Characterization of the prokaryotic community associated with the giant barrel sponge, Xestospongia muta across the Caribbean

Cara L. Fiore

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation
Fiore, Cara L., "Characterization of the prokaryotic community associated with the giant barrel sponge, Xestospongia muta across the Caribbean" (2013). Doctoral Dissertations. 733.
https://scholars.unh.edu/dissertation/733

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
CHARACTERIZATION OF THE PROKARYOTIC COMMUNITY ASSOCIATED
WITH THE GIANT BARREL SPONGE, XESTOSPONGIA MUTA ACROSS THE
CARIBBEAN

BY

CARA L FIORE
BS, SUNY at Cortland, 2003
MS, College of Charleston, 2006

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Microbiology

May, 2013
This dissertation has been examined and approved.

Michael Lesser
Dissertation Director, Dr. Michael P. Lesser
Research Professor of Marine, Estuarine and Freshwater Biology and Marine Sciences

Charles Walker
Dr. Charles Walker
Professor of Biochemistry, Molecular and Cellular Biology and Marine, Estuarine and Freshwater Biology and Marine Sciences

Elise R. Sullivan
Dr. Elise Sullivan
Clinical Assistant Professor of Biomedical Sciences and Marine Sciences

Marc Slattery
Dr. Marc Slattery
Professor of Pharmacognosy, University of Mississippi

Julie Olson
Dr. Julie Olson
Associate Professor of Biological Sciences, University of Alabama

4-1-13
Date
ACKNOWLEDGEMENTS

This work was supported by research grants from The National Geographic Society, the National Science Foundation, and the National Oceanic and Atmospheric Administration’s National Institute for Undersea Science and Technology and Ocean Exploration Program. My work was also supported by a dissertation year fellowship from the UNH Graduate School and grants from the UNH Marine Program. I am very grateful to have also received funding from the following sources to support travel to workshops and conferences: UNH Graduate School, UNH Microbiology Program, and the Department of Molecular, Cellular and Biomedical Sciences at UNH.

I am grateful for the support and guidance that I received from my advisor, Dr. Michael Lesser during the five years that I have worked on this project. While things did not always go according to plan, he was always ready to work with me to develop plans B, C, and sometimes D, and I appreciate his effort in seeing this project through to the end. I also greatly appreciate his support of my professional development by encouraging me to attend conferences and workshops relevant to my work. In addition, I had the support of a great committee, with two members, Dr. Julie Olson and Dr. Marc Slattery, who worked in the field with me. Their guidance and patience in the field, whether I was learning to dock a boat or setting up a complex experiment on a remote tropical island, was invaluable. I want to thank the rest of my committee, Dr. Charles Walker and Dr. Elise Sullivan, for their valuable comments and suggestions on this work.

I am grateful for the help of many people in the lab and in the field. First and foremost, is my labmate and friend, Dr. Jessica Jarett. Through cold weather, hot weather, swarms of nasty bugs, late nights of processing samples followed by too early mornings of collecting more samples, and seasickness, she still was willing to help whenever I needed it, and without her help in the field I would not have been able to get all of this work done. Her cheer and her willingness to give me her opinion on even the most crazy of my ideas was instrumental in getting through these five years. Having her and former graduate student Nathan Olson in the lab was crucial in developing my understanding of the ecosystem that we all worked on, each from different angles. I have also been fortunate to work with some wonderful undergraduates who contributed either directly or by moral support to this project: Micheline Labrie (helped with nitrogen gene cloning and DNA extractions), Abbey Tedford (helped with lab work for continuing projects of stress experiments and depth series), Dan Merselis, and Andrew McDermitt.

Getting back to field work, I had help from many wonderful people including: Dr. Deb Gochfeld, Liz Kintzing (UNH dive safety officer/diving guru), Dr. Chris Freeman, Sylvester Lee, Dr. Erica Hunkin, and Dr. Cole Easson. I want to give a special thank you to Cole, who helped whenever needed in the field and was instrumental in getting the stress experiments setup and helped with running the isotope experiments, and provided much needed moral support and comic relief. I also want to thank the field stations and staff there for support: CCMI at Little Cayman, PIMS at Lee Stocking Island, Bahamas,
and Aquarius Reef Base, Florida Keys. In particular, I want to take this opportunity to thank Marc Slattery for the opportunity to participate in the Aquarius mission and to our group in Aquarius (Marc, Mauritius Bell, Cole Easson) and ARB support staff and our support divers who made sure everything ran smoothly. It was a once in a lifetime opportunity and we had a lot of fun and did some excellent science!

This research involved a lot of computational work and I am still staggered by the learning curve that I am finally making my way up. I would not have gotten as far as I have without the help of Dr. Kelley Thomas and grad students in his lab: Kazu Okamoto and John Gaspar, and a postdoc Dr. Holly Bik. Another grad student in Computer Science, Matthew Hatem was also very helpful and patient with answering my many, many programming questions. Peter Schmitt was helpful with getting programs running on Inquiry, and Dr. Chris Miller was gracious enough to help me for many months to get his program, EMIRGE, to run. I appreciate the support from Dr. Louis Tisa and Sheldon Hurst for allowing me to use (take over) their lab computer at “crunch time”. I also want to thank the wonderful faculty and instructors at the 2012 Strategies and Techniques for Analyzing Microbial Population Structures (STAMPS) course at MBL, Woods Hole. The workshop that they have developed is a much needed addition to grad student and early career researcher education at a time when you need to (very quickly) become a computer scientist in addition to a biologist.

In addition, I attended the Light and Photosynthesis on Coral Reefs workshop at UNAM, Puerto Morelos, Mexico, and am extremely grateful to the instructors there for helping me to gain invaluable experience and education about the photobiology of coral reefs. Their commitment to educating the next generation of scientists in correct methodology and fostering interest in photobiology of coral reefs is commendable and I hope to build future collaborations with the faculty and students who attended the course.

I also want to thank the MCBS faculty who were willing to take time to discuss ideas or issues with me when I needed help. And, a special thank you to those who helped improve my teaching or helped make teaching labs a little bit easier: Bob Mooney, Bernie Schultz, Dr. Thomas Pistole, Dr. Cheryl Whistler, Dr. Michael Lee, Dr. Chris Bauer, Dr. Jessica Bolker, and Dr. Estelle Hrabak.

While my research has basically been my life for the past five years, I was also somehow fortunate to become friends with many wonderful grad students and others that I met during my time here, who were key in providing moral support and comic relief and I hope to count on these friends and colleagues for many years to come.

Lastly, I want to acknowledge my family and my other friends who have supported me through every endeavor in life, and have been incredibly understanding and supportive during these five years. I especially counted on the support and help from my sweetheart and “other half”, Bob Swarthout (and our quirky but lovable furry four-legged kids), without which I would not be completing this dissertation.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** ........................................................................................................ iii

**LIST OF TABLES** .................................................................................................................... vii

**LIST OF FIGURES** ................................................................................................................... viii

**ABSTRACT** ................................................................................................................................ x

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong> ..........................................................</td>
<td>1</td>
</tr>
<tr>
<td><strong>I. NITROGEN FIXING SYMBIOTIC ASSOCIATIONS IN THE MARINE ENVIRONMENT</strong></td>
<td>5</td>
</tr>
<tr>
<td>Introduction ........................................................................</td>
<td>5</td>
</tr>
<tr>
<td>Marine Nitrogen-Fixing Symbioses ....................................</td>
<td>18</td>
</tr>
<tr>
<td>Evolution and Ecology of Marine Nitrogen-Fixing Symbiosis</td>
<td>32</td>
</tr>
<tr>
<td><strong>II. MOLECULAR EVIDENCE FOR DIFFERENT SYMBIOTIC PROKARYOTIC COMMUNITIES IN GEOGRAPHICALLY DISTANT POPULATIONS OF THE CARIBBEAN SPONGE, XESTOSPONGIA MUTA</strong></td>
<td>37</td>
</tr>
<tr>
<td>Introduction ........................................................................</td>
<td>37</td>
</tr>
<tr>
<td>Methods and Materials ....................................................</td>
<td>40</td>
</tr>
<tr>
<td>Results ...............................................................................</td>
<td>47</td>
</tr>
<tr>
<td>Discussion ...........................................................................</td>
<td>52</td>
</tr>
<tr>
<td><strong>III. NITROGEN BIOGEOCHEMISTRY IN THE CARIBBEAN SPONGE, XESTOSPONGIA MUTA: A SOURCE OR SINK OF DISSOLVED INORGANIC NITROGEN?</strong></td>
<td>74</td>
</tr>
<tr>
<td>Introduction ........................................................................</td>
<td>74</td>
</tr>
<tr>
<td>Methods and Materials ....................................................</td>
<td>78</td>
</tr>
<tr>
<td>Results ...............................................................................</td>
<td>90</td>
</tr>
<tr>
<td>Discussion ...........................................................................</td>
<td>95</td>
</tr>
</tbody>
</table>
IV. METATRANSCRIPTOME ANALYSIS OF THE SPONGE XESTOSPONGIA MUTA: INSIGHT INTO FUNCTION AND ACTIVITY OF THE PROKARYOTIC COMMUNITY

Introduction ................................................................. 127
Methods and Materials .................................................. 131
Results ........................................................................... 134
Discussion ...................................................................... 140
SYNTHESIS ..................................................................... 180
LIST OF REFERENCES .................................................... 191
LIST OF TABLES

TABLE 2.1....................................................................................................................................60
TABLE 2.2....................................................................................................................................61
TABLE 3.1..................................................................................................................................107
TABLE 4.1..................................................................................................................................158
TABLE 4.2..................................................................................................................................159
TABLE 4.3..................................................................................................................................160
LIST OF FIGURES

FIGURE 2.1 ............................................. 62
FIGURE 2.2 ............................................. 63
FIGURE 2.3 ............................................. 64
FIGURE 2.4 ............................................. 65
FIGURE 2.5 ............................................. 66
FIGURE 2.6 ............................................. 67
FIGURE 2.7 ............................................. 68
FIGURE 2.8 ............................................. 70
FIGURE 2.9 ............................................. 71
FIGURE 2.10 .......................................... 72
FIGURE 2.11 .......................................... 73
FIGURE 3.1 ........................................... 108
FIGURE 3.2 ........................................... 109
FIGURE 3.3 ........................................... 110
FIGURE 3.4 ........................................... 112
FIGURE 3.5 ........................................... 113
FIGURE 3.6 ........................................... 116
FIGURE 3.7 ........................................... 119
FIGURE 3.8 ........................................... 120
FIGURE 3.9 ........................................... 122
FIGURE 3.10 ......................................... 123
FIGURE 3.11 ......................................... 124
FIGURE 4.1 ........................................... 161
FIGURE 4.2 ........................................... 162
FIGURE 4.3 ........................................... 163
FIGURE 4.4 ........................................... 164
Sponges have long been known to be ecologically important members of the benthic fauna on coral reefs. Recently, it has been shown that sponges, and their symbiotic microbes, are also important contributors to the nitrogen biogeochemistry of coral reefs. Here, I investigate the ecology and physiology of the microbial community associated the ecologically dominant sponge, *Xestospongia muta*. A natural experiment was conducted with *X. muta* from three different locations (Florida Keys, USA; Lee Stocking Island, Bahamas, and Little Cayman, Cayman Islands) to compare nitrogen cycling and prokaryotic community composition. The dissolved inorganic nitrogen (DIN) fluxes of sponges were studied using nutrient analysis, stable isotope ratios, and isotope tracer experiments. Results showed that the fluxes of DIN were variable between locations but clearly showed that *X. muta* can be either a source or sink of DIN. Stable isotope values of sponge and symbiotic bacterial fractions indicate that the prokaryotic community is capable of taking up both $\text{NH}_4^+$ and $\text{NO}_3^-$, and there is potential for translocation of labeled N from the symbiotic bacteria to the host. The prokaryotic
community composition of *X. muta*, and the variability of this community across the Caribbean were quantified using 454 pyrosequencing of the 16S rRNA gene. Phlyogenetic analysis showed differences between the sponge prokaryotic community and the surrounding bacterioplankton. Additionally, both symbiont and bacterioplankton populations were different between locations. In addition to the recovery of many sequences from bacterial phyla commonly found in sponges, a diverse archaeal community was also recovered from *X. muta* including sequences representing the phyla Euryarchaeota and Thaumarchaeota. Transcriptomic analysis for *X. muta* and its symbionts revealed a similar prokaryotic community composition to the metagenetic analyses indicating an active and diverse symbiotic community. Additionally, gene specific analyses combined with preliminary metatranscriptome data indicate the presence of genes involved in nitrogen cycling including *nifH* (nitrogen fixation), *amoA* (ammonia oxidation), *norB* (denitrification), and *nirK* (denitrification). Nitrogen cycling in *X. muta* appears to be more complex than previous studies have shown. These results have important ecological implications for the understanding of host-microbe associations, and provide a foundation for future studies addressing the functional roles these symbiotic prokaryotes have in the biology of the host sponge and the nutrient biogeochemistry of coral reefs.
INTRODUCTION

Coral reefs are unique and important ecosystems, providing essential ecosystem services (e.g., protection of coastlines, source of income and nutrition for local people) for many people in tropical coastal regions, as well as providing effective mitigation of hurricane energy for coastlines (Moberg and Folke 1999). Coral reefs are also hotspots of biodiversity, estimated to harbor 25% of marine species (Spalding and Grenfell 1997). We now know that the success of coral reef organisms in oligotrophic tropical waters is largely due to symbiotic associations with microorganisms. Our understanding of this relationship between microbes and important coral reef organisms has rapidly gained attention in the scientific community. Several studies have shown that complex microbial communities are in symbiosis with key reef members, including sponges and corals (Fiore et al. 2010), and that these communities have functional roles in the ecology, biogeochemistry and overall health and functioning of coral reefs. Furthermore, it has been shown that these microbial communities respond to shifts in environmental conditions with potentially detrimental results to the host (Dinsdale et al. 2008; Ainsworth et al. 2010).

Sponge-microbe symbioses appear to be unique among many well-documented symbioses (e.g., squid-\textit{Vibrio fisheri} (Nyholm and McFall-Ngai 2004), legume-rhizobia (Pawlowski and Bisseling 1996)) in that there is a wide diversity of symbionts that likely exhibit a range of interactions with the sponge host. Abundance of microbes has been documented by several studies and has indicated the existence of high microbial
abundance (HMA) and low microbial abundance (LMA) sponges, although the underlying factors leading to these differences are not well known and both HMA and LMA sponges are common on coral reefs. Abundance of microbes in HMA sponges have been estimated at $10^9$ microbial cells per ml of sponge or as much as 40-60% of the sponge volume, while LMA sponges typically contain microbial abundances similar to that of seawater ($10^5$ cells per ml) (Webster and Hill 2001; Hentschel et al. 2003). Interestingly, while most sponge microbiology studies have focused on characterizing the community of HMA sponges, a recent study on several LMA sponge species indicated that these appear to have unique prokaryotic communities that differ from HMA sponges, although they are less diverse (Giles et al. 2013). Documenting the diversity of symbionts in any type of sponge has proved challenging due to the filter feeding ecology of sponges and the difficulty in culturing the associated microbes. Symbiotic prokaryotes within sponges have generally been documented throughout the entire sponge; however, in some sponges there may be differences in the community composition between the inner mesohyl and outer mesohyl layers of the sponge (e.g., phototrophic prokaryotes concentrated in the outermost layers) (Vacelet and Donady 1977; Webster et al. 2001; Usher et al. 2001; Brück et al. 2008).

The maintenance of the microbial communities of sponges is an area of active research; however, some insight has been gained as to the transmission of certain microbes. Vertical transmission is generally considered indicative of a close symbiotic relationship, although, corals for example, are well known for their symbiosis with dinoflagellates even though they have horizontal transmission. Current evidence for sponge microbial communities suggests a combination of vertical and horizontal...
transmission to maintain symbiosis (Taylor et al. 2007; Schmitt et al. 2008), but how the sponge recognizes potential symbiotic bacteria as opposed to food or pathogens is still not fully known. However, sponges do have a relatively complex and effective innate immune system (Müller and Müller 2003), and two mechanisms have been proposed for the differential uptake of bacteria by sponges: 1) symbionts are specifically recognized by the sponges and not digested or 2) bacteria use extracellular capsules to evade detection by sponge cells (Wilkinson et al. 1984 in Taylor et al. 2007). Evidence so far seems to support the latter. Capsules have been noted on sponge-bacteria in several studies (Wilkinson 1978, Friedrich et al. 1999), and sponges will preferentially digest seawater microbes rather than sponge derived microbes (Werhl et al. 2007). However, evidence for specific interaction between sponges and microbes, based on the presence eukaryotic-like domains such as ankryin repeats and tetracopeptide repeats in symbiont genomes, have been documented in two recent studies (Siegl et al. 2010, Thomas et al. 2010).

There is still much we do not know about the numerous roles of many of these microbes or the functional relationship between the microbes and their hosts or including a basic understanding of the full taxonomic and functional diversity of these microbial communities. Characterization of these communities will increase our understanding of their role in the ecology coral reefs and how reefs might respond to future environmental changes. Here, I quantify to examine the taxonomic and functional diversity of microbes associated with the prominent barrel sponge on Caribbean coral reefs, *Xestospongia muta*, with an emphasis on their role in nutrient cycling on coral reefs.

It is well documented that coral reefs around the world are being degraded due to multiple factors (Lesser 2004; Hoegh-Guldberg et al. 2007). In a report by the World
Resources Institute, it was documented that 75% of the world's coral reefs are considered threatened when potential local threats (e.g., coastal development, watershed-based pollution) and thermal stress are considered (Burke et al. 2011). One of the key findings of the report is that changes in climate and ocean chemistry (e.g., ocean acidification) represent significant and growing threats (Burke et al. 2011). These threats not only affect corals but other vitally important coral reef members such as sponges, which as the second most abundant taxon on coral reefs significantly influence the structure and function of coral reefs (Diaz and Rutzler 2001). *Xestospongia muta* is abundant on Caribbean reefs and due to the large size and the high abundance of this sponge, it is likely that microbially mediated nutrient cycling within *X. muta* influences nutrient availability on coral reefs in the Caribbean. A better understanding of the contribution of important reef members such as the sponge *X. muta*, and its symbiotic bacteria, to nutrient cycling on the coral reefs will increase our understanding of coral reef resilience to environmental changes.

In the following chapters I will first introduce marine symbioses, specifically those involving nitrogen fixation as nitrogen is a limiting nutrient on coral reefs; I then address three main questions in regards to the prokaryotic community of *X. muta*: 1) What is the composition of the prokaryotic community and does this change over large spatial scales? 2) What is the impact of *X. muta* and its prokaryotic community on the cycling of nitrogen on coral reefs? And, 3) What is the metabolic capability of the prokaryotic community and what are the potential interactions between the sponge host and their symbiotic prokaryotes?
CHAPTER I

NITROGEN-FIXING SYMBIOTIC ASSOCIATIONS IN THE MARINE ENVIRONMENT

Introduction

The ability to convert atmospheric nitrogen into ammonia is restricted to members of the Bacterial and Archaeal domains in a process known as nitrogen fixation. Biological nitrogen fixation (BNF) has a long evolutionary history (Raymond et al. 2004) and contributes significantly to the amount of "new" nitrogen available to a wide variety of terrestrial, aquatic and marine organisms (Falkowski 1997; Galloway et al. 1995). Nutrient availability, but especially nitrogen, influences the trophic biology and ecology of all organisms and ultimately their ecological distribution and abundance. In the marine environment, a ratio of 106:16:1 of carbon:nitrogen:phosphorus has been described for the open ocean planktonic primary producers. Known as the Redfield ratio,

1 An excerpt of this chapter has been published as: Fiore CL, Jarett JK, Olson ND, Lesser MP (2010) Nitrogen fixation and nitrogen transformations in marine symbioses. Trends in Microbiology 18 (10): 455-463
this ratio was long believed to reflect the requirements for phytoplankton growth (Redfield et al. 1963) with nitrogen often cited as a limiting macronutrient. The role of nitrogen as a limiting nutrient in the oceans over ecological timescales has generally been accepted as a paradigm. However, the oceanographic community has recently begun to grapple with the idea of multiple resource co-limitation on primary productivity in the worlds oceans instead of a simple Liebig’s Law of the Minimum approach where only a single resource limits productivity (Arrigo 2005). This is a fundamental shift in thinking about nutrient biogeochemistry in the world’s oceans and it is initiating many new studies on the physiology of phytoplankton that are essential for making accurate carbon flux calculations (Karl et al. 2002 Zehr and Ward 2002; Arrigo 2005). Recent research has also found that the Redfield ratio more accurately represents a global average of the planktonic community, rather than a specific requirement for the growth of phytoplankton (Klausmeier et al. 2004). Physical factors such as local oceanography (Li and Hansell 2008; Church et al. 2002), exogenous input from nutrient runoff, and aeolian deposition (Dong et al. 2000; Fanning 1989) are now known to influence deviations from this ratio, which still supports the growth of primary producers. Additionally, biological factors such as the microbially mediated nitrogen fixation, nitrification and denitrification have recently become more appreciated as processes influencing biogeochemistry and nutrient stoichiometry (Arrigo 2005; Ward et al. 2007; Capone et al. 2008).

In the marine environment, nitrogen fixation was underestimated by early studies (Capone and Carpenter 1982), but recent studies estimate fixation rates closer to that of terrestrial environments (90-130 Tg N yr\(^{-1}\)) (Galloway et al. 1995) or higher (Karl et al. 2002; Quigg et al. 2003). One early reason for this underestimation was the low
contribution estimated for the cosmopolitan genus of free-living marine nitrogen-fixing bacteria *Trichodesmium* spp. (Capone et al. 1997) but most of the gap in nitrogen budgets was filled with the discovery of large numbers of oceanic unicellular cyanobacteria using molecular approaches (Montoya et al. 2004; Zehr et al. 2007). These discoveries, and the realization of the importance of marine sources of nitrogen in the global nitrogen budget has highlighted the need for continuing research into the complex cycling of nitrogen in the marine environment (Zehr and Ward 2002; Ward et al. 2007). While nitrogen cycling is also influenced by anthropogenic impacts (Vitousek et al. 1997) and physical forcing (Zehr and Ward 2002), microbial transformations of nutrients provides the fundamental underpinning to understand these processes and to improve our knowledge of the dynamics of microbial nitrogen cycling (Ward et al. 2007). The implications of nutrient transformations by marine microbes on seawater nutrient composition, global nutrient cycling, and plankton population distributions have been reviewed and highlight the complexity and importance of microbes in nitrogen cycling (Arrigo 2005; Ward et al. 2007). Many studies have focused on free-living diazotrophs, which fix nitrogen and are responsible for most of the new nitrogen in marine ecosystems. There are, however, many gaps in our knowledge of the diversity of bacteria that contribute to nitrogen fixation. For example, nitrogen fixation by symbiotic bacteria with zooplankton and phytoplankton (Braun et al. 1999; Zehr et al. 2000; Thompson et al. 2013) and non-planktonic organisms (Carpenter and Culliney 1975; Mohamed et al. 2008) are new discoveries and could potentially contribute to filling in the unknowns for nitrogen budgets in specific habitats. Newly discovered symbiotic nitrogen fixers now occur
frequently, and these symbioses influence host ecology and may have potentially large impacts on global nutrient cycling.

Environmental nitrogen limitation occurs in habitats such as mid-ocean gyres and shallow tropical waters, because of deep, stable thermoclines that prevents the upwelling of nutrients into the euphotic zone, or in the deep sea where nutrient poor food sources such as wood (Carpenter and Culliney 1975) can be found. However, a number of organisms have coevolved into symbiotic relationships with diverse group of prokaryotes that can perform biological nitrogen fixation (BNF) to overcome this limitation.

Nitrogen fixation is the transformation of dinitrogen gas ($N_2$) the most abundant form, but biologically unavailable, of nitrogen on the planet into biologically available ammonia ($NH_3$). Symbiotic bacteria that can fix nitrogen are not limited to the marine environment and many terrestrial nitrogen-fixing symbioses are often utilized in agronomy for their ability to replenish nutrients in the soil. Due to the important agricultural role nitrogen-fixing plant symbiosis play around the world, these provide some of the best-understood models of bacterial symbiosis. The processes of initiation, recognition, infection and biochemical communication between the host plant and nitrogen-fixing bacteria are well described (e.g. Garg and Geetanjali 2007). In freshwater systems, aquatic plants such as Azolla, which host heterocystous cyanobacteria as symbionts, as well as free-living cyanobacteria are major sources of fixed nitrogen and are well-documented participants in nitrogen cycling in aquatic ecosystems (Kalff 2002). Additionally, a majority of the early work on the physiology and genetics of nitrogen fixation was conducted on culturable free-living terrestrial bacteria in the genera Klebsiella, Azotobacter and Clostridium (reviewed in Klipp et al. 2004). Conversely, diversity and physiological
ecology of marine diazotrophic symbioses are not as well understood. This review encompasses historical and recent studies on the nitrogen fixation process with a particular focus on marine symbiotic associations. We also address where significant gaps exist in our knowledge of nitrogen fixation in these unique symbioses.

**Biological Nitrogen Fixation**

Nitrogen fixation is an energetically expensive process in which the two nitrogen atoms of a dinitrogen molecule are each reduced to ammonia. This process requires two enzymes, which form the nitrogenase complex. The Fe protein is a homo-dimer encoded by the \( nifH \) gene, and has a 4Fe:4S core. The Fe protein is reduced by the electron mediator, ferredoxin, and then reduces the second enzyme involved in the fixation process, the Mo-Fe protein. The Mo-Fe protein is a four-subunit \( \alpha_2\beta_2 \) protein with \( \alpha \beta \) dimers that are coded for by the \( nifD \) and \( nifK \) genes respectively, and contains a unique Fe/Mo-cofactor. The Mo-Fe protein reduces the \( \text{N}_2 \) molecule and is then subsequently reduced by the Fe-protein, completing the cyclic process.

In addition to the \( nif \)HDK genes, researchers have identified 17 other genes in the nitrogenase gene operon; the function of a number of these still remains unknown. Some of the additional genes of known function encode for protein cofactors or products necessary for the biosynthesis of inorganic cofactors, or proteins involved in electron transport (Dixon and Kahn 2004). The rest of the suite of \( nif \) genes with known functions include those that code for regulators of transcription and translation of the nitrogenase complex, as well as the overall fixation process. The reduction of the dinitrogen is an energetically costly process, requiring 16 ATP molecules and 8 reduced electron carriers
per molecule of dinitrogen reduced, necessitating tight regulation of this process and the biosynthesis of associated enzymes (Tripplett 2000). Expression of nitrogen regulating genes is tightly controlled at the transcriptional and post-transcriptional levels, and varies from species to species, often involving numerous genes on multiple operons (Merrick 2004). Global regulation of nitrogen metabolism is controlled by several regulatory proteins (Ntr) such as NtrC, B, L, rpoN (σ54), GlnK and GlnD that interact either directly or indirectly with the nif operon. Additionally, both negative (nifL) and positive (nifA) regulators of transcription are present within the nif operon itself (Klipp et al. 2004).

Ammonia (NH₃) and oxygen (O₂) are two main factors influencing the expression of nif genes and nitrogen fixation (Tripplett 2000). Sufficient concentrations of NH₃ obviate the need for nitrogen fixation. When NH₃ is low, the NtrC protein is activated and promotes transcription of nifA, which in turn promotes transcription of the nif operon, enabling nitrogen fixation to occur. When NH₃ is in excess and therefore nitrogen fixation is unnecessary, the activity of NtrC and NtrL proteins is inhibited, resulting in inhibition of nif transcription. The NifL protein is regulated by glnD and glnK gene products, while NtrC activity is regulated by NtrB (an enzyme that is both a kinase and phosphatase), which in turn, is regulated by the nitrogen status of the cell (Merrick 2004).

The other main controlling factor of nif gene expression is O₂, which down regulates nitrogenase transcription when present, and oxidizes and therefore irreversibly degrades the Fe-S cofactors of the nitrogenase enzyme (Merrick 2004). The NifL protein contains a flavin cofactor (FAD) that is critical for O₂ sensing by the protein (Merrick 2004). When O₂ levels are high, the NifL protein suppresses NifA activity so that the
synthesis of $O_2$-sensitive nitrogenase is shut down at transcription. Regulation of the $nif$ operon and other operons involved in nitrogen fixation is extremely complex and can vary among species, and research into the genetics of nitrogen fixation has predominantly been conducted in free-living bacteria (for review see Klipp et al. 2004), but how well this represents symbiotic bacterial gene expression of nitrogen-fixing genes is largely unknown. To illustrate this point, a recent study on the well-known *Azolla*-cyanobacteria symbiosis, collected the first molecular data on post-translational modification of NifH in a symbiotic system (Ekman et al. 2008). They found two forms of NifH, a modified form that was likely inactive in the cyanobiont, and an unmodified active form. The concentration of the active form of the protein NifH was over 9 times higher in the cyanobiont than the free-living cyanobacteria (NifK was also 2.5 times more abundant). This correlated very well with the observation of increased rates of fixation when the bacteria are *in hospite* as compared to free-living (Watanabe 1987). The mechanism for this post-translational modification appears to differ between symbiotic and free-living systems; the mechanism and the identity of the modification remain to be determined.

**Protection of nitrogenase from oxygen**

The transcriptional and translational regulation of nitrogenase synthesis described above is important in regulating the process of nitrogen fixation. Many diazotrophs, however, currently live and have evolved in environments where $O_2$ is present at atmospheric concentrations or under conditions where oxygenic phototrophs, such as cyanobacteria, evolve $O_2$. Under such conditions nitrogenase can be quickly degraded which decreases, or stops, nitrogen fixation and requires the energetically expensive
replacement of the proteins such that these organisms have evolved several strategies for protecting nitrogenase from $O_2$. Broadly speaking these strategies involve reducing $O_2$ concentrations in the cell, conformational changes of the nitrogenase protein, or spatial or temporal separation of the fixation process from oxygen. In symbioses, the host, the symbiont, or both partners may contribute to the protection of nitrogenase from oxygen, often using the same basic mechanisms utilized by free-living diazotrophs or adapting them to life in hospite. Additionally, some hosts create protected microaerobic, or anaerobic, environments for their nitrogen-fixing symbionts, such as root nodules in legumes or the gut in termites (Breznak 1982; Udvardi and Day 1977).

Reducing the concentration of oxygen in the cell is a common oxygen protection strategy and can be accomplished in several ways. Aerobic respiration both scavenges $O_2$ and produces ATP and reductant that can be used in nitrogen fixation (Hochman and Burris 1981). Hydrogenase enzymes are present in all aerobic nitrogen-fixing organisms (Fay 1992), and oxidize the $H_2$ produced by nitrogenase in the oxyhydrogen or Knalgas reaction, thereby removing an inhibitor of nitrogen fixation ($H_2$) (Mortenson 1978), conserving reducing power, and consuming $O_2$. Antioxidants such as catalase, peroxidase, and particularly superoxide dismutase play important roles in reducing the reactive oxygen species that are produced by photosynthesis and aerobic respiration, and can inactivate nitrogenase in addition to damaging other cellular components (Fay 1992).

