ELOVL4: Elongation-of-very-long-chain-fatty-acid protein 4 (mouse)	Q9EQC4	1e-23	4012/1	4.2
3	FA synthesis	Δ^5 fatty-acid desaturase (*Mortierella alpina*)	O74212	6e-39	120701/1	6.2
4	FA synthesis	Δ^6 fatty acid desaturase (human)	O95684	6e-46	92492/1	3.5
5	Lipid storage	DHAPAT: dihydroxyacetone phosphate acyltransferase (human)	O15228	1e-95	8091/1	1.4
6	Lipid storage	2-acylglycerol *O*-acyltransferase 2-a (*Xenopus*)	Q2KHS5	8e-75	10118/2	2.7
7	Lipid storage	2-acylglycerol *O*-acyltransferase 2-a (*Xenopus*)	Q2KHS5	4e-89	15365/1	-2.0
8	Lipid storage	2-acylglycerol *O*-acyltransferase 2-b (*Xenopus*)	Q5M7F4	9e-81	78512/1	2.4
9	Lipid storage	Diacyglycerol *O*-acyltransferase (*Mycobacterium tuberculosis*)	O06795	1e-19	76581/1	2.1
10	Lipid storage	AGPAT 1: 1-acyl-*sn*-glycerol-3-phosphate acyltransferase alpha (human)	Q99943	2e-23	67491/1	-1.4
11	Lipid storage regulation	Lipid-droplet surface-binding protein 2 (*Drosophila*)	Q9VX7	2e-08	45451/1	5.8
12	Lipase	HSL: Hormone-sensitive lipase (human)	Q05469	2e-84	13988/1	1.8
13	Lipase	ATGL: Adipose triglyceride lipase (mouse)	Q8BJ56	9e-62	16411/1	-3.3
14	FA transport	FATP1: Fatty-acid transport protein 1 (human)	Q6PCB7	3e-96	122320/1	-1.9
15	FA transport	FATP4: Long-fatty-acid transport protein 4 (orangutan)	Q5RDY4	8e-92	122313/8	-3.1
16	FA transport	SRB1: Scavenger receptor class B member 1 (human; CD36-related protein)	Q8WTV0	9e-65	77179/1	28
17	FA-CoA-ligase	ACSL4: Long-chain-fatty-acid ligase 4 (human)	O60488	0	89704/1	5.7
18	FA-CoA-ligase	ACSL5: Long-chain-fatty-acid ligase 5 (human)	Q9ULC5	5e-116	106694/1	2.9
19	FA β-oxidation	Organic cation (carnitine) transporter (*Drosophila*)	Q9VCA2	6e-35	125065/1	44
20	FA β-oxidation	CPT1: Carnitine *O*-palmitoyltransferase 1 (human)	P50416	0	66644/1	2.4
21	FA β-oxidation	CPT2: Carnitine *O*-palmitoyltransferase 2 (mouse)	P52825	0	12043/1	1.6
22	FA β-oxidation	CACT: Carnitine acylcarnitine carrier protein (bovine)	Q08DK7	2e-60	13918/1	n.s.
23	FA β-oxidation	VLCAD: Very-long-chain-specific acyl-CoA dehydrogenase (bovine)	P48818	0	108541/1	1.6

Table S2.3 (Continued)

24	FA β-oxidation	MTP: Trifunctional enzyme (pig)	Q29554	0	33057/3	1.4
25	FA β-oxidation	MCAD: Medium-chain-specific acyl-CoA dehydrogenase (bovine)	Q3SZB4	5e-165	92556/1	n.s.
26	FA β-oxidation	SCAD: Short-branched-chain-specific acyl-CoA dehydrogenase (*Dictyostelium*)	Q54RR5	2e-121	127382/1	n.s.
27	FA β-oxidation	Crotonase: Short-chain enoyl-CoA hydratase (*Dictyostelium*)	Q1ZXF1	4e-67	117452/1	n.s.
28	FA β-oxidation	M/SCHAD: Medium and short-chain l-3-hydroxyacyl-CoA dehydrogenase (human)	Q16836	5e-29	34949/1	n.s.
29	FA β-oxidation	DCI: Enoyl-Δ isomerase (human)	P42126	7e-33	55782/1	n.s.
30	FA β-oxidation	MCKAT: 3-ketoacyl-CoA thiolase (rat)	P13437	5e-145	56206/1	2.4
31	Glyoxylate cycle	Isocitrate lyase (*Bacillus halodurans*)	Q9K9H0	3e-164	101012/1	3.9
32	Glyoxylate cycle	Malate synthase (*Myxococcus xanthus*)	P95329	2e-141	22622/1	n.s.

[a] Genes encoding proteins putatively involved in lipid metabolism are arranged in groups by biological process (see Figure 2.3). FA, fatty acid.

[b] Because of the likelihood that the starvation conditions in Experiment 2 would affect lipid metabolism, the fold-change values from Experiment 1 are shown. Positive fold-changes, expression higher in symbiotic anemones; negative fold-changes, expression higher in aposymbiotic anemones. n.s., no significant differential expression observed.

Table S2.4 Presence or absence in the *Aiptasia* transcriptome of genes encoding the enzymes involved in the synthesis of particular amino acids. [a]

Line	Amino acid	Enzyme	UniProt accession number	*Aiptasia* locus #/ transcript #
1	Gln	Glutamine synthetase	P32288	104234/1
2	Glu	Glutamate synthase	Q12680	60857/1
3	Glu/Pro	NADP-specific glutamate dehydrogenase [b]	Q9C8I0	3911/1
4	Glu/Pro	NADP-specific glutamate dehydrogenase 2 [b]	P39708	95925/1
5	Glu/Pro	NAD-specific glutamate dehydrogenase [c]	P33327	99746/1
6	Glu/Pro	Glutamate dehydrogenase 2 [c]	Q38946	7229/1
7	Met	MTHFR: Methylenetetrahydrofolate reductase	Q9WU20	15095/1
8	Met	MS: Methionine synthase (cobalamin-dependent)	Q99707	50131/1
9	Met	Methionine synthase (cobalamin-independent) [d]	P05694	55393/1
10	Met	BHMT: Betaine-homocysteine *S*-methyltransferase 1	Q93088	45257/1
11	Cys/SAM [e]	MAT: Methionine adenosyltransferase 1	Q91X83	140402/1
12	Cys	SAHH: *S*-adenosyl-L-homocysteine hydrolase	P27604	91092/1
13	Cys	CBS: Cystathionine β-synthase	P32582	98284/1
14	Cys	CGL: Cystathionine γ-lyase	P31373	7792/1
15	Met/Cys/Thr/Ile/Lys	Aspartokinase/homoserine dehydrogenase	Q9SA18	
16	Met/Cys/Thr/Ile/Lys	Aspartokinase	P10869	
17	Met/Cys/Thr/Ile	Homoserine dehydrogenase	P31116	
18	Met/Cys	HAT: Homoserine *O*-acetyltransferase	P08465	70690/1
19	Met/Cys	CGL: Cystathionine γ-synthase	P47164	974/1
20	Met	Cystathionine β-lyase	P43623	
21	Met	Homocysteine *S*-methyltransferase 3	Q8LAX0	4396/6
22	Ser	D-3-phosphoglycerate dehydrogenase 1	P40054	294/1
23	Ser	Phosphoserine aminotransferase	P33330	57256/1
24	Ser	Phosphoserine phosphatase	P42941	122485/1
25	Ser	Catabolic L-serine/threonine dehydratase	P25379	21690/1
26	Ser/Gly	Serine hydroxymethyltransferase, mitochondrial	P37292	11787/1
27	Ser/Gly	Serine hydroxymethyltransferase, cytosolic	P37291	11787/1
28	Gly	Alanine-glyoxylate aminotransferase 1	P43567	47533/1
29	Gly	Serine-glyoxylate aminotransferase	Q56YA5	47531/1
30	Gly	Low specificity L-threonine aldolase	P37303	109186/1
31	Asp/Glu/Asn	Aspartate aminotransferase, mitochondrial	Q01802	23248/1
32	Asp/Glu/Asn	Aspartate aminotransferase, cytoplasmic	P46646	111366/1
33	Asn	Asparagine synthetase	P49089	51175/1
34	Ala	Alanine aminotransferase 1	P52893	89107/1
35	Pro	γ-glutamyl phosphate reductase	P54885	128220/1
36	Pro	Pyrroline-5-carboxylate reductase	P32263	115939/1
37	Arg	Caramoyl-phosphate synthetase	P31327	21357/1
38	Arg	Ornithine carbamoyltransferase	P00480	116500/1

Table S2.4 (Continued)

39	Arg	Argininosuccinate synthetase	P22768	53174/1
40	Arg	Argininosuccinate lyase	P04076	29236/1
41	Arg	Arginase-1	P05089	118787/1
42	Arg	*N*-acetylglutamate synthase	Q8N159	19094/1
43	Arg	Acetylglutamate kinase	Q01217	
44	Arg	Ornithine acetyltransferase	Q04728	
45	Phe/Tyr	Aromatic/aminoadipate aminotransferase 1	P53090	109224/1
46	Tyr	Tyrosine aminotransferase	Q9LVY1	58220/1
47	Tyr	Phenylalanine 4-hydroxylase	P00439	37855/1
48	Phe/Trp	Class-II DAHP synthetase-like protein	Q9SK84	
49	Phe/Trp	Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited	P32449	
50	Phe/Trp	Pentafunctional AROM polypeptide	P08566	
51	Phe/Trp	Chorismate mutase	P32178	
52	Phe/Trp	Chorismate synthase	P28777	
53	Phe/Trp	Anthranilate synthase component 1	P00899	
54	Phe/Trp	Anthranilate phosphoribosyltransferase	P07285	
55	Trp	Tryptophan synthase	Q42529	
56	His	ATP phosphoribosyltransferase	P00498	
57	His	Imidazole glycerol phosphate synthase hisHF	P33734	
58	His	Histidinol-phosphate aminotransferase	P07172	
59	His	Histidine biosynthesis trifunctional protein	P00815	
60	His	Histidinol dehydrogenase	Q9C5U8	
61	Val/Leu/Ile	Acetolactate synthase	P07342	8385/1
62	Val/Leu/Ile	Ketol-acid reductoisomerase, mitochondrial	P06168	
63	Val/Leu/Ile	Dihydroxy-acid dehydratase, mitochondrial	P39522	127954/1
64	Val/Leu/Ile	Branched-chain-amino-acid aminotransferase, cytosolic	P47176	85088/1
65	Leu	2-isopropylmalate synthase	P06208	
66	Leu	3-isopropylmalate dehydratase	P07264	
67	Leu	3-isopropylmalate dehydrogenase	P04173	
68	Ile	Threonine dehydratase, mitochondrial	Q9ZSS6	57366/1
69	Lys	Homocitrate synthase, mitochondrial	Q12122	
70	Lys	Kynurenine/α-aminoadipate aminotransferase,mitochondrial	Q8N5Z0	
71	Lys	Homoaconitase, mitochondrial	P49367	
72	Lys	Homoisocitrate dehydrogenase, mitochondrial	P40495	
73	Lys	L-aminoadipate-semialdehyde dehydrogenase	P07702	127184/1
74	Lys	Saccharopine dehydrogenase [NADP(+), L-glutamate-forming]	P38999	1580/1
75	Lys	Saccharopine dehydrogenase [NAD(+), L-lysine-forming]	P38998	76032/1
76	Lys	4-hydroxy-tetrahydrodipicolinate synthase 2, chloroplastic	Q9FVC8	
77	Lys	Dihydrodipicolinate synthase	Q0WSN6	
78	Lys	Diaminopimelate decarboxylase 1	Q949X7	37096/1? [f]
79	Thr	Threonine synthase	P16120	10016/1

[a] The UniProt Accession Number shown is for the seed sequence used to identify the *Aiptasia* transcript. For an *Aiptasia* transcript to be listed, its best reciprocal BLAST hit (to the same species as the seed sequence) had to be the seed sequence itself or to a sequence encoding a paralogous protein. Where no transcript is listed, no *Aiptasia* homologue of the seed sequence could be identified with confidence.

[b] Downregulated (transcript 3911/1) and upregulated (95925/1) in symbiotic relative to aposymbiotic anemones (see Figure 2.4). Both proteins had a *Bacterioides thetaiotaomicron* NAD(P)-utilizing glutamate dehydrogenase (UniProt P94598) as their top BLAST hit.

[c] No significant differential expression in symbiotic relative to aposymbiotic anemones.

[d] In contrast to transcript 50131/1 (MS in Figure 2.5), transcript 55393/1 showed no differential expression in symbiotic *vs.* aposymbiotic anemones.

[e] *S*-adenosyl-methionine (see Figure 2.5).

[f] Although this *Aiptasia* transcript met the formal criterion for inclusion (footnote a), the number of genomic reads mapping to it barely exceeded our cut-off for calling a sequence cnidarian (see Table 2.2), so that it may represent a contaminant rather than an *Aiptasia* gene encoding a homologue of this typically bacterial and plant enzyme.