Nitrogenase can be protected from the damaging effects of $O_2$ by temporarily “shutting off” and associating with a 2Fe-2S protein (Shetna’s protein, Scherings et al. 1977; Scherings et al. 1983). This mechanism has been studied in Azotobacter spp.; in this system the 2Fe-2S protein oxidizes the Fe protein (encoded by nifH), and an oxygen-
tolerant complex containing the oxidized Fe protein, the Mo-Fe protein, and the 2Fe-2S protein is formed. Although this complex has not been directly observed in diazotrophs other than *Azotobacter*, results from other studies imply that a similar mechanism may exist in the cyanobacteria *Oscillatoria* sp. (Stal and Krumbein 1985) and *Anabaena* sp. (Pienkos et al. 1983).

Spatial separation of nitrogen fixation from oxygen is accomplished in several ways. Organisms may live in environments such as sediments that are anoxic or microaerobic. Additionally, some filamentous cyanobacteria form specialized cells (heterocysts) for nitrogen fixation, which have thick glycolipid layers that limit the diffusion of oxygen into the cells (Thiel et al. 2004). Differentiation of heterocysts is a highly complex and tightly regulated process, and only occurs during nitrogen starvation (Wolk et al. 1994). One of the drawbacks of heterocysts is their high metabolic cost; their maximum frequency in a filament of cells is only about 25%, and it is estimated that about 4-5 vegetative cells are required to support each heterocyst with reductant and carbon skeletons for assimilation of nitrogen (Thiel et al. 2004). Some symbiotic diazotrophs, such as the cyanobiont of *Azolla*, also form heterocysts (Gebhardt and Nierzwicki-Bauer 1991). It is presumed that the host provides some metabolic support or regulatory cues, or both, because heterocysts occur with much greater frequency in symbiosis than in the free-living state (Adams 2000).

Other diazotrophs utilize a temporal separation of nitrogen fixation from oxygenic photosynthesis. In free-living cyanobacteria, nitrogen fixation typically occurs at night in non-specialized, photosynthetically competent cells (Bergman et al. 1997). For most of the cyanobacteria in which these patterns have been studied the cycles of photosynthesis
and nitrogen fixation are endogenous and persist even under constant environmental regimes (Bergman et al. 1997). Temporal separation was also documented in the symbiotic diazotrophs of a scleractinian coral. Cyanobacterial symbionts of the coral *Montastraea cavernosa* display higher fixation rates in early morning and in the evening, when oxygen concentrations in coral host tissue are lower, and are likely utilizing products of photosynthesis as an energy source (Lesser et al. 2007). Because of the complexity of the processes of nitrogen fixation and photosynthesis, many possible mechanisms exist for control and separation; in symbiotic lifestyles, this gives both host and symbiont many methods of regulation and cooperation.

A small number of free-living cyanobacterial species use a combination of spatial and temporal separation to protect nitrogenase from oxygen. The most notable of these is *Trichodesmium*, a non-heterocystous marine filamentous cyanobacterium that fixes nitrogen during daylight hours in a small number of cells that contain nitrogenase (Bergman and Carpenter 1991; Janson et al. 1994). Although nitrogen fixation is localized, these cells differ from heterocysts in that they appear to contain all the necessary photosynthetic machinery and are otherwise undifferentiated. Maximal nitrogen fixation occurs in the middle of the light period, when quantum yields of chlorophyll fluorescence are low and carbon fixation decreases transiently, effectively separating the periods of highest nitrogen fixation and oxygen evolution (Berman-Frank et al. 2001). Although neither the spatial nor temporal separation in *Trichodesmium* is very efficient independently, in combination they allow nitrogen fixation to take place. The use of both temporal and spatial separation as in *Trichodesmium* presents another
mechanism that could be at work in symbiotic nitrogen fixation associations, though it has yet to be reported.

A novel resolution of the co-occurring processes photosynthesis and nitrogen fixation was recently documented in a new strain of UCYN-A cyanobacteria, where the bacteria have a photoheterotrophic lifestyle (Bothe et al. 2010). This means they do not evolve O$_2$ from Photosystem II; rather they generate energy from light using cyclic photophosphorylation. Furthermore, UCYN-A was recently shown to live in symbiosis with a unicellular alga (Thompson et al. 2013). In this relationship, the cyanobacterium gains fixed carbon from the alga in exchange for fixed nitrogen (Thompson et al. 2013). Additionally, the genome of UCYN-A is highly streamlined, likely as a result of living in symbiosis with the alga (Tripp et al. 2010). Further investigation into the UCYN-A-prymnesiophyte symbioses may shed light on the evolution of symbiosis, particularly in regards to plastid evolution, and will have likely have important implications in our understanding of nitrogen cycling in the oceans. Additionally, it remains to be seen if a similar mechanism underpinning the relationship between UCYN-A and the alga is employed in other diazotrophic symbioses as well.

**When are nitrogen-fixing symbionts advantageous?**

The ability to fix nitrogen, either endogenously or via a symbiont, allows organisms to escape ecological limitations in a variety of habitats. Organisms may be able to expand their spatial niches into low-nutrient environments, or their dietary niches to resources that have low combined nitrogen content, and can thus reduce competition or gain a competitive advantage. Organisms harboring symbiotic diazotrophs in mutualisms
with highly evolved mechanisms for the efficient transfer of fixed nitrogen products are potentially more likely to gain these benefits.

Nitrogen-fixing organisms are ubiquitous in aquatic and terrestrial environments with limited inorganic nitrogen resources (Zehr et al. 2000; Vitousek et al. 2002). Lichens may be the best terrestrial example; they are common not only in desert soil crusts (Belnap 2002; Eskew and Ting 1978), but also on rocks (Seneviratne and Indrasena 2006) and lava flows, where they help create soil and begin the process of primary ecological succession (Crews et al. 2001; Kurina and Vitousek 2001). In all of these environments, both competitors and predators are scarce or absent. Similarly, the associations between higher plants and rhizobia (e.g., in legumes) allow these plants to thrive even in nitrogen-poor soils, a clear advantage over their competitors. Coral reefs are a marine ecosystem where it is advantageous to host nitrogen-fixing symbionts. The growth of and high productivity of coral reefs in these oligotrophic waters has been seen as a paradox since Darwin’s voyage on the Beagle. It is now believed that both plankton predation (Hamner 1995; Hamner et al. 2007), which transfers bioavailable nitrogen through the food web, and nitrogen fixation, coupled with efficient nutrient cycling, on and around reefs contribute significantly to the nitrogen requirements of reefs (Webb et al. 1975). Diazotrophs are common in the water column (Hewson et al. 2007), on the substrate (Charpy et al. 2007; Larkum et al. 1988), and in associations with invertebrates such as sponges (Mohamed et al. 2008), corals (Lesser et al. 2004) and possibly tunicates (Odintsov 1991; Pael 1984). Other classically nitrogen-limited habitats include mid-ocean gyres; areas with low rates of water exchange, such as the Red Sea and the Persian Gulf; and places where there is strong seasonal stratification of the water column. In
these habitats, high C:N ratios limit the growth of phytoplankton, with the exception of nitrogen-fixing species, such as *Trichodesmium* spp., which commonly form large blooms (Sellner 1997), and various diatoms which host nitrogen-fixing symbionts (Foster et al. 2009; Foster and Zehr 2006). The abundance of these organisms and the large contribution they make to the nitrogen budget (Carpenter and Romans 1991; Zehr et al. 2000) of the ocean clearly illustrate the ecological advantages of nitrogen fixation in an open ocean environment.

Environments that are nitrogen sufficient overall may still contain food sources that are nitrogen-poor and organisms that can subsist on these resources occupy a less competitive dietary niche. Wood-eating termites, which harbor symbiotic diazotrophs in their guts, are a typical terrestrial example (Breznak 1982). Nitrogen-fixing microbes of various types are present in the guts of many other arthropods as well (Kim et al. 2001; Lilburn et al. 2001; Nardi et al. 2002). Like termites, marine shipworms also consume wood, which has a very high C:N ratio, as a primary food source. They harbor dense cultures of nitrogen-fixing bacteria in a specialized organ called the gland of Deshayes (Waterbury et al. 1983), which provide the host with fixed nitrogen and allow the shipworm to survive on a wood diet (Lechene et al. 2007; Luyten et al. 2006). Diazotrophic symbionts can also allow their hosts to adapt to fluctuating food resources. Herbivorous sea urchins in temperate regions consume kelps and other seaweeds in which nitrogen content varies seasonally, and the diazotrophic bacteria in their guts regulate nitrogen fixation accordingly to supply sufficient nitrogen resources to the urchin year round (Guerinot and Patriquin 1981).
Marine nitrogen-fixing symbioses

The advantages of hosting diazotrophic symbionts are clear, but specifics such as the degree of integration between host and symbiont physiologies, transmission of symbionts, and significance of nutrient fluxes to hosts and/or symbionts vary from species to species. For some of the better-studied marine symbiotic systems such as shipworms, sea urchins, diatoms and dinoflagellates, as well as reef building corals and sponges some or all of these aspects of the symbiotic system are known.

Shipworms

The wood boring mollusks of the family Teredinidae, commonly known as shipworms, have caused significant destruction of man-made wooden structures placed in the ocean. Referred to as the termites of the sea, these mollusks bore holes in wood and utilize the cellulose as their sole nutrient source (Distel 2003). Woody plants contain only 0.03% to 0.1% nitrogen, thus these shipworms require an additional source of nitrogen to supplement their diet. Acetylene reduction assays of three species, *Psiloteredo megotar*, *Lyodrus pedicellatus*, and *Teredo navalis*, revealed that the shipworms benefit from microbially fixed nitrogen (Carpenter and Culliney 1975). For these three species, fixation rates were inversely correlated with dry weight, with higher fixation rates for juveniles when compared to the adults. This difference was attributed to juveniles having small gills making them inefficient filter feeders, and *T. navalis* collected from the oligotrophic Sargasso Sea exhibited a fixation rate 20 times higher compared to the other species which were found only in coastal habitats. The higher
fixation rates for the juveniles and Sargasso Sea samples indicates that the symbiotic bacteria are able to adjust their nitrogen contribution to the system depending on the host's nitrogen deficiency (Carpenter and Culliney 1975). Axenic cultures of nitrogen-fixing bacteria (family Spirillaceae) isolated from the cecum of *L. pedicellatus* fix nitrogen anaerobically, but are also capable of aerobic growth (Carpenter and Culliney 1975). Further analysis of this symbiosis was based on isolates obtained from five additional species of shipworms: *Bankia gouldi, T. navalis, Teredo furcifera, Teredo bartischii,* and *Psiloteredo healdi.* The symbionts identified were endosymbiotic, residing within bacteriocytes in the gland of Deshayes, which is part of the host gill tissue. The same isolate was obtained from all six species and was not only capable of nitrogen fixation but also possessed cellulolytic activity. The ability of the symbiotic bacteria to digest cellulose indicates an additional host derived benefit of the symbiosis: the bacteria assist in the digestion of the host's sole food source (Waterbury et al. 1983).

Ribotype analysis of the symbionts resulted in the description of a novel genus of γ-proteobacteria named *Teredinibacter,* with the original isolate given the species name *turnerae* (Distel et al. 1991; Distel et al. 2002b). Additional symbionts within this genus have been identified and all host bacteriocytes contain a single dominant symbiont ribotype, as well as a second less abundant ribotype (Distel et al. 2002a; Distel et al. 1991; Luyten et al. 2006). Although multiple symbionts were identified in this work the functional role of each is unknown. The incorporation of fixed nitrogen was found to vary between bacterial types within the bacteriocytes, indicating the potential for individual ribotypes to contribute different amounts to the total fixed nitrogen, or that the observed bacterial types may be at different developmental stages (Lechene et al. 2007).
It is also speculated that the presence of multiple symbionts allows the host to digest multiple wood types or produce various celluolytic enzymes (Distel et al. 2002a). Interestingly, a recent study sequenced the genome of *T. turnerai* and did not find genomic adaptations one would expect to see in an obligate symbionts such as reduced genome size or loss of % G + C content (Yang et al. 2009). In fact, features common to free-living bacteria, such as a large portion of genes for producing secondary metabolites and genes for protection from bacteriophage infection, led the authors to conclude that *T. turnerai* is a facultative endosymbiont that was recently or is currently existing as a free-living organism.

Investigations into the transmission of symbionts in teredinids indicate they may have the ability to acquire symbionts both vertically and horizontally (Sipe et al. 2000). Symbionts have yet to be cultured from juvenile shipworms, and the bacteria appear well suited to a free-living life style, implying a horizontal mode of transmission for a number of species (Greene and Freer 1986; Imam et al. 1990). However, a PCR based investigation of the species *Bankia setacea* revealed a novel species within the *Teredinibacter* genus as well as its presence in host eggs and ovaries, indicating vertical transmission (Sipe et al. 2000).

**Sea Urchins**

Another type of marine organism that has evolved a symbiotic association with nitrogen-fixing bacteria due to nitrogen limitation in its diet are some species of sea urchins. Nitrogen-fixing bacteria associated with sea urchins occur worldwide and in multiple urchin species, though the presence and abundance of the symbionts varies
depending on season, and food being consumed by the host. The urchin

*Strongylocentrotus droebachiensis* contains nitrogen-fixing symbionts in its intestinal tract, but only in individuals that are fed on kelp, which is a nitrogen deficient food source in the late spring and summer when it is primarily consumed (Guerinot et al. 1977). Further investigation using multiple species from different habitats (Nova Scotia kelp bed, Barbados eelgrass, rocks and shallow reef and Canada Northwest territory shallow water) found a significant inverse relationship between kelp nitrogen content and urchin nitrogenase activity (Guerinot and Patriquin 1981), and a similar relationship for eelgrass but with a lower monthly average. Nitrogen fixation in tropical urchins was highly variable, probably as a result of the differences in diet; *Tripneustes ventricosus* had the highest fixation rates and feeds on eelgrass which has low nitrogen content, while *Diadema antillarum* and *Echinometra lacunter* had much lower fixation rates and feed on coral polyps, encrusting algae, diatoms and macroalgae, which have comparatively higher levels of nitrogen. Variation within the temperate urchins was also observed and appeared to fluctuate in response to the level of nitrogen in the food source. Up to 5-fold seasonal fluctuations in the nitrogen content of kelp were documented. Correspondingly, seasonal variation in fixation rates in urchins suggested that in the summer and fall when nitrogen content of kelp is low, nitrogen fixation could be an important supplement of nutrients. Guerinot and Patriquin (1981) suggested that there could be suppression of nitrogenase by metabolic products because some urchins that tested negative for nitrogen fixation still contained nitrogen-fixing bacteria.

The nitrogen-fixing bacteria appeared to reside specifically in the gut and were identified as *Vibrio spp.* based on cultures, however, we are not aware of any molecular
identification of nitrogen-fixing bacteria associated with the gut of any sea urchin.

Transmission of nitrogen fixing bacteria in *S. droebachiensis* was initially investigated using feeding experiments (Guerinot and Patriquin 1981). Groups of urchins were starved for three months and nitrogen fixation rates and bacterial counts were quantified. These urchins were then fed kelp or placed in tanks of filtered seawater with antibiotics, followed by putting the urchins in unfiltered seawater and provided kelp again. The first group of urchins was initially negative for nitrogen fixation activity and only low numbers of nitrogen-fixing bacteria were present in the gut, however, after two weeks of being fed excess kelp, nitrogen fixation was observed and these urchins contained significantly higher numbers of nitrogen-fixing bacteria. Urchins exposed to antibiotics did not exhibit measurable rates of nitrogen fixation and no nitrogen-fixing bacteria were detected. Urchins from the antibiotic treated seawater that were placed in seawater with no antibiotics and fed kelp started to exhibit nitrogen fixation activity and nitrogen-fixing bacteria were detected after 18 days. While the described experiments indicate there is environmental acquisition of nitrogen-fixing bacteria by the urchin *S. droebachiensis*, there has been no documentation of vertical transmission for this system. Additionally, there is little information on the diversity of these symbionts, which could easily be assessed using molecular techniques, and is thus a potential topic for future investigations into this system, particularly given the commercial and ecological importance of many urchin species.
**Dinoflagellates**

The unicellular protists, diatoms and dinoflagellates, often inhabit nutrient poor ecosystems such as mid ocean gyres and thus symbiotic associations with nitrogen-fixing bacteria are potentially advantageous. Blooms of dinoflagellates and diatoms can often be detected remotely from space and due to their high abundance they play an important role in nutrient cycling in oceanic and freshwater systems. Many of these planktonic protists harbor symbiotic cyanobacteria and the association can vary from obligate endosymbiont to epibiont. The dinoflagellate genera *Ornithocercus*, *Histioneis* and *Citharistes* are common in deep waters of tropical and subtropical seas, and their abundance fluctuates seasonally in the Gulf of Aqaba's surface waters, peaking during the autumn when stratification and nutrient limitation are high (Gordon et al. 1994). These heterotrophic dinoflagellates have symbionts (phaeosomes) of the genera *Synechococcus* and *Synechocystis*, which reside within the horizontal groove in *Ornithocercus* and *Histioneis* and in a special chamber within the cells in *Citharistes* (Gordon et al. 2004).

Gordon et al. (1994) proposed that the hosts are likely receiving fixed nitrogen and fixed carbon from the symbionts, providing the host species with an advantage over other heterotrophic plankton during times of nutrient limitation. This hypothesis is supported by the spatial and temporal distribution of these dinoflagellates, which is correlated with ambient nutrient levels. Dinoflagellate populations are homogeneous throughout their depth of occurrence in early spring following winter mixing, but their numbers increase in surface waters when summer stratification begins and peak in the fall when nitrate levels at the surface reach their lowest concentration (Gordon et al. 2004). Similarly, symbiotic associations of cyanobacteria and dinoflagellates in the Bay of Bengal were
more common in the spring than during the summer and winter monsoon; the spring intermonsoon is also a period of nitrate limitation in surface waters (Jyothibabu et al. 2006). Oxygen levels in the seawater are also high at this time, thus the host may provide a unique low oxygen environment for nitrogen fixation (Gordon et al. 1994; Jyothibabu et al. 2006).

The dinoflagellates *Ornithocercus magnificus* and *O. steinii* were investigated further to determine if nitrogen fixation is occurring in the cyanobacterial symbionts. Using the acetylene reduction technique no nitrogen fixation was observed, though there is still the potential for carbon transport between the host and symbionts (Janson et al. 1995). The diversity of these symbionts has only been investigated at a phenotypic level through electron microscopy and immunolabeling. Lucas et al. (1991) initially described four cell types of symbiotic cyanobacteria associated with marine dinoflagellates, and immunolabeling of these symbionts indicated nitrogenase expression (Foster et al. 2006), but further investigation into these symbioses is required in order to gain a better understanding of the diversity of symbionts and their mode of transmission from one host generation to the next.

**Diatoms**

Diatoms, another common member of the phytoplankton community, are well-documented hosts of cyanobacterial symbionts, but nitrogen fixation has only been investigated in a few species. Observations on the freshwater species *Rhopalodia gibba* have revealed the presence of 1 to 10 obligate spheroid bodies capable of nitrogen fixation per cell (Deyoe et al. 1992; Prechtl et al. 2004). This symbiosis is highly specific,
indicating a highly evolved relationship between the host and symbiont, to the extent that the symbionts are referred to as spheroid bodies in the evolutionary process of becoming organelles (Prechtl et al. 2004). In support of this theory, a genome study of the symbiont revealed extensive reduction in size, similar to that of the mitochondria and chloroplasts when compared to their free-living relatives (Kneip et al. 2008).

Nitrogen-fixing cyanobacteria (*Richelia intracellularis*) are known symbionts of the marine diatom, *Hemiaulus huckii* (Villareal 1991). These cyanobacteria are able to reside on the surface or internally in the host (Janson 1999; Venrick 1974), and provide a major source of fixed nitrogen that supports growth in the host (Carpenter et al. 1999). Additionally, *R. intracellularis* is known to synthesize different heterocyst glycolipids than free-living heterocystous cyanobacteria (Schouten et al. 2013). The enodsymbiotic *R. intracellularis* produced C-5 sugars (ribose) as opposed to C-6 sugars characteristic of free-living cyanobacteria, and produced longer chain lengths. The differences in glycolipids between symbiotic and free-living cyanobacteria may be an adaptation to high O₂ environment in the host, and the authors suggest that this may provide a novel marker for identifying these symbioses (Schouten et al. 2013).

Heterocystous cyanobacteria such as *R. intracellularis* and *Calothrix* spp. have been observed in other open ocean diatoms including *Rhizosolenia* spp., *Chaetoceros* spp. (Janson 1999) and *Bacteriastrum* spp. (Venrick 1974). Cyanobacteria related to the nitrogen-fixing *Cyanothece* spp. have been observed in association with the tropical marine diatom *Climacodium frauenfeldianum* (Carpenter and Janson 2000), though as with many of these associations, much of the physiology of the symbiosis is purely speculative as is the mechanism of transmission of symbionts from diatom to diatom. A
recent study however, showed that the growth rates of diatoms and cyanobacteria in symbiosis are higher than their free-living counterparts (Foster et al. 2011). Additionally, nitrogen fixation rates for cells in symbiosis were up to over 400 fold higher than the free-living cells, indicating that there is some sort of feedback between the host and symbiont (Foster et al. 2011). Janson et al. (1999) and Pretchel et al. (2004) both suggested that based on the high degree of specificity between the host diatom and the cyanobacterial symbiont, vertical transmission is likely for *Rhizoselenia* spp., *Hemiaulus* spp. and *Rhopalodia gibba*; but this has not been confirmed with molecular techniques.

**Corals**

Coral reefs provide an opportunity to study relatively newly discovered associations between nitrogen-fixing bacteria and hosts such as corals and sponges that provide important ecosystem services (Moberg and Folke 1999). These symbioses have allowed hosts to be ecologically successful in the low combined nitrogen of these tropical coastal environments. Nitrogen-fixing bacteria have been hypothesized to be symbionts of reef building corals since Williams et al. (1987) identified nitrogen fixation in the skeleton of *Acropora variabilis*. This preliminary research did not investigate the diversity of symbionts or whether the host or symbiont gained any benefit from the association. However, based on the results from light/dark acetylene reduction assays it was hypothesized that the symbionts were photosynthetic cyanobacteria (Williams et al. 1987). Similarly, Odinstov (1987) observed nitrogenase activity in the hydrocoral *Millepora* and zooxanthellae, unicellular green algae and bacteria were noted in the skeleton as well as the tissue. It was not clear, however, which member of the
consortium was fixing nitrogen and how similar this association might be to that of hermatypic corals (Odintsov et al. 1987). Further investigation into the coral-diazotroph symbiosis identified nitrogen-fixing bacteria of the class γ-proteobacteria associated with the skeleton of *Favia favus* (Shashar et al. 1994), which displayed higher rates of fixation under illumination, and with exposure to glucose-enriched seawater. These results indicate that the nitrogen-fixing bacteria are utilizing glucose from either the host or algal co-symbiont as an energy source, implicating a potential benefit for the symbionts in this relationship.

Terminal restriction length polymorphisms (T-RFLP) 16s rRNA analysis identified a diverse and dynamic community of coral bacterial symbionts (Rohwer et al. 2002), including many possible nitrogen-fixing bacterial associates of reef building corals. However, it was not until 2004 that an endosymbiotic nitrogen-fixing bacteria symbiont was conclusively identified (Lesser et al. 2004). The cyanobacterial symbionts transfer fixed nitrogen to the coral’s algal symbionts, and fixed carbon in the form of glycerol has been hypothesized be transferred to the cyanobacteria to fuel nitrogen fixation (Lesser et al. 2007). Corals with cyanobacteria also appear to benefit through increased feeding deterrence (Jarett 2012), and the effects of cyanobacteria on growth and reproduction are currently being investigated. Initial analysis of the bacterial diversity for this system indicates that this symbiosis is comprised of a single cyanobacterial species (Lesser et al. 2004). In comparison a much greater diversity of nitrogen-fixing bacteria, including cyanobacteria and proteobacteria, were observed in close association with tissue of the Hawaiian corals *Montipora capitata* and *Montipora flabellata*. The benefits to each partner in these symbioses as well as whether the bacteria are endo- or epibionts remains
unknown; however, a correlation was found between *Vibrio nifH* transcript number and algal symbiont abundance for *M. capitata*, suggesting a close relationship (Olson et al. 2009). Members of the Vibrionaceae have previously been documented to be capable of nitrogen fixation (Shieh and Lin 1994; Chimetto et al. 2008). A diverse community of bacteria were identified but a conserved phylogenetic cluster of bacteria in the Vibrionaceae family were found only in association with *M. capitata*, and a less conserved cluster of γ-proteobacteria were only associated with *M. flabellata*. This symbiont specificity may indicate coevolved, highly specific, symbiotic associations. Nitrogen-fixing bacteria have also been found in association with the mucus of the Brazilian coral *Mussimilia hisipda*. Cultured *Vibrio* spp. on nitrogen free media were identified as being capable of nitrogen fixation through acetylene reduction assays, while the benefit to the host and symbiont in this association remains unknown (Chimetto et al. 2008).

For most of these associations, it is unknown if there is direct transfer of fixed nitrogen to the coral, if the symbiosis is a mutualism, parasitism, or neither, or even how widespread this symbiosis is among coral taxa and how this impacts the health and growth of the coral host. Whether symbionts are acquired by horizontal or vertical transmission is also largely unknown. For tissue-associated microbes in particular, the transmission of photosynthetic dinoflagellate symbionts of reef-building corals may offer clues. *Symbiodinium* are transmitted both vertically and horizontally, but vertical transmission appears to create or facilitate more specific symbioses (Stat et al. 2008), and is more common in areas geographically isolated from other reefs, such as Hawai‘i (Lajeunesse et al. 2004). The diversity of nitrogen-fixing bacteria that have been
identified with different coral species indicates the potential for a range of types of symbiotic associations between the coral host, its alga, and nitrogen-fixing symbionts.

**Sponges**

Microbial nitrogen transformations have also been observed in sponges, another prominent coral reef organism. Sponges are an important component in reef ecosystems as they consolidate rubble, have high biodiversity and biomass, and influence constituents of the surrounding water (Diaz and Rutzler 2001; Reiswig 1973; Taylor et al. 2007). The microbial communities associated with sponges are phylogenetically diverse (Simister et al. 2012), and are believed to be sponge-specific in many cases and have been shown to be important in nitrogen cycling (Taylor et al. 2007; Hoffman et al. 2009). Recently, denitrification, anammox and nitrification were conclusively observed in the cold water sponge *Geodia baretti* (Hoffmann et al. 2009), while nitrification and nitrogen fixation have been documented in multiple studies on sponges (Wilkinson and Fay 1979; Southwell et al. 2008). The association of nitrogen-fixing bacteria with sponges was first reported by Wilkinson and Fay (1979) using the acetylene reduction assay on sponges from the Red Sea. This activity was attributed to cyanobacteria because the sponges that tested positive all contained cyanobacteria, while the sponge that tested negative contained no cyanobacteria. The acetylene reduction assay was also used to detect nitrogen fixation in the sponge *Halichondria* sp. from the coast of Taiwan (Shieh and Lin 1994). However, it was found that $^{15}$N$_2$ tracer studies were more reliable in the detection of nitrogen fixation (Wilkinson et al. 1999), and tracer or natural abundance stable
isotopes have been used several times since then to examine nitrogen fixation in sponges (Southwell et al. 2007; Mohamed et al. 2008; Chapter 3).

The first, and so far only, study to use molecular genetic techniques to examine nitrogen fixation in sponges was by Mohamed et al. (2008); \textit{nifH} was detected in sponges that exhibited low $\delta^{15}$N$_2$ values indicative of nitrogen fixation (\textit{Ircinia strobilina} and \textit{Mycale laxissma} from the Florida Keys). Mohamed et al. (2008) also provided some insight into the diversity of symbiotic nitrogen-fixing bacteria using \textit{nifH}, uncovering a surprisingly diverse community. While nitrogen fixation in sponges was first attributed to cyanobacteria, heterotrophic bacteria such as Vibronaceae species are also associated with this process (Shieh and Lin 1994). Similarly, Mohamed et al. (2008) detected \textit{nifH} genes from $\alpha$- and $\gamma$-proteobacteria, cyanobacteria, and \textit{Desulfovibrio} spp; however, expression of \textit{nifH} was detected from cyanobacteria only. The authors noted that this may be due to time of sampling (day time) and expression of \textit{nifH} genes by other bacteria may take place at other times. Interestingly, most of the sequences obtained from the samples were novel, based on low similarity scores of the \textit{nifH}-deduced amino acid sequences, relative to known sequences in the database. Furthermore, sponge-derived \textit{nifH} DNA sequences were not observed in the water column, whereas sequences closest to free-living cyanobacteria such as \textit{Trichodesmium thiebaudii} were found in the water column samples but not in the sponges, indicating that there is likely a coevolved symbiosis between these bacteria and the host sponge.

Transmission of some microbial symbionts in sponges has been shown to occur vertically, with some sponge gametes or larvae harboring a single bacterial species or a complex assemblage including Cyanobacteria, Proteobacteria, Actinobacteria,
Bacteroidetes, and Planctomycetes (Oren et al. 2005; Usher et al. 2008). Much of this information has been based on electron microscopy, which provides visual confirmation of the microbial community but generally does not provide phylogenetic or metabolic information. Recent studies using 16s rRNA have identified a diverse community of bacteria in the larvae and gametes of sponges (Steindler et al. 2005), indicating that at least some of the symbiotic community is vertically transmitted in sponges. However, analysis of functional genes such as \textit{nifH} is needed to determine if nitrogen-fixing bacteria are transmitted vertically and if fixation is occurring in the larvae. Recent molecular studies of sponge symbionts have provided a base of experimental investigation into sponge-associated diazotrophs, but additional work is needed. For example, there is little known about the transfer of fixed nitrogen products to the host (Wilkinson et al. 1999). We are only beginning to understand which bacteria are fixing nitrogen, how they are transmitted, how prevalent this symbiosis is, and how this affects nutrient cycling within the sponge and on the reef.

The data from studying urchins as well as some of the other well studied symbiotic systems such as dinoflagellates, shipworms and plants indicate the main benefit to the host is receiving a source of fixed nitrogen, while the benefit to the symbionts, if these are mutualisms, is often less clear. As suggested by various studies the biggest advantage for the symbionts may be that the host provides a suitable microaerobic environment for the symbionts to fix nitrogen (Gordon et al. 1994, Jyothibabu et al. 2006), although in sponges the syymbiots may also benefit from the active pumping of water (and therefore nutrients) through the sponge. Many of the nitrogen-fixing bacteria are also photosynthetic and so the host may provide a suitable environment for this process as
well, depending on where the bacteria are found on or within the host. However, for some associations such as in urchins, suppression of nitrogen fixation appears to occur, indicating that perhaps the symbiont is also benefiting from the association as it continues to reside in the host. In the corals, it has been speculated that the symbiotic cyanobacteria, because they are capable of minimal photosynthesis, are living heterotrophically and gain carbon in the form of glycerol from the host (Lesser et al. 2007). Hosts appear to provide a safe refuge from predation, which may be an added benefit to these symbionts; however, more research into the specificity and evolution of these associations is needed before these questions can be answered.

Evolution and Ecology of Marine Nitrogen Fixing Symbiosis

Symbiotic associations between nitrogen-fixing bacteria and their eukaryotic hosts cover a continuum of symbiotic associations. These continuums may be characterized in a number of ways three of which are: (1) spatial and temporal aspects of the association, (2) specificity of members involved, and (3) the necessity of the association to the individual members (Starr 1975).

The temporal and spatial continuum is apparent when looking at nitrogen fixing symbiosis involving diatoms and dinoflagellates. In the summer months when environmental biologically available nitrogen levels are low, diatoms and dinoflagellates are more abundant and more often associated with nitrogen-fixing symbionts (Jyothibabu et al. 2006). There is also variation in symbiont location on and in the host in these symbioses, wherein symbionts are sometimes found as epibionts and sometimes as
endobionts as with the diatoms *Chaetoceros spp.* and *Rhodophilia gibba* respectively (Foster et al. 2006; Pretchel et al. 2004).

The host and symbiont specificity of the association also varies among host species and may also fall along the commensal, mutualistic, parasitic continuum. Highly specific symbioses are generally considered more evolved symbiotic associations and are largely mutualistic in nature (Douglas 1995). A high diversity of nitrogen-fixing bacteria is associated with tropical marine sponges with bacterial symbiont representatives from all major nitrogen-fixing bacterial taxonomic groups (Mohamed et al. 2008). In contrast to this diverse symbiosis only a few bacterial symbiont types are present in marine shipworms (Distel et al. 2002a). For these symbioses it is clear that the symbiosis is more evolved for shipworms when compared to that of sponges.

The third continuum is the necessity of the symbiosis to the partners involved. For some members the association is obligatory in that they are unable to survive aposymbiotically or in a free-living state. When examining the *Montastraea cavernosa* symbiosis a colony of an aposymbiotic host can be found neighboring a symbiotic host, indicating that this association although maybe beneficial to both host and/ or symbiont is not obligatory (Lesser et al. 2004). On the other end of the continuum is the intracellular spheroid body symbiont of the diatom *Rhopalodia gibba* where neither host nor symbiont has been observed or has cultured without the presence of the other member of the association (Pretchel et al. 2004). Where each of the individual marine nitrogen-fixing symbiosis falls along these continuums can reveal insight into the ecological and evolutionary aspects of these associations. Because of this, it is necessary to address gaps
in our understanding of these symbioses in regards to their placement along these continuums.

**Future research directions**

The paucity of knowledge about most marine nitrogen-fixing symbioses is important to recognize, and address, because we do not know the impact that these associations have on the ecology of either the hosts or of the larger ecosystems of which they are a part of. The presence of symbiotic diazotrophs can raise fundamental questions about our understanding of their hosts. It is thought that the photosynthetic dinoflagellate symbionts of corals are nitrogen limited, and that this limitation controls their growth and prevents them from multiplying faster than the host cells (Falkowski et al. 1993). However, in the *Montastraea cavernosa*/*cyanobacterial* symbiosis, nitrogen fixed by the cyanobacteria is transferred to the dinoflagellates in quantities sufficient to generate a stable isotope signal consistent with the utilization nitrogen fixation without causing a significant increase in zooxanthellae growth rates (Lesser et al. 2007), challenging this classical understanding.

On a larger scale, nitrogen budgets have been calculated for many ecosystems, but marine nitrogen budgets seldom take into account diazotrophic symbioses that appear to be uncommon. We do not yet know the magnitude of their contribution, but it may be substantial, particularly considering that many symbioses may remain undiscovered. For instance, sponges have not typically been considered major players in nitrogen cycling on coral reefs, but in light of recent studies (Mohamed et al. 2008) this should be reconsidered. Similarly, the contributions of symbionts of dinoflagellates and diatoms
have not always been considered in the estimation of global and oceanic nitrogen budgets.

The potential impact of changing environmental conditions, such as eutrophication, ocean acidification, and climate change on these symbioses is also unknown. Biologically available nitrogen is often abundant in eutrophic areas; this could obviate the need for nitrogen fixation and make these symbioses between nitrogen fixing bacteria and a variety of hosts less common. Symbioses that are obligate for each partner, such as the *Rhopalodia gibba* and spheroid body partnership, might have greatly reduced fitness in such environments and perhaps be extirpated from these areas. High atmospheric CO$_2$ concentrations that cause ocean acidification increase nitrogen fixation rates in the free-living cyanobacterium *Trichodesmium* (Hutchins et al. 2007; Levitan et al. 2007), but living inside a host-mediated environment may impose different challenges and restrictions on symbionts, or none at all if the host buffers against these changes.

Ocean acidification poses threats to calcifying organisms (Doney et al. 2009; Jokiel et al. 2008), some of which (shipworms, corals, sea urchins, and some sponges) are hosts to diazotrophic symbionts. Reductions in abundance or species richness of these hosts may lead to a loss of diversity in symbionts as well. Rising sea surface temperatures due to climate change also affect many of the host organisms discussed above or their other symbionts, particularly corals, sponges, and plankton (Hays et al. 2005; Hoegh-Guldberg et al. 2007; Webster et al. 2008). Future research aimed at advancing our understanding of the evolutionary history and the potential global impact of these important symbioses should focus on three broad goals: (1) characterization of known symbioses, particularly with respect to nutrient and energy exchange between partners, to determine if
relationships are mutualisms; (2) discovery of new diazotrophic symbioses; (3) describing the impacts of these symbioses on their respective hosts and the ecosystems of which they are a part.
Chapter II

MOLECULAR EVIDENCE FOR DIFFERENT SYMBIOTIC PROKARYOTIC COMMUNITIES IN GEOGRAPHICALLY DISTANT POPULATIONS OF THE CARIBBEAN SPONGE, *XESTOSPONGIA MUTA*²

**Introduction**

Sponges comprise one of the most basal lineages of metazoans, dating back to *ca.* 600 million years ago (Bergquist 1978). Additionally, symbioses between prokaryotes and sponges are almost as old as the sponges themselves (Wilkinson 1984; Thiel et al. 1999). As a result the symbiotic prokaryotes of sponges have played a significant role in the evolutionary ecology and physiology of sponges (Thacker 2005; Freeman and Thacker 2011). To better understand host-microbe evolution as well as the functional roles that sponges have in their ecosystems, we need additional taxonomic and functional characterizations of the sponge prokaryotic symbiotic community in a wide range of

---

² Excerpts from this chapter have been submitted for publication as: Fiore CL, Jarett JK, Lesser MP. Molecular evidence for different symbiotic prokaryotic communities in geographically distant populations of the Caribbean sponge, *Xestospongia muta*. MicrobiologyOpen. submitted.
sponge species. It was first suggested by Hentschel et al. (2002) that there are widespread, sponge-specific microbial communities that are distinct from the surrounding environment. Furthermore, they hypothesized that there was a core community of prokaryotes common to all sponges, particularly high microbial abundance sponges (HMA, *sensu* Hentschel et al. 2003). Following this study, two subsequent analyses of publically available sequence databases confirmed many of the original sponge specific clusters discovered (Taylor et al. 2007; Simister et al. 2011). The most recent assessment of sponge specific sequence clusters has shown clusters within 14 bacterial phyla and within the Archaea as well as eukaryotic Fungi (Simister et al. 2011). Additional support for the occurrence of taxonomically similar prokaryotic communities from geographically distant sponges was recently reported in a metagenetic study of 32 sponge species (Schmitt et al. 2011). Interestingly, this study also showed that there was evidence for a tropical clade of sponge microbes that are distinct from temperate and cold water sponge prokaryotic communities, potentially indicating the existence of subpopulations of sponge symbionts defined by environmental factors (Schmitt et al. 2011).

The most abundant groups present in sponges are of bacterial origin and include representatives from the phyla Actinobacteria, Chloroflexi, Proteobacteria, Cyanobacteria, Acidobacteria, and the candidate phylum ‘Poribacteria’ (Taylor et al. 2007; Schmitt et al. 2011; Simister et al. 2011). Taxonomically, sponge prokaryotes of lower rank are also diverse, as the first pyrosequencing studies of sponge symbiont 16S rRNA genes revealed a higher number of operational taxonomic units (OTUs) than previously reported at the 'species level' (Lee et al. 2010) and 'genus level' (Webster et al. 2011).
The use of high throughput sequencing methods to quantify symbiotic prokaryotic communities is increasing (Sogin et al. 2006; Uroz et al. 2010; Barott et al. 2011), largely due to the difficulty of culturing many prokaryotes and advances in sequencing technologies. The short reads (approx. 300 bp) generated by these next generation technologies have been shown to contain sufficient taxonomic information (Liu et al. 2007b; Quince 2009; Kunin et al. 2010) and can provide insight into rare members of sponge prokaryotic communities (Webster et al. 2010a; b; Lee et al. 2010; Schmitt et al. 2011). However, full-length 16S rRNA gene sequences are still important for establishing phylogenetic “guide” trees (Ludwig et al. 1998; Taylor et al. 2007) and to populate the databases that unknown sequences can then be compared to.

The aim of this study is to understand and characterize the taxonomic variability of the prokaryotic community of the ecologically dominant giant barrel sponge, *Xestospongia muta*, also known as the 'redwood of the reef' (McMurray et al. 2008), across the Caribbean basin. *Xestospongia muta* is a prominent member on coral reefs throughout the Caribbean and when collected from the same depth and conditions from multiple sites can be used to quantify and compare the taxonomic composition of the prokaryotic symbionts in different sponge populations as in a “natural experiment” (sensu Diamond (1986)). Most of the previous work on *X. muta* has been conducted in one location (Florida Keys) and this study includes that site and from a comparative perspective expands our understanding of the prokaryotic communities associated with this important coral reef sponge. Furthermore, *X. muta* is known to harbor Archaea (Lopez-Legentil et al. 2010), a group increasingly recognized as having important roles in nutrient cycling (Hallam et al. 2006; Hatzenpichler 2012), however, little is known about
the taxonomic composition of Archaea in this sponge. We performed pyrosequencing of 16S rRNA genes from X. muta samples collected from reefs near Key Largo, Florida, Lee Stocking Island, Bahamas, and Little Cayman, Cayman Islands. In this study I hypothesized that not only will the sponge symbionts be significantly different from bacterioplankton populations in the overlying water column but there will also be location-specific differences in the prokaryotic community composition of X. muta within the Caribbean.

**Materials and methods**

**Sample collection**

Replicate sponges (n=6) were sampled at approximately 15 m depth from each of three locations: Rock Bottom Reef, Little Cayman, Cayman Islands (LC) (19°42'7.36" N, 80°3'24.94" W), North Perry Reef, Lee Stocking Island (LSI) (23°47'0.03" N, 76°6'5.14" W), Bahamas, and Conch Reef, Key Largo, FL (FL) (24°57'0.03" N, 80°27'11.16" W). All populations were sampled during the late spring and summer of 2011 where the maximum photosynthetically active radiation (PAR; 400-700 nm) irradiance at noon for these depths at all three locations is ~500-600 μmol quanta m⁻² s⁻¹ (Lesser unpublished). Sponge pieces were cut from the top rim of the sponge ("pie slice" of pinacoderm and mesohyl) and placed in a plastic bag with minimal amounts of seawater and placed on ice until reaching the laboratory. Each sponge sample was then placed in DNA buffer (20% DMSO, 0.25 M EDTA, saturated NaCl, (Seutin et al. 1991). All samples were kept frozen until reaching the University of New Hampshire where they
were maintained at -70°C. Water samples (n=3, 4 L each) were collected contemporaneously and filtered onto 0.22 μm filters (Whatman, USA) and frozen in DNA buffer.

**DNA extraction, PCR, and sequencing**

CTAB (hexadecyltrimethylammonium bromide) DNA extractions were performed on all sponge and seawater samples as follows. A section of each sponge sample that included both the outer pigmented layers and inner mesohyl of the sponge were cut into smaller pieces with a razor blade for processing. Filters were cut in half (half saved for later use) and also cut into smaller pieces for processing. Samples for both sponges and filters were placed in 600 μl of 2x CTAB mixture (Tris, pH 8.0 (0.0121 g ml⁻¹), NaCl (0.0818 g ml⁻¹, EDTA (0.00744 g ml⁻¹), CTAB (0.002 g ml⁻¹)) and homogenized with a pestle and with brief sonication. Proteinase k (5 μl of 20 mg ml⁻¹) was added and samples were incubated at 64°C for 3 h. Equal volume of chloroform was added to the samples followed by centrifugation at 12 000 x g for 10 min. The aqueous layer was transferred to a new tube and DNA was precipitated with equal volumes of 100% ethanol. The samples were spun again for 10 min followed by two washes with 70% ethanol, and then the pellet was allowed to dry before being resuspended in 30 μl of molecular grade water. Extractions were checked for quality and concentration using a NanoDrop spectrophotometer (2000c, Thermo Fisher, USA). In some cases, a phenol:chloroform:isoamyl alcohol (25:24:1) extraction was performed to clean the samples. Briefly, the suspension was brought to 100 μl volume, and 1/10 volume of KOAc was added followed by equal volume of phenol:chloroform:isoamyl alcohol and vortexed
for 5 s. Samples were centrifuged for 2 min at maximum speed and the aqueous layer was transferred to a clean tube. Three volumes of 100% ethanol were added to each sample and vortexed for 5 s, and then centrifuged for 10 min at maximum speed. The pellet was then washed twice with 500 µl of 70% ethanol (2 min of centrifuge in between), and then allowed to air dry.

The 16S rDNA of each sample was amplified and barcoded for multiplexed pyrosequencing using Titanium adapter sequences A (forward primer) and B (reverse primer), and a 10 bp barcode sequence added to the PCR primers. Primers designed to amplify Bacteria and Archaea (hypervariable V6 region) were used, consisting of the forward primer U789F (5'-TAGATACCSSGTAGTCC-3') and the reverse primer U1068R ('-CTGACGRCRGCCATGC-3') (Baker et al. 2003; Lee et al. 2010). Three reactions of 25 µl were performed for each sample and pooled prior to electrophoresis. The PCR consisted of 0.25 µl of 50x Titanium Taq polymerase (Clontech, Mountain View, CA, USA), 2.5 µl of 10x Titanium Taq buffer, 0.2 mM dNTPs (Promega, Madison, WI, USA), 0.4 µM of each barcoded primer and 25 ng of genomic DNA template. Reactions were performed with a Thermocycler (Eppendorf Mastercycler, Germany) using the following protocol: initial denaturation for 5 min at 95°C, 26 cycles of 95°C for 30 s, 53°C for 30 s, and 72°C for 45 s, followed by 6 min at 72°C. PCR products were then electrophoresed on a 1% agarose gel and purified with Qiaquick gel extraction kit (Qiagen, Valencia, CA, USA). Samples were then purified with Agentcourt AMPure XP bead kit (Beckman Coulter, Danvers, MA, USA) and quantified with a DyNA Quant 200 fluorometer (Hoefer, Holliston, MA, USA) per manufacturer's protocol prior to combining all samples in equilmolar concentration. Samples were
pyrosequenced on the ROCHE/454 GS FLX+ platform (Roche, Branford, CT, USA) at the University of Illinois W.M. Keck Center for Comparative and Functional Genomics (Urbana-Champaign, IL, USA).

**Taxonomic assignment and diversity estimations of OTUs**

The Quantitative Insights Into Microbial Ecology (QIIME) pipeline (Caporaso 2010) on the Amazon Elastic Compute Cloud (EC2) was used for all analyses except where noted. Sequences from sponge and water samples, as well as coral samples (Jarett, 2012) were analyzed together up to the step of clustering into operational taxonomic units (OTUs), following which, sponge and water samples were analyzed separately. Raw sequence reads were filtered for quality by discarding short reads (<200 bp), or reads with more than two mismatches with the primer sequence, or with ambiguous nucleotides, or with an average quality score less than 25. A custom Perl script based on the QIIME script "split_libraries.py", was used to trim primers from the sequences, assign reads to their sample of origin (based on MID tags), and reverse complement the reads originating from the B adapter (reverse reads) (changes from QIIME default: -l 300 -M 2 -b 10 -z). Trie clustering (QIIME team, unpublished, http://qiime.org) was used to collapse reads that are prefixes of each other into clusters and discard singleton reads. This technique has been used to rapidly and easily remove erroneous reads from pyrosequencing and provides an alternative 'noise' removal method to denoising, which is computationally intensive (Behnke et al. 2010). The uclust algorithm (Edgar 2010) was used to cluster the remaining reads into operational taxonomic units (OTUs) at 97% similarity (settings: --max_accepts 20, --max_rejects 500, --stepwords20, and --
word_length 12) and the most abundant sequences was selected as the representative sequence for each cluster. ARB (Ludwig 2004) was used to align representative sequences to the SILVA non-redundant reference database, release 108 (Pruesse et al. 2007) using the SINA plug-in. Aligned sequences were re-formatted in QIIME and with custom Perl scripts for use in the QIIME pipeline. The alignment was filtered and poorly aligned sequences (at least 50 consecutive nucleotides without gaps), flagged as a potential chimeric sequence by ChimeraSlayer (Haas et al. 2011), or with significant BLAST (Altschul et al. 1990) matches (E-value $< 1 \times 10^{-10}$ and at least 97% identity) to a custom database of likely contaminants (18S rDNA of demosponges, alveolates from SILVA reference database) were discarded.

Taxonomy was assigned to representative sequences using the RDP classifier and a minimum confidence cutoff of 0.8 (Wang et al. 2007) in QIIME. Assignments of "Root" or "Root:Bacteria" were rechecked by BLAST against the NCBI nr database, and was discarded if the top hit was not 16S rDNA. In QIIME, OTU tables were rarefied to equalize sampling depth across samples (step size=100, total seqs=7500, num-reps=20), then the alpha diversity metrics of observed species, Simpson diversity index, and Shannon diversity index were calculated and rarefaction curves were composed. The Chao1 metric was not used because it is based on the ratio of singleton to doubleton sequences and we removed singleton sequences. An approximately maximum-likelihood phylogenetic tree was built using Fasttree 2 (Price and Dehal 2010) in QIIME and was used to calculate weighted and unweighted UniFrac distance values and perform the UniFrac Monte Carlo significance test (Hamady et al. 2009). To assess the similarity between microbial communities in the sponge or water, Bray-Curtis distance values were
calculated based on rarefied OTU tables for non-phylogenetic diversity comparisons. Principle coordinates were then generated and used to create 2 dimensional plots (jackknifed) and to perform jackknifed sample clustering by UPGMA. Additionally, multidimensional scaling (MDS) and analysis of similarity (ANOSIM) were performed on the OTU table (created in QIIME) in the program PRIMER v6 (Clarke 1993; Clarke and Gorley). The OTU table was square root transformed prior to analysis and Bray-Curtis similarity metric was used. The R values produced by pairwise comparisons in the ANOSIM is the best indicator of differences between groups and an R of 0.5 was used as a critical threshold with values equal to or higher than 0.5 indicating a difference between group means (Clarke and Gorley).

ANOVA and the G test of independence were each utilized to examine potentially significant OTUs between two sample types (e.g., water vs sponge, LC sponges vs LSI sponges) using the OTU category significance tool in QIIME. The use of ANOVA allows for the determination of whether OTU relative abundance is different between categories (i.e., sample type or location), while the G test determines whether the presence or absence of an OTU is associated with a category.

**Phylogenetic Analyses**

Representative OTU sequences for specific taxonomic groups were selected for treeing. Sequences and their closest matches from GenBank using the blastn tool were aligned to the SILVA non-redundant SSU reference database (108) in ARB using the SINA plug-in and sequences were added to the tree using the parsimony quick add tool in ARB. The alignment was also used to build neighbor-joining and maximum parsimony
trees in the program MEGA version 5 (Tamura et al. 2011) and the three treeing methods were compared based on tree topology. A neighbor-joining tree is presented here and nodes are marked that are consistent with the other treeing methods. Sequences resulting from pyrosequencing were submitted to the CAMERA (Cyberinfrastructure for Microbial Ecology Research and Analysis, http://camera.calit2.net/) website under project accession CAM_P_0000957.

**Pigment analysis**

Lyophilized *X. muta* samples (*n* = 6) were cut into two ½ cm² pieces consisting of the outermost layer of tissue. One piece was used for chlorophyll extraction and the other was used for phycoerythrin (PE) extraction. Chlorophyll was extracted by homogenizing in 100% acetone (in a darkened laboratory), then gently mixed in 400 μl of MgCO₃ (1 g MgCO₃ in 14 ml double distilled water), followed by addition of 3 ml 100% acetone, vortexed and incubated at 4°C overnight. The mixture was then centrifuged at 1800 x g for 5 min and the supernatant was analyzed on a NanoDrop spectrophotometer (2000c, Thermo Fisher, USA) at 630, 647, 663, and 664 nm (90% acetone as blank).

Trichromatic equations (Jeffrey and Humphrey 1975) were used to calculate chlorophylls (*a, b, c*) (in μg ml⁻¹). Phycoerythrin was extracted following methods of (Wyman 1992). Briefly, the sponge samples were homogenized in 500 μl of 50 mM phosphate buffer pH 7.0 (PB) and incubated at 4°C overnight. The PB mixture was then sonicated for 15 s (duty 50%) with glass beads and centrifuged at 3220 x g for 8 min. The supernatant was then centrifuged at 100 000 x g for 60 min and the resulting supernatant was analyzed on a NanoDrop spectrophotometer (2000c, Thermo Fisher, USA) at 492 and 542 nm.
Assumptions of ANOVA were tested and only Chl $a$ : PE ratios did not pass the assumptions and could not be transformed to meet the assumptions. Differences between pigment concentrations at each location were determined using ANOVA with nontransformed data, followed by Tukey’s HSD. The non-parametric Kruskal-Wallis (KW) test was used to determine differences between pigment ratios at each location.

**Results**

**Phylogeography of X. muta prokaryotic communities using 454 pyrosequencing**

Pyrosequencing of the 16S rRNA genes yielded 323,542 sequence reads (average read length 289 nucleotides). Following quality filter steps and removal of singleton reads there were 233,469 reads that were then clustered into OTUs at 97% similarity. A total of 1,664 OTUs remained following removal of chimeric sequences and contaminants (i.e., 18S rRNA gene sequences, chloroplast sequences). 407 OTUs were recovered from the sponge samples and included 17 phyla (Acidobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Deferribacteres, Deinococcus-Thermus, Firmicutes, Gemmatimonadetes, Nitrospirae, Planctomycetes, Proteobacteria, Spirochaetes, Verrucomicrobia, Crenarchaeota, and candidate phyla ‘Poribacteria’, ‘TM7’, and ‘SBR1093’). Seawater samples yielded 1,458 OTUs and included 27 phyla. Additionally, sponges contained 185 unique OTUs, while 222 OTUs were shared with seawater OTUs.

A rarefaction analysis revealed that seawater samples were approaching an asymptote at around 7,000 sequences per sample but were not sufficiently sampled to capture the total diversity of the community (Fig. 2.1 A). The sponge samples did reach
an asymptote around 7,000 sequences per sample based on rarefaction analysis, and had a range of 7,648 to 14,247 sequences per sample (Table 2.1). Most of these sequences clustered into a few OTUs for both sponge and seawater samples (Fig. 2.1 B). Up to 279 OTUs in one sponge sample were recovered, and up to 514 OTUs per water sample was recovered. The alpha diversity estimates of observed species were similar to observed 'species-level' (97%) OTU number for each sample (Table 2.1) and the average number of observed species was estimated to be 201 for sponges, and 342 for water samples. The Shannon and Simpson diversity metrics showed the sponges to be more diverse than the water samples, and showed FL and LSI to be more diverse than LC for both water and sponge samples (Table 2.1).

The prokaryotic community composition of the seawater was significantly different from those of the sponges (Unifrac p-test, p < 0.05). Multidimensional scaling (MDS) showed that the sponges grouped together by location and the same was observed for the water samples (Fig. 2.2). Analysis of similarity (ANOSIM) revealed a significant difference between locations for sponges and water respectively (R = 0.5, p = 0.04 (sponges); R = 1, p = 0.04 (water)), with significant differences resulting from pairwise comparisons between locations using the combined sponge or water samples from each location (R = 0.5 – 0.6, p = 0.1 (all sponges); R = 1, p = 0.1 (all water)). Pairwise comparisons of individual sponges using the Unifrac p-test also yielded significant differences: XmLC1 is significantly different from XmFL1, XmLSI1, and XmLSI2 (Unifrac p-test, p < 0.05). Lastly, XmLSI3 is significantly different from XmLC2 and XmLSI2 (Unifrac p-test, p < 0.05).
Clustering based on weighted Unifrac distance (Fig. 2.3) also showed that sponges from each location were generally most similar to each other, with the exception of XmLSI1 and XmLC3. When unweighted Unifrac distance was used, which does not take into account OTU abundance, the clusters were generally the same, however, XmLSI3 and XmLC1 were more dissimilar compared to the other sponge samples (Fig. 2.4). Clustering based on Bray-Curtis distance yielded similar results to Unifrac based clustering with water and sponges forming separate clusters and within the sponge cluster, the sponge LC samples are generally more different from FL and LSI sponge samples than FL and LSI are from each other (Fig. 2.5). Principle coordinates analysis (PCoA) based on weighted and unweighted Unifrac distances also demonstrated more of a separation of samples by location with weighted Unifrac distance, particularly for XmLC samples, and clearly separated sponge and water samples (Fig. 2.6).

Specific differences in relative abundance of OTUs between sponge and water samples were observed for several major groups including Proteobacteria, Choroflexi, Poribacteria, and Gemmatimonadetes (Fig. 2.7 A). Within the individual sponge samples, the most obvious difference was that LC sponges generally harbored more cyanobacterial and proteobacterial OTUs and fewer OTUs within the Chloroflexi than sponge samples from other locations (Fig. 2.7 B). Comparison of class-level taxonomy highlighted even more differences between sponge and water samples (Fig. 2.8). For example, the OTUs classified as Proteobacteria in the water samples were almost all further classified as Alphaproteobacteria, while in the sponge samples they were roughly split between Alpha- and Deltaproteobacteria. Also, many of the OTUs classified as Bacteroides from the water samples were identified as Flavobacteria while Bacteroides OTUs from sponge
samples were dominated by Sphingobacteria. Lastly, differences in classification of
archaeal OTUs at the phylum and class level were also observed between sponge and
water samples (Fig. 2.8).

As most sequences fell within a low number of OTUs, when the top 90% of OTUs
are examined from sponges in each location, subtle differences in representative taxa are
revealed (Fig. 2.9). LC sponges have 12 taxonomic groups in the top 90% of OTUs,
whereas FL and LSI sponges have 17 groups, with Gammaproteobacteria and
Synechococcophycidae containing a higher number of OTUs in the LC sponges (Fig.
2.9). The classes, Gemmatimonadetes, EC214 (phylum SBR1093), Acidobacteria, and
Anaerolineae (phylum Chloroflexi) were present in the top 90% of OTUs in FL and LSI
sponges but not LC sponges. FL and LSI sponges contain the same taxonomic groups in
the top 90% of OTUs but the number of OTUs in several of the groups differ as can be
seen based on order of the listed groups in figure 5. In contrast, the seawater samples
contain only five taxonomic groups in the top 90% of OTUs (Fig. 2.9).

The OTU category significance test in QIIME identified OTUs that were
significantly more abundant in one group of samples than another group (Table 2.2). Of
particular note, OTUs identified as more abundant in sponges relative to seawater and
particularly so in FL and LSI sponges, included members of the ‘Poribacteria’,
Acidobacterium, Syntrophobacteraceae, Entotheonellaceae, Gemmatimonadetes,
Chromatiales, and Chloroflexi-4. In addition, OTUs identified as more abundant in
sponges, but particularly in LC samples, included taxa of Synechococcaceae and
Alphaproteobacteria.
Taxonomic Composition of Archaea in *X. muta*

Representative sequences from OTUs classified as Archaea were recovered from sponge and water samples and fell into three distinct clades (Fig. 2.10). Clades I and II were comprised of sequences representing the phylum Euryarchaeota. Clade I contained OTU sequences from both water samples and *X. muta* samples and were related to euryarchaeotes recovered from sediments and microbial mats. Clade II contained only OTU sequences from water samples and was most closely related to marine group II and III euryarchaeotes, although also included one sequence from the sponge *Axechina raspailoides* (Fig. 2.10). Clade III contained only OTU sequences from *X. muta* samples, and these were related to the previously described thaumarchaeotes *Nitrosopumilus maritimus* and *Candidatus Cenarchaeum symbiosum* of the group I.Ia Archaea. Some non-sponge derived sequences from other datasets were also included in clade III, preventing it from being classified as a sponge specific cluster. Other OTU sequences from sponge and water samples in the current study fell into the clade of Thaumarchaeota but were distinct from the clade III group of mostly sponge-derived sequences (Fig. 2.10).

Characterization of the photoautotrophic community by pigment analysis

Chlorophyll *a* and PE were detected in the outer layers of *X. muta* at all locations. Chl *a* was significantly different between locations (ANOVA, $F_{2,10} = 68.949$, $p < 0.05$), and pairwise comparisons showed FL to be significantly higher than both LC (TukeyHSD, $p = 1.0 \times 10^{-6}$) and LSI (TukeyHSD, $p = 6.8 \times 10^{-4}$) while LSI was significantly higher than LC (TukeyHSD, $p = 0.002$) (Fig. 2.11 A). Phycoerythrin was
also significantly different between locations (ANOVA, $F_{2,10} = 21.966$, $p = 0.0002$), and pairwise comparisons showed that FL was significantly higher than LC (TukeyHSD, $p = 0.0002$) and LSI was significantly higher than LC (TukeyHSD, $p = 0.002$) (Fig. 2.11 A). The ratio of Chl $a$ to PE was not significantly different between locations (KW, chi-squared = 3.6, df = 2, $p = 0.16$) (Fig. 2.11 B).

**Discussion**

The results of this study show that *X. muta* has a prokaryotic community that is distinct from the surrounding seawater, and that there is a significant effect of location for both the symbiotic prokaryotic community of the sponges and the bacterioplankton community. While the community composition overall was similar to that documented by Montalvo and Hill (2011), four additional phyla (Thaumarchaeota, Euryarchaeota, Spirochaetes, and the candidate phylum ‘Poribacteria’) were recovered from the sponge in the current study, and one phylum (candidate phylum ‘TM6’) was only recovered from seawater in the current study. This supports previous findings that the 300-400 bp products generated using 454 pyrosequencing have sufficient taxonomic information for diversity studies because of the coverage across hypervariable regions and the development of appropriate bioinformatic approaches (Liu et al. 2007a; Quince 2009; Kunin et al. 2010).