A

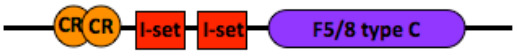
```
Apr2       TKTPESPMSSSKDPGKPYSCSHTKFFIVTFLSHTLLLCAGIVPLYITTNHRLVTVEDRLT 60
Apr1       ------------------------------------------------------------

Apr2       VHDMELNSCCLVKEGYGNFQDHEETKVTLNREKYVQTQFIDRVRRNTPYISRDVLNSIRM 120
Apr1       ------------------------------------------------------------

Apr2       EVRNHILNLTASQFCQVPDKICTRGAPGKRGLRGRRGSRGRRGRPGHKGIKGLPGKYGKQ 180
Apr1       --------------------RQRGPPGPPGLPGKSGPRGSIGPQGPK------------ 27
                          Box I ──►**.**  ** *: *.**  *  * *

Apr2       GLRGFPGQKGQKGDIGNRGPPGLPGPKGERGKEVTEPSVFISPSILTVTENQTATFHCNA 240
Apr1       ------GLAGKKGDIGRPGLPCKPLIIN------YPPKVSLPVGPIYVKEGDNLILSCHV 75
                 *  *:*****. * *  * Box 2 ──►*.* :. . : *.*.:.  : *:.

Apr2       HGYPKPQITWKMGSGKQIDFGKTRIDKSAGLLEISNVSEKDTGNYTCSAKSVLGEDANTV 300
Apr1       TGYPKPKVTWSKVMDS---LPSKRSFITTNRLKVLSVQKQDSGLYVCAGSNTLGSAVETI 132
            *****:::**.   ..   : ..*   ::. *:: .*.::*:* *.*:....**. .:*:

Apr2       SLLVKFPPRFTEVQKPFQTILQGSTVNLKCAALGYPPPIITWTKMLGSLPTKRSQQNGGK 360
Apr1       KVIVVSAPKFISTPPQQVNKNTCEKLTLDCQARGDPPAVITWSKEKGRLPTDRTQLINGR 192
           .:::*  .*::* ..     .    ..:.*.* * * **.:***:*  * ***.*:*  .*:

Apr2       LTITRFQQSDSGSYQCEAVNSVGKNIFYTTLSFGACDDALGMQSKAIRDSQITASSSYSS 420
Apr1       LTITGMTTSDAGKYTCTAVSAGVATSKSVTRVTVKETGKLKFSRDSVTDYIVVKR----- 247
           **** :  **:*.* * **.:   .   .*      . * :. .:: *  :.
                                                  Box3 ◄──

Apr2       AYLPYYGRLNIVLGYGGWLAKSNTKGQWIQVDLLQATRITAIATQGTSKYDEWTTSYSLQ 480
Apr1       -KLPAMARLTVCL----WMMTSKKNSVLISYAVPGSINEILLLDVG-KRLSVWLGDVSWD 301
             **  .**.: *    *: .*:.:.  *.  :  : .   :   * .: . *  . * :

Apr2       YSYDGTSFRDYEGGKTLPGNSDRSTVVKNNLDPAIAARYIRLLPKTYHSYMVIRMELYGC 540
Apr1       SGVHVTDGQWHHICATWDNSAGQTILYKDGVRRAPSSTRSR------------------- 342
            . . *. : :.   *  ..:.:: : *:.:  * ::   *

Apr2       QL 542
Apr1       --
```

B

Human peroxidasin 1

Human peroxidasin 2

***Aiptasia* peroxidasin-related 2**



***Aiptasia* peroxidasin-related 1**

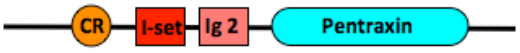


Figure S2.2 Distinct but related genes whose products may be involved in host tolerance of the symbiont. The transcripts differentially expressed between symbiotic and aposymbiotic anemones included two whose top blastx hit in SwissProt was a human peroxidasin and three whose top blastx hit was a mammalian plasma kallikrein (Figure 2.6; Table S2.5). (A,B) The two *Aiptasia* peroxidasin-related proteins (Apr1 and Apr2) appear to represent distinct gene products with limited domain homology both to each other and to human peroxidasins. (A) ClustalW sequence alignment of the two *Aiptasia* proteins shows interspersed identical and different amino acids as expected from distinct gene products rather than from alternative splice products or misassembled contigs. Boxes show regions of sequence similarity between the *Aiptasia* proteins but not the human ones (Box 1) or among all four proteins (Boxes 2 and 3), as diagrammed in B. Stars denote amino acid residues that are identical, two dots denotes conserved substitutions, and one dot denotes semi-conserved substitutions. (B) Schematic diagram comparing protein domains found in human peroxidasins and the *Aiptasia* peroxidasin-related proteins using Pfam. LRR, leucine-rich repeat; I-set, immunoglobulin I-set; Peroxidase, domain with similarity to canonical peroxidises; VWC, von Willebrand factor type-C; CR, collagen triple-helix repeat; Ig 2, immunoglobulin; Pentraxin, domain with similarity to pentraxin pattern-recognition receptors displaying Ca^{2+}-dependent ligand binding; F5/8 type C domain (or discoidin domain), with cell-adhesion functions. (C) ClustalW sequence alignment of three *Aiptasia* plasma-kallikrein homologs shows interspersed identical and different amino acids as expected from distinct gene products. The box shows regions of sequence similarity between *Aiptasia* proteins representative of the shared trypsin domain. Symbol representations are the same as described in A.

Table S2.5 Genes potentially involved in host tolerance of the symbiont that are differentially expressed between symbiotic and aposymbiotic anemones. [a]

Protein (from top BLAST hit)	UniProt accession number	Locus #/ transcript #	BLAST-hit E-value	Fold-Change [b]
A. Response to oxidative stress				
Catalase	P04040	100968/1	0	-4.7
ADAM (disintegrin and metalloproteinase domain-containing protein) 9	Q13443	123296/1	2e-62	-3
Transient receptor potential cation channel (subfamily M, member 2)	Q91YD4	125627/1	7e-16	-2.9
Peroxidasin-related protein 1	Q92626	99631/1	4e-06	-2.5
Dual oxidase 2	Q8HZK2	17080/2	5e-27	-2
Allene oxide synthase-lipoxygenase	O16025	9291/1	4e-38	-1.8
Soluble guanylate cyclase 88E	Q8INF0	7254/1	1e-171	1.9
Peroxidasin-related protein 2	A1KZ92	57146/1	7e-25	2.9
B. Inflammation/tissue remodeling/response to wounding				
Transmembrane serine protease 6 [c]	Q9DBI0	81296/1	3e-50	-6.1
Plasma kallikrein [d]	P14272	12789/1	7e-47	-4
Mannan-binding lectin serine peptidase 1 (MASP-1) [e]	Q8CHN8	36375/1	5e-12	-3
Plasminogen [f]	P00747	21286/1	3e-47	-3
Plasma kallikrein [d]	P03952	84752/1	1e-48	-2.9
Ephrin type-a receptor 3 [g]	P29320	6695/1	2e-63	-2.8
Phospholipase A2 (isoform 4) [h]	Q6T179	3740/5	1e-26	-2.4
Arachidonate 5-lipoxygenase [i]	P48999	55879/1	5e-28	-1.7
Plasma kallikrein [d]	P26262	73922/1	8e-40	2.1
Ficolin 2 [j]	Q15485	62279/1	1e-42	2.2
Vanin-I [k]	Q58CQ9	48344/1	8e-118	2.6
Discoidin, CUB, and LCCL domain containing 2 [l]	Q91ZV2	80843/1	3e-12	2.7
Hepatocyte nuclear factor 4 (alpha) [m]	P22449	34830/4	1e-113	3.6
Adenosine A2b receptor [n]	O13076	66307/1	2e-17	4.2
Scavenger receptor class B member 1 [o]	Q8WTV0	77179/1	9e-65	28
C. Apoptosis/cell death				
Transcription factor E2F2	P56931	1568/1	5e-07	-4.2
Receptor-binding cancer antigen expressed on SiSo cells	Q865S0	42283/1	4e-06	-2.7
Tumor protein p73	Q9JJP2	88973/1	2e-35	-2.4
Paired box protein Pax-3	P23760	46973/1	3e-43	-1.8
Apoptosis-inducing factor 1 (mitochondrial)	Q9JM53	21845/2	1e-171	-1.8

Table S2.5 (Continued)

TNF (Tumor Necrosis Factor) receptor-associated factor 3	Q13114	30586/1	2e-74	1.8
TNF superfamily member 12	O43508	18277/1	1e-07	1.9
Kruppel-like factor 11	O14901	58173/1	8e-51	2.6
G1 to S phase transition 1	P15170	25564/1	0	2.8
Growth arrest and DNA damage-inducible protein (GADD45 gamma)	Q9Z111	55453/1	6e-09	5.1
Ribonucleoside-diphosphate reductase (small chain C)	Q9LSD0	18748/1	1e-132	12
Organic cation transporter	Q9VCA2	88336/1	6e-35	44
TNF receptor superfamily member 27	Q8BX35	94982/1	9e-10	60

[a] The set of all transcripts displaying differential expression by RNA-Seq was analyzed to identify biological processes (based on GO terms) that were overrepresented in this set relative to the background transcriptome (see Materials and Methods). The sets of processes identified here (A, B, and C) emerged from this analysis and may be involved in host tolerance of the symbiont.

[b] In all but one case, the arithmetic mean of the values from the two RNA-Seq experiments is shown. For transcript 77179/1 (last line of section B), the value from Experiment 1 is shown for reasons explained in Supplementary Table 2, footnote b. Positive fold-changes, expression higher in symbiotic anemones; negative fold-changes, expression higher in aposymbiotic anemones.

[c] Hydrolyzes a range of proteins including type I collagen, fibronectin, and fibrinogen and may play a role in matrix-remodeling processes (Hooper et al., 2003).

[d] Serine proteases activated by tissue injury or microbial invasion; they activate the release of potent pro-inflammatory cytokines that ultimately result in the release of effector molecules such as nitric oxide and tumor necrosis factor-α and can stimulate the complement innate-immunity system (Lalmanach et al., 2010; Moreau et al., 2005).

[e] Plays a role as an amplifier of the complement cascade, potentially via the activation of MASP-2 (Takahashi et al, 2013).

[f] The zymogen of plasmin; it can be activated via plasma kallikrein and functions in the breakdown of fibrin in fibrinolysis, the activation of proteases, and the modulation of cell adhesion (Li et al., 2003).

[g] A receptor tyrosine kinase that binds membrane-bound ephrin family ligands residing on adjacent cells and regulates cell-cell adhesion, cytoskeletal organization, and cell migration (Smith et al., 2004).

[h] Releases arachidonic acid from cellular membrane phospholipids, leading to its conversion to pro-inflammatory prostaglandins via arachidonate 5-lipoxygenase (Moreau et al., 2005).

[i] See note h.

[j] A lectin whose binding to microbial surface glycans can initiate activation of the complement pathway [8]; it also appears to bind to cell-surface glycans of *Symbiodinium* (Logan et al., 2010).

[k] Hydrolyzes pantetheine to pantothenic acid and cysteamine, the latter of which can lead to acute and chronic epithelial inflammation (Martin et al., 2004).

[l] Thought to play a role in cell adhesion and wound healing (Kobuke et al., 2001).

[m] A transcriptional regulator that is decreased in inflammatory bowel disease and protects against chemically-induced colitis in mice (Darsigny et al., 2001).

[n] Its activation appears to result in inhibition of pro-inflammatory cytokine production, and mice deficient in A2b receptors are more susceptible to intestinal inflammation (Gessi et al, 2011).

[o] See text.

Supplementary Materials and Methods for Chapter 2

Identification and Optimization of qPCR Standards for *Aiptasia*

Six housekeeping genes were selected as potential qPCR standards based on their prior use in coral studies. Gene names used here are those assigned to the *Aiptasia pallida* genes and differ in most cases from those used in the other organisms. The genes encoding ribosomal protein L11 (*RPL11*), NADH-dehydrogenase subunit 5 (*NDH5*), and glyceraldehyde-3-phosphate-dehydrogenase (*GPD1*) were reported to be stable in *Porites astreoides* during heat stress, settlement induction, and metamorphosis (Kenkel et al*.,* 2011). The genes encoding ribosomal protein S7 (*RPS7*) and adenosylhomocysteinase (*AHC1*) were used as standards during studies of thermal stress in *Acropora aspera* (Leggat et al*.,* 2011). The β-actin gene (*ACT1*) was used to explore modulation of host-gene expression (Rodriguez-Lanetty et al., 2006) and was used as the standard for early qPCR studies in our lab.

Primers were developed and tested for these six potential standard genes. The aposymbiotic *A. pallida* transcriptome (Lehnert et al., 2012) was searched using tblastx with sequences from *Porites lobata* for *NDH5*, *P. astreoides* for *RPL11*, *Urticina eques* for *GPD1*, *Acropora millepora* for *RPS7*, and *Nematostella vectensis* for *AHC1*. The loci identified in the *A. pallida* transcriptome were searched using blastx in NCBI and all top hits were indeed the genes of interest. The identified loci were then translated using ORFPredictor and the longest ORFs were used to identify conserved sequences by performing protein alignments in

MacVector with sequences available from NCBI. Conserved sequences were then used to develop primers using PrimerQuest from Integrated DNA Technologies (IDT).

Primers were tested on *A. pallida* cDNA and gDNA. Primers that spanned an exon-intron junction were preferentially identified for further use (Table S2.6). PCR products were cloned into a TA cloning vector and electroporation-competent *E. coli* cells were transformed with the plasmids. Transformed cells were plated on Ampicillin/X-Gal plates and white/light-blue colonies were selected for colony PCR using M13 forward and reverse primers. PCR products were sequenced, and the sequences were aligned with the expected sequences from the transcriptome. All primer pairs accurately selected the sequences of interest.

Table S2.6 Primer sequences used for potential qPCR standards

Gene	Primer sequences
RPL11	F: AGCCAAGGTCTTGGAGCAGCTTA R: TTGGGCCTCTGACAGTACAGTGAACA
RPS7	F: ACTGCAGTCCACGATGCTATCCTT R: GTCTGTTGTGCTTTGTCGAGATGC
NDH5	F: AGCAGTTGGTAAGTCTGCACAA R: GTAACCATGGTAGCAGCATGAA
GPD1	F: AACAGCTTTGGCAGCACCTGTAGA R: TGCTTTCACAGCAACCCAGAAGAC
AHC1	F: CCATTACAGCAACAACACAGGCCA R: GCATCAAACGTTGGCAGATGAAGC
ACT1	F: ACACCGTCTTGTCAGGAGGTTCAA R: TCCACATCTGTTGGAAGGTGGACA

The six genes were then tested for their expression levels across 11 experimental conditions (Table S2.7). RNA was extracted from 3-4 medium-sized anemones from each condition using a Trizol/RNeasy hybrid protocol (details available upon request). RNA integrity was checked both by using a Nanodrop and by running samples on a 2% agarose gel. For all RNA samples used, 260/280 readings were >1.9, and two clear rRNA bands were visible. For

each condition, 300 ng of RNA was reverse transcribed using the Maxima® First Strand cDNA-synthesis kit for RT-qPCR (Fermentas). 17 µL of RT product was then diluted with 23 µL of H_2O. 2 µL of this cDNA solution was then used for the qPCR reaction. Each qPCR well had 2 µL of cDNA, 2 µL of H_2O, 5 µL of *Power* SYBR® Green PCR Master Mix (Applied Biosystems), and 1 µL of a primer mix containing 1.5 µM forward (F) primer and 1.5 µM reverse (R) primer.

The primer efficiency of each primer pair was tested across a dilution series of 1:1, 1:10, 1:100, 1:1000, and 1:10000 cDNA; the calculated efficiencies were 95-105%. Possible gDNA contamination in RNA samples was tested by running RNA-only controls; these samples showed no amplification. Standard qPCR settings were used, and an additional dissociation stage was added to test for the presence of multiple products. The dissociation stage showed only one clear peak in every case.

Table S2.7 Experimental conditions used to test gene-expression levels by qPCR

Conditions [a]	CC7 Sym [b]	CC7 Apo [c]
Room Temperature (27ºC)	x	x
1 h heat shock (35ºC)	x	x
1.5 h heat shock (37ºC)	x	x
1 h cold shock (8ºC)	x	x
1 h incubation with 500 µg/mL dsRNA[d] (27ºC)	x	x
Kept in the dark for 1 month (27ºC)	x [e]	not done

[a] Except for the sample incubated in the dark, all anemones were incubated on a 12L:12D cycle with 25 µmol photons m^{-2} s^{-1} from Cool White fluorescent bulbs, and the manipulations indicated were performed during the light period.

[b] Symbiotic anemones (containing the endogenous population of Clade A *Symbiodinium*) of the CC7 clonal line of *Aiptasia* (Sunagawa et al., 2009).

[c] Aposymbiotic CC7 animals that had been cured of their endogenous *Symbiodinium* by a combination of cold shock, DCMU treatment, and extended growth in the dark (Lehnert et al., 2012). All anemones were screened for absence of dinoflagellates prior to use in these experiments.

[d] dsRNA (477 bp) synthesized for *A. pallida* nematogalectin gene knockdown.

[e] Represents a partially aposymbiotic condition.

Ct values for each of the six genes under each of the 11 conditions were analyzed using geNorm (Vandesompele et al., 2002) to determine the relative expression stabilities of the prospective standard genes; the M-values are inversely proportional to the stabilities of the genes (Table S2.8). *ACT1* (M = 0.625) and *AHC1* (M = 0.775) were considerably less stable in expression than the four genes shown in the table.

Statistical analysis of the qPCR results also indicates that *ACT1* should not be used as an expression standard in the study of symbiosis in *Aiptasia* due to the large expression difference between aposymbiotic and symbiotic animals: there was a significant (p = 0.002) up-regulation in *ACT1* expression in aposymbiotic (or mostly aposymbiotic) anemones compared to symbiotic anemones across all conditions. This was determined by normalizing qPCR Ct values with the two most stable standard genes (*RPL11* and *RPS7*) and performing a Mann-Whitney statistical test on *ACT1* expression levels in aposymbiotic and symbiotic anemones.

Table S2.8 Assessment of gene-expression stability under various conditions [a]

Gene	Protein encoded	geNorm M	Product Sequence	Product Length	Primer Efficiency
RPL11	Component of the 60S ribosomal subunit	0.357	AGCCAAGGTCTTGGAGCAGCT TACAGGCCAACAGCCTGTGTT TTCAAAAG **(INTRON – 236 bp)** CTCGCTACACTGTGAGATCTTT TGGAATCAGAAGGAACGAGA AGATCTCTGTTCACTGTACTGT CAGAGGCCCAA	cDNA 125 bp gDNA 361 bp	98%
RPS7	Component of the 40S ribosomal subunit	0.380	ACTGCAGTCCACGATGCTATC CTTGAAGATCTTGTCTTTCCTA GTGAAATTGTTGGCAAAAGGA TAAGAGTTAAACTTGATGGTT CACGTCTCGTTAAAGTG **(INTRON – 411 bp)** CATCTCGA CAAAGCACAACAGAC	cDNA 125 bp gDNA 536 bp	97%
NDH5	NADH-dehydrogenase subunit 5	0.423	AGCAGTTGGTAAGTCTGCACA ATTAGGCTTACACACTTGGTT ACCGGATGCAATGGAAGGT **(INTRON – 1729 bp)** CCAACTC CGGTGTCTGCCTTGATTCATGC	cDNA 105 bp gDNA 1834 bp	95%

			TGCTACCATGGTTAC		
GPD1	Glyceraldehyde-3-phosphate-dehydrogenase	0.530	AACAGCTTTGGCAGCACCTGT AGAGGCTGGGATGATATTCTG ATTGGCACCTCTACCATCACG CCATTTCT **(INTRON – 567 bp)** TCCCACTAGGTCCATCTACAG TCTTCTGGGTTGCTGTGAAAG CA	cDNA 114 bp gDNA 681 bp	95%

[a] Tested across the 11 experimental conditions described in Table S2.7.

Table S3.1 Transcripts used to validate RNA-Seq data via RT-qPCR and the fold change data from RNASeq and RT-qPCR datasets. Fold change is expressed in pathogen-exposed anemones relative to controls.

Locus #/contig #	Top SwissProt BLAST Hit	Primer Sequences (F/R)	Product length	Avg. efficiency	Fold-Change RNASeq	Fold-Change RT-qPCR
32435/1	Cytochrome P450 3A24	GGCCACACTCTGTCTATCT CTTCCTTCTGGCTGGTTATG	117	92	22.6	22.7
21472/1	Probable E3 ubiquitin-protein ligase ARI9	GCATTTGAGGGTGATTTCTTG AGACTGGGATGAGGAAGATAG	141	93	9.7	9.0
122298/5	Histidine ammonia-lyase	GGCTTTGGTGAGGTATTAAGG CGAGGGTGTGTAAGTCAATTAG	131	95	5.5	4.9
31324/1	TNF receptor-associated factor 3	TGAAGCAATGTTCATCGAGC ACAGCCCTCCATTTTACGTG	272	90	4.9	5.4
22981/1	Tyrosine kinase receptor Cad96Ca	CAGAGAAAGACGCTTACGATAC GAGACGGACAAGAAGGAAATAG	99	95	4.3	3.1
21357/1	Carbamoyl-phosphate synthase	GGTGAGAAACAGGCAGATAC CACCCATAGACAGCAGTAATC	106	95	4.3	4.3
84775/1	*60S ribosomal protein L10*	ACGTTTCTGCCGTGGTGTCCC CGGGCAGCTTCAAGGGCTTCAC	175	90	1.2	1.1
58671/1	Glyceraldehyde-3-phosphate dehydrogenase	AACAGCTTTGGCAGCACCTGTAGA TGCTTTCACAGCAACCCAGAAGAC	114	89	1.1	1.1
82926/1	*40S ribosomal protein S7*	ACTGCAGTCCACGATGCTATCCTT GTCTGTTGTGCTTTGTCGAGATGC	125	91	1.0	-1.0
99523/1	*Polyadenylate-binding protein 1*	GTGCAAGGAGGCGGACAGCG TGGGCTGATTGCGGGTTGCC	150	88	-1.1	-1.0
12311/1	60S ribosomal Protein L11	AGCCAAGGTCTTGGAGCAGCTTA TTGGGCCTCTGACAGTACAGTGAACA	125	89	-1.2	-1.0
84201/1	*Cytochrome c oxidase*	AGCAGTTGGTAAGTCTGCACAA	105	93	-1.4	-1.0

GTAACCATGGTAGCAGCATGAA

Table S3.1 (Continued)

127561/1	Transcription factor SOX-14	GACATATCAGAGCTTCCATTCA CGAGCGTGAGTCAGTATTAAG	122	95	-2.0	-2.3
76790/1	E3 ubiquitin-protein ligase DZIP3	AGAAGGGCATCGGTTAGA ACTGATGGAGAAGGAGTAGAG	110	90	-2.5	-2.6
101692/1	Calponin-1	CCCTGGATGATGGTGTTATTT GACTTCTGACATGCGTCTATAA	142	90	-2.6	-2.6
15861/1	Protein NLRC5	ATTGGCAGGTGTTGGTAAG AGGGTAGATCAGTGGAGTAAG	120	94	-2.6	-2.5
60730/1	Sushi domain-containing protein 2	CGAATGGGTGTGGAGTAAAG GTAGGAGGGCATGCTATAAAC	140	92	-2.7	-1.1
118088/1	Forkhead box protein D1	CATGTTCGA AAA CGG GAGTT AAATACGGCGGTAGAGCAGA	213	93	-3.6	-3.8

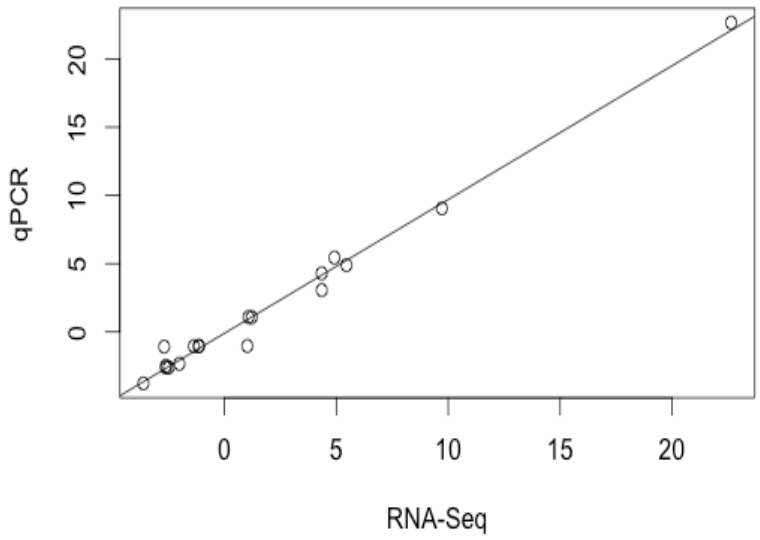


Figure S3.1 Validation of the RNA-Seq results with RT-qPCR. Fold change expression data of 18 genes obtained via RNA-Seq plotted against the corresponding expression data obtained with RT-qPCR (see also Table S3.1). A Pearson correlation coefficient of 0.99 ($p<2.22e-16$) demonstrates the high level of consistency between the two data sets.

Table S3.2 Significantly enriched clusters via DAVID functional clustering analysis. P-values are from a modified Fisher's Exact Test to identify overrepresented functions in differentially expressed genes as compared to the background transcriptome. The enrichment score is the –log geometric mean of the p-values in each cluster.

GO ID	GO term	Count	%	p-value
Annotation Cluster 1	***Enrichment Score: 3.67***			
GO:0045184	establishment of protein localization	119	6.54	2.04E-05
GO:0008104	protein localization	130	7.15	2.18E-05
GO:0015031	protein transport	116	6.38	3.67E-05
GO:0034613	cellular protein localization	61	3.35	1.72E-04
GO:0070727	cellular macromolecule localization	61	3.35	3.39E-04
GO:0006886	intracellular protein transport	54	2.97	7.57E-04
GO:0046907	intracellular transport	82	4.51	2.53E-02
Annotation Cluster 2	***Enrichment Score: 3.66***			
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	37	2.03	4.24E-05
GO:0007167	enzyme linked receptor protein signaling pathway	44	2.42	7.48E-05
GO:0007166	cell surface receptor linked signal transduction	104	5.72	3.38E-03
Annotation Cluster 3	***Enrichment Score: 4.64***			
GO:0006278	RNA-dependent DNA replication	13	0.71	9.15E-06
GO:0015074	DNA integration	10	0.55	5.82E-05
Annotation Cluster 4	***Enrichment Score: 4.40***			
GO:0016265	death	77	4.23	2.91E-07
GO:0008219	cell death	76	4.18	4.69E-07
GO:0012501	programmed cell death	69	3.79	7.45E-07
GO:0006915	apoptosis	64	3.52	2.88E-06
GO:0042981	regulation of apoptosis	66	3.63	6.42E-04
GO:0043067	regulation of programmed cell death	66	3.63	8.21E-04
GO:0010941	regulation of cell death	66	3.63	9.65E-04
GO:0043065	positive regulation of apoptosis	30	1.65	4.40E-02
GO:0006917	induction of apoptosis	22	1.21	5.18E-02
GO:0012502	induction of programmed cell death	22	1.21	5.63E-02
GO:0043068	positive regulation of programmed cell death	30	1.65	5.78E-02

Table S3.2 (Continued)

Annotation Cluster 5	***Enrichment Score: 2.45***			
GO:0045449	regulation of transcription	174	9.57	5.45E-04
GO:0006350	transcription	142	7.81	1.42E-03
GO:0006355	regulation of transcription, DNA-dependent	101	5.55	1.20E-02
GO:0051252	regulation of RNA metabolic process	106	5.83	1.68E-02
Annotation Cluster 6	***Enrichment Score: 2.30***			
GO:0008203	cholesterol metabolic process	15	0.82	6.78E-04
GO:0008202	steroid metabolic process	23	1.26	1.17E-02
GO:0016125	sterol metabolic process	16	0.88	1.57E-02
Annotation Cluster 7	***Enrichment Score: 2.42***			
GO:0010627	regulation of protein kinase cascade	26	1.43	3.28E-04
GO:0010740	positive regulation of protein kinase cascade	17	0.93	1.88E-03
GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB cascade	11	0.60	1.29E-02
GO:0043122	regulation of I-kappaB kinase/NF-kappaB cascade	11	0.60	2.73E-02
Annotation Cluster 8	***Enrichment Score: 1.95***			
GO:0006446	regulation of translational initiation	11	0.60	1.59E-03
GO:0010608	posttranscriptional regulation of gene expression	33	1.81	1.74E-02
GO:0032268	regulation of cellular protein metabolic process	45	2.47	2.05E-02
GO:0006417	regulation of translation	24	1.32	2.68E-02
Annotation Cluster 9	***Enrichment Score: 1.9***			
GO:0016192	vesicle-mediated transport	77	4.23	7.55E-03
GO:0016044	membrane organization	50	2.75	7.68E-03
GO:0006897	endocytosis	31	1.70	1.81E-02
GO:0010324	membrane invagination	31	1.70	1.97E-02
Annotation Cluster 10	***Enrichment Score: 1.88***			
GO:0032869	cellular response to insulin stimulus	13	0.71	1.63E-04
GO:0032868	response to insulin stimulus	14	0.77	2.25E-03
GO:0043434	response to peptide hormone stimulus	17	0.93	4.83E-03
GO:0008286	insulin receptor signaling pathway	8	0.44	8.32E-03
GO:0010033	response to organic substance	57	3.13	4.08E-02

Table S3.2 (Continued)