Diversity estimates for *Xestospongia muta* in this study were similar to the diversity estimates reported by a previous study Montalvo and Hill (2011). The Diversity metrics of Shannon and Simpson showed that sponge samples were more diverse than seawater
samples. This is likely a result of the water samples being dominated by a few OTUs, whereas the sponge samples exhibited a similar trend but had more OTUs and a greater number of sequences in those OTUs. The dominant phyla recovered from *X. muta*, based on relative abundance of OTUs, are similar to those recovered from studies on sponges in diverse locations and habitats (Taylor et al. 2005; Simister et al. 2011), including representatives from the Proteobacteria, Chloroflexi, Cyanobacteria, and 'Poribacteria'. Additionally, OTUs classified as Actinobacteria, Bacteroidetes, Acidobacteria, Gemmatimonadetes, and Thaumarchaeota also contributed to the diversity of the symbiotic prokaryotic community composition in *X. muta*.

All of the major groups of prokaryotes associated with sponges, including Archaea, have been shown to contain monophyletic sponge-specific clusters (Simister et al. 2011). Additionally, in a study comparing the prokaryotic community of the congeners *X. muta* from the Florida Keys and *Xestospongia testudinaria* from the Pacific, some species-specific clusters within the Chloroflexi, Acidobacteria, and Cyanobacteria were observed based on cloned 16S rRNA gene sequences from *X. muta* (Montalvo and Hill 2011). This suggests some specificity of the symbiotic community at the genus level for the host. 'Poribacteria' and Archaea however, were not recovered from *X. muta* or *X. testudinaria* (Montalvo and Hill 2011), probably due to primer bias issues (Suzuki and Giovannoni 1996) and the use of eubacterial primers (27F/1492R) by Montalvo and Hill (2011). 'Poribacteria' are considered a sponge-specific group, although sequences representing this group have also been documented in low abundance from water column samples (Webster et al. 2010a). Recently, the 'Poribacteria' genome was characterized using single-cell genomics, providing insight into the functional role of this ubiquitous
group of bacteria (Siegl et al. 2010). Genome characterization revealed a mixotrophic lifestyle for the ‘Poribacteria’, with the potential for carrying out denitrification, and the authors suggest that these are commensal bacteria within the sponge (Siegl et al. 2010). The ‘Poribacteria’ are also one of the most diverse groups of sponge symbionts (Schmitt et al. 2011) and was a relatively diverse group (11 OTUs, 9377 sequences) in the current study. Future characterization of additional ‘Poribacteria’ genomes or transcriptomes should yield considerable insight into the taxonomic and functional diversity that likely exists in this group of sponge symbionts.

Archaeal sequences recovered from X. muta were classified as Euryarchaeota or Thaumarchaeota, groups now commonly represented in sponge prokaryotic 16S rRNA gene libraries (Lee et al. 2010; Simister et al. 2011). Sponge-specific clusters have been documented for both phyla, although the Thaumarchaeota has received more attention as it contains the prevalent sponge symbiont ‘Cenarchaeum symbiosum’ (Schleper et al. 1998; Hallam et al. 2006). The Thaumarchaeota also contain ammonia-oxidizing archaea (AOA), the discovery of which has had a significant impact of our understanding of marine nitrogen cycling (Herndl et al. 2005; Hallam et al. 2006) Interestingly, the physiology of the AOA (such as Nitrosopumilus maritimus) and what are now considered amoA-encoding archaea (AEA), such as ‘C. symbiosum’ (Hatzenpichler 2012), is not well understood and more genetic and physiological studies are necessary to elucidate the mechanisms and regulation of ammonia oxidation in these organisms (Hatzenpichler 2012). We have shown here, however, that diverse Archaea are present in X. muta, providing additional support for AOA in X. muta that has been suggested in previous studies (Lopez-Legentil et al. 2010; Chapter 3).
The significant differences in the prokaryotic community composition observed between the seawater and sponges samples is consistent with previous studies (Hentschel et al. 2002; Lee et al. 2010), and provides additional support for the existence of sponge-specific prokaryotic communities observed in other studies (Hentschel et al. 2002; Taylor et al. 2007). We also hypothesized that there would be location-specific differences in the prokaryotic community of *X. muta*. Based on the results of the ANOSIM and UniFrac p-test we observed significant differences in the prokaryotic community for both the sponge and the water column samples between locations. The differences we observed are surprising given that several studies have documented that the symbiotic prokaryotic communities collected from sponges in different marine habitats are similar (Taylor et al. 2007; Schmitt et al. 2011; Simister et al. 2011), and one previous study that characterized sponge prokaryotic communities from different locations in the Red Sea also did not find significant differences between locations (Lee et al. 2010). The sponges in the current study do in fact have similar taxonomic composition with several sponge-specific groups, but differ in the abundance of these taxonomic groups while the water samples are also similar to each other, differ in the abundance of their shared taxonomic groups, are less diverse than the sponges and contain none of the sponge-specific groups. Further support for the higher similarity of the prokaryotic community within populations from each location than between them is provided by the PCoA and cluster analyses. The samples generally grouped more clearly by location with the weighted analyses, which take into account OTU abundance, supporting the notion that differences in the abundance of taxonomic groups is important to the observed significant differences. The fact that two sponges, XmLSI1 and XmLC3 did not group with their respective locations
cannot be explained based on current analyses alone, but warrants further investigation into potential differences of individual sponge hosts. There is apparently enough distance and/or differences in habitat between the three locations to allow for differences in the symbiotic community of *X. muta* from these locations. However, the biological significance of these differences is unclear and different types of meta-analysis studies in the future could help address this question.

While several studies have reported similar microbial communities from sponges in different marine habitats, one recent characterization of sponge prokaryotic communities in 32 sponge species provided support for a tropical clade of sponge symbiotic prokaryotes, indicating that subpopulations of sponge symbionts may be selected for based on specific environmental factors (Schmitt et al. 2011). It is possible that environmental factors have an important role in structuring sponge prokaryotic communities in the Caribbean and elsewhere. In this study the bacterioplankton communities differed by location, and environmental factors have been shown to influence bacterioplankton community structure (Martiny et al. 2006; Pommier et al. 2007). Furthermore, given the proposed model for sponge symbiont transmission, where both vertical (parent to offspring) and horizontal (acquired from the environment) contribute to the community composition (Schmitt et al. 2008), it is possible that there is a common prokaryotic community that is vertically transmitted in *X. muta*, but that bacterioplankton from the water column also contribute to the symbiotic community. Some of the recovered prokaryotic 16S rRNA gene sequences may be from transient organisms passing through the sponge as it was feeding at the time of sampling, but as
the differences between the sponge and seawater community are so distinct, it is likely that transient prokaryotes comprise a negligible proportion of recovered gene sequences.

Further insight into the differences in the prokaryotic community of *X. muta* and seawater between locations is provided by the results of the OTU category significance test. Synechococcaceae and Alphaproteobacteria OTUs were more abundant in the LC sponges, both of which contain common bacterioplankton members, although these OTUs were particularly more abundant in sponges than seawater. FL and LSI sponges had similar abundances of OTUs that were significantly more abundant in sponges, and the lineages of these OTUs were taxa that are known to contain sponge symbionts. To some extent the OTU abundances in the sponges followed the same trend of abundance in the water samples. For example, OTU 1790, classified as *Poribacteria*, is significantly more abundant in LSI and FL sponges relative to the LC sponges, and within the water column samples, OTU 1790 is most abundant in the LSI samples followed by LC and FL. The extent that regional environmental factors (e.g., region-specific oceanography, temperature, nutrients) or localized factors (i.e., environment immediately surrounding individual sponges), in addition to vertical transmission, might play a role in structuring the prokaryotic communities of *X. muta* is still unknown, but these data highlight the need for further investigation into the taxonomic structure of sponge microbial communities across a geographic range such as the Caribbean basin.

The pigment analyses also provide some insight into the structure of the photoautotrophic community specifically in *X. muta* at each location. The differences in concentrations of chlorophyll *a* (chl-*a*) and phycoerythrin (PE) between locations provide further support for location specific differences in the structure of the prokaryotic
community. For example, the concentration of PE is higher than chl-α in FL sponges but the opposite is true in the LSI sponges. This may indicate differences in the photoautotrophic community composition with FL sponges potentially having more cyanobacteria, which are known to contain PE. Alternatively, the differences may be a result of the physiological state of the cyanobacteria and other photoautotrophs at each location (i.e., growth rate, state of photosystems based on irradiance levels (Kana and Gilbert 1987; Subramaniam et al. 1999)), particularly as FL and LSI appeared to have about the same number of sequences matching to cyanobacterial OTUs. Additionally, the concentrations of phycoerythrin and chlorophyll α are lowest for the LC sponges, which is counterintuitive given that the LC sponges had significantly greater abundance of sequences matching to Synechococcaceae than LSI or FL had. However, there may be other factors influencing pigment concentration, such as nutrient availability and irradiance (Tao et al. 2006), which may vary on a microscale within the sponge host. Differences in light levels at the three described locations have not been described, but there may be differences in irradiance within the sponge host as a result of light scattering by siliceous spicules, similar to the effect that the skeleton has on light scatter in corals (Enríquez et al. 2005), which could influence pigment concentrations. These data provide an interesting point of future research on the photobiology of this dominant Caribbean sponge.

Results of the current study indicate there is a sponge specific community within X. muta that is distinct from the surrounding water column. Additionally, there are some region specific differences in the prokaryotic community composition, but does this translate to functional differences between locations? While 16S rRNA gene sequences
cannot be used to infer functionality, there are a few exceptions where a strong connection between taxonomy and function are well established such as with the nitrifying bacteria *Nitrospira* (Bayer et al. 2007) and potentially the Chromatiales and Syntrophobacteraeae, which are well known to be involved in sulfur cycling (Hoffmann et al. 2005b; Taylor et al. 2007 and references therein). These groups were recovered from *X. muta* in the current study, and nitrification in particular has been documented for *X. muta* previously (Southwell et al. 2008a; Chapter 3). The potential for anaerobic processes has also recently been documented in *X. muta* based on nutrient analyses and stable isotopic studies (Chapter 3); however, further research will be needed to connect the taxonomic diversity of these sponges with functional diversity. The results presented here provide a base for designing future metatranscriptomic investigations as well as other experimental studies on sponge prokaryotic ecology.
Table 2.1. Sampling depth, number of OTUs, and the diversity estimate (based on a rarefied OTU table) for each sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sampling Depth</th>
<th>no. OTUs (97%)</th>
<th>Observed species</th>
<th>Shannon</th>
<th>Simpson</th>
</tr>
</thead>
<tbody>
<tr>
<td>XmFL.1</td>
<td>11,595</td>
<td>244</td>
<td>231</td>
<td>6.4</td>
<td>0.98</td>
</tr>
<tr>
<td>XmFL.2</td>
<td>12,064</td>
<td>251</td>
<td>233</td>
<td>6.4</td>
<td>0.98</td>
</tr>
<tr>
<td>XmFL.3</td>
<td>12,818</td>
<td>256</td>
<td>238</td>
<td>6.4</td>
<td>0.98</td>
</tr>
<tr>
<td>XmLC.1</td>
<td>10,369</td>
<td>121</td>
<td>113</td>
<td>3.6</td>
<td>0.83</td>
</tr>
<tr>
<td>XmLC.2</td>
<td>8,477</td>
<td>163</td>
<td>160</td>
<td>4.7</td>
<td>0.90</td>
</tr>
<tr>
<td>XmLC.3</td>
<td>7,648</td>
<td>187</td>
<td>180</td>
<td>6.0</td>
<td>0.97</td>
</tr>
<tr>
<td>XmLSI.1</td>
<td>11,027</td>
<td>218</td>
<td>205</td>
<td>6.3</td>
<td>0.98</td>
</tr>
<tr>
<td>XmLSI.2</td>
<td>13,451</td>
<td>222</td>
<td>204</td>
<td>6.1</td>
<td>0.98</td>
</tr>
<tr>
<td>XmLSI.3</td>
<td>14,247</td>
<td>279</td>
<td>248</td>
<td>6.5</td>
<td>0.98</td>
</tr>
<tr>
<td>H2OFL.1</td>
<td>23,082</td>
<td>497</td>
<td>337</td>
<td>5.5</td>
<td>0.93</td>
</tr>
<tr>
<td>H2OFL.2</td>
<td>22,838</td>
<td>458</td>
<td>315</td>
<td>5.3</td>
<td>0.93</td>
</tr>
<tr>
<td>H2OFL.3</td>
<td>21,438</td>
<td>375</td>
<td>264</td>
<td>5.1</td>
<td>0.91</td>
</tr>
<tr>
<td>H2OLC.1</td>
<td>19,050</td>
<td>498</td>
<td>327</td>
<td>5.4</td>
<td>0.92</td>
</tr>
<tr>
<td>H2OLC.2</td>
<td>21,468</td>
<td>483</td>
<td>349</td>
<td>5.2</td>
<td>0.91</td>
</tr>
<tr>
<td>H2OLC.3</td>
<td>21,773</td>
<td>514</td>
<td>237</td>
<td>5.3</td>
<td>0.91</td>
</tr>
<tr>
<td>H2OLS1.1</td>
<td>26,961</td>
<td>576</td>
<td>357</td>
<td>5.6</td>
<td>0.94</td>
</tr>
<tr>
<td>H2OLS1.2</td>
<td>21,664</td>
<td>451</td>
<td>312</td>
<td>5.6</td>
<td>0.94</td>
</tr>
<tr>
<td>H2OLS1.3</td>
<td>19,920</td>
<td>437</td>
<td>314</td>
<td>5.6</td>
<td>0.94</td>
</tr>
</tbody>
</table>
Table 2.2. Results of the OTU category significance test in QIIME, the corrected p-value, means at each location, and a consensus lineage is given for each significant OTU. Relative abundance (more abundant in sponges or water samples) is also given.

<table>
<thead>
<tr>
<th>OTU</th>
<th>Bonferroni corrected p value</th>
<th>XmLC mean</th>
<th>XmFL mean</th>
<th>LXmSI mean</th>
<th>Consensus lineage</th>
<th>OTU relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>3777</td>
<td>1.86E-09</td>
<td>0.00102</td>
<td>0.00196</td>
<td>0.00000</td>
<td>0.098458</td>
<td>Synechococcaceae</td>
</tr>
<tr>
<td>1057</td>
<td>1.10E-08</td>
<td>0.00169</td>
<td>0.00273</td>
<td>0.00104</td>
<td>0.253016</td>
<td>Rickettsiales</td>
</tr>
<tr>
<td>3795</td>
<td>1.51E-08</td>
<td>0.00084</td>
<td>0.00022</td>
<td>0.00293</td>
<td>0.036044</td>
<td>Rickettsiales</td>
</tr>
<tr>
<td>98</td>
<td>2.80E-08</td>
<td>0.00000</td>
<td>0.00722</td>
<td>0.00573</td>
<td>0.03944</td>
<td>Synechococcaceae</td>
</tr>
<tr>
<td>3922</td>
<td>4.20E-08</td>
<td>0.00173</td>
<td>0.00000</td>
<td>0.00036</td>
<td>0.036209</td>
<td>koll13(Actinobacteria)</td>
</tr>
<tr>
<td>2548</td>
<td>6.48E-07</td>
<td>0.00044</td>
<td>0.00000</td>
<td>0.00053</td>
<td>0.016122</td>
<td>Rhodospirillaceae</td>
</tr>
<tr>
<td>726</td>
<td>1.59E-06</td>
<td>0.00044</td>
<td>0.00089</td>
<td>0.00078</td>
<td>0.07033</td>
<td>Rhodospirillaceae</td>
</tr>
<tr>
<td>3071</td>
<td>1.58E-05</td>
<td>0.00044</td>
<td>0.00062</td>
<td>0.00104</td>
<td>0.05478</td>
<td>OM60(Oceanospirillales)</td>
</tr>
<tr>
<td>48</td>
<td>2.13E-05</td>
<td>0.00751</td>
<td>0.02707</td>
<td>0.031216</td>
<td>0.00018</td>
<td>PAUC34f</td>
</tr>
<tr>
<td>1219</td>
<td>9.98E-05</td>
<td>0.00968</td>
<td>0.01002</td>
<td>0.018478</td>
<td>0.00040</td>
<td>Rhodospirillaceae</td>
</tr>
<tr>
<td>979</td>
<td>0.000105489</td>
<td>0.00664</td>
<td>0.01734</td>
<td>0.021082</td>
<td>0.000113</td>
<td>HTCC2089(Oceanospirillales)</td>
</tr>
<tr>
<td>3296</td>
<td>0.000105742</td>
<td>0.00044</td>
<td>0.00064</td>
<td>0.00096</td>
<td>0.031387</td>
<td>Rhodobacteraceae</td>
</tr>
<tr>
<td>3108</td>
<td>0.000151058</td>
<td>0.01676</td>
<td>0.011938</td>
<td>0.014380</td>
<td>0.000038</td>
<td>Gammaproteobacteria</td>
</tr>
<tr>
<td>3467</td>
<td>0.000355209</td>
<td>0.04796</td>
<td>0.01856</td>
<td>0.010031</td>
<td>0.00064</td>
<td>Sva0725(Acidobacteria)</td>
</tr>
<tr>
<td>1790</td>
<td>0.000495</td>
<td>0.03368</td>
<td>0.09283</td>
<td>0.07496</td>
<td>0.000324</td>
<td>Poribacteria</td>
</tr>
<tr>
<td>1065</td>
<td>0.000778068</td>
<td>0.05618</td>
<td>0.01949</td>
<td>0.015587</td>
<td>0.000011</td>
<td>Acidobacterium</td>
</tr>
<tr>
<td>3813</td>
<td>0.00080136</td>
<td>0.01615</td>
<td>0.02896</td>
<td>0.032433</td>
<td>0.000118</td>
<td>Cerarchaeaceae</td>
</tr>
<tr>
<td>2140</td>
<td>0.001116066</td>
<td>0.02759</td>
<td>0.02420</td>
<td>0.041158</td>
<td>0.000122</td>
<td>Syntrophobacteraceae</td>
</tr>
<tr>
<td>875</td>
<td>0.001156522</td>
<td>0.06547</td>
<td>0.015227</td>
<td>0.04487</td>
<td>0.000022</td>
<td>Entotheonellaceae</td>
</tr>
<tr>
<td>1727</td>
<td>0.002477021</td>
<td>0.01371</td>
<td>0.02947</td>
<td>0.026204</td>
<td>0.000200</td>
<td>Acidimicrobiales</td>
</tr>
<tr>
<td>2232</td>
<td>0.00295262</td>
<td>0.04987</td>
<td>0.013018</td>
<td>0.012166</td>
<td>0.000016</td>
<td>Gemmatimonadetes</td>
</tr>
<tr>
<td>818</td>
<td>0.003632372</td>
<td>0.01276</td>
<td>0.030204</td>
<td>0.027478</td>
<td>0.000044</td>
<td>Chloroflexi</td>
</tr>
<tr>
<td>2438</td>
<td>0.005846369</td>
<td>0.03678</td>
<td>0.008158</td>
<td>0.006513</td>
<td>0.000000</td>
<td>Chloroflexi</td>
</tr>
<tr>
<td>487</td>
<td>0.008933073</td>
<td>0.02011</td>
<td>0.003169</td>
<td>0.003853</td>
<td>0.000000</td>
<td>Alphaproteobacteria</td>
</tr>
<tr>
<td>852</td>
<td>0.011178918</td>
<td>0.01123</td>
<td>0.009938</td>
<td>0.011067</td>
<td>0.000000</td>
<td>Gemmatimonadetes</td>
</tr>
<tr>
<td>268</td>
<td>0.012621738</td>
<td>0.06000</td>
<td>0.012011</td>
<td>0.010413</td>
<td>0.000044</td>
<td>Chromatiales</td>
</tr>
<tr>
<td>3456</td>
<td>0.02172279</td>
<td>0.13690</td>
<td>0.09356</td>
<td>0.043222</td>
<td>0.000160</td>
<td>Synechococcaceae</td>
</tr>
<tr>
<td>3114</td>
<td>0.024486898</td>
<td>0.04064</td>
<td>0.003473</td>
<td>0.012044</td>
<td>0.000091</td>
<td>Rhodobacteraceae</td>
</tr>
<tr>
<td>439</td>
<td>0.033869602</td>
<td>0.03178</td>
<td>0.013078</td>
<td>0.026087</td>
<td>0.000118</td>
<td>Rhodothermaceae</td>
</tr>
<tr>
<td>3484</td>
<td>0.039020148</td>
<td>0.00607</td>
<td>0.010533</td>
<td>0.011187</td>
<td>0.000027</td>
<td>Chloroflexi-4</td>
</tr>
<tr>
<td>2443</td>
<td>0.041488301</td>
<td>0.014173</td>
<td>0.023651</td>
<td>0.025024</td>
<td>0.000027</td>
<td>Chloroflexi-4</td>
</tr>
</tbody>
</table>
Figure 2.1. Rarefaction curves for the OTU (97% similarity) sequences of 16S rRNA genes from *X. muta* and seawater based on Observed Species estimates (A). Average rank abundance curves based on the number of sequences in each OTU for *X. muta* and for seawater (B).
Figure 2.2. Multidimensional scaling analysis for sample location using Bray-Curtis similarity for the OTU table (97% similarity) generated in QIIME (square root transformation) for sponge samples (A) and water samples (B).
Figure 2.3. UPGMA clustering using weighted Unifrac distance. Jackknifing was performed (7500) on the tree and support values for all nodes were 75-100%, except where noted by a diamond (50-75%) or star (25-50%).
Figure 2.4. UPGMA clustering using unweighted Unifrac distance. Jackknifing was performed (7500) on the tree and support values for all nodes were 75-100%, except where noted by a diamond (50-75%) or star (25-50%).
Figure 2.5. UPGMA clustering using Bray-Curtis distance. Jackknifing was performed (7500) on the tree and support values for all nodes were 75-100%, except where noted by the diamond (50-75%).
Figure 2.6. Principle-coordinate analysis plots based on weighted Unifrac distance comparing PC1 and PC2 (A) and unweighted Unifrac distance comparing PC1 and PC2 (B). Locations are color-coded: red = XmFL, purple = XmLSI, orange = XmLC, blue = H2OFL, brown = H2OLSI, green = H2OLC.
Figure 2.7 (previous page). Average relative abundance of OTUs (97% similarity) at the phylum level for each sponge location and the water column (A), relative abundance of OTUs for individual samples (B), and relative abundance of OTUs (97% similarity) at the class level (as determined using RDP taxonomy) for each *X. muta* sample at each location and seawater at each location (C). The "Other" category in (B) contains unassigned Archaea, unassigned Bacteria, Chlamydia, PAUC34f, Spirochaetes, ZB2 (seawater only), XB3 (seawater only), and Lentisphaerae (seawater only), which all had low relative abundance (< 0.5%). The “Other” category in (C) contains unclassified Archaea, unclassified Bacteria, unclassified Bacteroidetes, Chlamydiae, 2 Firmicute classes, Fusobacteria, 2 GN02 classes, Gemmatimonadetes, Lentisphaerae, Nitrospira, PAUC34f, 3 Planctomycete classes, Epsilonproteobacteria, Gammaproteobacteria, 1 SAR406 class, 2 SBR1093 classes, 2 Spirochaete classes, 1 TM6 class, 2 TM7 classes, 1 Tenericutes class, 5 Verrucomicrobia classes, ZB2, ZB3, and unclassified Proteobacteria.
Figure 2.8. Relative abundance of OTUs (97% similarity) at the class level (as determined using RDP taxonomy) for each X. muta sample at each location and seawater at each location. The "Other category contains unclassified Archaea, unclassified Bacteria, unclassified Bacteroidetes, Chlamydiae, 2 Firmicute classes, Fusobacteria, 2 GN02 classes, Gemmatimonadetes, Lentisphaerae, Nitrospira, PAUC34f, 3 Planctomycete classes, Epsilonproteobacteria, Gammaproteobacteria, 1 SAR406 class, 2 SBR1093 classes, 2 Spirochaete classes, 1 TM6 class, 2 TM7 classes, 1 Tenericutes class, 5 Verrucomicrobia classes, ZB2, ZB3, and unclassified Proteobacteria.
Figure 2.9. Average relative abundance of OTUs (97% similarity) comprising the top 90% of OTU abundance for each location. Taxonomy is provided at the phylum level and class level if available (based on RDP taxonomy).
Figure 2.10. Neighbor-joining tree based on archaeal 16S rRNA genes of OTU sequences from sponge and water samples. Filled in diamonds represent nodes also observed in the maximum parsimony tree and open diamonds represent nodes also observed in the ARB tree following addition of OTU sequences to the SILVA nr SSU dataset. The scale bar represents 4% sequence divergence.
Figure 2.11. Average Chlorophyll-a and Phycoerythrin concentrations normalized to surface area of *X. muta* sponges at each location (A). LC sponges were significantly lower in Chl-a relative to the other two sites and lower in PE relative to the other two sites. Average ratios of Chl-a : PE are also shown for each location (B). Standard error bars are shown.
Chapter III

NITRIGEN BIOGEOCHEMISTRY IN THE CARIBBEAN SPONGE, XESTOSPONGIA MUTA: A SOURCE OR SINK OF DISSOLVED INORGANIC NITROGEN?

Introduction

Sponges are an ecologically dominant component in many marine ecosystems, including coral reefs, where they contribute to the consolidation of reefs, prevent erosion, filter large quantities of seawater and provide habitat and food for many invertebrates and fishes (Reiswig 1971; Diaz and Rutzler 2001; Ribeiro et al. 2003). Because of their ability to efficiently filter picoplankton, sponges can also contribute significantly to the coupling of productivity in the overlying water column to the benthos (Reiswig 1971; Ribes et al. 2005; Lesser 2006). More recently, sponges and their prokaryotic symbionts have become an important area of research to quantify the fluxes of DIN by sponges and the biogeochemical cycling of nutrients on coral reefs (Ribes et al. 2005; Weisz et al.

---

3 Excerpts from this chapter have been submitted for publication as: Fiore CL, Baker DM, Lesser MP. Nitrogen biogeochemistry in the Caribbean sponge, Xestospongia muta: a source or sink of dissolved inorganic nitrogen? PLoS ONE, submitted.
Nitrogen cycling on tropical coral reefs is particularly important, as nitrogen is a limiting nutrient and the success of several coral reef taxa (e.g., corals) is dependent on their symbiotic partners and the efficient recycling of nitrogen between host and symbionts (Muscatine and Porter 1977). Compared to ambient seawater the excurrent water of actively pumping sponges is often enriched in DIN such as NO$_3^-$ (or NO$_2^- +$ NO$_3^-$) as a result of nitrification (Fiore et al. 2010). This has been documented for sponges from coral reefs (Corredor et al. 1988; Southwell et al. 2008), mangroves, seagrass beds (Diaz and Ward 1997), as well as sponges from temperate and cold-water environments (Eroteida and Ribes 2007; Bayer et al. 2008; Hoffmann et al. 2009). In fact, coral reef sponges have been documented to have rates of nitrification that are significantly higher (5.8 – 16 mmol m$^{-2}$ d$^{-1}$ NO$_3^-$) (Diaz and Ward 1997; Southwell et al. 2008) than what has been reported for benthic habitats such as microbial mats (up to 1.4 mmol m$^{-2}$ d$^{-1}$ NO$_3^-$ (Bonin and Michotey 2006) or coral reef sediment (1.68 mmol m$^{-2}$ d$^{-1}$ NO$_3^-$ (Capone et al. 1992)).

All pathways of nitrogen biogeochemistry have been reported to occur in sponges (Fiore et al. 2010) including nitrogen fixation which was initially measured using the acetylene reduction method (Wilkinson and Fay 1979; Shieh and Lin 1994; Wilkinson et al. 1999). Stable isotope tracer studies (i.e., $^{15}$N$_2$) later confirmed the presence of nitrogen fixation, albeit at low rates, in several species of sponges from coral reefs (Wilkinson et al. 1999; Southwell 2007). Recently, the first sponge-derived nifH gene sequences and transcripts, which encode for the iron protein component of the nitrogenase enzyme responsible for nitrogen fixation, were documented in two sponge species from the Florida Keys (Mohamed et al. 2008). With the presence of anaerobic
microhabitats in sponges (Hoffmann et al. 2005c; Fiore et al. 2010), other anaerobic nitrogen transformations including sulfate reduction, denitrification and anaerobic ammonium oxidation (anammox) have been observed and quantified using stable isotopic tracer methods, radiolabeled isotopes and recovery of gene specific sequences for key enzymes (Hoffmann et al. 2005c; Mohamed et al. 2009; Hoffmann et al. 2009). Interestingly, genomic analysis of the candidate phylum Poribacteria, which is found in sponges from numerous marine habitats (Taylor et al. 2007; Schmitt et al. 2011), suggests that Poribacteria may also be capable of denitrification (Siegl et al. 2010).

Nitrification in sponges, which produces the bulk of the DIN released (Diaz and Ward 1997; Southwell et al. 2008), has been documented using field collected samples and fixed volume incubation methods with nutrient analyses (Corredor et al. 1988; Diaz and Ward 1997; Hoffmann et al. 2009). Southwell et al. (Southwell et al. 2007; Southwell et al. 2008 b) was the first to use an in situ method to identify actively nitrifying sponges and estimate the flux of DIN onto the adjacent coral reef, and found no significant difference in flux of DIN between the incubation and in situ methods.

Nitrifying prokaryotes in sponges include, ammonia-oxidizing bacteria (AOB) such as those in the genus *Nitrosospira*, as well as nitrite-oxidizing bacteria such as *Nitrospira* spp. These groups have been recovered from sponges and identified based on 16S rRNA genes and genes for the ammonia monooxygenase enzyme, typically subunit A (*amoA*) (Diaz 1997; Mohamed et al. 2009; Ribes et al. 2012). Recently, ammonia-oxidizing archaea (AOA) have been reported in sponges and were recovered and identified based on genes and transcripts for *amoA* (Steger et al. 2008; Lopez-Legentil et al. 2010). As in other environments, it may be the case with sponges, that AOA are more abundant than
AOB, and although no studies have examined this specifically yet, we do know that archaeal symbionts, and putative AOA specifically, are at least present and widespread in sponges (Steger et al. 2008; Simister et al. 2011).

The interest in nutrient fluxes mediated by sponges and their symbionts as well as the nutrient biogeochemistry of coral reefs has resulted in a surge of research into the prokaryotic community composition of sponges and the processes they mediate. Recently, this has included the use of high throughput sequencing methods, such as 454 pyrosequencing of the 16S rRNA gene (Webster et al. 2010b; Lee et al. 2010; Schmitt et al. 2011; Simister et al. 2011), which have increased our understanding of the prokaryotic composition of many sponges species from different marine habitats. This genetic information has provided useful complementary insight for characterizing prokaryotic mediated nutrient cycling in these sponges.

On Caribbean coral reefs *Xestospongia muta* is an ecologically dominant member of the benthic community, and on Conch Reef, FL (USA) the number of *X. muta* has been shown to be significantly increasing over time (McMurray et al. 2008). *Xestospongia muta* is also characterized as a high microbial abundance sponge (Hentschel et al. 2003) but little was known about the composition of this community other than it contained Cyanobacteria (Steindler et al. 2005) until recent studies documented a diverse prokaryotic community in the *Xestospongia muta* and other members of this genus (Lopez-Legentil et al. 2010; Montalvo and Hill 2011). Additionally, *X. muta* outside of the Florida Keys have not been as well studied generally until recently. Using 454 pyrosequencing the prokaryotic symbionts of *X. muta* from the Florida Keys, Cayman Islands and Bahamas has recently been completed (Chapter 2).
The primary goal of this study was to quantify, and compare, the DIN fluxes in *X. muta* from the three geographically separated populations where a 454 pyrosequencing study of sponge symbionts was carried out (Chapter 2). Here I ask whether sponges from these same populations in the Caribbean have different fluxes of DIN, and potentially how any differences in DIN fluxes, may be related to the taxonomy of their symbiotic prokaryotes using a comparative approach and a natural experiment (Diamond 1986). Additionally, molecular markers of nitrogen cycling (i.e., *nif, nir, nos, nor, amo* genes) were examined from sponge samples from all locations to better understand potential prokaryotic nitrogen transformations in *X. muta*.

**Materials and methods**

**Sample Locations**

Replicate sponges (n=6) were sampled at approximately 15 m from each of three locations: Rock Bottom Reef, Little Cayman, Cayman Islands (LC) (19°42'7.36'' N, 80°3'24.94'' W), North Perry Reef, Lee Stocking Island, Bahamas (LSI) (23°47'0.03'' N, 76°6'5.14'' W), and Conch Reef, Key Largo, FL (FL) (24°57'0.03'' N, 80°27'11.16'' W). All populations were sampled during the late spring and summer of 2011 where the maximum irradiance of photosynthetically active radiation (PAR; 400-700 nm) at noon for all three locations is ~500-600 μmol quanta m⁻² s⁻¹ (M.P. Lesser personal communication). Necessary permits were obtained for all three locations: the Marine Conservation Board, Cayman Islands; Department of Marine Resources, Bahamas; NOAA ONMS permit number FKNMS-2011-066 for Conch Reef, Florida Keys.
Nutrient analyses and rates of sponge pumping

Ambient and excurrent water samples for nutrient analysis were collected from individual sponges (n = 6) at each location for nutrient analysis by slowly filling 100 ml syringes and placing all water samples on ice for transport to shore. Ambient water was obtained by filling the syringe adjacent to each sponge (within 20 cm of the sponge body wall) and excurrent water was obtained by placing weighted Tygon® tubing inside the sponge close to the base of the spongocoel that was attached to a 100 ml syringe and drawing water into the syringe slowly (~1 ml s⁻¹). Syringes were then purged of approximately 10 ml and then 40 ml was saved and frozen for NH₄⁺ and NO₂⁻ + NO₃⁻ analysis (NOₓ⁻). For the nutrient analyses the water samples were thawed and filtered (0.22 μm, Whatman, USA) to remove particulate matter then re-frozen and sent to the Nutrient Analytical Facility at Woods Hole Oceanographic Institute (WHOI, Woods Hole, MA, USA) where analyzed using a Lachet QuickChem 8000 (flow injection analysis system) according to standard protocols to determine concentration of NH₄⁺ and NOₓ⁻. Instrumental errors associated with the measurements were calculated as relative standard deviation (RSD) and includes: NH₄⁺ -0.6% measured RSD, and NO₂⁻ + NO₃⁻ -0.59% measured RSD. At all locations sampling was performed in the morning (~9 AM) and the evening (~5 PM). Sponges were marked near their base with labeled flagging tape attached to nails embedded in the substrate to facilitate repeated measurements on the same sponges.

The volume flow or pumping rates for each individual sponge was determined as previously described (Trussell et al. 2006). A small amount (~1 ml) of fluorescein dye was injected using a syringe and 16 gauge needle into the sponge just below the base of
the spongocoel and the time(s) that the dye front took from its first appearance at the base of the spongocoel to the top of the spongocoel was recorded to obtain the centerline fluid velocity to calculate volume flux or pumping rate. We understand that unlike previous studies on tubular sponges where plug flow can be reasonably assumed (e.g., Trussell et al. 2006) the morphology of \textit{X. muta} likely creates more complicated excurrent plumes where the velocity across the oscula is not uniform (Weisz et al. 2007 a). This is easily observed using the timing of multiple dye tracks on \textit{X. muta} injected in different locations with dye tracks closer to the sponge wall being slower than the centerline flow (M.P. Lesser personal communication). As a result we recognize that our measurements of volume flow or pumping rates are expected to be an overestimate. That said our estimates of volume flow are in agreement with the results of Southwell et al. (2008 b) using similar techniques for \textit{X. muta}. Additionally, in our hands we have never observed cessation of pumping, or other artifacts, as a result of exposure to fluorescein in both thin walled and thick walled sponges (Trussell et al. 2006; Lesser 2006). Both spongocoel and total sponge volume were calculated by measuring sponge height, base circumference, osculum diameter, and spongocoel depth and inner diameter with a measuring tape (to ± 1.0 mm) and volume calculated as previously described (McMurray et al. 2008). The mass (kg) of individual sponges was then calculated by multiplying the individual total sponge volume (l), obtained as described above, by the average density of \textit{X. muta} sponges (0.617 g cm\(^{-3}\)) which was determined from direct measurements of the displacement volume and mass of pieces (n=5) of sponge (including both mesohyl and pinacoderm). The flux of nutrients was then calculated by multiplying the ΔDIN (the difference in nutrient concentration between the ambient and excurrent water in μmol l\(^{-1}\))
by the flow rate (cm s\(^{-1}\)) and normalized to both sponge volume and mass for comparisons between sites.

Pumping rates and nutrient flux data were tested for assumptions of ANOVA and if the data failed either normality or homoscedasticity a constant integer to all values was added followed by log transformation. The transformed data passed Bartlett’s test (Bartlett 1937) for homoscedasticity but often failed the Shapiro-Wilk's (Royston 1995) test for normality. Because Bartlett’s test is sensitive to deviations in normality (Sokal and Rohlf 1995) we choose to proceed with ANOVA, which is known to be robust to deviations from normality (Schmider et al. 2010), on the transformed data. To determine if time of day was a significant factor in the flux of DIN, a two-way ANOVA with interaction was performed using the statistical program R (R core team) with time (AM and PM) and location as fixed factors for the flux of \(\text{NO}_x^-\), \(\text{NH}_4^+\), total DIN and pumping rate as the response variables. A repeated measures ANOVA was not performed because the general requirement of this approach is three time points. Since the effect of time and interaction of time with location was not significant the flux of \(\text{NO}_x^-\), \(\text{NH}_4^+\), total DIN and pumping rate, \(\Delta\text{DIN}\) and pumping rates for each location were calculated by averaging the AM and PM values for each individual sponge. Collapsing the design to a single factor analysis to examine differences between locations was then assessed using a one-way ANOVA with location as a fixed factor (Sokal and Rohlf 1995). Flux values of volume normalized DIN were extrapolated to estimate the flux of DIN to the coral reef per m\(^2\) based on abundance of \textit{X. muta} on each reef (Conch reef (FL), unpublished data; South Perry reef (LSI), and Rock Bottom reef (LC) D. Gochfeld personal communication). This was done by multiplying the average volume (l) of sponges on
each reef, by the average volume-normalized flux of DIN ($\mu$mol l$^{-1}$ h$^{-1}$), and by the average density of $X.\ muta$ (m$^{-2}$). Relative standard deviation of each parameter was used to estimate standard deviation and standard error of the reef flux calculations.

**Flow Cytometry**

Ambient and excurrent water samples were collected as described above for the nutrient analyses for another set of sponges (n=4) from LSI only. Approximately 3 ml from each collected water sample were fixed in electron microscopy grade paraformaldehyde at a final concentration of 0.5% in filtered (0.22 $\mu$m) seawater and frozen at -50°C. Frozen water samples were sent to the Bigelow Laboratory for Ocean Sciences J.J. MacIsaac Aquatic Cytometry Facility where they were stored in liquid nitrogen until analysis. Each sample was analyzed for cell abundances using a Becton Dickinson FACScan flow cytometer with a 30 mW, 488 nm laser. Simultaneous measurements of forward light scattering (FSC, relative size), 90° light scatter (SSC), chlorophyll fluorescence (>650 nm), and phycoerythrin fluorescence (560 - 590 nm) were made simultaneously on each sample as previously described (Lesser 2006). Calculations of cyanobacteria, prochlorophyte, and heterotrophic cell concentration and filtering efficiency were performed as previously described (Lesser 2006). Technical replicates (n=2) were averaged for each sample and the cell abundance of heterotrophic bacteria was determined using PicoGreen (Molecular Probes), a dsDNA specific dye, which stains all prokaryotes (emission fluorescence 515-525 nm). Subtraction of the chl $a$ containing picoplankton from the total prokaryotes yielded the heterotrophic bacterial component of the community while cyanobacterial and prochlorophyte cells were
differentiated by the presence or absence, respectively, of phycoerythrin fluorescence. All filtered cells were converted to carbon and nitrogen equivalents using the following conversions; heterotrophic bacteria: 20 fg C cell$^{-1}$ (Ducklow et al. 1993), *Prochlorococcus*: 53 fg C cell$^{-1}$ (Morel et al. 1993), *Synechococcus*: 470 fg C cell$^{-1}$ (Campbell et al. 1994), heterotrophic bacteria: 3.3 fg N cell$^{-1}$ (FaggerBakke et al. 1996), *Prochlorococcus*: 9.4 fg N cell$^{-1}$ (Bertilsson et al. 2003), *Synechococcus*: 35 fg N cell$^{-1}$ (Bertilsson et al. 2003). Data were log transformed or arcsin transformed as necessary and an ANOVA followed by Tukey’s HSD were performed to test for significant differences in the number of filtered cells between cell types (cyanobacteria, prochlorophytes, heterotrophic bacteria and total cells), filtration efficiency and total particulate carbon (POC) and nitrogen (PON) consumed by sponges.

**Stable isotopic analyses and tracer experiments**

Sponge samples that were frozen without buffer (n=3 each location) were later lyophilized, ground to a powder with a mortar and pestle, and then acid treated with 1 M HCl to remove carbonate and rinsed with distilled water and allowed to dry. An analysis of samples from FL, separated into the outer pigmented layers of the sponge and the non-pigmented inner tissues (containing the pinacoderm and outer mesohyl respectively), showed no significant differences in stable isotope signatures (Fiore and Lesser, unpublished data) so whole cross-sections of sponge samples, consisting of both pinacoderm and mesohyl, from all locations were analyzed. Samples were then sent to the Marine Biological Laboratory (MBL) for the analysis of particulate C and N as well as the natural abundance of the stable isotopes $\delta^{15}$N and $\delta^{13}$C. Samples were analyzed
using a Europa ANCA-SL elemental analyzer-gas chromatograph attached to a continuous-flow Europa 20-20 gas source stable isotope ratio mass spectrometer. The carbon isotope results are reported relative to Vienna Pee Dee Belemnite and nitrogen isotope results are reported relative to atmospheric air and both are expressed using the delta (δ) notation in units per mil (%o). The analytical precision of the instrument is ± 0.1%, and the mean precision of sample replicates for δ13C was ± 0.4 %o and δ15N was ± 0.2 %o. A one-way ANOVA was used to test for significant differences between locations for δ13C, δ15N and C:N ratios followed by the post hoc multiple comparison Tukey's HSD test as needed.

Two stable isotope tracer experiments were conducted during the summer of 2011 at LSI: the first used Na15NO3 (5 mg l⁻¹ final concentration) plus H13CO3 (50 mg l⁻¹) and the second used 15NH4 (0.31 mg l⁻¹) as tracers (Sigma-Aldrich, USA). The method was the same for each experiment: 11 individual X. muta (average volume 172 ± 77 ml or mass 0.106 ± 0.048 kg; mean ± SD) were collected by cutting through the base of the sponge but keeping the floor of the sponge intact from approximately 12 m at North and South Perry reefs at LSI and held in a large holding tank with flow through seawater for 5 d to recover from being removed from the reef. Care was taken to ensure that sponges were never exposed to air and that light levels were maintained at the same levels as found at ~12 m using neutral density screens over the outdoor flowing seawater tanks. Sponges were checked for pumping activity using fluorescein dye and incubated statically with the tracer compound(s) for 4 h. Subsequently, T0 sponges (n=3) were then removed and stored for analysis.
The remaining sponges were placed in individual aquaria with flow through seawater and sponges were sampled at 3 h (n=2), 6 h (n=3) and at 12 h (n=3) to quantify the uptake by the prokaryotic community and potential transfer to the host sponge. Frozen samples were initially processed by separating the bacteria and sponge fractions following the methods of Freeman and Thacker (Freeman and Thacker 2011) except for two steps: an initial centrifugation was performed at 520 x g for 4 min, and the resulting sponge pellet was rinsed an additional two times. The purity of the sponge and bacterial fractions were assessed using light and epifluorescence light microscopy as described by Freeman and Thacker (2011). The sponge fractions always contained at least 85% (per field) of large cells (8 – 10 μm diameter) and low natural fluorescence, whereas the bacterial fractions contained only small cells (< 1 – 2 μm diameter) and high natural (=phycoerythrin) fluorescence. While efforts were made to purify the sponge fraction as much as possible from all prokaryotic cells, it is possible that some additional prokaryotes (e.g., intracellular and non-cyanobacterial (not autofluorescent)) were not detected in the sponge fraction. These methods have been shown to be effective for other sponge species and were optimized for use with X. muta; however, the fact that some contamination was unavoidable was considered when interpreting the results of these experiments. Samples were then lyophilized and ~1.0 mg was weighed and placed into silver capsules (Costech, CA, USA) and acidified three times with 20 μl of 12 M HCl. Samples were allowed to dry in between acidifications, then oven dried at 50°C for 48 h. Samples were combusted in a Carlo-Erba NC2500 elemental analyzer, and the resulting gas was analyzed in a Thermo Delta V isotope ratio mass spectrometer via a Conflo III open-split interface. The analytical precision of the instrument was ± 0.2 ‰, and the
mean standard deviation of sample replicates for δ¹³C was ± 0.4 %o and for δ¹⁵N it was ± 0.8 %o for enriched samples and ± 0.1 %o and ± 0.1 %o for natural abundance samples, respectively. For the tracer experiments the data were log transformed as necessary to meet the assumptions of parametric statistics and a two-way ANOVA with interaction, with fraction and time as fixed factors, was used to assess treatment effects.

Gene specific analyses

CTAB (hexadecyltrimethylammonium bromide) DNA extractions were performed on all sponge and seawater samples as follows. A section of each sponge sample that included both the outer pigmented layers and inner layers of the sponge were cut into smaller pieces with a razor blade for processing. Filters were cut in half (half saved for later use) and also cut into smaller pieces for processing. Samples for both sponges and filters were placed in 600 μl of 2x CTAB mixture (Tris, pH 8.0 (0.0121 g ml⁻¹), NaCl (0.0818 g ml⁻¹, EDTA (0.00744 g ml⁻¹), CTAB (0.002 g ml⁻¹)) and homogenized with a pestle and with brief sonication. Proteinase k (5 μl of 20 mg ml⁻¹) was added and samples were incubated at 64°C for 3 h. Equal volume of chloroform was added to the samples followed by centrifugation at 12 000 x g for 10 min. The aqueous layer was transferred to a new tube and DNA was precipitated with equal volumes of 100% ethanol. The samples were spun again for 10 min followed by two washes with 70% ethanol, and then the pellet was allowed to dry before being resuspended in 30 μl of molecular grade water. Extractions were checked for quality and concentration using a NanoDrop spectrophotometer (2000c, Thermo Fisher, USA). In some cases, a phenol:chloroform:isoamyl alcohol (25:24:1) extraction was performed to clean the
samples. Briefly, the suspension was brought to 100 μl volume, and 1/10 volume of KOAc was added followed by equal volume of phenol:chloroform:isoamyl alcohol and vortexted for 5 s. Samples were centrifuged for 2 min at maximum speed and the aqueous layer was transferred to a clean tube. Three volumes of 100% ethanol were added to each sample and vortexed for 5 s, and then centrifuged for 10 min at maximum speed. The pellet was then washed twice with 500 μl of 70% ethanol (2 min of centrifuge in between), and then allowed to air dry.

Genes involved in nitrogen transformations including, ammonium oxidation and nitrification, denitrification, anammox, and nitrogen fixation, were investigated using with gene-specific primers. Attempts to amplify genes using primers specific to bacterial amoA and amoB genes (ammonia monoxygenase), nirS genes, and 16S rRNA genes of known anammox bacteria were not successful. Listed below are the references that were used for the primer selection and PCR protocols (amoA (Rotthauwe et al. 1997), amoB (Calvó and Garcia-Gil 2004; Junier et al. 2008), nirS (Braker et al. 1998), 16S rRNA (Schmid et al. 2005; Shu and Jiao 2008). Primers specific to archaeal amoA (Arch-amoAF (5'- STAA AT GTCTGGCTTAGACG- 3') and Arch-amoAR (5'- GCGGCCATCCATCGTATGT- 3'), (Francis et al. 2005)) genes were also utilized based on methods described by (Lopez-Legentil et al. 2010). PCR reactions for archaeal amoA gene amplification were as described for pyrosequencing preparation and the thermocycler protocol was the same as described in (Lopez-Legentil et al. 2010). PCR for nifH genes (component of nitrogenase) was performed using two primer sets: OF1 (5'-ATXGTCCGXTGXGAXCCYACG-3') and OR2 (5'- ATGGTGTGGGCGGCGRTAZAKYGCCATCAT-3') (X = C or T, Y = G or C, R = G or
A, Z = C, G, or A, and K = G or T) (Olson et al. 1998), and a nested set modified from (Zehr and McReynolds 1989): nifH3 (5'-ATRITRGNCRTA-3') and IGK (5'-AARGGNNGGNNATHGNGAA-3') followed by nifH1 (5'-GAYCCNAARGCNAG-3') and nifH2 (5'-ANDGCCATCATYTCNCC-3') (Y = T or C, R = A or G, D = A, G, or T, and N = A, C, G, or T) (Olson 2010). Triplicate reactions were made for each sample using the same reaction mixture as described for the pyrosequencing preparation, however, for the second step of the nifH nested protocol 1 μl of the first PCR step was used as template for the second step. The thermocycler protocol for OF1/OR2 primers was as follows: 95°C 5 min, 35 cycles of 95°C 30 s, 62°C 30 s, and 72°C 30 s, followed by 15 min at 72°C. The thermocycler protocol for nifH3/IGK and nifH1/nifH2 primers was as follows: 95°C 5 min, 35 cycles of 95°C 1 min, 55°C 1 min, and 72°C 1 min, followed by 5 min at 72°C.

The triplicate reactions were pooled and electrophoresed on a 1% agarose gel and extracted using Qiaquick gel extraction kit (Qiagen, Valencia CA). Products from the pinacoderm and mesohyl were pooled for the amoA gene since a band was observed in both and preliminary analysis suggested that there was little difference in the archaeal community between the two sponge sections, nifH gene products were not combined). Gel extracted products were cloned in *Escherichia coli* using the pGEM-T easing vector system (Promega, Madison, WI, USA) and JM109 chemically competent cells according to manufacturers instructions (Promega). White colonies were hand picked and added to a 96 well plate with 900 μl of LB liquid media with ampicillin. The PureYield plasmid miniprep system (Promega) was used on a subset of colonies for each sample following the manufacturers instructions (Promega) and then sequenced on an ABI3130 genetic
analyzer to confirm presence of insert. The 96 well plate was incubated overnight at
37°C, centrifuged at 1500 x g for 15 min, the supernatant was decanted and the plate with
pellets was frozen and sent to Functional Biosciences Inc. (Madison, WI, USA) for
sequencing on an ABI 3730xi DNA sequencer.

Chromatograms resulting from Sanger sequencing were manually checked for
quality, and vector sequence was trimmed in the program Geneious v4.8.4 (Biomatters
Ltd., http://www.geneious.com/). Trimmed sequences were then BLASTed against the
NCBI nr database using Geneious and matching sequences were compiled. Pairwise
alignments in Geneious were also performed for each sequence and its closest BLAST
match, which were then translated and manually checked for errors. Sequences from this
study and closest matches from the BLAST were aligned (BLOSUM55 with default
settings in Geneious) and the alignment was used to create a neighbor-joining tree in
Geneious. \textit{nifH} sequences were also aligned to the most recent version (updated
February 17, 2012) of the \textit{nifH} ARB database maintained by the Zehr Laboratory (Zehr et
al. 2003), which has alignments of DNA and amino acid sequences (using Hidden
Markov Models (HMM) for the latter) from GenBank, and several different neighbor-
joining trees based on amino acid alignments. In ARB, sequences from the current study
were translated and aligned to database amino acid sequences using the ARB Integrated
aligner and added to the tree using the QuickAdd Parsimony tool. The tree selected for
use was a neighbor-joining tree built with representative sequences from CD-HIT along
with genomic sequences and putative chimeras removed
(\textquoteleft\textit{tree\_AA\_RepSeqDec2011\_plusAllGenomeSeqs\_noPutChimeras\_MASK1}\textquoteright). \textit{nifH} and
amoA gene sequences were deposited to GenBank (accession numbers JQ912215 - JQ912238 (nifH), JQ912113 - JQ912214 (amoA)).

**Results**

**Stable Isotopic Signatures of Sponges**

The values of δ¹³C from each location were not significantly different from each other (ANOVA, $F_{2, 6} = 1.16, p = 0.38$). The δ¹³C of sponge samples ranged from -19.1 to -18.4 ‰ (Fig. 3.1). The δ¹⁵N of sponge samples ranged from 4.0 to 4.4 ‰ (Fig. 3.1), and were not significantly different between locations (ANOVA, $F_{2, 6} = 0.52, p = 0.62$). The ratios of C:N were significantly different between locations (ANOVA, $F_{2, 6} = 22.57, p = 0.002$), with post hoc pairwise comparisons showing that FL sponges had significantly higher C:N ratios than LSI (Tukey’s HSD, $p < 0.05$), and that LC sponges significantly higher than LSI (Tukey’s HSD, $p < 0.05$). There was no significant difference between LC and FL (Tukey’s HSD, $p > 0.05$). Despite the significant results for C:N ratios the mean values did not vary greatly, ranging from 4.33 – 4.83.

**Inorganic Nitrogen Fluxes in Xestospongia muta**

The difference in nutrient concentration between the ambient and excurrent of NH₄⁺, NO₃⁻ + NO₂⁻ (NOₓ⁻) and total DIN for X. muta varied considerably between individual sponges as expected for sponges over a large size range (Table 3.1). The ΔNH₄⁺ values were not significantly different between locations (ANOVA, $F_{2, 15} = 3.19, p = 0.07$) as were the ΔNOₓ⁻ values (ANOVA, $F_{2, 15} = 2.58, p = 0.11$). ΔDIN values,
however, were significantly different between sites (ANOVA, $F_{2,15} = 6.82$, $p = 0.008$) with post hoc pairwise comparisons showing that FL sponges were significantly lower than both LSI and LC sponges LSI (Tukey’s HSD, $p < 0.05$) which were not significantly different than each other (Tukey’s HSD, $p > 0.05$). No measurements of ambient NO$_x$ exceeded 4 μM eliminating the potential for ambient nutrient concentrations to be confounded by oceanographic features such as internal waves (Leichter et al. 1996; 2003). Sponge pumping rates varied with size (Table 3.1) and did not differ significantly with location (ANOVA, $F_{2,15} = 1.61$, $p = 0.23$) (Fig. 3.2).

The volume and mass normalized fluxes of NH$_4^+$ (Table 1, Fig. 3.3 A, B) were not significantly different between locations (ANOVA, $F_{2,15} = 0.45$, $p = 0.65$ (volume); $F_{2,15} = 0.85$, $p = 0.45$ (mass)). The fluxes of NO$_x^-$ normalized to sponge volume (Table 1, Fig. 3.3 B) were significantly different between locations (ANOVA, $F_{2,15} = 4.89$, $p = 0.02$) with FL sponges significantly lower than LSI and LC not significantly different than either FL or LSI (Fig. 3.3 A) but when normalized to mass did not show a significant effect of location (ANOVA, $F_{2,15} = 3.56$, $p = 0.054$). The flux of total DIN (NO$_x^-$+NH$_4^+$) normalized to volume and to mass were not significantly different among locations (ANOVA, $F_{2,15} = 1.24$, $p = 0.32$ (volume); $F_{2,15} = 2.09$, $p = 0.16$ (mass)). The flux of total DIN extrapolated to the reef based on X. muta abundance was highest for LSI and lowest for FL (Fig. 3.4).

**Feeding Study**

*Xestospongia muta* (n=4) from LSI instantaneously filtered an average of $1.5 \times 10^7$ cells ml$^{-1}$ and there was a significant effect of cell type (ANOVA, $F_{3,12} = 4.45$, $p = 0.03$)
(Fig. 3.5 A). The number of both total cells and heterotrophic bacteria filtered was significantly higher than that of prochlorophytes (Tukey’s HSD, p < 0.05) but not cyanobacteria (Tukey’s HSD, p > 0.05), while the number of cyanobacterial cells filtered was indistinguishable (Tukey’s HSD, p > 0.05) from the prochlorophyte or total cell and heterotrophic cell groupings (Fig. 3.5 A). For the filtration efficiency of each cell type there were no significant differences (ANOVA, $F_{3,12} = 0.45$, $p = 0.72$) (Fig. 3.5 B). The total amount of POC for each retained cell type was greatest for cyanobacteria, but there was no significant difference between cell types (ANOVA, $F_{3,12} = 2.77$, $p = 0.09$). Differences between the amount of PON for each retained cell type was significant (ANOVA, $F_{3,12} = 4.68$, $p = 0.02$) and greatest for total cells and heterotrophic bacteria compared to prochlorophytes (Tukey’s HSD, $p = 0.03$) but not cyanobacteria (Tukey’s HSD, $p > 0.05$), while the PON of cyanobacterial cells was indistinguishable (Tukey’s HSD, $p > 0.05$) from the prochlorophyte or total cell and heterotrophic cell groupings (Fig. 3.5 C).

**Nitrogen tracer experiment: Nitrate and Bicarbonate**

While sponge and bacterial fractions became more enriched from 3 to 6 hours there was no significant effect of enrichment of $^{15}$N from the NO$_3^-$ tracer in the experimental sponges, neither the sponge or bacterial fraction, or over time (ANOVA, $F_{7,14} = 0.84$, $p = 0.57$) with all effects tests being non-significant (fraction, $p = 0.36$, time, $p = 0.36$ and interaction term, $p = 0.84$) (Fig. 3.6 A). Differences of experimental samples in enrichment from control samples are shown by Delta ($\Delta$) values in figure 3.7 A. At the same time the enrichment of $^{13}$C from the bicarbonate tracer experiment was significant
(ANOVA, $F_{7,14} = 20.9$, $p < 0.0001$) with fraction being non-significant ($p = 0.13$) and time being significant ($p < 0.001$) but no significant interaction term, ($p = 0.056$) (Fig. 3.6 B). As a result post-hoc multiple comparison tests were only performed for time. Generally, the bacterially fraction became more enriched in $^{13}\text{C}$ then the sponge fraction over time with all sampling periods being significantly different than $T_0$ (Tukey's HSD $p < 0.05$) but not significantly different (Tukeys' HSD $p > 0.05$) from each other (Fig. 3.6 B).

**Nitrogen tracer experiment: Ammonium**

There was significant enrichment of $^{15}\text{N}$ from the $\text{NH}_4^+$ tracer in the experimental sponges (ANOVA, $F_{7,15} = 3.38$, $p = 0.03$) with the effects tests for fraction ($p = 0.03$) and time ($p = 0.02$) being significant and the interaction term non-significant ($p = 0.19$) (Fig. 3.6 C). Post-hoc multiple comparison testing for time revealed a significant (Tukey's HSD $p < 0.05$) increase in enrichment over time with the bacterial fraction exhibiting greater enrichment (Fig. 3.6 C). Differences of experimental samples in enrichment from control samples are shown by $\Delta$ values in figure 3.7 B. The control samples (natural abundance isotope values) yielded isotopic values similar to those observed in the FL samples, although the bacterial fractions were slightly more depleted in $^{15}\text{N}$ and $^{13}\text{C}$ isotopes (sponge: $\delta^{15}\text{N} \ 5.7\%o \ (\pm 0.36)$, $\delta^{13}\text{C} \ -20.4\%o \ (\pm 0.35)$; bacteria: $\delta^{15}\text{N} \ 3.6\%o \ (\pm 0.27)$, $\delta^{13}\text{C} \ -21.7\%o \ (\pm 0.22)$).
Characterization of nitrogen fixation and archaeal ammonia oxidation genes

Gene sequences encoding for the nitrogen fixing enzyme nitrogenase (\textit{nifH}) \((n = 10\) from sponges, \(n = 14\) from seawater) and for archaeal ammonia monooxygenase (\textit{amoA}) \((n = 84\) from sponges, \(n = 23\) from seawater) were successfully amplified from \textit{X. muta} and from the water column for LC and LSI. Sequences from FL were obtained, but were all low matches to the database and did not align with the other \textit{nifH} sequences so were left out of analysis. \textit{nifH}-deduced amino acid sequences from \textit{X. muta} fell into cyanobacterial and proteobacterial groups (\textit{nifH} Cluster I, Chien & Zinder 1996) (Fig. 3.8). The cyanobacterial \textit{nifH}-deduced amino acid sequences from the current study were similar to, but distinct from those documented previously in sponges and corals (78 – 92%) (Mohamed et al. 2008; Olson et al. 2009) (Fig. 3.8 A). The cyanobacterial \textit{nifH}-deduced amino acid sequences were also similar to \textit{Xenococcus} sp. (87 - 89%), \textit{Myxosarcina} sp. (87 - 89%) and the cyanobacterium UCYN-A (83 - 85%). Proteobacterial \textit{nifH}-deduced amino acid sequences from the current study were most similar to either Alphaproteobacteria such as \textit{Bradyrhizobium japonicum} (80 – 87%), or to Gammaproteobacteria such as \textit{Methylcrostys} sp. (89 – 100%), \textit{Vibrio} spp. (96 – 97%), and \textit{Azotobacter chroococcum} (95 – 96%) (Fig. 3.8 B). Some matches to \textit{Desulfovibrio} spp. (Deltaproteobacteria) were also observed (80 – 83%). Seawater and sponge \textit{nifH}-deduced amino acid sequences were distinct although they shared some higher-level lineages (Fig. 3.8). Similar results were obtained when \textit{nifH}-deduced amino acid sequences were clustered with a NifH database in ARB (Zehr et al. 2003), with most sequences forming unique clusters similar to either Proteobacteria (Deltaproteobacteria
and Alphaproteobacteria (Fig. 3.9), Gammaproteobacteria not shown) or Cyanobacteria (Fig. 3.10).