GO:0032870	cellular response to hormone stimulus	16	0.88	4.29E-02
GO:0009725	response to hormone stimulus	30	1.65	1.24E-01
GO:0009719	response to endogenous stimulus	30	1.65	2.69E-01
Annotation Cluster 11	***Enrichment Score: 1.78***			
GO:0051090	regulation of transcription factor activity	13	0.71	2.39E-04
GO:0051101	regulation of DNA binding	14	0.77	5.92E-04
GO:0051098	regulation of binding	16	0.88	1.51E-03
GO:0043392	negative regulation of DNA binding	8	0.44	2.70E-03
GO:0043433	negative regulation of transcription factor activity	7	0.38	6.21E-03
GO:0051100	negative regulation of binding	8	0.44	1.14E-02
GO:0051091	positive regulation of transcription factor activity	5	0.27	1.03E-01
GO:0051099	positive regulation of binding	6	0.33	1.15E-01
GO:0043388	positive regulation of DNA binding	5	0.27	1.79E-01
GO:0051092	positive regulation of NF-kappaB transcription factor activity	3	0.16	3.39E-01
GO:0051052	regulation of DNA metabolic process	3	0.16	9.97E-01
Annotation Cluster 12	***Enrichment Score: 1.75***			
GO:0046649	lymphocyte activation	19	1.04	2.59E-03
GO:0002521	leukocyte differentiation	15	0.82	5.26E-03
GO:0045321	leukocyte activation	20	1.10	5.44E-03
GO:0030099	myeloid cell differentiation	12	0.66	6.73E-03
GO:0001775	cell activation	21	1.15	7.64E-03
GO:0042110	T cell activation	13	0.71	8.23E-03
GO:0030097	hemopoiesis	26	1.43	1.47E-02
GO:0048534	hemopoietic or lymphoid organ development	26	1.43	2.59E-02
GO:0030098	lymphocyte differentiation	11	0.60	3.23E-02
GO:0002520	immune system development	27	1.48	4.49E-02
GO:0042113	B cell activation	9	0.49	5.27E-02
GO:0030183	B cell differentiation	6	0.33	9.61E-02
GO:0030217	T cell differentiation	6	0.33	2.30E-01
Annotation Cluster 13	***Enrichment Score: 1.67***			

Table S3.2 (Continued)

GO:0043038	amino acid activation	17	0.93	4.05E-03
GO:0043039	tRNA aminoacylation	17	0.93	4.05E-03
GO:0006418	tRNA aminoacylation for protein translation	17	0.93	4.05E-03
GO:0006399	tRNA metabolic process	22	1.21	8.86E-02
GO:0034660	ncRNA metabolic process	30	1.65	7.33E-01
Annotation Cluster 14	***Enrichment Score: 1.57***			
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	22	1.21	4.64E-03
GO:0010605	negative regulation of macromolecule metabolic process	59	3.24	5.15E-03
GO:0010629	negative regulation of gene expression	42	2.31	8.74E-03
GO:0031327	negative regulation of cellular biosynthetic process	42	2.31	1.28E-02
GO:0010558	negative regulation of macromolecule biosynthetic process	41	2.25	1.32E-02
GO:0009890	negative regulation of biosynthetic process	42	2.31	1.72E-02
GO:0006357	regulation of transcription from RNA polymerase II promoter	47	2.58	2.95E-02
GO:0045892	negative regulation of transcription, DNA-dependent	28	1.54	4.36E-02
GO:0016481	negative regulation of transcription	32	1.76	6.11E-02
GO:0051253	negative regulation of RNA metabolic process	28	1.54	7.56E-02
GO:0045934	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	33	1.81	1.97E-01
GO:0051172	negative regulation of nitrogen compound metabolic process	33	1.81	2.21E-01
Annotation Cluster 15	***Enrichment Score: 1.50***			
GO:0006044	N-acetylglucosamine metabolic process	5	0.27	1.38E-02
GO:0006041	glucosamine metabolic process	5	0.27	1.38E-02
GO:0006040	amino sugar metabolic process	5	0.27	2.22E-02
GO:0009225	nucleotide-sugar metabolic process	5	0.27	6.31E-02
GO:0006047	UDP-N-acetylglucosamine metabolic process	3	0.16	1.33E-01
Annotation Cluster 16	***Enrichment Score: 1.40***			
GO:0007040	lysosome organization	9	0.49	1.18E-03
GO:0007033	vacuole organization	10	0.55	1.29E-02
GO:0050885	neuromuscular process controlling balance	5	0.27	2.38E-01
GO:0050905	neuromuscular process	5	0.27	6.77E-01

Table S3.2 (Continued)

Annotation Cluster 17	***Enrichment Score: 1.39***			
GO:0051338	regulation of transferase activity	26	1.43	9.67E-03
GO:0044093	positive regulation of molecular function	38	2.09	1.32E-02
GO:0043549	regulation of kinase activity	25	1.37	1.47E-02
GO:0042325	regulation of phosphorylation	31	1.70	2.73E-02
GO:0045859	regulation of protein kinase activity	23	1.26	2.98E-02
GO:0043085	positive regulation of catalytic activity	33	1.81	3.46E-02
GO:0019220	regulation of phosphate metabolic process	31	1.70	3.97E-02
GO:0051174	regulation of phosphorus metabolic process	31	1.70	3.97E-02
GO:0033674	positive regulation of kinase activity	16	0.88	7.14E-02
GO:0051347	positive regulation of transferase activity	16	0.88	7.83E-02
GO:0045860	positive regulation of protein kinase activity	14	0.77	1.51E-01
GO:0032147	activation of protein kinase activity	8	0.44	3.41E-01
Annotation Cluster 18	***Enrichment Score: 1.34***			
GO:0006906	vesicle fusion	7	0.38	1.31E-02
GO:0048284	organelle fusion	8	0.44	3.16E-02
GO:0016050	vesicle organization	11	0.60	5.07E-02
GO:0006944	membrane fusion	10	0.55	7.41E-02
GO:0001845	phagolysosome formation	3	0.16	1.33E-01
Annotation Cluster 19	***Enrichment Score: 1.31***			
GO:0030099	myeloid cell differentiation	12	0.66	6.73E-03
GO:0048872	homeostasis of number of cells	10	0.55	2.92E-02
GO:0030218	erythrocyte differentiation	6	0.33	7.91E-02
GO:0034101	erythrocyte homeostasis	6	0.33	9.61E-02
GO:0048821	erythrocyte development	3	0.16	1.83E-01

Table S3.3 Structure/grouping of the 14 transcripts that were differentially expressed with top blastx hits to mammalian TRAF proteins.

Locus #/Contig #	Top Uniprot blastx hit	MATH Domain	unique MATH Sequence	Tree Grouping
46907/1	TNF receptor-associated factor 2	x	same as 46912/1 and 46914/1	TRAF1/2/3
46888/1	TNF receptor-associated factor 2	x	x	TRAF1/2/3
86236/1	TNF receptor-associated factor 2	x	x	TRAF1/2/3
67642/1	TNF receptor-associated factor 3	x	same as 89050/1	TRAF1/2/3
89049/1	TNF receptor-associated factor 3	x	x	TRAF1/2/3
89050/1	TNF receptor-associated factor 3	x	same as 67642/1	TRAF1/2/3
78485/1	TNF receptor-associated factor 3		x	TRAF1/2/3
46912/1	TNF receptor-associated factor 3	x	same as 46907/1 and 46914/1	TRAF1/2/3
46914/1	TNF receptor-associated factor 3	x	same as 46907/1 and 46912/1	TRAF1/2/3
114598/1	TNF receptor-associated factor 3	x	x	TRAF1/2/3
31324/1	TNF receptor-associated factor 3	x	x	TRAF1/2/3
30592/1	TNF receptor-associated factor 4	x	x	unique lower metazoan
32870/1	TNF receptor-associated factor 5		x	TRAF1/2/3
109431/1	TNF receptor-associated factor 6	x	x	TRAF6

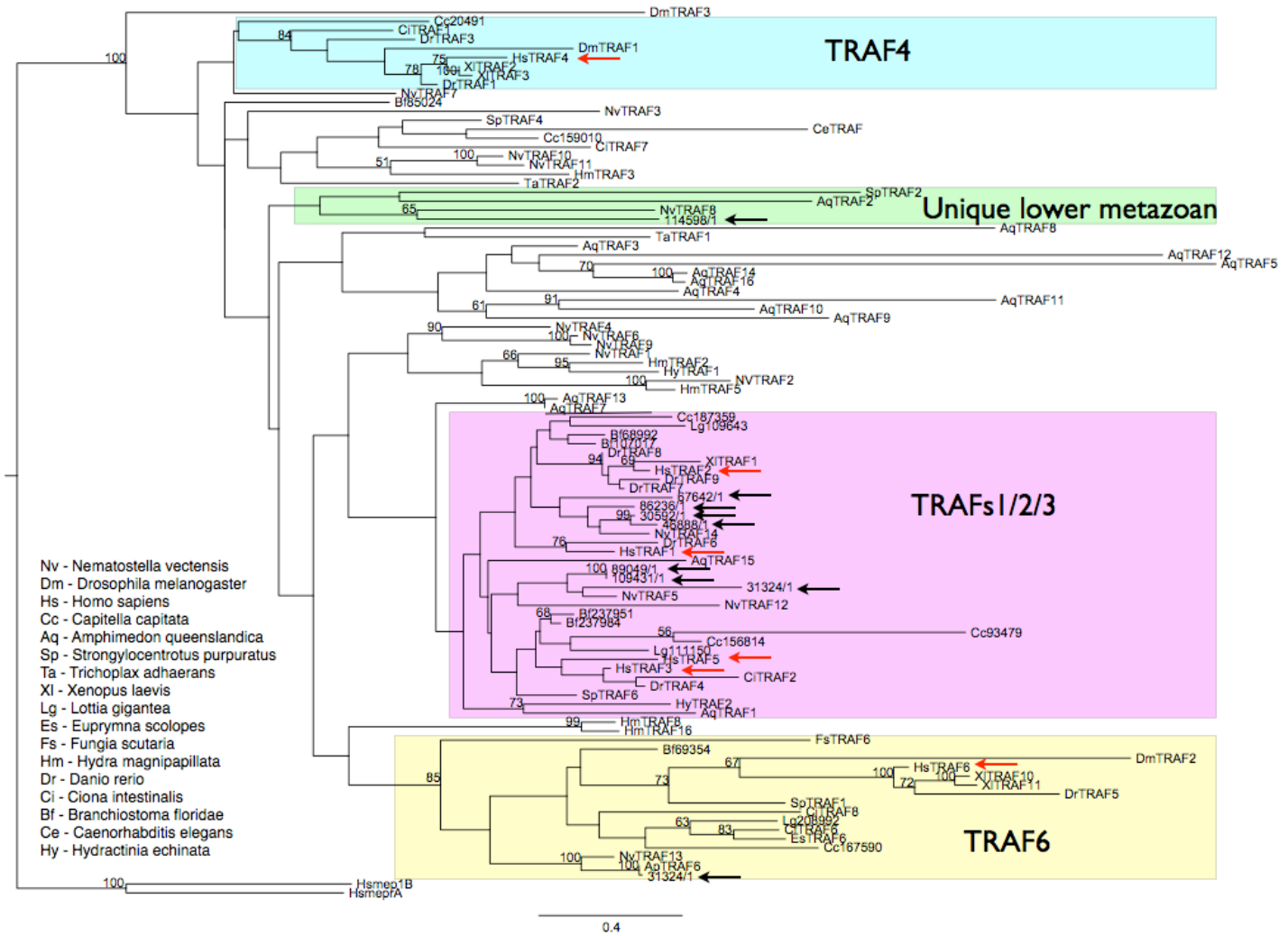


Figure S3.2. Phylogenetic tree of MATH domains of TRAFs from both vertebrate and invertebrate taxa. Boxes are labeled according TRAF nomenclature in humans. *Aiptasia* sequences are highlighted with black arrows and human sequences are highlighted with red arrows. Branches with bootstrap values greater than 50 are labeled.

Table S3.4 Differentially expressed transcripts with functions in apoptosis based on GO Biological Process terms. Fold change is expressed in pathogen-exposed anemones relative to controls.

Fold-change	p-value*	Top blastx hit to SwissProt	Accession	e-value
4.9	4.73E-50	TNF receptor-associated factor 3	Q60803	5.00E-75
4.64	1.11E-19	TNF receptor-associated factor 5	P70191	6.00E-19
4.32	6.34E-112	NADPH-cytochrome P450 reductase	P37040	0
4.19	2.19E-20	Cation transport regulator-like protein 1	Q5SPB6	6.00E-42
4.16	2.19E-19	TNF receptor-associated factor 3	Q60803	1.00E-12
3.18	2.60E-59	Death ligand signal enhancer	P60924	1.00E-29
3.15	6.03E-41	TNF receptor-associated factor 2	P39429	2.00E-88
3.13	4.50E-36	Growth arrest and DNA damage-inducible protein GADD45 gamma	Q9Z111	6.00E-09
2.77	3.81E-38	SAM pointed domain-containing Ets transcription factor	Q9WTP3	7.00E-16
2.62	1.76E-41	DnaJ homolog subfamily A member 3, mitochondrial	Q99M87	5.00E-98
2.6	1.99E-23	Programmed cell death protein 2	Q16342	1.00E-75
2.58	2.59E-14	Protein BTG1	Q63073	2.00E-18
2.34	1.22E-08	TNF receptor-associated factor 3	Q13114	6.00E-45
2.25	1.45E-17	TNF receptor-associated factor 3	Q60803	7.00E-85
2.12	4.54E-34	Eukaryotic translation initiation factor 2-alpha kinase 3	Q9Z2B5	1.00E-134
2.11	6.38E-05	P2X purinoceptor 7	Q99572	6.00E-12
2.07	2.20E-03	TNF receptor-associated factor 2	P39429	1.00E-86
2	5.58E-12	Sphingosine kinase 2	Q9JIA7	2.00E-69
1.99	2.49E-02	Tumor necrosis factor receptor superfamily member 27	Q8BX35	1.00E-09
1.96	3.19E-24	WD repeat-containing protein 92	Q96MX6	1.00E-155
1.96	4.75E-13	Immediate early response 3-interacting protein 1	P85007	2.00E-15
1.95	4.18E-25	Myocyte-specific enhancer factor 2C	A4UTP7	5.00E-64
1.94	2.37E-21	Growth hormone-inducible transmembrane protein	Q5XIA8	2.00E-54
1.88	1.28E-02	TNF receptor-associated factor 4	Q9BUZ4	2.00E-28
1.82	1.77E-17	E3 ubiquitin-protein ligase Topors	Q80Z37	9.00E-51
1.79	1.43E-14	Cullin-5	Q93034	0