Archaeal amoA-deduced amino acid sequences from *X. muta* were similar to those documented previously in *X. muta* by (Lopez-Legentil et al. 2010) (99% identity, E = 0.0) and other sponges and other environments, including, marine sediments, groundwater, and freshwater (88 – 99% identity, E = 0.0) (Fig. 3.11). Sequences were largely similar between seawater and sponge samples. Additionally, the amoA-deduced amino acid sequence from the known ammonia-oxidizer *Nitrosopumilus maritimus* is more similar to the amoA-deduced amino acid sequences recovered from sponge and seawater samples from the current study than the amoA-deduced amino sequence from common sponge symbiont *Cenarchaeum symbiosum* (Fig. 3.11).

**Discussion**

This study of the inorganic nitrogen fluxes in *Xestospongia muta* is the first to show differences in nitrification and denitrification both within and between populations of this ecologically dominant sponge on Caribbean coral reefs. Due to the high, and increasing, abundance of *X. muta* (McMurray et al. 2008), understanding the nitrogen biogeochemistry in this sponge is important for understanding the nitrogen biogeochemistry on reefs in the Caribbean.
Nitrogen biogeochemistry

The stable isotopic values measured for both C and N in *X. muta* tissue, comprising both the host tissue and prokaryotic biomass were not significantly different between sites. Additionally, for the samples from Conch Reef the isotopic values of C and N are similar to those documented in previous studies on *X. muta* (Southwell 2007; Mohamed et al. 2008). For *X. muta* the possibility remains that a mixed isotopic N signal is observed because of heterotrophic feeding on particulate organic matter (POM) combined with the isotopic signatures of nitrogen fixation and/or other processes that fractionate N such as NH$_4^+$ or NO$_3^-$ uptake by the symbiotic prokaryotes as has been observed in other host-symbiont systems which results in more variable $\delta^{15}$N values (Tanaka et al. 2006; Weisz et al. 2007b). In particular, the fractionation of N during nitrogen fixation yields an average $\delta^{15}$N of approximately 0.0 ‰ (Mariotti 1983; Peterson and Fry 1987), and trophic enrichment typically results in a +2.2 to +3.5 ‰ increase per trophic level for $\delta^{15}$N (Vander Zanden and Rasmussen 2001; McCutchan et al. 2003). Therefore, several studies have used a cutoff of ≤ 2.0 ‰ to indicate N from a fixed source (Carpenter et al. 1997; Montoya et al. 2002; Mohamed et al. 2008). Using this value there is no stable isotopic evidence that nitrogen fixation was occurring in *X. muta* (Fig. 3.1) although *nifH* genes sequences were recovered from the current study.

Additionally, carbon fixation by marine phytoplankton typically results in $\delta^{13}$C values of about -19 to -24 ‰ (Fry 2006), with an average of +0.5 to +1.0 ‰ enrichment per trophic level (Michener and Schell 1994). Based on previous studies that have used stable isotope analysis to investigate the relationship between sponges and their symbionts (Weisz et al. 2007b; Southwell et al. 2007; Freeman and Thacker 2011), a
cutoff for $\delta^{13}$C of -18%o or lower was used as an indication of photoautotrophic carbon fixation for *X. muta*. Using this cutoff sponges from all locations show evidence of harboring photoautotrophic symbionts. In this regard Freeman and Thacker (2011) demonstrated that high microbial abundance (HMA) sponges, such as *X. muta*, can obtain either C or N, or both, from their symbionts. Any interpretation of tissue stable isotopic signatures must, however, include heterotrophy on picoplankton for sponges. The measurements of cells cleared from LSI sponges show that these sponges utilize significant POM resources and this is consistent with studies on individual sponge species from the Caribbean (Trussell et al. 2006; Lesser 2006).

The isotopic signature of $\delta^{13}$C for both zooplankton and POM ranges from -14 to -25%o and for $\delta^{15}$N it is +4 to +6%o (Land et al. 1975; Owens 1987; Peterson and Fry 1987) and mean values of -19.9%o for $\delta^{13}$C and -3.6%o for $\delta^{15}$N POM for waters adjacent to the Florida Keys (Lamb and Swart 2008). While we did not examine the stable isotope values for POM in the current study, POM values from other studies on sponges (Pile 1997; Trussell et al. 2006; Weisz et al. 2007b; Freeman and Thacker 2011) have suggested that transfer of C from symbionts is species dependent and that POM is the primary source of C for most of these active suspension feeders. Based on the feeding study of LSI sponges, and despite the high abundance of resident bacteria, sponges are actively and non-selectively filtering most of the bacteria from the ambient water, which would supply significant amounts of POC and PON that could be potentially used by the host. Our results may suggest some site related differences on this point with sponges from the more open ocean sites of LSI and LC being more dependent on photoautotrophic sources of C and coastal FL sponges being more dependent on POM.
for their C requirements (Fig. 1). Not considered here, and rarely quantified, is the fact that sponges are known to take-up and use significant amounts of dissolved organic matter, which is believed to be mediated by the symbiotic bacteria themselves and can be an important source of both C and N (Maldonado et al. 2012).

In a survey of nitrification in sponges on Conch Reef, Florida Southwell et al. (2007; Southwell et al. 2008 a, b) reported evidence of nitrification in nine out of twelve sponge species, including X. muta. The rates of nitrification measured in these sponges varied, but overall they were at least two orders of magnitude higher than other habitats (e.g., benthos, coral rubble). In comparison, nitrification was detected in the majority of X. muta sponges in the current study but unlike Southwell et al. (2007; Southwell et al. 2008 b), X. muta from Conch Reef exhibited a negative flux of NOx−, indicating that either denitrification or anammox processes were taking place or possibly dissimilatory nitrate reduction. Our results from LSI and LC are, however, consistent with Southwell et al. (Southwell et al. 2007; 2008 b) although a previous study of the same population of X. muta from LSI also showed both positive and negative fluxes of NOx− (data not shown).

Xestopongia muta were actively pumping for every measurement taken during this study, which was not significantly different over time of day or between locations. This is also similar to measurements of X. muta pumping activity reported for Conch Reef sponges by Southwell et al. (2008 b). While pumping rates were not significantly different the observed variability in the unidirectional pumping of sponges has the potential to create microhabitats where both anaerobic nitrogen transformations (e.g., denitrification) and aerobic nitrogen transformations (e.g., nitrification) could occur
(Hoffmann et al. 2005c; 2009; Schläppy et al. 2010b; Fiore et al. 2010). Additionally, recent 16S rRNA gene characterization studies of *X. muta* (Montalvo and Hill 2011; Chapter 2) have shown many sequences representing bacterial groups that are capable of denitrification and anammox (i.e., Burkholderiales, Pseudoalteromonadaceae, Poribacteria, Planctomycetes).

Interestingly, Southwell et al. (2008 a) found that NO$_X$ made up the majority of the DIN pool from *X. muta* and that NO$_X$ was almost entirely NO$_3^-$. In this study we observed, in addition to positive net NO$_3^-$ fluxes, a greater net efflux of NH$_4^+$ for all samples of *X. muta*. For some *X. muta* populations (i.e., LSI and LC) the flux of NH$_4^+$ had a significant impact on total DIN. These observed differences in the fluxes of NH$_4^+$, probably generated from the utilization of nitrogen rich POM by the sponge host, is unusual given there is an active nitrifying community (Muro-Pastor et al. 2005; Southwell et al. 2007; Southwell et al. 2008 a) and a prokaryotic photosynthetic community (Southwell 2007) that could readily utilize NH$_4^+$ in this sponge (Muro-Pastor et al. 2005; Chapter 2).

The tracer experiments provide some insight into what may be occurring. It appears that the symbiotic prokaryotic community utilizes NH$_4^+$ as would an active nitrifying community (Fig. 3.6 C) and that NH$_4^+$ availability is not limiting rates of nitrification. As discussed above the host sponge is the likely source of NH$_4^+$ that is then taken up by symbiotic prokaryotes. Additionally, NH$_4^+$ is also a likely substrate for aerobic ammonia oxidation by the crenarcheote community, which is supported by the recovery of crenarchaeal *amoA* genes in *X. muta* (Muro-Pastor et al. 2005); current
study), expression of *amo*A genes in previous studies (Muro-Pastor et al. 2005) and previous tracer studies (Southwell et al. 2008).

The small increase in sponge host $^{15}$N during the $\text{NH}_4^+$ experiment may also indicate direct uptake of $\text{NH}_4^+$ by the sponge, as has been demonstrated for corals (Yellowlees et al. 1994). Although, as in the $\text{NO}_3^-$ tracer experiment, the accumulation of tracer in the sponge fraction is more likely to be explained by transfer of N from the bacteria to the host. $\text{NO}_3^-$ can also be utilized by the prokaryotic community, as demonstrated by the increased $\delta^{15}\text{N}$ of the bacterial fractions incubated with $^{15}\text{NO}_3^-$ (Fig 4 A). Because $\text{NH}_4^+$ is energetically cheaper to use for amino acid biosynthesis ($\text{NO}_3^-$ has to first be reduced to $\text{NH}_4^+$), it is likely that $\text{NO}_3^-$ would be utilized during times of increased energy production, such as during photosynthesis. Photosynthetically driven $\text{NO}_3^-$ uptake has been demonstrated in planktonic communities (Maguer et al. 2011), and may explain the increase in $\text{NO}_3^-$ uptake at the 6 h time point, which is when the sponges were exposed to natural solar radiation, while it was dark at the 12 h sampling point (Fig. 3.7). However, it should be noted that heterotrophic bacteria take up $\text{NO}_3^-$ as well (Kirchman 1994). As the symbiotic nitrifiers in the sponge produce $\text{NO}_3^-$ it would provide a source of $\text{NO}_3^-$ for uptake by the photosynthetic community as well as a substrate for denitrification. If assimilatory and dissimilatory processes are competing for $\text{NO}_3^-$, $\text{NO}_2^-$, and $\text{NH}_4^+$, which has been documented in other communities (Mackey et al. 2011), then this may have a significant role in nitrogen cycling in the sponge holobiont and further research is necessary to tease apart these processes.

A factor that would limit the rate of nitrification is the variability of $\text{O}_2$ concentration in the sponge tissues (Schläppy et al. 2010 b). This could explain the
higher net effluxes of NH$_4^+$ compared to NO$_x^-$ fluxes, as oxygen is required for nitrification to occur. Additionally, H$_2$S is known to inhibit nitrification and denitrification (Caffey et al. 1993; Purubsky et al. 2009) and may have a role in regulating these processes in X. muta. H$_2$S may be present in X. muta, as bacteria involved in sulfur cycling have been recovered from this sponge (Chromatiales, Syntrophobacteraceae, Chapter 2). For example, we know that nitrification occurs in X. muta (Southwell et al. 2007, this study) and denitrification may occur in hypoxic/anoxic microzones (Hoffmann et al. 2005 b; Schläppy et al. 2010 a, b; Fiore et al. 2010) that is dependent on sponge pumping and would likely be coupled to NO$_3^-$ production by nitrification; if H$_2$S inhibits these processes, whether partially or completely, then there would be excess NH$_4^+$ available in the sponge and explain the observed net effluxes of NH$_4^+$. If nitrification and denitrification are tightly coupled, then variations in H$_2$S or O$_2$ concentrations may indeed influence the rates of these processes and the net fluxes of various species of DIN.

The anaerobic oxidation of NH$_4^+$ (anammox) is another process that may utilize both NH$_4^+$ and the NO$_2^-$ generated from nitrification. The rates, however, of anammox are relatively low in the water column (Kuypers et al. 2003), as is the only documented rate for anammox in sponges (Hoffmann et al. 2009). It is possible that anammox may occur within anoxic microhabitats of X. muta, and support for this is provided by the presence of planctomycete bacteria in X. muta (Montalvo and Hill 2011; Chapter 2) but the nutrient flux data clearly show that a net efflux of NH$_4^+$ is still occurring in all sponges suggesting an abundance of this substrate for either nitrification or anammox.
Clearly, further studies are necessary to determine the rate of anammox, if it is occurring, and quantify the role it has in NH$_4^+$ utilization and nitrogen cycling in this sponge.

Characterization of nitrogen genes

We amplified cyanobacterial *nif*H genes from *X. muta*, which are similar to, but distinct from the cyanobacterial *nif*H sequences amplified from *Ircinia strobilina* and *Mycale laxissima* (Mohamed et al. 2008). Mohamed and colleagues (2008) observed potential host-specific groups of diazotrophic bacteria. The fact that the *nif*H sequences in the current study are similar to the Mohamed et al. (2008) sequences but still grouped separately, may indicate that there are host-species specific nitrogen fixing symbionts in *X. muta* as well. Further comparative studies are needed, however, before such host-specific groups can be confirmed. Non-cyanobacterial *nif*H sequences were also amplified and were most closely matched to different genera within the Proteobacteria, such as *Azotobacter*, *Bradyrhizobium*, and *Vibrio*, which represent the same non-cyanobacterial groups recovered by Mohamed et al. (2008). However, unlike the cyanobacterial sequences in the current study, these proteobacterial *nif*H sequences were very distinct from previously published sponge sequences (Mohamed et al. 2008) and most other environmental *nif*H sequences (Zehr et al. 2003), further suggesting sponge specific and perhaps species specific nitrogen fixing symbionts. Additionally, the recovered *nif*H sequences from *X. muta* were distinct from the surrounding seawater *nif*H sequences, suggesting that these come from sponge specific bacteria as described by Mohamed et al. (2008) for *Ircinia strobilina* and *Mycale laxissima*. 
Since the sponge provides a constant source of ammonia as a waste product, we would not necessarily expect the energetically costly process of nitrogen fixation to occur, but there may be competition for the ammonia produced by the host. We know that there is an active nitrifying community within *X. muta* (Southwell et al. 2008 a, b) and that some of the ammonium is hypothesized to be provided by the sponge (Southwell et al. 2008 a). Additionally, some *X. muta* has been shown to be a sink for ammonium (Southwell et al. 2008 b; current study). Furthermore, *X. muta* hosts ammonia oxidizing archaea, which have been shown to be active based on *amoA* transcripts (Lopez-Legentil et al. 2010). Archaeal *amoA* genes were also recovered from all sponges in the current study and were identical to those recovered by Lopez-Legentil and colleagues (2010) from *X. muta* in FL, and were similar to *amoA* gene sequences from the known ammonia-oxidizer *N. pumilus*. Interestingly, complete removal of ammonia by ammonium oxidizers within the sponge *Geodia barretti* was reported by Hoffmann *et al.* (2009), which could potentially create microhabitats within the sponge where ammonia is limiting. Furthermore, active uptake of ammonia by photosynthetic cells during the daytime (Maguer et al. 2011) may provide another avenue for creating ammonia-limited microhabitats in the sponge, in which case, the ability to fix nitrogen may be advantageous. The analysis of the bacterial fractions from stable isotope tracer experiments provides evidence that the prokaryotic community of *X. muta* is capable of taking up ammonium and may therefore create microniches low in DIN where nitrogen-fixation can occur. In any event a recent review of the literature on the relationship between DIN and nitrogen fixation clearly shows that diazotrophic bacteria can still fix nitrogen in the presence of elevated concentrations of DIN (≥ 1 μM) Knapp 2012).
The process of nitrification consists of both ammonia oxidation to nitrite and nitrite oxidation to nitrate. As discussed earlier, active ammonia oxidizing archaea are present in *X. muta* (Lopez-Legentil et al. 2010), however, it is not clear if bacterial ammonia oxidizers are present and/or active. Attempts to amplify bacterial *amoA* genes were unsuccessful; suggesting that most of the ammonia oxidation performed in *X. muta* is by Archaea. However, the potential role of bacterial ammonia oxidizers cannot be ruled out. Additionally, attempts to amplify the *nirS* gene, which is involved in denitrification, were unsuccessful, indicating that either these genes are not present or they are divergent enough from the primer sequences used here that they did not amplify. Molecular support for anaerobic nitrogen transformations is provided by preliminary work to amplify genes encoding for nitrate reductase subunit K, nitric oxide reductase subunits, and nitrous oxide reductase subunits (data not shown), although further investigations are necessary to examine gene sequences and transcripts for these genes.

**Conclusions**

From a coral reef biogeochemistry perspective, studies on multi-species sponge assemblages, or coral reef communities dominated by active suspension feeding sponges, has shown the significant role of active suspension feeding and the coupling of POC and PON from the water column to the benthos (Ribes et al. 2003; Perea-Blázquez et al. 2012). Additionally, the composition of DIN released into the water column by sponges would influence how it might be utilized, and who utilizes it in the surrounding environment, as NH$_4^+$ is more readily incorporated into biomass than NO$_3^-$ which can then potentially support local increases in planktonic community production (O'Neil and
Capone 2008). Finally, and as discussed by Southwell et al. (2008 b), excess inorganic nutrients, such as release of DIN by sponges, may have detrimental effects on coral reef ecosystems by stimulating an increase in the growth of fleshy algae in the absence of herbivores (Southwell et al. 2008 b). It is important that further research be done to determine what the ecosystem level effects of DIN release by sponges, and particularly from *X. muta* in regards to Caribbean coral reefs as it is believed to be a primary contributor of DIN released by sponges (Southwell et al. 2008 b).

We have shown that the flux of DIN from *X. muta* is highly variable on Caribbean coral reefs, which may have a significant impact on the availability of DIN on coral reefs given the high abundance of these sponges. The may be of particular importance on coral reefs near Little Cayman, Cayman Islands, and Lee Stocking Island, Bahamas, as net flux of DIN was greater than for reefs near the Florida Keys. Nitrification had been previously demonstrated to occur in *X. muta*, and we show here that other nitrogen transformations including denitrification and/or anammox may occur in these sponges as well as the importance of active suspension feeding on the nitrogen rich pool of picoplankton.

Genes involved in nitrogen fixation have also been documented for the first time in *X. muta*, providing further support for complex nitrogen cycling in this sponge. Further work is needed to better characterize the nitrogen transformations and flux of DIN from *X. muta* and other sponges on Caribbean coral reefs. This will require additional investigations on the functional activity of the symbiotic prokaryotic community of sponges using a combination of experimental and molecular approaches (i.e., transcriptomics) that will yield insight into the taxonomy and function of this community,
and how this impacts nutrient fluxes and biogeochemical cycling on Caribbean coral reefs.
Table 3.1. Calculated volume, mass, and flux parameters for samples of *Xestospongia muta* at each location.

<table>
<thead>
<tr>
<th>Location</th>
<th>Spongocoel (L)</th>
<th>Volume (L)</th>
<th>Mass (kg)</th>
<th>Flow rate (L h⁻¹)</th>
<th>( \Delta \text{DIN NH}_4^+ ) (µmol L⁻¹)</th>
<th>( \Delta \text{DIN NO}_3^- ) (µmol L⁻¹)</th>
<th>Flux NH₄⁺ (µmol h⁻¹ L⁻¹)</th>
<th>Flux NO₃⁻ (µmol h⁻¹ L⁻¹)</th>
<th>Flux DIN (µmol h⁻¹ L⁻¹)</th>
<th>Flux NH₄⁺ (µmol h⁻¹ kg⁻¹)</th>
<th>Flux NO₃⁻ (µmol h⁻¹ kg⁻¹)</th>
<th>Flux DIN (µmol h⁻¹ kg⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL</td>
<td>4</td>
<td>43</td>
<td>26.8</td>
<td>4050</td>
<td>-0.15</td>
<td>-0.99</td>
<td>-1.13</td>
<td>-1.03</td>
<td>-1.16</td>
<td>-0.21</td>
<td>-0.167</td>
<td>-0.188</td>
</tr>
<tr>
<td>FL</td>
<td>20</td>
<td>101</td>
<td>62.2</td>
<td>18090</td>
<td>0.00</td>
<td>0.65</td>
<td>0</td>
<td>117</td>
<td>13</td>
<td>117</td>
<td>0</td>
<td>189</td>
</tr>
<tr>
<td>FL</td>
<td>39</td>
<td>36</td>
<td>22.4</td>
<td>6372</td>
<td>0.35</td>
<td>-0.30</td>
<td>668</td>
<td>-496</td>
<td>172</td>
<td>904</td>
<td>-0.804</td>
<td>100</td>
</tr>
<tr>
<td>FL</td>
<td>33</td>
<td>111</td>
<td>68.2</td>
<td>21938</td>
<td>0.10</td>
<td>-0.09</td>
<td>26</td>
<td>-26</td>
<td>1</td>
<td>43</td>
<td>-0.42</td>
<td>1</td>
</tr>
<tr>
<td>FL</td>
<td>14</td>
<td>76</td>
<td>47.0</td>
<td>15377</td>
<td>-0.90</td>
<td>0.26</td>
<td>-188</td>
<td>51</td>
<td>-137</td>
<td>-387</td>
<td>82</td>
<td>-305</td>
</tr>
<tr>
<td>FL</td>
<td>7</td>
<td>37</td>
<td>22.9</td>
<td>6480</td>
<td>-0.05</td>
<td>-0.01</td>
<td>-2</td>
<td>-9</td>
<td>-9</td>
<td>-5</td>
<td>-3</td>
<td>-7</td>
</tr>
<tr>
<td>LC</td>
<td>12</td>
<td>64</td>
<td>39.3</td>
<td>4656</td>
<td>5.10</td>
<td>2.60</td>
<td>328</td>
<td>116</td>
<td>444</td>
<td>1653</td>
<td>1037</td>
<td>2690</td>
</tr>
<tr>
<td>LC</td>
<td>5</td>
<td>12</td>
<td>7.4</td>
<td>2649</td>
<td>1.80</td>
<td>-0.30</td>
<td>2665</td>
<td>-181</td>
<td>2464</td>
<td>1313</td>
<td>-122</td>
<td>1191</td>
</tr>
<tr>
<td>LC</td>
<td>4</td>
<td>50</td>
<td>31.1</td>
<td>5423</td>
<td>0.35</td>
<td>0.90</td>
<td>38</td>
<td>92</td>
<td>130</td>
<td>681</td>
<td>1887</td>
<td>2567</td>
</tr>
<tr>
<td>LC</td>
<td>6</td>
<td>49</td>
<td>30.3</td>
<td>7109</td>
<td>0.40</td>
<td>0.35</td>
<td>81</td>
<td>71</td>
<td>151</td>
<td>333</td>
<td>1021</td>
<td>1354</td>
</tr>
<tr>
<td>LC</td>
<td>4</td>
<td>17</td>
<td>10.5</td>
<td>2145</td>
<td>3.30</td>
<td>0.25</td>
<td>213</td>
<td>41</td>
<td>254</td>
<td>4279</td>
<td>324</td>
<td>4603</td>
</tr>
<tr>
<td>LC</td>
<td>14</td>
<td>63</td>
<td>38.9</td>
<td>8463</td>
<td>0.45</td>
<td>0.20</td>
<td>60</td>
<td>31</td>
<td>90</td>
<td>350</td>
<td>272</td>
<td>622</td>
</tr>
<tr>
<td>LSI</td>
<td>5</td>
<td>22</td>
<td>13.6</td>
<td>6912</td>
<td>0.36</td>
<td>1.23</td>
<td>-180</td>
<td>438</td>
<td>258</td>
<td>-292</td>
<td>710</td>
<td>418</td>
</tr>
<tr>
<td>LSI</td>
<td>10</td>
<td>45</td>
<td>27.6</td>
<td>16640</td>
<td>0.11</td>
<td>0.53</td>
<td>40</td>
<td>66</td>
<td>106</td>
<td>64</td>
<td>107</td>
<td>172</td>
</tr>
<tr>
<td>LSI</td>
<td>29</td>
<td>81</td>
<td>50.2</td>
<td>28800</td>
<td>0.50</td>
<td>0.63</td>
<td>213</td>
<td>159</td>
<td>372</td>
<td>344</td>
<td>258</td>
<td>603</td>
</tr>
<tr>
<td>LSI</td>
<td>2</td>
<td>18</td>
<td>11.0</td>
<td>11040</td>
<td>-0.12</td>
<td>0.87</td>
<td>-162</td>
<td>1202</td>
<td>1040</td>
<td>-262</td>
<td>1948</td>
<td>1686</td>
</tr>
<tr>
<td>LSI</td>
<td>3</td>
<td>23</td>
<td>14.0</td>
<td>2687</td>
<td>-0.19</td>
<td>0.30</td>
<td>12</td>
<td>15</td>
<td>27</td>
<td>19</td>
<td>24</td>
<td>43</td>
</tr>
<tr>
<td>LSI</td>
<td>2</td>
<td>14</td>
<td>8.6</td>
<td>11220</td>
<td>2.53</td>
<td>0.67</td>
<td>4313</td>
<td>793</td>
<td>5106</td>
<td>6990</td>
<td>1286</td>
<td>8276</td>
</tr>
</tbody>
</table>
Figure 3.1. δ¹⁵N and δ¹³C values (mean ± SE) for Xestospongia muta (n=3) for each location. FL= Florida Keys LC=Little Cayman, LSI=Lee Stocking Island, Bahamas.
Figure 3.2. The average flow rate (l h$^{-1}$ ± SE) of sponges from each location. FL= Florida Keys LC= Little Cayman, LSI= Lee Stocking Island, Bahamas
Figure 3.3. The average flux (mean ± SE) of NH$_4^+$, NO$_x^-$, and DIN for each location normalized to sponge volume (previous page) (A) and mass (B). Treatment groups with similar superscripts are not statistically different from each other. FL= Florida Keys LC=Little Cayman, LSI=Lee Stocking Island, Bahamas
Figure 3.4. The average flux of total DIN for each location (± SE) based on volume-normalized fluxes. FL = Florida Keys, LC = Little Cayman, LSI = Lee Stocking Island, Bahamas.
Filtered Cells (cells m$^{-1}$, mean ± SE)

- Cyanobacteria
- Heterotrophic
- Prochlorophytes
- Total Cells

**Cell Type**

*Note: The diagram shows different cell types with corresponding values, but specific details are not transcribed here.*
Filtration Efficiency (% mean ± SE)

Cyanobacteria  Heterotrophic  Prochlorophytes  Total Cells

Cell Type
Figure 3.5. Number of filtered cells (page 118) (A), filtration efficiency (previous page) (B), and particulate organic matter as carbon and nitrogen (C) available from filtered cells for *X. muta* from LSI (n = 4). Treatment groups (mean ± SE) with similar superscripts are not statistically different from each other.
A

\[ \delta^{15}\text{N} \% (\text{mean} \pm \text{SE}) \]

- **Sponge**
- **Bacterial**

![Graph showing the change in \( \delta^{15}\text{N} \% \) over time (h) for Sponge and Bacterial samples.](image)

- Time (h) range from 0 h to 12 h.
- The graph indicates a significant increase in \( \delta^{15}\text{N} \% \) over time for both Sponge and Bacterial samples.
B

-17

-18

-19

-20

-21

-22

$\delta^{13}C$ % (mean ± SE)

Time (h)

0 h 3 h 6 h 12 h

Sponge
Bacterial

• b
□ b

-17

-18

-19

-20

-21

-22

$\delta^{13}C$ % (mean ± SE)

Time (h)

0 h 3 h 6 h 12 h

Sponge
Bacterial

• b
□ b
Figure 3.6. A) $\delta^{15}$N over time for $^{15}$N nitrate (page 109) and B) $\delta^{13}$C over time for $^{13}$C bicarbonate enriched sponge and bacterial fractions (previous page), C) $\delta^{15}$N over time for $^{15}$N ammonium enriched sponge and bacterial fractions. Samples collected at 3 h were under low irradiances while samples collected at 6 h had been exposed to sunlight (Under screens representative of light levels at collection depth) and samples from 12 h were collected at night (for both experiments). Treatment groups (mean ± SE) with similar superscripts are not statistically different from each other.
Figure 3.7. $\Delta^{15}\text{N}$ at each time point for nitrate enriched (A) and ammonium enriched (B) samples. Diamonds represent sponge fractions and squares represent bacterial fractions. The gray bar represents low light during incubation with tracer, the white bar represents exposure to sunlight, and the black bar represents nighttime up to the point of sampling. Standard error is shown.
Figure 3.8. (Previous page) Neighbor-joining tree based on *nifH* deduced amino acid sequences. *Desulfovibrio desulfuricans* was used as the outgroup for the Cyanobacteria (A) and *Trichodesium erythraeum* was the outgroup for the Proteobacteria (B). Sequences from this study are labeled as SW for seawater followed by clone number and Xm for *Xestospongia muta* followed by the clone number. Bootstrap (n = 100) values above 50% are shown.
Figure 3.9. Portions of a neighbor-joining tree created in ARB using the nfH ARB database (Zehr et al., 2003) showing sequences from the current study that fell into groups with Deltaproteobacteria (A) and Alphaproteobacteria (B). Sequences from this study have a square by them and are labeled as Xm or SW followed by clone number (X. muta and seawater respectively) and GenBank accession number. Sequences from the ARB database have the database identification number in parentheses. Collapsed branches show the number of sequences in the cluster and due to space limitation ID numbers can be provided upon request. Arrow(s) indicates where this is connected to a large tree and scale bar represents 10% sequence divergence.
Figure 3.10. Portions of a neighbor-joining tree created in ARB using the nifH ARB database (Zehr et al., 2003) showing sequences from the current study that fell into two different groups with Cyanobacteria, one with seawater sequences only (A) and one with one X. muta sequence (B). Sequences from this study have a square by them and are labeled as Xm or SW followed by clone number (X. muta and seawater respectively) and GenBank accession number. Sequences from the ARB database have the database identification number in parentheses. Collapsed branches show the number of sequences in the cluster and due to space limitation ID numbers can be provided upon request. Arrow(s) indicates where this is connected to a large tree and scale bar represents 10% sequence divergence.
**Bootstrap values**

<table>
<thead>
<tr>
<th>Node</th>
<th>NJ</th>
<th>ML</th>
<th>Bootstrap values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>71</td>
<td>31</td>
<td>Xm5.15 (JQ912197)</td>
</tr>
<tr>
<td>2</td>
<td>96</td>
<td>96</td>
<td>Xm4.7 (JQ912174)</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>-</td>
<td>SW6.3 (JQ912132)</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>-</td>
<td>SW1.4 (JQ912114)</td>
</tr>
<tr>
<td>5</td>
<td>75</td>
<td>-</td>
<td>Xm4.4 (JQ912171)</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>53</td>
<td>Xm5.10 (JQ912192)</td>
</tr>
<tr>
<td>7</td>
<td>64</td>
<td>61</td>
<td>Xm6.10 (JQ912208)</td>
</tr>
<tr>
<td>8</td>
<td>64</td>
<td>-</td>
<td>Uncultured cnarcnachaeote clone 25 3 Aplysina insularis (EU049839)</td>
</tr>
</tbody>
</table>