Table S3.4 (Continued)

1.73	7.40E-05	Paired box protein Pax-3	P23760	3.00E-43
1.73	5.61E-10	Caspase-8	O89110	2.00E-33
1.7	9.14E-11	Lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog	Q9JLJ0	2.00E-12
1.7	5.48E-10	Tumor necrosis factor receptor superfamily member 27	Q8BX35	1.00E-10
1.69	3.95E-02	Toxin CaTX-A	Q9GNN8	2.00E-10
1.68	1.65E-13	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	1.00E-107
1.68	2.64E-18	Transitional endoplasmic reticulum ATPase	P55072	0
1.68	3.79E-15	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit	Q27HV0	2.00E-13
1.66	5.32E-10	TNF receptor-associated factor 3	Q13114	4.00E-78
1.65	6.33E-13	Serine/threonine-protein kinase LMTK1	Q80YE4	2.00E-50
1.62	1.41E-06	Transcription factor E2F2	P56931	5.00E-07
1.62	3.66E-05	Cell death protein 3	P45436	1.00E-28
1.59	1.88E-02	Spastin	Q05AS3	1.00E-136
1.58	7.94E-14	Nuclear factor NF-kappa-B p105 subunit	P19838	3.00E-96
1.58	6.24E-07	Migration and invasion enhancer 1	Q9CQ86	2.00E-09
1.57	7.67E-04	Sacsin	Q9NZJ4	2.00E-15
1.56	2.66E-05	Ubiquitin carboxyl-terminal hydrolase CYLD	Q80TQ2	1.00E-121
1.56	1.08E-05	Apoptosis regulator R1	Q91827	2.00E-20
1.54	1.54E-02	Tumor necrosis factor receptor superfamily member 16	P18519	7.00E-09
1.54	6.88E-11	Peptide methionine sulfoxide reductase	P08761	5.00E-43
1.53	6.70E-10	Interferon-induced helicase C domain-containing protein 1	Q8R5F7	1.00E-103
1.53	3.81E-10	Bcl-2-like protein 2	P70345	2.00E-08
1.51	2.47E-05	TNF receptor-associated factor 3	Q60803	3.00E-91
1.5	2.52E-03	Caspase-8	Q14790	1.00E-27
1.49	5.27E-03	28S ribosomal protein S29, mitochondrial	Q9ER88	2.00E-58
1.48	2.91E-07	Tyrosine-protein kinase ABL1	P00519	0
1.47	9.03E-05	Kelch-like protein 20	Q5R7B8	9.00E-25
1.47	9.75E-07	DNA damage-regulated autophagy modulator protein 2	Q9CR48	4.00E-24

Table S3.4 (Continued)

1.46	1.43E-05	Kelch-like protein 20	Q08DK3	1.00E-143
1.46	2.72E-02	DNA topoisomerase 1	P11387	0
1.45	2.12E-04	Palmitoyl-protein thioesterase 1	P50897	1.00E-98
1.44	3.96E-02	Egl nine homolog 3	Q9H6Z9	6.00E-57
1.41	1.26E-02	Fibroblast growth factor receptor 2	Q8JG38	4.00E-28
1.41	1.26E-05	DNA excision repair protein ERCC-6	Q03468	0
1.41	3.56E-03	Sacsin	Q9NZJ4	0
1.4	2.50E-03	Apoptosis inhibitor 5	Q5ZMW3	1.00E-116
1.4	1.83E-04	26S protease regulatory subunit 10B	P62333	0
1.4	1.02E-04	Kelch-like protein 20	Q08DK3	1.00E-107
1.38	2.51E-05	Proto-oncogene tyrosine-protein kinase receptor Ret	P07949	5.00E-50
1.37	2.38E-03	Organic cation transporter protein	Q9VCA2	2.00E-77
1.37	1.24E-05	Baculoviral IAP repeat-containing protein 6	Q9NR09	5.00E-56
1.36	4.82E-02	5'-AMP-activated protein kinase catalytic subunit alpha-2	Q5RD00	1.00E-175
1.36	6.01E-06	Neurofilament heavy polypeptide	P12036	4.00E-17
1.35	9.73E-04	Sacsin	Q9NZJ4	0
1.34	2.07E-04	Protein kinase C delta type	P28867	0
1.34	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.34	1.80E-02	Ribosomal protein S6 kinase alpha-2	Q9WUT3	0
1.34	8.96E-03	Induced myeloid leukemia cell differentiation protein Mcl-1 homolog	Q7YRZ9	8.00E-10
1.34	1.05E-04	Bcl-2-like protein 1	P53563	3.00E-13
1.34	8.10E-04	Ribosomal protein S6 kinase beta-1	Q6TJY3	1.00E-151
1.34	1.76E-05	Ubiquilin-1	Q8R317	1.00E-66
1.33	2.81E-04	Dual specificity protein phosphatase 6	Q16828	6.00E-73
1.33	8.30E-03	Ribonuclease ZC3H12A	Q5D1E8	5.00E-63
1.33	1.10E-05	Intersectin-1	Q9Z0R4	0
1.32	4.31E-02	Protein AATF	Q9QYW0	4.00E-81
1.32	2.34E-05	Leucine-rich repeat serine/threonine-protein kinase 2	Q5S006	0
1.32	8.51E-03	Tau-tubulin kinase 2	Q6IQ55	1.00E-105

Table S3.4 (Continued)

1.31	1.05E-04	Bcl-2-like protein 1	P53563	3.00E-13
1.3	5.47E-05	Large proline-rich protein bag6	A4IH17	3.00E-35
1.3	2.44E-04	Histone deacetylase 2	Q92769	0
1.29	8.51E-03	Cyclin-dependent kinase 11B	P21127	1.00E-160
1.29	4.58E-02	Ras-related GTP-binding protein C	Q9HB90	4.00E-50
1.29	3.07E-04	Zinc finger FYVE domain-containing protein 26	Q68DK2	1.00E-174
1.29	1.52E-02	Gamma-aminobutyric acid receptor-associated protein	P60517	4.00E-60
1.28	2.60E-02	Bifunctional apoptosis regulator	Q9NZS9	2.00E-69
1.28	1.85E-03	Proteasome subunit alpha type-4	P25789	1.00E-114
1.28	1.19E-03	HEAT repeat-containing protein 1	Q7SY48	0
1.28	3.00E-02	Protein JTB	O88824	4.00E-10
1.28	4.08E-03	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	1.00E-101
1.28	7.56E-03	Double-stranded RNA-specific adenosine deaminase	Q99MU3	1.00E-103
1.28	6.71E-04	Transcriptional repressor p66-alpha	Q86YP4	1.00E-24
1.28	8.87E-04	Apoptosis regulator BAX	Q07813	8.00E-21
1.27	4.72E-03	Sphingosine-1-phosphate lyase 1	Q8CHN6	1.00E-163
1.27	3.28E-03	Tumor necrosis factor receptor superfamily member 16	P18519	3.00E-09
1.27	2.18E-02	RB1-inducible coiled-coil protein 1	Q9ESK9	4.00E-58
1.26	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.25	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69
1.25	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.25	1.85E-03	Proteasome subunit alpha type-4	P25789	1.00E-114
1.25	7.10E-03	Baculoviral IAP repeat-containing protein 6	Q9NR09	1.00E-61
1.25	2.01E-02	TFIIH basal transcription factor complex helicase XPB subunit	Q4G005	0
1.25	8.96E-03	Induced myeloid leukemia cell differentiation protein Mcl-1 homolog	Q7YRZ9	8.00E-10
1.25	4.31E-02	Protein AATF	Q9QYW0	4.00E-81
1.25	4.58E-02	Ras-related GTP-binding protein C	Q9HB90	4.00E-50
1.25	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.24	2.52E-02	26S protease regulatory subunit 6B	P43686	0

Table S3.4 (Continued)

1.24	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69
1.24	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.24	1.85E-03	Proteasome subunit alpha type-4	P25789	1.00E-114
1.24	4.87E-02	26S proteasome non-ATPase regulatory subunit 7	P51665	1.00E-114
1.24	1.05E-04	Bcl-2-like protein 1	P53563	3.00E-13
1.24	8.96E-03	Induced myeloid leukemia cell differentiation protein Mcl-1 homolog	Q7YRZ9	8.00E-10
1.24	1.79E-02	Fibroblast growth factor receptor 2	P21802	1.00E-100
1.24	3.03E-02	Lysine-specific histone demethylase 1A	Q6ZQ88	0
1.24	4.82E-02	5'-AMP-activated protein kinase catalytic subunit alpha-2	Q5RD00	1.00E-175
1.24	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.24	2.18E-02	RB1-inducible coiled-coil protein 1	Q9ESK9	4.00E-58
1.24	1.05E-04	Bcl-2-like protein 1	P53563	3.00E-13
1.24	4.31E-02	Protein AATF	Q9QYW0	4.00E-81
1.24	2.07E-04	Protein kinase C delta type	P28867	0
1.23	1.19E-02	Transcription factor HES-4-A	Q90Z12	2.00E-18
1.23	1.33E-02	Uncharacterized protein C9orf72	Q96LT7	5.00E-61
1.23	4.87E-02	26S proteasome non-ATPase regulatory subunit 7	P51665	1.00E-114
1.23	1.49E-02	Tumor protein p73	Q9JJP2	8.00E-46
1.23	2.07E-04	Protein kinase C delta type	P28867	0
1.23	2.61E-02	Uncharacterized protein C3orf38 homolog	Q3TTL0	7.00E-39
1.22	8.31E-03	Serine/threonine-protein kinase/endoribonuclease IRE1	O75460	0
1.22	2.57E-02	FAS-associated factor 1	Q924K2	1.00E-104
1.22	8.77E-03	Bifunctional protein NCOAT	Q9EQQ9	1.00E-110
1.22	2.52E-02	26S protease regulatory subunit 6B	P43686	0
1.22	4.87E-02	26S proteasome non-ATPase regulatory subunit 7	P51665	1.00E-114
1.22	7.13E-03	Apoptotic chromatin condensation inducer in the nucleus	Q9UKV3	1.00E-33
1.21	3.28E-02	TNF receptor-associated factor 6	B6CJY4	6.00E-62
1.21	6.26E-03	Beta-hexosaminidase subunit beta	P07686	1.00E-141
1.2	1.80E-02	Ribosomal protein S6 kinase alpha-2	Q9WUT3	0

Table S3.4 (Continued)

1.2	1.80E-02	Ribosomal protein S6 kinase alpha-2	Q9WUT3	0
1.2	2.18E-02	RB1-inducible coiled-coil protein 1	Q9ESK9	4.00E-58
1.2	2.07E-04	Protein kinase C delta type	P28867	0
1.2	1.22E-02	Ataxin-2	Q99700	1.00E-29
1.2	1.52E-02	Gamma-aminobutyric acid receptor-associated protein	P60517	4.00E-60
1.2	9.57E-03	Baculoviral IAP repeat-containing protein 6	O88738	0
1.2	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69
1.19	1.79E-02	Fibroblast growth factor receptor 2	P21802	1.00E-100
1.19	4.49E-02	U4/U6.U5 tri-snRNP-associated protein 1	O43290	1.00E-116
1.19	1.63E-02	Cullin-1	Q9WTX6	0
1.18	4.72E-02	Dynamin-like 120 kDa protein, mitochondrial	Q5U3A7	0
1.18	4.42E-02	Caspase-7	P55214	1.00E-54
1.17	4.02E-02	TNF receptor-associated factor 3	Q13114	1.00E-109
1.17	4.56E-02	E3 ubiquitin-protein ligase synoviolin	Q9DBY1	1.00E-150
-1.17	4.32E-02	Casein kinase II subunit alpha	Q60737	1.00E-127
-1.17	4.60E-02	Unconventional myosin-XVIIIa	Q92614	2.00E-08
-1.18	4.17E-02	Single-stranded DNA-binding protein 3	Q9BWW4	9.00E-47
-1.19	3.28E-02	Tumor necrosis factor ligand superfamily member 10	P50591	7.00E-15
-1.19	3.97E-02	Fibroblast growth factor receptor 1	P16092	3.00E-70
-1.2	1.81E-02	Aldehyde dehydrogenase, mitochondrial	P11884	0
-1.2	1.43E-02	Unconventional myosin-XVIIIa	Q92614	0
-1.2	4.08E-02	Glyceraldehyde-3-phosphate dehydrogenase	Q05025	1.00E-137
-1.2	3.51E-02	Serine/threonine-protein kinase TAO1	O88664	3.00E-78
-1.21	2.88E-02	Ribosomal protein S6 kinase alpha-2	Q9WUT3	0
-1.21	2.66E-02	Protein Wnt-4	P56705	2.00E-63
-1.22	6.70E-03	TNF receptor-associated factor 3	Q60803	2.00E-98
-1.22	1.61E-02	Apoptosis-inducing factor 3	Q96NN9	5.00E-94
-1.22	4.72E-03	Double-stranded RNA-specific adenosine deaminase	P55265	2.00E-80
-1.22	3.09E-02	Fibroblast growth factor receptor 3	Q9I8X3	8.00E-42

Table S3.4 (Continued)