**Uncultured cnarcnachaeote isolate B11(1:A3) Xestospongia muta (GQ485668)**

**Uncultured cnarcnachaeote isolate B9N(1:F7) Xestospongia muta (GQ485742)**

**Uncultured cnarcnachaeote isolate B9N(1:F7) Xestospongia muta (GQ485742)**

**Uncultured cnarcnachaeote clone 25 3 Aplysina insularis (EU049839)**

---

**Uncultured cnarcnachaeote isolate B9N(1:F7) Xestospongia muta (GQ485742)**

**Uncultured cnarcnachaeote isolate B9N(1:F7) Xestospongia muta (GQ485742)**

**Uncultured cnarcnachaeote clone PlakorGrA Plakortis sp. (EU049828)**

**Uncultured cnarcnachaeote clone PlakorGrA Plakortis sp. (EU049828)**

**Uncultured cnarcnachaeote clone PlakorGrA Plakortis sp. (EU049828)**

**Uncultured cnarcnachaeote clone PlakorGrA Plakortis sp. (EU049828)**

**Uncultured cnarcnachaeote clone Cliona sp. (EU049834)**

**Uncultured cnarcnachaeote clone MA7-28 coral (EF382527)**

**Uncultured cnarcnachaeote clone MA7-28 coral (EF382527)**

**Uncultured cnarcnachaeote clone LuffGrA Luffariella sp. (EU049831)**

---

**SW5.1 (JQ912126)**

**SW5.4 (JQ912128)**
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Uncultured crenarchaeote clone Aplysia insularis (EU049838)</td>
</tr>
<tr>
<td>SW6.6</td>
<td>(JQ912133)</td>
</tr>
<tr>
<td>Xm2.13</td>
<td>(JQ912155)</td>
</tr>
<tr>
<td>Xm4.8</td>
<td>(JQ912175)</td>
</tr>
<tr>
<td>Xm6.1</td>
<td>(JQ912199)</td>
</tr>
<tr>
<td>Xm2.8</td>
<td>(JQ912151)</td>
</tr>
<tr>
<td>Uncultured crenarchaeote clone Aplysia insularis (EU049838)</td>
<td></td>
</tr>
<tr>
<td>Xm1.2</td>
<td>(JQ912135)</td>
</tr>
<tr>
<td>Xm2.15</td>
<td>(JQ912157)</td>
</tr>
<tr>
<td>Xm4.12</td>
<td>(JQ912178)</td>
</tr>
<tr>
<td>Xm6.3</td>
<td>(JQ912201)</td>
</tr>
<tr>
<td>Uncultured crenarchaeote clone Plakortis sp. (EU049828)</td>
<td></td>
</tr>
<tr>
<td>Xm1.8</td>
<td>(JQ912139)</td>
</tr>
<tr>
<td>Xm6.15</td>
<td>(JQ912213)</td>
</tr>
<tr>
<td>SW2.4</td>
<td>(JQ912149)</td>
</tr>
<tr>
<td>SW5.3</td>
<td>(JQ912185)</td>
</tr>
<tr>
<td>SW6.1</td>
<td>(JQ912199)</td>
</tr>
<tr>
<td>Xm1.4</td>
<td>(JQ912134)</td>
</tr>
<tr>
<td>Xm1.5</td>
<td>(JQ912136)</td>
</tr>
<tr>
<td>Xm1.9</td>
<td>(JQ912140)</td>
</tr>
<tr>
<td>Xm1.11</td>
<td>(JQ912142)</td>
</tr>
<tr>
<td>Xm1.14</td>
<td>(JQ912145)</td>
</tr>
<tr>
<td>Xm2.2</td>
<td>(JQ912147)</td>
</tr>
<tr>
<td>Xm2.3</td>
<td>(JQ912148)</td>
</tr>
<tr>
<td>Xm2.4</td>
<td>(JQ912149)</td>
</tr>
<tr>
<td>Xm2.12</td>
<td>(JQ912154)</td>
</tr>
<tr>
<td>Xm2.14</td>
<td>(JQ912156)</td>
</tr>
<tr>
<td>Xm3.1</td>
<td>(JQ912158)</td>
</tr>
<tr>
<td>Xm3.4</td>
<td>(JQ912160)</td>
</tr>
<tr>
<td>Xm3.5</td>
<td>(JQ912161)</td>
</tr>
<tr>
<td>Xm4.1</td>
<td>(JQ912169)</td>
</tr>
<tr>
<td>Xm6.3</td>
<td>(JQ912170)</td>
</tr>
<tr>
<td>Xm4.10</td>
<td>(JQ912176)</td>
</tr>
<tr>
<td>Xm4.14</td>
<td>(JQ912180)</td>
</tr>
<tr>
<td>Xm4.15</td>
<td>(JQ912181)</td>
</tr>
<tr>
<td>Xm5.3</td>
<td>(JQ912185)</td>
</tr>
<tr>
<td>Xm5.5</td>
<td>(JQ912187)</td>
</tr>
<tr>
<td>Xm5.7</td>
<td>(JQ912189)</td>
</tr>
<tr>
<td>Xm5.9</td>
<td>(JQ912191)</td>
</tr>
<tr>
<td>Xm5.12</td>
<td>(JQ912194)</td>
</tr>
<tr>
<td>Xm5.14</td>
<td>(JQ912196)</td>
</tr>
<tr>
<td>Xm1.13</td>
<td>(JQ912144)</td>
</tr>
<tr>
<td>Xm6.2</td>
<td>(JQ912200)</td>
</tr>
<tr>
<td>Xm4.6</td>
<td>(JQ912173)</td>
</tr>
<tr>
<td>Xm6.11</td>
<td>(JQ912209)</td>
</tr>
<tr>
<td>Xm4.13</td>
<td>(JQ912179)</td>
</tr>
<tr>
<td>Xm6.12</td>
<td>(JQ912200)</td>
</tr>
<tr>
<td>Xm5.31</td>
<td>(JQ912193)</td>
</tr>
<tr>
<td>Xm6.9</td>
<td>(JQ912207)</td>
</tr>
<tr>
<td>Xm5.4</td>
<td>(JQ912186)</td>
</tr>
<tr>
<td>Xm6.14</td>
<td>(JQ912212)</td>
</tr>
<tr>
<td>Xm6.5</td>
<td>(JQ912203)</td>
</tr>
<tr>
<td>Xm6.16</td>
<td>(JQ912214)</td>
</tr>
<tr>
<td>Xm3.10</td>
<td>(JQ912165)</td>
</tr>
<tr>
<td>SW1.2</td>
<td>(JQ912123)</td>
</tr>
</tbody>
</table>
Figure 3.11. Neighbor-joining (NJ) tree based on archaeal amoA gene sequences (including previous two pages). Sequences from this study are labeled as Xm or SW followed by clone number (X. muta and seawater respectively). Bootstrap values above 50% are shown for nodes on the NJ version and maximum likelihood (ML) version of the tree.
Chapter IV

METATRANSCRIPTOMIC ANALYSIS OF THE SPONGE XESTOSPONGIA MUTA: FUNCTIONAL ACTIVITY OF THE SPONGE AND THE PROKARYOTIC COMMUNITY

Introduction

Sponges are widespread and ecologically important members of many benthic habitats and of ecosystems such as coral reefs (Diaz and Rutzler 2001). In particular, sponges are known to be an important link in coupling water column processes to benthic habitats as they actively pump water through the sponge, modifying water column constituents as they do so (Lesser 2006). In high microbial abundance (HMA) sponges, the water passing through the sponge is often modified as a result of physiological processes of symbiotic prokaryotes (Diaz and Ward 1997; Southwell et al. 2008 b). However, many of these processes are poorly understood and there is a need to better understand the structure and function of sponge symbiotic prokaryotic communities, as this is vital to better understanding the ecology and function of important coral reef member such as sponges.
Studies investigating the community composition of sponge prokaryotic symbionts have often used non culture-based methods for bacterial characterization as opposed to culture-based methods, typical of classical microbiology, due to difficulty in culturing sponge bacteria (Olson et al. 2000); however, there have been some notable exceptions where sponge bacteria have been successfully cultured (Wilkinson 1978; Shieh and Lin 1994; Olson et al. 2000; Webster and Hill 2001; Taylor et al. 2005). Non culture-based methods have typically included measuring chlorophyll a or other photosynthetic pigment concentrations over an environmental gradient (Erwin and Thacker 2007; Southwell 2007) or in response to experimental manipulation (Erwin and Thacker 2008), or the use of stable isotopes in experimental manipulations (Wilkinson et al. 1999; Weisz et al. 2007a; Hoffmann et al. 2009; Freeman and Thacker 2011). One approach has been the use of microscopy techniques (Rutzler 1990; Webster et al. 2001; Usher et al. 2004) that have continued to mature with the use of antibody and nucleotide probes (e.g., FISH, Manz et al. 2000). Few extensive physiological experiments have been done but those that have been performed have led to key insights into sponge-microbe relationships particularly in regards to photosymbionts (Wilkinson and Fay 1979; Arillo et al. 1993). These studies have provided valuable insight into the relationship between sponges and their symbionts, and the distribution of key symbionts among sponge hosts.

Now, the vast majority of sponge-prokaryote studies have utilized DNA sequencing methods often in combination with other experimental approaches. Historically, such sequencing studies have utilized Sanger sequencing following cloning (Fieselar et al. 2004; Steindler et al. 2005; Mohamed et al. 2008; Hentschel et al. 2002), denaturing gradient gel electrophoresis (DGGE) (Webster et al. 2001; Webb and Maas 2002;
Webster et al. 2004), or restriction fragment length polymorphism (RFLP) (Hill et al. 2006; Lopez-Legentil et al. 2010; Negandhi et al. 2010), and have added valuable genetic information concerning identification and characterization of the prokaryotic symbiotic communities of sponges. With the advances in DNA sequencing technologies, more recent studies are utilizing high throughput sequencing (HTS) as a technique to better quantify the structure and function of sponge prokaryotic communities (Thomas et al. 2010; Siegl et al. 2010; Webster et al. 2010a; Lee et al. 2010; Schmitt et al. 2011). It is important to note however, that such methods are in fact complementary methods and physiological and ecological experiments are still needed to provide a complete picture of important biological processes. That said; HTS provides a unique avenue for both characterizing sponge symbiotic prokaryotic community structure based on 16S rRNA genes and characterizing function based on genomic and transcriptomic techniques.

Using HTS, the *Amphimedon queenslandica* genome has been sequenced, as well as the partial genome of a symbiotic member of the cosmopolitan sponge specific phylum *Candidatus (Ca)* Poribacteria. A study by Thomas and colleagues (2010) used metagenomics to compare genes expressed by a sponge microbial community and by the water column community. Differences in the types of genes expressed and the differential expression of genes were observed between water column bacteria and the sponge bacteria datasets, providing further evidence for the existence and persistence of a sponge-specific bacterial community (Thomas et al. 2010). Another recent study delved further into the function of sponge symbionts by sequencing the metatranscriptome of the sponge *Geodia barretti* (Radax et al. 2012). Important metabolic processes such as nitrification were quantified as transcripts that encode for key enzymes required for this
and other related processes. Additionally, rRNA was recovered from the same samples and showed that the dominant active bacteria in *G. barretti* were from the Chloroflexi, *Ca* Poribacteria, and Acidobacteria phyla (Radax et al. 2011). Another landmark study in sponge microbiology compared the metagenomes of six different sponge hosts to examine and describe coevolution of sponge microbial communities (Fan et al. 2012). The use of HTS methods in sponge microbiology allows us to gain unprecedented insight into the composition and function of sponge prokaryotic symbiotic communities.

On Caribbean coral reefs, the giant barrel sponge *Xestospongia muta* is abundant and ecologically important (McMurray et al. 2008; Chapters 2, 3). *X. muta* is also a high microbial abundance sponge with a diverse symbiotic prokaryotic community (Chapter 2). Furthermore, experimental studies have shown that *X. muta* and its symbiotic prokaryotic community significantly influence nitrogen cycling on coral reefs and more specifically, that these sponges can serve as both a source and sink for dissolved inorganic nitrogen (Chapter 3). To better understand the microbially-mediated processes underlying the dynamics of nitrogen cycling in this important sponge, we have sequenced the host transcriptome and symbiont metatranscriptome from one *X. muta* individual collected from coral reefs at each of three locations in the Caribbean: Florida Keys, USA, Little Cayman, Cayman Islands, and Lee Stocking Island, Bahamas. Here, I characterize the sponge transcriptome as well as the metatranscriptome of the symbiotic prokaryotic community, both of which highlight key physiological processes occurring in the sponge holobiont.
Methods

Sample Collection

Sample collection was as described in Chapter 2. Briefly, Replicate sponges (n=6) were sampled at approximately 15 m from each of three locations: Rock Bottom Reef, Little Cayman, Cayman Islands (LC) (19°42'7.36" N, 80°3'24.94" W), North Perry Reef, Lee Stocking Island (LSI) (23°47'0.03" N, 76°6'5.14" W), Bahamas, and Conch Reef, Key Largo, FL (FL) (24°57'0.03" N, 80°27'11.16" W). Sponge pieces were cut from the top rim of the sponge (“pie slice” of pinacoderm and mesohyl) and placed in a plastic bag with minimal amounts of seawater and placed on ice until reaching the laboratory. Each sponge sample was then placed in RNAlater (Ambion), and were kept frozen until reaching the University of New Hampshire where they were maintained at -70°C. While six sponges were collected and preserved at each location, only one sample from each location was used for RNA extractions and sequencing and these will be referred to as XmFL, XmLC, XmLSI.

RNA extraction and sequencing

Total RNA was extracted from sponge samples using the RNeasy mini kit (Qiagen, Valencia, CA) and DNase treated using DNA free (Ambion, Grand Island, NY). The resulting RNA concentration and quality was determined using a NanoDrop 2000 and a bionanalyzer (Agilent, Santa Clara, CA). Samples with a RIN of at least 7.5 were selected for sequencing (all samples had concentrations of 90-230 ng μl⁻¹) and were
immediately sent to the University of Illinois (Urbana-Champaign) for library construction and sequencing. Three samples were processed and sent for sequencing (XmFL, XmLC, XmLSI). At the University of Illinois, eukaryotic rRNA was removed from 1 μg of total RNA with Ribominus Eukaryote kit (Invitrogen) and library construction was performed using TruSeq RNA Sample Prep kit (Illumina, San Diego, CA) per manufacturers instructions. The library containing the three sponge samples plus six coral samples (Jarett 2012) were pooled in equimolar concentration, quantitated by qPCR, amplified with 10 cycles of PCR, and sequenced using one lane on an Illumina HiSeq2000 for paired-end reads approximately 100 nt in length and an average insert size of 240 nt (yielding approximately ~380 M reads).

Raw reads were quality filtered using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit) and Prinseq (Schmieder and Edwards 2011). In FASTX-Toolkit, reads were filtered based on an average quality score of 30, a minimum length of 50, and trim adapter sequences using scripts by De Wit et al. (2012). Because almost all reads were observed to contain an “N” as the first nucleotide, the first nucleotide was trimmed from all reads in Prinseq. Following this, reads from paired files were check to remove any unpaired reads in Fast-x using the paired-end script from De Wit et al. (2012).

**Assembly of putative mRNA reads**

Quality filtered reads were mapped against the SILVA 111 SSU and LSU nr reference databases (Quast et al. 2012) in CLC Genomics Workbench 6.0 (CLC bio). Default settings of mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction
= 0.5, similarity fraction = 0.8, were used and map randomly was chosen for non-specific match handling. Unmapped reads were saved as putative mRNA reads. The putative mRNA reads were assembled in Velvet (Zerbino and Birney 2008) with a kmer of 61 in velveth and flags –exp_cov auto –cov_cutoff auto in velvetg, and then assembled in Metavelvet (Namiki et al. 2012) with the flags: -ins_length 240 -exp_covs 65.5_58.5_49.5_25.5. Running velveth with different kmer values and comparing the N50 values and contig numbers from each run allowed for optimization of kmer size used in velveth. Coverage in Metavelvet was estimated by first running the program with default settings and then using the Velvet script scriptEstimateCovMulti.py.

Putative sponge mRNA reads were binned by mapping quality filtered reads to the *Amphimedon queenslandica* genome with low complexity regions masked (obtained from Ensemble Metazoa (EMBL-EBI, metazoan.ensemble.org)) in CLC Genomics Workbench 6.0. Mapping settings were as listed above, and unmapped reads were then used with the filter_fasta.py script in the Quantitative Insights Into Microbial Ecology (QIIME) pipeline (Caporaso 2010) to filter the reads by and collect reads that mapped to the *A. queenslandica* genome. The collected sponge reads were then assembled *de novo* in CLC Genomics Workbench 6.0.

**Taxonomic and functional assignment of putative mRNA contigs**

Contigs from Metavelvet were then used in a BLASTx search against the NCBI RefSeq protein database (Sayers et al. 2010) using the Community Cyberinfrastructure for Advanced Microbial Ecology Research and Analysis (CAMERA) portal (Sun et al. 2010). The top 20 matches with an E-value of $10^{-4}$ were saved and this setting was used
in the rest of the BLAST searches. MEGAN (Huson et al. 2007) was then used to visualize the BLASTX search results to obtain taxonomic information on the reads as well as classification by SEED subsystems (Overbeek et al. 2005) and KEGG pathways (Kanehisa and Goto 2000; Kanehisa et al. 2011).

Contiguous sequences (contigs) from the sponge assembly were analyzed similarly by a BLASTx search against the NCBI refseq protein database (Sayers et al. 2010) in CAMERA (Sun et al. 2010). MEGAN (Huson et al. 2007) was used to visualize the BLASTX search results to obtain taxonomic information on the reads as well as classification by SEED subsystems (Overbeek et al. 2005) and KEGG pathways (Kanehisa and Goto 2000; Kanehisa et al. 2011). Additionally, sponge contigs were imported to Blast2GO (Conesa and Götz 2008) and another BLASTX was performed against NCBI non-redundant protein database (Sayers et al. 2010), annotated with Gene Ontology (GO) terms and then searched against the KEGG databases and visualized in KEGG pathways (Kanehisa and Goto 2000; Kanehisa et al. 2011).

Results

Illumina sequencing, quality filtering and assembly

A total of over 386 million reads were obtained from one lane of Illumina sequencing, of this ~18 million reads corresponded to each read 1 and read 2 for XmFL, ~23 million reads corresponded to each read 1 and read 2 for XmLC, and ~29 million reads corresponded to each read 1 and read 2 for XmLSI. The proportion of reads removed and trimmed during quality filtering steps for each sample is shown in Table
4.1. Mapping to SILVA removed 16 - 23% of reads prior to performing the velvet and metavelvet assembly. The metavelvet assembly yielded 36,061 contigs with an N50 = 835 (50% of contigs are this size or larger). For the sponge only reads, 4 to 6% reads mapped to the *A. queenslandica* genome and when assembled yielded 890 contigs with an N50 = 293.

**Analysis of the assembly of putative sponge mRNA reads**

The majority of putative sponge mRNA contigs were taxonomically assigned to *A. queenslandica* or other sponges including *X. muta* using both MEGAN (Figure 4.1A) and Blast2GO (Figure 4.1B) (MEGAN, 447 contigs; Blast2GO, ~700 contigs). However, some (45 – 60 contigs) were assigned to other eukaryotic organisms, most of which could have been transiently located in the sponge tissue at the time of collection, or were unassigned (142 contigs) (Figure 4.1). Based on the SEED subsystem classification most contigs were unassigned, and the rest were split between the subsystems of: carbohydrates, stress response, protein metabolism, RNA metabolism, and respiration (Figure 4.2). More detail was obtained from the KEGG classification, where contigs were split roughly evenly into: metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, and human disease (Figure 4.3).

Within metabolism, most contigs matched to enzymes involved in oxidative phosphorylation, followed by carbon fixation in photosynthetic organisms, methane metabolism, nitrogen metabolism and sulfur metabolism (Figure 4.4). The enzymes involved in photosynthetic pathways included enzymes also common to carbohydrate and
amino acid metabolisms such as aminotransferases, pyruvate kinase, fructose-1,6-bisphosphatase. The enzymes involved in methane metabolism also included those common to general metabolic processes such as catalase, enolase, fructose-1,6-bisphosphatase, and glycine hydroxymethyltransferase. Enzymes involved in metabolism of major amino acid groups were also present in the sponge (Figure 4.5A). Enzymes involved in carbohydrate metabolism represented diverse pathways including the citrate cycle (TCA cycle), glycolysis, propanoate and butanoate metabolism, pyruvate metabolism, and the metabolism of a variety of sugars (Figure 4.5B). Contigs matching to enzymes and other proteins involved in vitamin metabolism such as folate biosynthesis were recovered as well as for enzymes involved in chlorophyll and porphyrin metabolism (Figure 4.5C). Matches to enzymes and other proteins involved in lipid metabolism were mostly from pathways involved in fatty acid metabolism, followed by glycerolipid and sphingolipid metabolism (Figure 4.5D).

Genetic information processing included proteins involved in transcription (i.e., RNA polymerase, transcription factors, and spliceosomes), translation (i.e., ribosomes and aminoacyl-tRNA biosynthesis), and folding, sorting and degradation (i.e., molecules involved in protein transport, proteolysis, and RNA degradation). Environmental information processing grouping was dominated by signal transduction, followed by membrane transport and signal molecules and interaction (Figure 4.6A). Within the signal transduction grouping were matches to proteins in a variety of common signal transduction pathways in eukaryotes, the largest proportion of which represented the MAPK signaling pathway followed by the Wnt signaling pathway (Figure 4.6B).
The cell-processing category included four major groupings: cell communication, cell growth, regulation of actin cytoskeleton, and transport and catabolism (Figure 4.7A). Proteins involved in cell communication were divided into adherens junction, focal adhesion, gap junction, and tight junction (Figure 4.7B). Within the cell growth category were proteins involved in major regulatory pathways such as apoptosis and p53 signaling pathway, but other cell cycle pathways such as oocyte meiosis were present as well (Figure 4.7C). The transport and catabolism categories were equally dominated by endocytosis and lysosome proteins, followed by peroxisome proteins (Figure 4.7D).

The largest category in organismal systems was the immune system, followed by the nervous system and endocrine system groups and several more regulatory system groupings (Figure 4.8A). Contigs matching to proteins in the immune system category typically included kinases, protein subunits of complexes such as nuclear factor NF-kappa-B, factor proteins such as TNF receptor-associated factor 3, as well as inhibitors and substrates of a variety of regulatory enzymes, and represented major and diverse signaling pathways characteristic of innate immunity (Figure 4.8B). Molecules involved in the endocrine system category included enzymes for protein degradation, fatty acid metabolisms and signaling (i.e., peroxisome proliferator-activator receptor (PPAR) and adipocytokine pathways), phosphatases and kinases specific to carbohydrate signaling (i.e., insulin pathway) and lipid signaling pathways, GTPase, and calmodulin, among other regulatory proteins (Figure 4.8C). Excretory pathways present in the sponge include ion regulation (i.e., aldosterone-regulated sodium reabsorption pathway, proximal tubule bicarbonate reclamation pathway), and osmotic control (i.e., vaspressin-regulated water reabsorption pathway) (Figure 4.8D). The nervous system pathways include a variety of
enzymes and other proteins that have homologs in characterized human pathways for neuron signal transduction and potentiation, and cell differentiation (Figure 4.8E).

Proteins involved in secondary metabolite and xenobiotic degradation and metabolism were also recovered (Figure 4.9). Additionally, proteins involved in mediating disease processes were recovered including proteins associated with infectious disease, cancers, and neurodegenerative disease in particular (Figure 4.10A). Because infectious disease was the largest category and the most relevant category for interactions between the host and sybimonts, infectious disease was further characterized to include proteins involved in *Helicobacter pylori* infection pathways as characterized in humans, as well as *Escherichia coli* and *Vibrio cholerae* (Figure 4.10B).

### Taxonomic and functional analysis of the assembly of total putative mRNA reads

Of the contigs that were classified taxonomically as eukaryotes, most (5,326) were classified as Porifera, while the rest were classified as other eukaryotes that could have been in the sponge or would be genetically similar to organisms in the sponge at the time of sampling (Figure 4.11, 4.12A). The portion of contigs classified as prokaryotes included six bacterial phyla (n=685 contigs including unclassified Bacteria) and one archaeal phylum, Thaumarchaeota (n=129 contigs) (Figure 4.12B). Significant portions of the contigs were also classified as low complexity likely due to repeat regions or homopolymer runs (n=4,053).

Using SEED classification, the contigs were divided into 15 subsystems, although significant portions of the contigs were not classified (n= 8,011) (Figure 4.13). A comparison of subsystems present in the sponge only versus the metatranscriptome
assembly is shown in Table 4.2. Respiration included proteins mainly involved in oxidative phosphorylation and electron transport, including cytochrome B6-F complex, soluble cytochromes and other electron carriers, and formate hydrogenase. The stress subsystem included three main categories of pathways: heat shock dnaK gene cluster, Hfl operon, and oxidative stress. The virulence subsystem included five subcategories of virulence related proteins (Figure 4.14).

Using KEGG classification, more contigs were classified than with SEED and divided into the same major groups as observed with the sponge contigs (Figure 4.15). Within metabolism, the largest portions were classified as carbohydrate metabolism, amino acid metabolism, energy metabolism, and lipid metabolism (Figure 4.16). Several categories were further examined, including energy metabolism, which contained diverse metabolic pathways for energy conservation (Figure 4.17A). Vitamin metabolism also contained more diverse subcategories than the sponge contigs (Figure 4.17B). Proteins involved in the metabolism of various xenobiotics were also recovered (Figure 4.17C). Differences in the categories of metabolism of secondary metabolites between the sponge and the metatranscriptome assemblies are shown in Table 4.3.

Genetic processing information included proteins mostly involved in folding, sorting and degradation, as well as translation, transcription, and replication and repair processes. For environmental processing, signal transduction pathways contained majority of matches to contigs (~75%), with the rest split between membrane transport and singling molecules and interaction pathways. Cell processing pathways included cell communication, cell growth and death, transport and catabolism, as well as cell motility (Figure 4.18A). Organismal systems also included pathways common to all major
systems in vertebrate animals, as well as the general categories of development and environmental adaptation (Figure 4.18B).

The category of human disease included major pathways involved in neurodegenerative diseases, infectious diseases, and cancer (Figure 4.19A). Further characterization of immune system pathways revealed diverse proteins and enzymes involved in a variety of immune responses (Figure 4.19B). Additionally, the infectious disease category also contained a variety of pathways characterized by specific pathogenic bacteria in humans (Figure 4.19C).

**Discussion**

This study has provided important insight into the activity and functional complexity in the abundant Caribbean sponge, *Xestospongia muta*, and its diverse symbiotic prokaryotic community. While one previous metatranscriptomic study included minimal information on sponge transcripts or genome in addition to the prokaryotic transcripts or genomes (Radax et al. 2012), this is the first study, to my knowledge, to simultaneously characterize both the sponge transcriptome and the metatranscriptome of the prokaryotic community in depth. An active and diverse prokaryotic community is present in *X. muta*, and the sponge exhibits several potential mechanisms for interacting with its symbiotic community, in addition to demonstrating tight networks of genes involved in the regulation of cell growth, cell differentiation, sensory pathways, and immune responses.
Sponge activity and function

The results of the sponge assembly are similar to previous studies of sponge genomes, transcriptomes, or express sequence tag (EST) in terms of functional categories, and largely represent basic, and highly conserved, animal cellular and metabolic processes (Srivastava et al. 2010; Harcet et al. 2010). For example, EST analysis of the sponge *Suberites domuncula* revealed that the largest functional group of transcripts was categorized in cellular processes and signaling (36%), followed by metabolism (18%), information storage and processing (16%), and then poorly or uncharacterized transcripts (Harcet et al. 2010). For *X. muta* the main difference in its transcriptome is that the metabolism and cellular processes were roughly equally represented. It is difficult what, if any, differences in metabolism between the two species of sponge may be resulting from, as further characterization of metabolic pathways in the *S. domuncula* study was not performed. It is possible that there are some contaminating prokaryotic contigs classified under the sponge metabolism in *X. muta*, such as methane metabolism and carbon fixation in photosynthetic organisms.

In general, the heterotrophic metabolism of the sponge was apparent because of the number of transcripts matching to proteins involved in oxidative phosphorylation, such as ATPase, reductase (H+- translocating), oxidase, and a reductase. Also, components of the TCA cycle, as well as glycolysis and the pentose phosphate pathway were recovered. Nitrogen metabolism in the sponge was characterized by enzymes involved in regulating glutamine and glutamate levels which is the primary mechanism of cellular control for
enzymes involved in nitrogen assimilation such as glutaminase I, dehydrogenase (NAD(P)+), dehydrogenase, and reductase. Several contigs also matched to enzymes involved in nitrification and denitrification pathways, likely indicating some prokaryotic contamination. However, these were a reductase and two oxidases, and it is also possible that they are similar enough to the prokaryotic reductases in these pathways to be considered a match, but may be eukaryotic enzymes from the sponge.

The recovery of enzymes and other proteins involved in the metabolism of amino acids, vitamins, lipids were as expected for a functioning eukaryotic organism, with amino acids and other molecules being synthesized and recycled. The metagenomic and transcriptomic study of sponges is a rapidly expanding area of research due to the long evolutionary history of sponges and their potential for containing bioactive compounds (Taylor et al. 2007). However, there are only a handful of studies that have characterized the functional roles of sponge transcripts in a similar way to what has been done for the prokaryotic portion, and certainly even rarer when done simultaneously, leaving few studies for comparison with the results presented here. Several, interesting comparisons can still be made which highlight future areas of research. One study that characterized the microbial community transcripts from the sponge *Cymbastela concentrica*, and from the water column indicated that the sponge was likely gaining vitamin B12 via transfer from its prokaryotic symbionts rather than from its food (Thomas et al. 2010). In *X. muta*, folate biosynthesis and the “one carbon pool by folate” (whereby folate adds single carbon species (e.g., methyl, methylene, etc.) to other molecules, often purines and pyrimidines), pathways dominate the vitamin metabolism. Folate, also known as vitamin B9, may be obtained from the ingestion of food or transferred by symbionts as with
vitamin B12 in C. concentrica. In the X. muta metatranscriptome, diverse pathways of vitamin and cofactors synthesis were observed, such as vitamin B6 metabolism, and any of these could be transferred to the sponge host. Further investigation of the folate, and other vitamin molecules and pathways, could be an interesting avenue of future research as it relates to the integration of the sponge host and its microbial symbionts.