-1.23	4.84E-03	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit DAD1	Q5E9C2	2.00E-38
-1.23	3.28E-03	Fibroblast growth factor receptor 2	Q8JG38	4.00E-46
-1.24	2.28E-02	Krueppel-like factor 11	O14901	8.00E-51
-1.24	4.21E-02	Protein FAM188A	Q0IIH8	2.00E-95
-1.25	3.25E-03	Nesprin-1	Q8NF91	1.00E-106
-1.25	4.48E-03	Fibroblast growth factor receptor 3	Q91287	1.00E-111
-1.26	1.19E-02	Kelch-like protein 20	Q6DFF6	1.00E-112
-1.27	5.35E-03	Neurogenic locus Notch protein	P07207	6.00E-75
-1.28	4.35E-04	Protein unc-13 homolog B	O14795	0
-1.29	1.53E-02	Fibroblast growth factor receptor 3	Q91287	1.00E-150
-1.3	1.27E-02	Caspase-2	Q98943	4.00E-25
-1.31	8.37E-04	Rho guanine nucleotide exchange factor 4	Q9NR80	1.00E-126
-1.32	1.01E-03	Organic cation transporter protein	Q9VCA2	9.00E-57
-1.35	2.62E-03	Ras-related protein ced-10	Q03206	3.00E-82
-1.35	1.80E-03	Fibroblast growth factor receptor 3	Q61851	5.00E-66
-1.36	6.31E-05	Mucosa-associated lymphoid tissue lymphoma translocation protein 1	Q9UDY8	6.00E-26
-1.39	5.96E-04	Mucosa-associated lymphoid tissue lymphoma translocation protein 1	Q9UDY8	2.00E-07
-1.39	3.15E-02	Serine/threonine-protein kinase/endoribonuclease IRE1	O75460	3.00E-24
-1.44	1.42E-06	Serine/threonine-protein kinase PAK 7	D4A280	1.00E-115
-1.44	2.38E-07	E3 ubiquitin-protein ligase Mdm2	P56951	7.00E-13
-1.45	3.36E-02	Protein BTG1	P62324	2.00E-23
-1.49	1.54E-05	Lipopolysaccharide-induced tumor necrosis factor-alpha factor	Q99732	1.00E-20
-1.5	4.15E-02	Inhibitor of growth protein 3	Q498T3	5.00E-10
-1.51	1.56E-02	Dipeptidase 1	P31429	1.00E-73
-1.56	3.91E-04	Quinone oxidoreductase PIG3	Q53FA7	2.00E-90
-1.63	1.24E-06	Tyrosine-protein kinase PR2	Q9I7F7	5.00E-39
-1.63	9.76E-03	Programmed cell death protein 10	Q6DF07	1.00E-51
-1.79	1.49E-03	Sodium/myo-inositol cotransporter 2	Q8K0E3	1.00E-125

Table S3.4 (Continued)

-1.82	2.88E-07	Tripartite motif-containing protein 2	D2GXS7	1.00E-22
-1.89	4.92E-08	Organic cation transporter protein	Q9VCA2	8.00E-63
-2.03	2.22E-02	Tripartite motif-containing protein 2	D2GXS7	9.00E-14
-2.53	1.25E-05	Organic cation transporter protein	Q9VCA2	2.00E-76
-2.74	7.15E-03	Fibroblast growth factor receptor 2	P18461	8.00E-37
-3.5	5.92E-08	Bcl-2-related ovarian killer protein	Q9I8I2	8.00E-15

*Adjusted for false discovery rate.

Table S3.5 Differentially expressed transcripts with functions in protein ubiqutination based on GO Biological Process terms. Fold change is expressed in pathogen-exposed anemones relative to controls.

Locus #/Contig #	Fold-change	p-value*	Top blastx hit to SwissProt	Accession	e-value
Ubiquitin-activating enzymes (E1)					
19910/1	2.30	3.00E-18	Ubiquitin-like modifier-activating enzyme 6	A0AVT1	5.00E-85
12587/1	1.48	4.87E-10	Ubiquitin-like modifier-activating enzyme 1	Q29504	0.00E+00
115498/1	1.25	7.98E-03	Ubiquitin-like modifier-activating enzyme 6	Q8C7R4	1.00E-170
Ubiquitin/Ubiquitin-like modifer-conjugating enzymes (E2)					
106552/1	24.29	6.58E-236	NEDD8-conjugating enzyme ubc12	O74549	8.00E-19
33636/1	1.37	1.24E-05	Baculoviral IAP repeat-containing protein 6	Q9NR09	5.00E-56
44839/1	1.33	2.59E-03	Ubiquitin-conjugating enzyme E2-22 kDa	P52486	2.00E-74
18031/1	1.25	7.10E-03	Baculoviral IAP repeat-containing protein 6	Q9NR09	1.00E-61
123514/3	1.24	1.53E-02	Ubiquitin-conjugating enzyme E2 O	Q9C0C9	1.00E-80
8306/1	1.20	9.57E-03	Baculoviral IAP repeat-containing protein 6	O88738	0.00E+00
30020/1	-1.23	1.07E-02	Ubiquitin-conjugating enzyme E2Q-like protein 1	A0PJN4	9.00E-52
22826/1	-1.31	9.96E-05	Ubiquitin-conjugating enzyme E2 Q1	Q7TSS2	2.00E-75
E3 ligases (E3)					
21472/1	9.66	1.13E-129	Probable E3 ubiquitin-protein ligase ARI9	Q9SKC3	7.00E-26
31324/1	4.9	4.73E-50	TNF receptor-associated factor 3	Q60803	5.00E-75
32870/1	4.64	1.11E-19	TNF receptor-associated factor 5	P70191	6.00E-19
78485/1	4.16	2.19E-19	TNF receptor-associated factor 3	Q60803	1.00E-12
46907/1	3.15	6.03E-41	TNF receptor-associated factor 2	P39429	2.00E-88
75320/1	2.38	8.67E-34	E3 ubiquitin-protein ligase DTX3L	Q8TDB6	1.00E-44
89049/1	2.34	1.22E-08	TNF receptor-associated factor 3	Q13114	6.00E-45
21886/1	2.29	3.30E-03	Probable E3 ubiquitin-protein ligase ARI5	Q8L829	4.00E-25
46912/1	2.25	1.45E-17	TNF receptor-associated factor 3	Q60803	7.00E-85
21895/1	2.22	7.42E-03	Probable E3 ubiquitin-protein ligase ARI5	Q8L829	4.00E-25
82159/1	2.19	4.60E-29	E3 ubiquitin-protein ligase CBL-B	Q6DFR2	1.00E-169
93098/1	2.11	1.67E-23	E3 ubiquitin-protein ligase RNF180	Q5RAK3	2.00E-08
46888/1	2.07	2.20E-03	TNF receptor-associated factor 2	P39429	1.00E-86

Table S3.5 (Continued)

30592/1	1.88	1.28E-02	TNF receptor-associated factor 4	Q9BUZ4	2.00E-28
26488/1	1.82	1.77E-17	E3 ubiquitin-protein ligase Topors	Q80Z37	9.00E-51
21049/1	1.69	7.51E-06	Ubiquitin-protein ligase E3A	Q05086	3.00E-85
89050/1	1.66	5.32E-10	TNF receptor-associated factor 3	Q13114	4.00E-78
129278/1	1.65	1.08E-03	Probable E3 ubiquitin-protein ligase RNF144A-A	Q5RFV4	7.00E-63
10173/1	1.60	1.31E-04	RING finger and CHY zinc finger domain-containing protein 1	Q9CR50	1.00E-29
10178/1	1.56	4.68E-03	RING finger and CHY zinc finger domain-containing protein 1	Q9CR50	1.00E-29
79224/1	1.54	3.08E-03	E3 ubiquitin-protein ligase TRIM71	E7FAM5	4.00E-26
46914/1	1.51	2.47E-05	TNF receptor-associated factor 3	Q60803	3.00E-91
8352/1	1.33	1.34E-05	E3 ubiquitin-protein ligase arih1	B1H1E4	0.00E+00
115717/3	1.33	3.12E-04	E3 ubiquitin-protein ligase TRIM33	Q99PP7	2.00E-07
109181/1	1.31	2.62E-03	E3 ubiquitin-protein ligase RNF213	E9Q555	0.00E+00
91194/1	1.30	1.45E-02	E3 ubiquitin-protein ligase HECTD1	Q9ULT8	4.00E-20
41964/1	1.28	9.22E-03	E3 ubiquitin-protein ligase MIB2	Q5ZIJ9	1.00E-167
120599/1	1.27	3.07E-04	E3 ubiquitin-protein ligase HECTD1	Q9ULT8	0.00E+00
109431/1	1.21	3.28E-02	TNF receptor-associated factor 6	B6CJY4	6.00E-62
24734/1	1.20	1.72E-02	E3 ubiquitin-protein ligase MARCH8	Q5T0T0	9.00E-74
86236/1	1.2	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69
57301/1	1.18	2.53E-02	E3 ubiquitin-protein ligase HUWE1	Q7Z6Z7	0.00E+00
54135/1	1.18	2.76E-02	NEDD4-like E3 ubiquitin-protein ligase WWP1	Q8BZZ3	0.00E+00
67642/1	1.17	4.02E-02	TNF receptor-associated factor 3	Q13114	1.00E-109
37983/3	1.17	4.56E-02	E3 ubiquitin-protein ligase synoviolin	Q9DBY1	1.00E-150
3696/1	-1.17	4.15E-02	E3 ubiquitin-protein ligase TTC3	O88196	6.00E-38
25704/1	-1.21	2.76E-02	E3 ubiquitin-protein ligase RNF103	O00237	1.00E-102
114598/1	-1.22	6.70E-03	TNF receptor-associated factor 3	Q60803	2.00E-98
126251/2	-1.26	1.79E-02	E3 ubiquitin-protein ligase UHRF1	B6CHA3	0.00E+00
906/4	-1.44	2.38E-07	E3 ubiquitin-protein ligase Mdm2	P56951	7.00E-13

Table S3.5 (Continued)

51805/1	-1.45	5.93E-03	E3 ubiquitin-protein ligase TRIM33	Q56R14	1.00E-34
94460/1	-1.46	1.27E-02	E3 ubiquitin-protein ligase TRIM71	F6QEU4	3.00E-26
89564/1	-1.47	1.41E-02	E3 ubiquitin-protein ligase DZIP3	Q7TPV2	3.00E-10
28390/1	-1.78	1.79E-02	E3 ubiquitin-protein ligase SHPRH	Q149N8	1.00E-85
64989/1	-1.82	2.88E-07	Tripartite motif-containing protein 2	D2GXS7	1.00E-22
87899/1	-2.03	2.22E-02	Tripartite motif-containing protein 2	D2GXS7	9.00E-14
94117/1	-2.07	2.85E-02	E3 ubiquitin-protein ligase DZIP3	Q86Y13	2.00E-16
94131/1	-2.10	3.75E-06	E3 ubiquitin-protein ligase DZIP3	Q86Y13	8.00E-09
76790/1	-2.50	5.01E-30	E3 ubiquitin-protein ligase DZIP3	Q7TPV2	5.00E-07
51808/1	-3.32	1.12E-04	E3 ubiquitin-protein ligase TRIM33	Q56R14	3.00E-32
2499/1	-3.60	1.88E-05	Tripartite motif-containing protein 3	O75382	2.00E-22
95915/1	-5.78	2.29E-05	Tripartite motif-containing protein 59	Q922Y2	7.00E-14
Members of E3 ligase complexes					
6504/1	2.74	5.74E-15	F-box/LRR-repeat protein 13	Q8CDU4	3.00E-11
128850/1	2.41	4.73E-05	F-box/LRR-repeat protein 4	Q9C5D2	1.00E-07
35485/1	1.79	1.43E-14	Cullin-5	Q93034	0.00E+00
129330/1	1.65	6.04E-04	F-box/LRR-repeat protein 8	Q96CD0	1.00E-35
124974/1	1.31	4.97E-02	DCN1-like protein 4	Q5RHX6	3.00E-70
74655/1	1.21	6.82E-03	F-box/WD repeat-containing protein 7	Q969H0	7.00E-16
106905/1	1.21	1.85E-02	F-box/LRR-repeat protein 5	Q8C2S5	1.00E-61
32205/1	1.19	1.63E-02	Cullin-1	Q9WTX6	0.00E+00
11346/2	-1.26	2.77E-02	F-box only protein 36	Q5R796	3.00E-34
96190/1	-1.30	4.38E-04	F-box/LRR-repeat protein 20	Q9CZV8	1.00E-146
22443/1	-1.37	6.44E-03	F-box only protein 16	Q9QZM9	9.00E-58
2902/1	-1.50	5.82E-04	F-box/WD repeat-containing protein 10	Q5XX13	1.00E-108
129835/1	-1.55	1.61E-04	F-box/LRR-repeat protein 4	Q8BH70	2.00E-95
Deubiqutinating enzymes					
24068/1	1.62	4.30E-02	Ubiquitin carboxyl-terminal hydrolase CYLD	Q5RED8	1.00E-16

Table S3.5 (Continued)

24058/1	1.56	2.66E-05	Ubiquitin carboxyl-terminal hydrolase CYLD	Q80TQ2	1.00E-121
26492/1	1.39	1.49E-07	Ubiquitin carboxyl-terminal hydrolase 8	P40818	1.00E-112
89818/1	1.30	8.81E-04	TNFAIP3-interacting protein 1	Q15025	1.00E-16
Proteosome					
56614/1	1.67	1.08E-06	26S proteasome non-ATPase regulatory subunit 11A	F6P3G4	1.00E-148
36958/1	1.63	6.09E-06	Proteasome subunit alpha type-7-A	Q9PVY6	1.00E-101
123386/1	1.63	8.83E-04	Proteasome subunit alpha type-3	Q58DU5	1.00E-100
32657/1	1.62	9.08E-15	26S proteasome non-ATPase regulatory subunit 3	Q2KJ46	1.00E-169
29842/1	1.57	3.20E-09	26S proteasome non-ATPase regulatory subunit 2	Q5R9I6	0.00E+00
22959/3	1.54	2.55E-05	26S proteasome non-ATPase regulatory subunit 12	Q9D8W5	1.00E-160
77235/1	1.43	2.28E-04	Proteasome subunit alpha type-2	O73672	1.00E-113
123676/2	1.42	6.81E-08	26S proteasome non-ATPase regulatory subunit 1	Q3TXS7	0.00E+00
13825/1	1.40	1.83E-04	26S protease regulatory subunit 10B	P62333	0.00E+00
24532/3	1.37	8.45E-04	26S proteasome non-ATPase regulatory subunit 8	Q5RE15	8.00E-64
24241/1	1.36	8.87E-05	26S proteasome non-ATPase regulatory subunit 13	P84169	1.00E-109
39371/1	1.34	2.20E-03	26S protease regulatory subunit 6A	O88685	0.00E+00
9401/1	1.34	1.30E-04	Proteasome subunit alpha type-5	Q9Z2U1	1.00E-108
102797/1	1.30	3.92E-03	26S proteasome non-ATPase regulatory subunit 6	Q99JI4	1.00E-131
120706/2	1.28	1.85E-03	Proteasome subunit alpha type-4	P25789	1.00E-114
100449/1	1.24	2.52E-02	26S protease regulatory subunit 6B	P43686	0.00E+00
3281/1	1.23	4.87E-02	26S proteasome non-ATPase regulatory subunit 7	P51665	1.00E-114
123798/3	1.21	4.73E-02	26S protease regulatory subunit 7	P46472	0.00E+00
64936/1	1.20	2.63E-02	Proteasome subunit alpha type-1	Q9R1P4	1.00E-105
106278/1	-1.29	1.22E-02	Proteasome inhibitor PI31 subunit	Q3SX30	2.00E-28
Associated proteins					
25984/1	3.47	7.25E-44	Transient receptor potential cation channel subfamily A member 1	Q6RI86	1.00E-10
59641/1	2.09	9.50E-13	MFS-type transporter SLC18B1	D3Z5L6	1.00E-61

Table S3.5 (Continued)

51546/1	1.72	2.78E-04	Vacuolar protein sorting-associated protein VTA1 homolog	Q9CR26	6.00E-21
60163/1	1.70	9.14E-11	Lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog	Q9JLJ0	2.00E-12
33916/1	1.68	2.64E-18	Transitional endoplasmic reticulum ATPase	P55072	0.00E+00
41335/1	1.66	7.67E-13	Ran-specific GTPase-activating protein	P43487	2.00E-50
92031/1	1.47	9.75E-07	DNA damage-regulated autophagy modulator protein 2	Q9CR48	4.00E-24
51548/1	1.45	3.89E-03	Vacuolar protein sorting-associated protein VTA1 homolog	Q9CR26	1.00E-52
129510/1	1.4	2.75E-07	Tumor susceptibility gene 101 protein	Q99816	9.00E-95
85556/1	1.38	1.86E-05	Tribbles homolog 2	Q92519	2.00E-99
13541/1	1.36	2.15E-02	Kelch-like protein 12	Q6NRH0	5.00E-39
33812/1	1.35	2.04E-02	Nucleoporin NUP188 homolog	Q5SRE5	7.00E-80
125776/1	1.3	5.47E-05	Large proline-rich protein bag6	A4IH17	3.00E-35
85034/1	1.28	5.55E-04	Granulins	P28799	1.00E-63
11252/1	1.24	4.85E-03	RNA polymerase-associated protein CTR9 homolog	Q6DEU9	0.00E+00
15792/1	1.23	3.94E-02	Zinc finger protein 227	Q86WZ6	1.00E-62
68308/1	1.21	8.56E-03	Tankyrase-1	O95271	0.00E+00
62716/1	1.2	1.52E-02	Protein argonaute-2	Q9QZ81	0.00E+00
69212/1	1.2	1.95E-02	Kinesin-like protein KIF16B	B1AVY7	3.00E-70
16805/1	1.19	3.66E-02	Tankyrase-1	O95271	3.00E-14
16122/1	-1.18	2.86E-02	Peroxisomal multifunctional enzyme type 2	P51659	0.00E+00
29717/1	-1.21	4.69E-02	SPRY domain-containing SOCS box protein 3	Q3MHZ2	3.00E-61
48924/1	-1.23	2.71E-02	NACHT, LRR and PYD domains-containing protein 14	Q86W24	7.00E-33
117579/1	-1.25	3.25E-03	Nesprin-1	Q8NF91	1.00E-106
32137/1	-1.28	1.12E-02	Kelch-like protein 3	E0CZ16	2.00E-76
11893/1	-1.29	1.45E-02	LIM domain kinase 1	P53668	1.00E-118
21856/1	-1.31	3.52E-02	Helicase-like transcription factor	Q14527	1.00E-106
69227/1	-1.32	6.21E-03	U-box domain-containing protein 36	Q8GZ84	6.00E-13

Table S3.5 (Continued)

90534/1	-1.33	1.31E-04	FERM and PDZ domain-containing protein 4	Q14CM0	6.00E-93
82308/1	-1.36	6.31E-05	Mucosa-associated lymphoid tissue lymphoma translocation protein 1	Q9UDY8	6.00E-26
1073/1	-1.37	6.40E-05	Polycomb protein SUZ12	Q15022	1.00E-125
82314/1	-1.39	5.96E-04	Mucosa-associated lymphoid tissue lymphoma translocation protein 1	Q9UDY8	2.00E-07
18291/1	-1.39	1.70E-02	Indoleamine 2,3-dioxygenase 2	Q6ZQW0	8.00E-53
106694/1	-1.55	4.34E-02	Long-chain-fatty-acid--CoA ligase 5	Q9ULC5	1.00E-116
126855/1	-2.54	5.75E-05	Opioid-binding protein/cell adhesion molecule	Q14982	2.00E-20
83879/1	-2.59	1.27E-06	Nephrocystin-3	Q7Z494	1.00E-33
60717/1	-2.76	1.28E-02	Sushi domain-containing protein 2	Q9UGT4	1.00E-23

*Adjusted for false discovery rate.

Table S4.1A Correlation between RNA-Seq and RT-qPCR measurements of differential gene expression in pathogen-exposed relative to controls in both aposymbiotic and symbiotic anemones anemones.

Locus #/ Contig #	Top blastx hit to Swiss-Prot	Fold-Change RNASeq	Fold-Change qPCR	Fold-Change RNASeq	Fold-Change qPCR
32435/1	Cytochrome P450 3A24	22.6	22.7	5.5	5.1
21472/1	Probable E3 ubiquitin-protein ligase ARI9	9.7	9.0	2.7	2.4
122298/5	Histidine ammonia-lyase	5.5	4.9	1.8	3.0
31324/1	TNF receptor-associated factor 3	4.9	5.4	2.1	2.2
22981/1	Tyrosine kinase receptor Cad96Ca	4.3	3.1	2.0	1.2
21357/1	Carbamoyl-phosphate synthase	4.3	4.3	1.7	1.7
84775/1	60S ribosomal protein L10*+	1.2	1.1	1.2	1.1
58671/1	Glyceraldehyde-3-phosphate dehydrogenase+	1.1	1.1	-1.0	-1.0
82926/1	40S ribosomal protein S7*+	1.0	-1.0	1.2	-1.1
99523/1	Polyadenylate-binding protein 1 *	-1.1	-1.0	-1.1	1.1
12311/1	60S ribosomal Protein L11	-1.2	-1.0	-1.1	-1.4
84201/1	Cytochrome c oxidase*+	-1.4	-1.0	-1.2	-1.0
127561/1	Transcription factor SOX-14	-2.0	-2.3	-2.5	-2.2
76790/1	E3 ubiquitin-protein ligase DZIP3	-2.5	-2.6	-2.3	-2.0
101692/1	Calponin-1	-2.6	-2.6	-4.0	-4.3
15861/1	Protein NLRC5	-2.6	-2.5	-3.9	-3.4
60730/1	Sushi domain-containing protein 2	-2.7	-1.1	-4.1	1.0
118088/1	Forkhead box protein D1	-3.6	-3.8	-4.6	-6.5

*reference gene for aposymbiotic anemones

+reference gene for symbiotic anemones

Table S4.1B Primer sequences, product length, and average efficiency for RT-qPCR data. [a]

Locus #/ Contig #	Forward Primer	Reverse Primer	Product length	Avg. efficiency
32435/1	GGCCACACTCTGTCTATCT	CTTCCTTCTGGCTGGTTATG	117	92
21472/1	GCATTTGAGGGTGATTTCTTG	AGACTGGGATGAGGAAGATAG	141	93
122298/5	GGCTTTGGTGAGGTATTAAGG	CGAGGGTGTGTAAGTCAATTAG	131	95
31324/1	TGAAGCAATGTTCATCGAGC	ACAGCCCTCCATTTTACGTG	272	90
22981/1	CAGAGAAAGACGCTTACGATAC	GAGACGGACAAGAAGGAAATAG	99	95
21357/1	GGTGAGAAACAGGCAGATAC	CACCCATAGACAGCAGTAATC	106	95
84775/1	ACGTTTCTGCCGTGGTGTCCC	CGGGCAGCTTCAAGGGCTTCAC	175	90
58671/1	AACAGCTTTGGCAGCACCTGTAGA	TGCTTTCACAGCAACCCAGAAGAC	114	89
82926/1	ACTGCAGTCCACGATGCTATCCTT	GTCTGTTGTGCTTTGTCGAGATGC	125	91
99523/1	GTGCAAGGAGGCGGACAGCG	TGGGCTGATTGCGGGTTGCC	150	88
12311/1	AGCCAAGGTCTTGGAGCAGCTTA	TTGGGCCTCTGACAGTACAGTGAACA	125	89
84201/1	AGCAGTTGGTAAGTCTGCACAA	GTAACCATGGTAGCAGCATGAA	105	93
127561/1	GACATATCAGAGCTTCCATTCA	CGAGCGTGAGTCAGTATTAAG	122	95
76790/1	AGAAGGGCATCGGTTAGA	ACTGATGGAGAAGGAGTAGAG	110	90
101692/1	CCCTGGATGATGGTGTTATTT	GACTTCTGACATGCGTCTATAA	142	90
15861/1	ATTGGCAGGTGTTGGTAAG	AGGGTAGATCAGTGGAGTAAG	120	94
60730/1	CGAATGGGTGTGGAGTAAAG	GTAGGAGGGCATGCTATAAAC	140	92
118088/1	CATGTTCGA AAA CGG GAGTT	AAATACGGCGGTAGAGCAGA	213	93

[a] Transcripts are listed in the same order as in Table S4.1A

Table S4.2. List of differentially expressed genes with apoptotic functions as determined by GO Biological Process terms that were uniquely expressed in *S. marcescens*-exposed aposymbiotic or symbiotic anemones or expressed in both types of anemones relative to controls.

Locus #/ Contig #	Fold-Change	p-value	Protein name	SwissProt Accession	e-value	pro/anti apoptotic
Unique in aposymbiotic anemones						
4948/2	4.23	1.36E-17	Cation transport regulator-like protein 1	Q5SPB6	6.00E-42	pro
78485/1	4.14	1.20E-13	TNF receptor-associated factor 3	Q60803	1.00E-12	pro
55453/1	3.16	7.15E-21	Growth arrest and DNA damage-inducible protein GADD45 gamma	Q9Z111	6.00E-09	pro
46907/1	3.15	6.03E-41	TNF receptor-associated factor 2	P39429	2.00E-88	both
27972/2	2.64	4.15E-23	DnaJ homolog subfamily A member 3, mitochondrial	Q99M87	5.00E-98	both
125725/1	2.13	1.74E-04	P2X purinoceptor 7	Q99572	6.00E-12	pro
24567/1	2.1	6.05E-03	Organic cation transporter protein	Q9VCA2	8.00E-65	pro
46888/1	2.09	5.04E-03	TNF receptor-associated factor 2	P39429	1.00E-86	both
56290/1	2.02	6.54E-08	Sphingosine kinase 2	Q9JIA7	2.00E-69	anti
95010/1	1.99	2.46E-02	Tumor necrosis factor receptor superfamily member 27	Q8BX35	1.00E-09	pro
21984/1	1.98	1.14E-12	WD repeat-containing protein 92	Q96MX6	1.00E-155	pro
23220/1	1.98	1.68E-09	Immediate early response 3-interacting protein 1	P85007	2.00E-15	pro
32905/1	1.95	9.24E-13	Growth hormone-inducible transmembrane protein	Q5XIA8	2.00E-54	pro
30592/1	1.9	4.59E-02	TNF receptor-associated factor 4	Q9BUZ4	2.00E-28	both
26488/1	1.83	2.76E-08	E3 ubiquitin-protein ligase Topors	Q80Z37	9.00E-51	pro
35485/1	1.81	1.32E-09	Cullin-5	Q93034	0	anti
51289/1	1.74	2.06E-04	Caspase-8	O89110	2.00E-33	pro
46973/1	1.73	3.18E-02	Paired box protein Pax-3	P23760	3.00E-43	anti
95000/1	1.71	9.69E-03	Tumor necrosis factor receptor superfamily member 27	Q8BX35	1.00E-10	pro
33916/1	1.7	1.69E-07	Transitional endoplasmic reticulum ATPase	P55072	0	anti

Table S4.2 (Continued)

60163/1	1.7	6.49E-06	Lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog	Q9JLJ0	2.00E-12	pro
4432/1	1.66	4.19E-05	Serine/threonine-protein kinase LMTK1	Q80YE4	2.00E-50	pro
8113/1	1.64	2.31E-03	Cell death protein 3	P45436	1.00E-28	pro
1568/1	1.62	6.76E-04	Transcription factor E2F2	P56931	5.00E-07	both
76372/1	1.6	2.42E-03	Migration and invasion enhancer 1	Q9CQ86	2.00E-09	anti
73699/1	1.59	6.71E-07	Nuclear factor NF-kappa-B p105 subunit	P19838	3.00E-96	anti
24058/1	1.57	8.68E-04	Ubiquitin carboxyl-terminal hydrolase CYLD	Q80TQ2	1.00E-121	pro
67926/1	1.54	4.46E-05	Bcl-2-like protein 2	P70345	2.00E-08	anti
54450/1	1.54	3.35E-04	Interferon-induced helicase C domain-containing protein 1	Q8R5F7	1.00E-103	pro
46914/1	1.52	5.74E-03	TNF receptor-associated factor 3	Q60803	3.00E-91	pro
65023/1	1.51	2.59E-02	Caspase-8	Q14790	1.00E-27	pro
117960/2	1.5	2.74E-02	28S ribosomal protein S29, mitochondrial	Q9ER88	2.00E-58	pro
821/1	1.49	1.17E-02	Tyrosine-protein kinase ABL1	P00519	0	pro
21172/1	1.48	2.17E-03	Kelch-like protein 20	Q5R7B8	9.00E-25	anit
92031/1	1.48	3.79E-04	DNA damage-regulated autophagy modulator protein 2	Q9CR48	4.00E-24	pro
12823/1	1.47	1.82E-03	Kelch-like protein 20	Q08DK3	1.00E-143	anti
52243/1	1.45	7.31E-03	Palmitoyl-protein thioesterase 1	P50897	1.00E-98	anti
127996/1	1.42	4.85E-02	Fibroblast growth factor receptor 2	Q8JG38	4.00E-28	pro
129849/1	1.42	1.36E-03	DNA excision repair protein ERCC-6	Q03468	0	pro
103533/1	1.41	2.31E-02	Apoptosis inhibitor 5	Q5ZMW3	1.00E-116	anti
53322/1	1.41	1.25E-02	Kelch-like protein 20	Q08DK3	1.00E-107	anti
124732/1	1.38	1.92E-02	Organic cation transporter protein	Q9VCA2	2.00E-77	pro
63295/1	1.35	1.37E-02	Ubiquilin-1	Q8R317	1.00E-66	anti
53971/1	1.35	3.47E-02	Ribosomal protein S6 kinase beta-1	Q6TJY3	1.00E-151	anti
86856/1	1.34	4.70E-03	Intersectin-1	Q9Z0R4	0	anti
23887/1	1.33	6.08E-03	Leucine-rich repeat serine/threonine-protein kinase 2	Q5S006	0	pro

Table S4.2 (Continued)

710/1	1.31	1.38E-02	Histone deacetylase 2	Q92769	0	anti
125776/1	1.31	8.57E-03	Large proline-rich protein bag6	A4IH17	3.00E-35	pro
90495/1	1.3	1.92E-02	Apoptosis regulator BAX	Q07813	8.00E-21	pro
13279/1	1.29	3.42E-02	HEAT repeat-containing protein 1	Q7SY48	0	anti
58544/1	1.28	4.49E-02	Sphingosine-1-phosphate lyase 1	Q8CHN6	1.00E-163	pro
17372/1	1.25	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69	both
100449/1	1.24	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69	both
109431/1	1.21	3.28E-02	TNF receptor-associated factor 6	B6CJY4	6.00E-62	pro
86236/1	1.2	3.74E-02	TNF receptor-associated factor 2	Q12933	4.00E-69	both
67642/1	1.17	4.02E-02	TNF receptor-associated factor 3	Q13114	1.00E-109	pro
114598/1	-1.22	6.70E-03	TNF receptor-associated factor 3	Q60803	2.00E-98	pro
81783/1	-1.27	4.81E-02	Protein unc-13 homolog B	O14795	0	pro
355/1	-1.3	4.23E-02	Rho guanine nucleotide exchange factor 4	Q9NR80	1.00E-126	pro
104611/1	-1.31	3.62E-02	Organic cation transporter protein	Q9VCA2	9.00E-57	pro
16165/1	-1.34	4.09E-02	Fibroblast growth factor receptor 3	Q61851	5.00E-66	pro
82308/1	-1.35	2.43E-02	Mucosa-associated lymphoid tissue lymphoma translocation protein 1	Q9UDY8	6.00E-26	anti
82314/1	-1.38	2.46E-02	Mucosa-associated lymphoid tissue lymphoma translocation protein 1	Q9UDY8	2.00E-07	anti
17181/1	-1.43	1.78E-03	Serine/threonine-protein kinase PAK 7	D4A280	1.00E-115	anti
906/4	-1.43	3.40E-04	E3 ubiquitin-protein ligase Mdm2	P56951	7.00E-13	anti
67712/1	-1.48	1.10E-03	Lipopolysaccharide-induced tumor necrosis factor-alpha factor	Q99732	1.00E-20	pro
119516/1	-1.55	1.01E-02	Quinone oxidoreductase PIG3	Q53FA7	2.00E-90	pro
119860/1	-1.77	1.20E-02	Sodium/myo-inositol cotransporter 2	Q8K0E3	1.00E-125	anti
100445/1	-2.51	3.69E-05	Organic cation transporter protein	Q9VCA2	2.00E-76	pro
88035/1	-2.75	2.08E-02	Fibroblast growth factor receptor 2	P18461	8.00E-37	pro
125622/1	-2.9	2.64E-05	Transient receptor potential cation channel subfamily M member 7	Q925B3	7.00E-11	pro

Table S4.2 (Continued)

Unique in symbiotic anemones						
15238/1	1.85	3.66E-06	Eukaryotic translation initiation factor 4 gamma 1	Q04637	1.00E-160	anti
27524/1	1.74	1.22E-04	Zinc finger protein SNAI2	O43623	4.00E-63	anti
61394/1	1.65	1.59E-03	Ribonuclease ZC3H12A	Q5D1E8	5.00E-63	pro
86435/1	1.56	1.44E-02	Apoptotic protease-activating factor 1	O88879	0	pro
25878/1	1.52	1.52E-02	Serine-protein kinase ATM	Q62388	4.00E-88	pro
6121/1	1.43	4.11E-02	Fibroblast growth factor receptor 2	P21802	1.00E-100	pro
34062/1	1.33	2.42E-02	DNA polymerase beta	Q27958	1.00E-123	anti
128224/1	1.30	3.93E-02	U4/U6.U5 tri-snRNP-associated protein 1	O43290	1.00E-116	pro
48680/1	-1.35	3.35E-02	Growth arrest-specific protein 2	O43903	3.00E-48	pro
26178/3	-1.38	1.52E-02	Serine/threonine-protein kinase 17A	Q9GM70	3.00E-70	pro
21333/1	-1.39	3.02E-02	Secreted frizzled-related protein 3	Q92765	2.00E-28	pro
126358/1	-1.43	3.00E-02	E3 ubiquitin-protein ligase NRDP1	Q5FWL3	1.00E-109	pro
11259/2	-1.43	3.54E-02	Transcription factor AP-4	Q01664	3.00E-33	pro
58683/1	-1.44	4.08E-02	Glyceraldehyde-3-phosphate dehydrogenase	Q05025	1.00E-127	pro
38578/1	-1.44	3.28E-02	Fibroblast growth factor receptor 3	Q91287	1.00E-150	pro
104739/1	-1.50	4.24E-02	TNF receptor-associated factor 4	Q9BUZ4	5.00E-27	both
18915/1	-1.57	1.42E-02	DEP domain-containing mTOR-interacting protein	Q8TB45	2.00E-53	pro
58173/1	-1.69	2.14E-05	Krueppel-like factor 11	O14901	8.00E-51	pro
115677/1	-1.71	2.26E-02	Deoxyribonuclease-1	P21704	6.00E-55	pro
125687/1	-1.81	1.92E-02	GTP cyclohydrolase 1	P30793	3.00E-69	pro
25092/1	-1.89	2.63E-02	Organic cation transporter protein	Q9VCA2	3.00E-73	pro
128001/1	-2.25	1.79E-02	Kelch-like protein 20	Q08DK3	2.00E-09	anti
38649/1	-2.37	2.63E-02	Transmembrane protein 173	Q3TBT3	4.00E-32	pro
8815/1	-3.01	8.10E-03	Transcription factor Dp-1	Q17QZ4	2.00E-23	pro
Expressed in both symbiotic states						
32870/1	3.03	9.07E-05	TNF receptor-associated factor 5	P70191	6.00E-19	anti

Table S4.2 (Continued)

31324/1	4.90	4.73E-50	TNF receptor-associated factor 3	Q60803	5.00E-75	pro
32870/1	4.64	1.11E-19	TNF receptor-associated factor 5	P70191	6.00E-19	anti
115532/1	2.75	1.54E-06	SAM pointed domain-containing Ets transcription factor	Q9WTP3	7.00E-16	pro
89049/1	2.34	1.22E-08	TNF receptor-associated factor 3	Q13114	6.00E-45	pro
46912/1	2.25	1.45E-17	TNF receptor-associated factor 3	Q60803	7.00E-85	pro
31324/1	2.13	1.43E-05	TNF receptor-associated factor 3	Q60803	5.00E-75	pro
119381/6	2.12	2.59E-07	Programmed cell death protein 2	Q16342	1.00E-75	pro
6274/1	1.95	2.44E-03	Protein BTG1	Q63073	2.00E-18	pro
46912/1	1.92	1.15E-03	TNF receptor-associated factor 3	Q60803	7.00E-85	pro
46907/1	1.90	1.67E-03	TNF receptor-associated factor 2	P39429	2.00E-88	both
16729/1	1.76	2.59E-07	NADPH--cytochrome P450 reductase	P37040	0	anti
89050/1	1.66	5.32E-10	TNF receptor-associated factor 3	Q13114	4.00E-78	pro
89049/1	1.64	4.93E-02	TNF receptor-associated factor 3	Q13114	6.00E-45	anti
33510/2	1.63	1.75E-04	Eukaryotic translation initiation factor 2-alpha kinase 3	Q9Z2B5	1.00E-134	pro
76995/1	1.61	2.92E-04	Myocyte-specific enhancer factor 2C	A4UTP7	5.00E-64	anti
125279/4	1.55	1.40E-03	Death ligand signal enhancer	P60924	1.00E-29	pro
30644/1	1.55	5.46E-04	Interferon-induced helicase C domain-containing protein 1	Q9BYX4	1.00E-107	pro
94757/1	1.51	2.76E-03	UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit	Q27HV0	2.00E-13	pro
96676/1	1.47	4.91E-02	Apoptosis regulator R1	Q91827	2.00E-20	pro
89050/1	1.43	1.18E-02	TNF receptor-associated factor 3	Q13114	4.00E-78	pro
104929/1	1.43	3.87E-02	Proto-oncogene tyrosine-protein kinase receptor Ret	P07949	5.00E-50	pro
11446/1	1.36	1.42E-02	Dual specificity protein phosphatase 6	Q16828	6.00E-73	pro
122906/1	-1.94	6.29E-05	Organic cation transporter protein	Q9VCA2	8.00E-63	pro
11742/1	-2.03	3.95E-02	Bcl-2-related ovarian killer protein	Q9I8I2	8.00E-15	anti
64989/1	-2.14	2.51E-05	Tripartite motif-containing protein 2	D2GXS7	1.00E-22	anti

Table S4.3 General linear model with binomial sampling examining the effect of bacterial treatment and symbiotic state on the proportion of time anemones spent in the extended behavioral state. Significant p-values are in bold.

	Estimate	Standard Error	z-value	p-value
Intercept	2.79	0.34	8.13	**4.14E-16**
State Sym	0.12	0.50	0.25	8.03E-01
Treatment low	0.12	0.50	-9.64	2.00E-16
Treatment Medium	-2.48	0.38	-6.54	**6.18E-11**
Treatment High	-3.73	0.39	-9.64	**2.00E-16**
State: Treatment Low	-1.80	0.65	-2.79	**5.33E-03**
State: Treatment Medium	0.17	0.55	0.31	7.54E-01
State: Treatment High	-0.47	0.56	-0.83	4.07E-01

SUPPLEMENNTARY REFERENCES

Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. Genome biology 11: R106.

Darsigny, M., Babeu, J.P., Dupuis, A.A., Furth, E.E., Seidman, E.G., et al. (2009) Loss of hepatocyte-nuclear-factor-4alpha affects colonic ion transport and causes chronic inflammation resembling inflammatory bowel disease in mice. PloS one 4: e7609.

Endo, Y., Matsushita, M., and Fujita, T. (2007) Role of ficolin in innate immunity and its molecular basis. Immunobiology 212: 371–379.

Gessi, S., Merighi, S., Fazzi, D., Stefanelli, A., Varani, K., et al. (2011) Adenosine receptor targeting in health and disease. Expert Opin Investig Drugs 20: 1591–1609.

Hooper, J.D., Campagnolo, L., Goodarzi, G., Truong, T.N., Stuhlmann, H., et al. (2003) Mouse matriptase-2: identification, characterization and comparative mRNA expression analysis with mouse hepsin in adult and embryonic tissues. The Biochemical journal 373: 689–702.

Kobuke, K., Furukawa, Y., Sugai, M., Tanigaki, K., Ohashi, N., et al. (2001) ESDN, a novel neuropilin-like membrane protein cloned from vascular cells with the longest secretory signal sequence among eukaryotes, is up-regulated after vascular injury. The Journal of biological chemistry 276: 34105–34114.

Kenkel, C.D., Traylor, M.R., Wiedenmann, J., Salih, A., and Matz, M.V. (2011) Fluorescence of coral larvae predicts their settlement response to crustose coralline algae and reflects stress. Proceedings of the Royal Society B: Biological Sciences 278: 2691–2697.

Lalmanach, G., Naudin, C., Lecaille, F., and Fritz, H. (2010) Kininogens: More than cysteine protease inhibitors and kinin precursors. Biochimie 92: 1568–1579.

Leggat, W., Seneca, F., Wasmund, K., Ukani, .L, Yellowlees, D., et al. (2011) Differential responses of the coral host and their algal symbiont to thermal stress. PloS one 6: e26687.

Lehnert, E.M., Burriesci, M.S., and Pringle, J.R. (2012) Developing the anemone *Aiptasia* as a tractable model for cnidarian-dinoflagellate symbiosis: the transcriptome of aposymbiotic *A. pallida*. BMC genomics 13: 271.

Li, W.Y., Chong, S.S.N., Huang, E.Y., and Tuan, T.L. (2003) Plasminogen activator/plasmin system: A major player in wound healing? Wound Repair and Regeneration 11: 239–247.

Logan, D.D.K., LaFlamme, A.C., Weis, V.M., and Davy, S.K. (2010) Flow-cytometric characterization of the cell-surface glycans of symbiotic dinoflagellates (*Symbiodiniym* spp.) Journal of Phycology 46: 525–533.

Martin, F., Penet, M.F., Malergue, F., Lepidi, H., Dessein, A., et al. (2004) Vanin-1(-/-) mice show decreased NSAID- and *Schistosoma*-induced intestinal inflammation associated with higher glutathione stores. Journal of Clinical Investigation 113: 591–597.

Moreau, M.E., Garbacki, N., Molinaro, G., Brown, N.J., Marceau, F., et al. (2005) The kallikrein-kinin system: current and future pharmacological targets. Journal of Pharmacological Sciences 99: 6–38.

Rodriguez-Lanetty, M., Phillips, W.S., and Weis, V.M. (2006) Transcriptome analysis of a cnidarian-dinoflagellate mutualism reveals complex modulation of host gene expression. BMC Genomics 7: 23.

Smith, F.M., Vearing, C., Lackmann, M., Treutlein, H., Himanen, J, et al. (2004) Dissecting the EphA3/Ephrin-A5 interactions using a novel functional mutagenesis screen. The Journal of Biological Chemistry 279: 9522–9531.

Sunagawa, S., Wilson, E.C., Thaler, M., Smith, M.L., Caruso, C., et al. (2009) Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone *Aiptasia pallida* and its dinoflagellate endosymbiont. BMC Genomics 10: 258.

Takahashi, M., Iwaki, D., Kanno, K., Xiong, J., Matsushita, M., et al. (2013) Mannose-Binding Lectin (MBL)-Associated Serine Protease (MASP)-1 Contributes to Activation of the Lectin Complement Pathway.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 3: Research0034.1-Research0034.11.