Another interesting observation was the identification of enzymes involved in butanoate metabolism in X. muta. Butarate is a product produced solely by obligate anaerobes as an end product of fermentation (Zeikus 1980). In human, such bacteria are part of the microbiome in the colon and butyric acid from the colon fauna is transferred to epithelial cells and processed in a butuanoate pathway in mitochondria (Roediger 1982; Miller and Wolin 1996). Butarate is also known to have varying effects on cell proliferation, apoptosis and regulation (Tanaka et al. 1990), and may have a similar role in X. muta. The observation of potentially highly similar microbiome members between humans and sponges provides an interesting notion that there may be some prokaryotes present in the microbiome of “higher taxa” such as vertebrates, that have been closely associated with eukaryotic organisms since the early evolution of metazoan. Clearly further research is needed in this regard, particularly as anaerobic pathways can differ in prokaryote-invertebrate relationships (Wrede 2012), but as more sponges are subject to genome and transcriptome sequencing and as we are able to further annotate the eukaryotic and prokaryotic portions of these genomes and metagenomes, we will likely be able to better piece together the identity and function of key symbionts and their role in the evolution of these organisms.
When the genome of *Amphimedon queenslandica* was sequenced, one of the primary goals was to determine whether all six hallmarks of multicellularity exist in the sponge and characterize their origin (Srivastava et al. 2010). The six hallmarks are: regulated cell cycling and growth, programmed cell death, cell-cell and cell-matrix adhesion, developmental signaling and gene regulation, allore cognition and innate immunity, and specialization of cell types. The *A. queenslandica* genome did indeed contain genes that encode for proteins important in major pathways of all six hallmarks. Furthermore, sponges possess a relatively complex genome encoding key regulatory and sensory molecules to allow for development and regulation of a multicellular organism and the ability to sense and respond to environmental stimuli (Srivastava et al. 2010). Many of the identified major molecules and pathways in *A. queenslandica* were also observed in *X. muta*. These include enzymes in the p53 tumor suppressor pathway, as well as the insulin-signaling pathway, the PPAR pathway, Wnt, MAPK, TGF-beta signaling pathway among others. These results indicate that these are highly conserved pathways of cell development, differentiation, and death.

Sponges are able to sense and respond to environmental stimuli as well as to non-self cells, be they potential pathogens, food, or symbionts (Richards et al. 2008; Srivastava et al. 2010). Additionally, *X. muta* and other sponges have an immunological repertoire that allows for identifying and clearing unwanted cells and pathogens. The immunological and sensory capabilities of *X. muta* and other sponges probably have important roles in initiating and maintaining symbiosis with its prokaryotic symbionts, possibly through a winnowing process similar to the *Euprymna-Vibrio* symbiosis (Nyholm et al. 2004). Interestingly, within the human disease category, one of the largest
components was that of infectious disease, perhaps indicating that either there is infection (or attempted infection) of sponge cells by pathogenic bacteria, or that harmless interactions between the bacteria and sponge may employ some of the same proteins. Furthermore, the KEGG categories for pathways involved in *Escherichia coli*, *Helicobacter pylori*, and *Vibrio cholera* infections were observed in *X. muta*, suggesting that there are in fact some pathogenic interactions occurring. Reports of sponge diseases have increased substantially within recent years, although only one case has confirmed a bacterial culprit based on fulfilling Koch’s postulates (Webster 2007 and references therein). With an estimated $10^5$ cells ml$^{-1}$ in seawater (Giovannoni and Stingl 2005) and the sponge pumping $\sim 530$ l of seawater per hour per kilogram of sponge on average (Chapter 3), it is likely that the sponge regularly encounters potentially harmful bacteria and would need to be able to rid itself of such pathogens, in which case the transcripts in the category would represent specific immune responses to bacteria at the time of collection. Further investigation into these transcripts might yield more insight into pathobiology of potential sponge pathogens.

How the sponge identifies food or pathogens versus symbionts is still not fully known at this point. The study of prokaryotic transcripts from *C. concentrica* included an overrepresentation of ankrin repeats (AR) and tetratricopeptide repeats (TPR) in the sponge transcripts relative to the seawater comparison (Thomas et al. 2010). AR and TPR are known to mediate protein-protein interactions and are involved in a variety of regulator functions (Das et al. 1998; Hryniewicz-Jankowska et al. 2002) and have been observed in the genomes of obligate and facultative intracellular symbionts of eukaryotes (Wu et al. 2004). The authors proposed that these might be involved in identification of
symbionts versus food by the sponge (Thomas et al. 2010). Additionally, many of the identified AR and TPR proteins had signal peptides for extracellular secretion in Gram-negative bacteria, indicating that the proteins have a role outside of the bacterial cell (Thomas et al. 2010). AR and TPR proteins have not been recovered from the X. muta metatranscriptome analysis so far but further investigation may yield a positive find of these molecules, or there may be other mechanisms for interactions. It is clear from the sponge transcripts that there are several pathways available for the sponge to use to interact with prokaryotes including immunological defenses, sensory system pathways, and probably more mechanisms will be identified upon further investigation. Clearly, further research into these potential pathways for signaling between the sponge and symbionts is needed.

**Taxonomic profile of the X. muta prokaryotic community**

The taxonomic composition of the prokaryotic community based on putative mRNA contigs showed a less diverse community of prokaryotes compared to the metagenetic analysis of 16S rRNA genes using 454 pyrosequencing (Chapter 2), but most of the major groups from that analysis were represented. The Proteobacteria contained the largest number of contigs, which was expected given their dominance in the 16S rRNA analyses (Chapter 2) and that they are known to be an important taxonomic group recovered in other sponges (Kamke et al. 2010; Thomas et al. 2010; Radax et al. 2012; Fan et al. 2013).

The Proteobacteria are highly diverse bacteria capable of a variety of metabolic activities, but more specific information can be gleaned by examining the classes and
orders within the proteobacterial phylum. Four proteobacterial classes were represented
with the Gamma- and Alphaproteobacteria containing the most contigs, followed by beta-
and Deltaproteobacteria. Within the gammaproteobacterial class, interestingly, were
orders often associated with disease in corals and sponges, such as Enterobacteriales,
Legionellales, Pseudomonadales, and Vibrionales (Sutherland et al. 2004; Cervino et al.
2006; Negandhi et al. 2010). However, these groups are also known to contain non-
pathogenic symbionts, particularly Altermonadales, Enterobacteriales, Pseudomonadales,
Xanthomonadales, and Vibrionales (Cárdenas et al. 2012). It is difficult to say for sure
based on taxonomy alone which of these, if any, may be pathogenic or potentially
pathogenic, but it is likely that most represent commensal or mutualistic associations
based on previous studies of sponge-microbe interactions (Taylor et al. 2007). However,
it is interesting to note that pathways involved in pathogenic infections were active in the
metatranscriptome analyses (see below).

Within the Alphaproteobacteria were the orders Rhizobiales, Rhodobacteriales, and
Rhodospirillales. Rhizobiales is a group noted for symbiotic nitrogen fixing bacteria and
methanotrophic bacteria (Carvalho et al. 2010; Stein et al. 2010), and members of the
Rhizobiales have been recovered from corals and sponges (Sharp et al. 2007; Cárdenas et
al. 2012), including some nitrogen-fixing Rhizobiales (Lema et al. 2012).
Rhodobacteriales and Rhodospirillales are known marine alphaproteobacteria and have
been documented in sponge and other symbioses previously (Howard et al. 2011; Gruber-
Vodicka et al. 2011; Schmitt et al. 2011) including in *X. muta* (Chapter 2). Beta and
Deltaproteobacteria were also present, and most likely active, in *X. muta*. Specifically,
burkholderiales were the only group of betaproteobacteria and these were also recovered
from 16S rRNA gene analysis providing further support for their presence and activity. There was no further taxonomic information for the Deltaproteobacteria, a group containing aerobic and anaerobic bacteria and known sulfate reducers and oxidizers, all of which have been documented in sponges previously (Taylor et al. 2007).

Another major bacterial group present and potentially very active in *X. muta* is cyanobacteria. The contigs within cyanobacteria matched to the order Chroococcales and genus *Synechococcus*. *Synechococcus spongiarum* has been documented in a variety of sponges from around the world and has been documented in *X. muta* (Erwin and Thacker 2007; Taylor et al. 2007). *Synechococcus spongiarum* and other cyanobacteria are assumed to be photoautotrophic, however, the extent to which there is a mutualistic relationship between the cyanobacteria and the host sponge, including transfer of fixed carbon or other molecules, varies (Erwin and Thacker 2007; Freeman and Thacker 2011) and for *X. muta* it is not well characterized (Erwin and Thacker 2007; Chapters 2, 3).

The other bacterial groups included Firmicutes, Actinobacteria, Bacteroidetes, and Chloroflexi, all of which contain previously documented sponge symbionts (Taylor et al. 2007; Simister et al. 2012). Within the Firmicutes, surprisingly, were matches to both aerobic and anaerobic spore forming bacteria: Bacilli and Clostridia, providing some evidence for anaerobic microhabitats within *X. muta*. Chloroflexi were the most abundant group in the sponge *Geodia barretti* based on ribosomal contigs from RNA sequencing analysis (Radax et al. 2012), and were a significant portion of reads based on 16S rRNA gene analyses in *X. muta* (Montalvo and Hill 2011; Chapter 2).

Another important and abundant group recovered from *X. muta* and other sponges, based on 16S rRNA gene analysis, is the archaeal group Thaumarchaeota (Simister et al. 2012).
2011; Chapter 2). In the current study, a large proportion of assigned prokaryotic contigs was classified as Thaumarchaeota and further classified as matching to the known sponge symbiont *Cenarchaeum symbiosum* (Preston et al. 1996; Hallam et al. 2006), as well as marine ammonia oxidizer *Nitrosopumilus maritimus* (Walker et al. 2010). Archaeal genes and transcripts involved in ammonium oxidation have been recovered from diverse sponges including *X. muta* (Lopez-Legentil et al. 2010; Fan et al. 2013; Chapter 2), although there is some debate as to whether or not *C. symbiosium* can perform ammonia oxidation (Hatzenpichler et al. 2012). While further investigation is needed to determine the taxonomic identification of, and functional roles of Thaumarchaeota present in sponges such as *X. muta*, it is clear that *X. muta* hosts an abundant, diverse and active archaeal community.

A diverse group of active bacteria, generally representative of the 16S rRNA gene analyses was also observed in the high microbial abundance sponge *Ancorina alata* (Kamke et al. 2010). Kamke and colleagues (2010) compared DNA and RNA libraries for 16S rRNA in *A. alata* as well as a low microbial abundance sponge *Polymastia* sp. While clone libraries were more diverse for *A. alata*, there was substantial overlap between DNA and RNA libraries for both sponges, with sponge-specific lineages within Actinobacteria, Chloroflexi, and Gemmatimonadetes being the most common (Kamke et al. 2010). The results of the current study as well as Kamke et al. (2010) indicate that much of the prokaryotic diversity that has been observed in sponges also represents metabolically active prokaryotic symbionts.
Functional profile of the *X. muta* prokaryotic community

Many of the differences observed between the SEED classification of sponge and metatranscriptome contigs were expected, including the addition of categories such as virulence, photosynthesis, sulfur metabolism, motility and chemotaxis, and cell wall and capsule synthesis in the metatranscriptome. Some other categories such as DNA metabolism, cell division and cell cycle, amino acids and derivatives, and fatty acids, lipids and isoprenoids were surprisingly not recovered from the sponge. This may be a result of higher activity of these pathways in the metatranscriptome portion or that there was low similarity or poor matching ability with the SEED groupings for the sponge contigs in these categories. The KEGG categories were roughly proportional between the sponge and metatranscriptome contigs, indicating that there are many common processes to both the sponge and prokaryotic communities, and/or that the major differences between the activity of the two communities lies in more specific categories within the broad groupings listed for KEGG.

We would expect there to be differences in metabolism in particular, between the sponge and the prokaryotic community. Even though there are likely heterotrophic bacteria present, prokaryotes are capable of diverse mechanisms for energy capture and use. A large portion of the contigs involved in energy metabolism in the metatranscriptome represent pathways in oxidative phosphorylation, which is to be expected as we know that the sponge as well as many prokaryotes would be utilizing this respiratory pathway. Also present in the metatranscriptome are enzymes involved in pathways for photoautotrophic organisms, including photosynthetic antenna proteins such as phycoerythrin and phycocyanin in addition to enzymes involved in photosynthesis. We
know that cyanobacteria containing chlorophyll a and phycoerythrin are present in *X. muta* (Chapter 2), and these results support the photoautotrophic nature of these symbionts.

Other autotrophic pathways are represented as well including the reductive carboxylate cycle for carbon fixation, and enzymes involved in sulfur, methane, and nitrogen metabolism for energy conservation. Sulfur metabolism included three contigs that matched to a sulfite reductase, a 3’-phosphoadenosine 5’-phosphosulfate synthase, both of which are involved in sulfate/sulfite reduction to hydrogen sulfide, and a homoserine O-acetyltransferase. The latter is involved in cysteine and methionine metabolism as well as sulfur metabolism. Bacteria involved in sulfur cycling have been recovered from sponges previously (Imhoff and Trüper 1976; Manz et al. 2000; Webster et al. 2001; Hoffmann et al. 2005 b) and in one case the authors were able to detect anaerobic microhabitats within the sponge, which allowed for both aerobic and anaerobic sulfur transformations to occur in the sponge (Hoffmann et al. 2005). Taxonomic matches to bacteria potentially involved in sulfur cycling include the Rhodospirillales and Chloroflexi (Imhoff et al. 1976; Webster et al. 2001), though further taxonomic identification is needed to determine if, and which, sulfur cycling bacteria are present and active in *X. muta*. These results do provide further support for anaerobic microhabitats to exist in *X. muta*, and is the first documentation of active genes involved in sulfur cycling in this sponge.

Transcripts classified in methane metabolism included enzymes involved in carbon fixation as well as methane oxidation, and as Rhizobiales were one of the active groups recovered in this sponge, it is possible that methanotrophs may be present and active in *X. muta*. 
muta. Further characterization is required, however, to determine which prokaryotes may be performing methane oxidation and their abundance in the symbiotic community. A unique symbiosis has been documented between cladorhizid sponges, which prey on zooplankton, and methanotrophic bacteria (Vacelet et al. 1996). The sponge also gains nutrition by feeding on the methanotrophic bacteria, which oxidize methane derived from a deep-sea mud volcano (Vacelet et al. 1995). In X. muta, methane for methane oxidation may be present in the water column or could be derived from methanogenic archaea, but these have so far not been documented in this sponge, and would be unlikely to be present. Further investigation into the metabolism of the methanotrophs and possible transfer of fixed carbon or “farming” of symbionts needs to be performed to determine the relationship of these bacteria with X. muta.

Transcripts classified in nitrogen metabolism included two matches to nitrate reductase enzymes, involved in denitrification. Interestingly, further investigation into these matches revealed that one nitrate reductase is likely of bacterial origin while the other is likely of archaeal origin. Furthermore, in the Blast2GO analyses, these enzymes also matched to two enzymes involved in anaerobic ammonium oxidation (anammox). Further work is needed to better characterize these transcripts and determine which pathways they are involved in. These results however, do provide additional support of anaerobic nitrogen transformations in X. muta, which were hypothesized to occur based on nutrient analyses of this sponge (Chapter 3). Enzymes involved in other respiration processes using nitrogen such as the two-step process of nitrification, were not recovered in this analysis, but genes and transcripts for ammonia oxidation by archaea have been recovered previously in this sponge (Lopez-Legentil et al. 2010; Chapter 3). Additionally,
genes involved in nitrogen fixation have been recovered from *X. muta* (Chapter 3), which could provide an additional source of nitrogen for the sponge (Chapter 3). Nitrogen uptake and assimilation are tightly regulated by the global regulator PII, and many other nitrogen regulatory proteins, the activity of which is controlled by the ratio of \( \alpha \)-ketoglutarate to glutamate (Merrick 2004; also see Chapter 1). Many of the other contigs involved in nitrogen cycling, not surprisingly are involved in the deamination of glutamate to \( \alpha \)-ketoglutarate, providing feedback on the nitrogen status of the cell. It should be noted here, however, that nitrogen regulation via glutamate and \( \alpha \)-ketoglutarate occurs in prokaryotes and eukaryotes and the contigs involved in this pathway may be from the sponge or the prokaryotes, or mostly likely, both. Based on the results presented here as well as previous work on nitrogen cycling in *X. muta* (Chapter 3), it is likely that complex nitrogen cycling exists within this sponge and further investigation is needed to determine the better characterize which nitrogen transformations occur in the sponge and how these influence the metabolism of the host sponge.

The metabolic activities of symbiotic prokaryotes are an important aspect of their relationship with the sponge, and while there are examples of transfer of fixed carbon or nitrogen from symbionts to the sponge host (Wilkinson and Fay 1979; Arillo et al. 1993; Freeman and Thacker 2011), it is generally assumed that for most sponges, the prokaryotes help prevent build up of toxic wastes such as ammonia or sulfides (Taylor et al. 2007; Hoffmann et al. 2009). Another commonly held view is that many symbiotic prokaryotes produce metabolites that act as deterrents to potential predators of sponges (Schmidt et al. 2000; Hentschel et al. 2001). Furthermore, multiple lines of evidence
indicate that secondary metabolites, in addition to deterring predators, also provide antifouling, inhibition of overgrowth, and UV protection to sponges (Thompson 1985; Waddell and Pawlik 2000; Kelly et al. 2003; Jones et al. 2005). Xestospongia species possess a suite of secondary metabolites including terpenoids, alkaloids, pentacyclic polyketides (notably xestosaprols), ene-yne tetrahydrofurans, and acetylenic acids (Quinn and Tucker 1991; Richelle-Maurer et al. 2003; Dai et al. 2010; Longeon et al. 2010). In other sponges some bioactive compounds are documented to be produced by the sponge itself, such as the alkyl-lipid derivatives of Suberites domuncula, which have antimicrobial properties (Müller et al. 2004 a). However, it appears that many compounds are produced by bacteria within the sponge (Bewley and Faulkner 1998; Müller et al. 2004 b).

Several more matches to secondary metabolite biosynthesis were recovered in the metatranscriptome assembly than the sponge assembly, including compounds that are almost certainly from the prokaryotes such as terpenoids, flavone, betalain, and streptomycin. It is difficult to say at this time whether the other pathways are active in the sponge or in the prokaryotes. Most of the compounds with active pathways in the sponge and metatranscriptome are considered biologically active compounds with antimicrobial or antitumor activity (Amagata et al. 2003; Vinothkumar and Parameswaran 2013). It is likely that at least some of these have a role, whether or not that is its original purpose, in deterring predators and/or fouling of the sponge. This hypothesis is based on the fact that previous studies have documented varying levels of deterrence of feeding from X. muta extracts (Jones et al. 2005). Furthermore, one study observed bioactive compounds
concentrated in the ectosome of *X. muta* (Richelle-Maurer et al. 2003), which would suggest antifouling, antimicrobial, and/or antipredatory roles.

Secondary metabolites may also provide a mechanism for interactions between their prokaryotic symbionts and the sponge host. Siegl et al. (2010) documented genes encoding for secondary metabolite production, polyketide synthases (PKSs), in the poribacterial genome that may be involved in host-microbe interactions. PKSs are responsible for producing a variety of complex polyketides (Schirmer et al. 2005), and a recent study of sponge metagenomes yielded the existence of a ubiquitous and highly sponge-specific group of PKSs (Fieseler et al. 2007). Homologs of PKS genes have been shown to be major symbiosis and virulence factors in other systems (Minnikin et al. 2002; Parada et al. 2006), and in Poribacteria, it has been suggested that products of PKS genes may be involved in host-symbiont interactions (Siegl et al. 2010).

Other symbiosis factors documented in recent sponge-microbe studies include proteins involved in host immunity. For example, the sponge *Suberites domuncula* was shown to have different immune responses to Gram-negative and Gram-positive bacteria (Böhm et al. 2001; Müller et al. 2004; Wiens et al. 2005). When the *S. domuncula* was exposed to the lipopolysaccharide (LPS), a bacterial endotoxin (representing Gram-negative bacteria), the sponge increased synthesis of compounds with antibacterial activity, and further investigation revealed an LPS receptor on the surface of sponge cells (Müller et al. 2004 a, Wiens et al. 2005). When the sponge was exposed to peptidoglycan, representing Gram-positive bacteria, endocytosis pathways are activated with the release of lysozyme (Thakur et al. 2004). The recovery of AR and TPR domains in sponge bacteria provide another potential mechanism for interaction with their eukaryotic host.
(Thomas et al. 2010; Siegl et al. 2010; Fan et al. 2012). Siegl and colleagues (2010) also observed cell surface proteins with Ig-like domains and laminin G domains, which the authors considered to be putative host-interaction factors. It is known that the sponge can select for non-sponge bacteria and these studies provide some mechanisms for how this recognition occurs, though there is still more to investigate in this regard. Additionally, it is now widely accepted that many resident bacteria evade the sponge immune system by use of capsules or other extracellular masking mechanisms (Wilkinson et al. 1981; Wilkinson 1984).

In *X. muta*, the SEED subsystem classification of transcripts involved in cell wall and capsule synthesis were recovered in the metatranscriptome assembly but not in the sponge assembly, providing some support for the hypothesis that some symbiotic bacteria use capsules to evade the sponge immune system. However, it should be noted that capsules are known to be important in multiple functions (Roberts 1996). Further investigation into the contigs classified into these subsystems might yield more insight into which bacteria this activity corresponds to and what molecules and pathways they are utilizing. Extracellular proteins such as integrin subunits as well as a focal adhesion kinase 2 (FAK2), which is involved in focal adhesion between cells and the ECM, and can be activated by integrin activity (Zachary and Rozengurt 1992), were recovered from the metatranscriptome. Also present in the metatranscriptome, are transcripts orthologous to genes encoding for proteins involved in T-cell and B-cell activation, and antigen presentation pathways, as well as to proteins involved in bacterial infections in humans. These active proteins may be an indication of sponge cells interacting with resident prokaryotic cells, and potentially with pathogenic bacteria.
Conclusion

The results presented here indicate that *Xestospongia muta*, like other sponges that have been subject to genome or transcriptome sequencing so far, possesses highly regulated pathways and mechanisms for interacting with the environment and with prokaryotic organisms that may be harmless or pathogenic. The prokaryotes, including both Bacteria and Archaea, are active in the sponge and carry out diverse metabolic processes, including anaerobic nitrogen and sulfur transformations. The extent to which the sponge exploits these processes, whether for waste removal, defenses, or nutrition remains poorly characterized and needs to be further explored. The work performed here provides information for generating further hypotheses in regards to sponge-microbe interactions as well as data for investigating many of these hypotheses.
Table 4.1. The starting number of reads, the number of reads removed at each quality filtering step, and the total reads left after filtering for each sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Starting no. reads</th>
<th>Q &lt; 30</th>
<th>min length 20</th>
<th>adapter only</th>
<th>unpaired</th>
<th>Total reads after filtering</th>
</tr>
</thead>
<tbody>
<tr>
<td>XmFL</td>
<td>37,432,452</td>
<td>0</td>
<td>2,110,863</td>
<td>47,511</td>
<td>2,054,040</td>
<td>33,220,038</td>
</tr>
<tr>
<td>XmLC</td>
<td>46,295,392</td>
<td>0</td>
<td>2,267,799</td>
<td>52,984</td>
<td>2,200,591</td>
<td>41,774,018</td>
</tr>
<tr>
<td>XmLSI</td>
<td>58,246,826</td>
<td>0</td>
<td>3,292,295</td>
<td>73,225</td>
<td>3,184,624</td>
<td>51,696,682</td>
</tr>
</tbody>
</table>
Table 4.2. Number of contiguous sequences (contigs) classified under SEED and KEGG categories for the metatranscriptome assembly (MV) and the sponge assembly (Sponge).

<table>
<thead>
<tr>
<th>SEED Category</th>
<th>no. contigs - MV</th>
<th>no. contigs - Sponge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein Metabolism</td>
<td>64</td>
<td>3</td>
</tr>
<tr>
<td>Virulence</td>
<td>30</td>
<td>NA</td>
</tr>
<tr>
<td>Stress Response</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>RNA Metabolism</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>DNA Metabolism</td>
<td>8</td>
<td>NA</td>
</tr>
<tr>
<td>Respiration</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Photosynthesis</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>Cell Division and Cell Cycle</td>
<td>2</td>
<td>NA</td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sulfur Metabolism</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Motility and Chemotaxis</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Amino Acids and Derivatives</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Clustering-based subsystems</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Fatty Acids, Lipids, and Isoprenoids</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Cell Wall and Capsule</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Not assigned</td>
<td>8011</td>
<td>713</td>
</tr>
<tr>
<td>KEGG Metabolism</td>
<td>554</td>
<td>112</td>
</tr>
<tr>
<td>Cellular Processes</td>
<td>515</td>
<td>111</td>
</tr>
<tr>
<td>Organismal Systems</td>
<td>401</td>
<td>93</td>
</tr>
<tr>
<td>Human Diseases</td>
<td>393</td>
<td>98</td>
</tr>
<tr>
<td>Genetic Information Processing</td>
<td>392</td>
<td>87</td>
</tr>
<tr>
<td>Environmental Information Processing</td>
<td>364</td>
<td>76</td>
</tr>
<tr>
<td>Not assigned</td>
<td>6450</td>
<td>371</td>
</tr>
</tbody>
</table>
Table 4.3. Number of contiguous sequences (contigs) assigned to the metabolite category in KEGG for the metatranscriptome assembly (MV) and the sponge assembly (Sponge).

<table>
<thead>
<tr>
<th>Metabolite category</th>
<th>no. contigs - MV</th>
<th>no. contigs - Sponge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biosynthesis of ansamycins</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Limonene and pinene degradation</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Phenylpropanoid biosynthesis</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Isoquinoline alkaloid biosynthesis</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Tropane, piperidine and pyridine alkaloid biosynthesis</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Caffeine metabolism</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Novobiocin biosynthesis</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Terpenoid backbone biosynthesis</td>
<td>7</td>
<td>NA</td>
</tr>
<tr>
<td>Flavone and flavonol biosynthesis</td>
<td>3</td>
<td>NA</td>
</tr>
<tr>
<td>Betalain biosynthesis</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Streptomycin biosynthesis</td>
<td>5</td>
<td>NA</td>
</tr>
</tbody>
</table>
Figure 4.1. Assignment of contiguous sequences from the sponge assembly using the program MEGAN (A), where the size of the circles are based on the number of contigs in that category (n = 142 Low complexity; n = 445 A. queenslandica; n = 7 Euteleostomi), and the program Blast2GO (B).
Figure 4.2. Assignment of putative mRNA contiguous sequences (contigs) using the SEED subsystem classification in MEGAN. Size of the circles are based on the number of contigs in that category. The number of contigs not assigned is 713.
Figure 4.3. Figure 4.2. Assignment of putative mRNA contiguous sequences (contigs) using the KEGG classification in MEGAN. Size of the circles are based on the number of contigs in that category. The number of contigs not assigned is 371.
Figure 4.4. Pie chart showing proportion of contiguous sequences (contigs) from the sponge assembly classified as involved in energy metabolism using KEGG databases in MEGAN. The number of total contigs in the piechart = 35.
Figure 4.5 Pie charts showing proportion of contiguous sequences (contigs) from the sponge assembly classified as involved in amino acid metabolism (A and B; n=41 and 19 contigs respectively), vitamin metabolism (n=11 contigs) (C), and fatty acid metabolism (n=10 contigs) (D) using KEGG databases in MEGAN.
Figure 4.6. Pie charts showing proportion of contiguous sequences (contigs) from the sponge assembly classified as involved in environmental processing (n=76 contigs) (A) signal transduction specifically (n=67 contigs) (B) using KEGG databases in MEGAN.
Figure 4.7. Pie charts showing proportion of contiguous sequences (contigs) from the sponge assembly classified as involved in cellular processing (n=111 contigs) (A), and the subsections of A are shown in (n=48 contigs) (B) cell communication (n=29 contigs), (C) cell growth and death (n=40 contigs), and (D) transport and catabolism using KEGG databases in MEGAN.
Figure 4.8 Pie charts showing proportion of contiguous sequences (contigs) from the sponge assembly classified as involved in organismal systems (n=93 contigs) (A), and the subsections of A are shown in (B) immunity (n=50 contigs), (C) endocrine (n=8 contigs), (D) excretory (n=8 contigs), and (E) nervous system (n=30 contigs) using KEGG databases in MEGAN.
Figure 4.9. Pie chart showing proportion of contiguous sequences from the sponge assembly that were assigned to the KEGG category: secondary metabolites and xenobiotics degradation (n=16 contigs), in MEGAN.
Figure 4.10. Pie charts showing proportion of contiguous sequences from the sponge assembly that were assigned to the KEGG category in MEGAN: human disease (n=98 contigs) (A), and the subsection infectious disease is further examined (n=40 contigs) (B).
Figure 4.11. Taxonomic assignment of contiguous sequences (contigs) from the metatranscriptome assembly, showing breakdown of prokaryotic taxa in the MEGAN tree (A) and a pie chart showing relative abundance of just the prokaryotic contigs (B). The number of contigs assigned to Bacteria was 685, and the number of contigs assigned to Archaea was 129. The number of contigs assigned to major bacterial phyla are: cyanobacteria, 56; actinobacteria, 12; bacteroidetes, 12; chloroflexi, 11; firmicutes, 39; proteobacteria 279; unclassified, 38.
Figure 4.12. Breakdown of prokaryotic taxa into lower taxonomic groupings (A) and breakdown of major eukaryotic groups of assigned contiguous sequences (contigs) (B). The number of contigs assigned to major bacterial phyla are: cyanobacteria, 56; actinobacteria, 12; bacteroidetes, 12; chloroflexi, 11; firmicutes, 39; proteobacteria 279; unclassified, 38. The number of contigs assigned to major eukaryotic groups are: Eukaryota, 7512; Metazoa, 6913; *Amphimedon*, 5155; low complexity, 4053; not assigned, 34.
Figure 4.13. Assignment of contiguous sequences (contigs) from the metatranscriptome assembly to the SEED subsystem categories in MEGAN. The number of contigs not assigned was 8011.
Figure 4.14. Assignment of contiguous sequences (contigs) to the SEED subsystem virulence, the largest assigned category (n=30 contigs).
Figure 4.15. Assignment of contiguous sequences (contigs) from the metatranscriptome assembly to the KEGG categories in MEGAN. The number of contigs not assigned was 6450.
Figure 4.16. Pie chart showing the assignment of contiguous sequences (contigs) to the KEGG category of metabolism in MEGAN (n=554 contigs).
Figure 4.17. Pie charts showing the breakdown of KEGG categories for energy metabolism (n=141 contigs) (A), vitamin metabolism (n= 40 contigs) (B), and xenobiotic degradation and metabolism (n=82 contigs) (C).
Figure 4.18. Pie charts showing the breakdown of KEGG categories for cellular processes (n=515 contigs) (A), and organismal systems (n= 401 contigs) (B).
Figure 4.19. Pie charts showing the breakdown of KEGG categories for human disease (n=393 contigs) (A), immune system (n=196 contigs) (B), and infectious disease (n=119 contigs) (C).
SYNTHESIS

Symbiosis has existed between eukaryotic organisms and prokaryotes likely since the emergence of the first eukaryotic cells over 1 billion years ago (McFall-Ngai et al. 2013 and references therein), and almost certainly since the emergence of the first multicellular animals over 600 million years ago (Li et al. 1998). In some cases, plants and animals, as well as the prokaryotic symbionts, have evolved mutualistic associations, whereby one cannot survive without the other (e.g., lichens, hydrothermal vent worms). The extent to which prokaryotic symbionts have influenced the development and functioning of complex animals including invertebrates and vertebrates was reviewed recently by McFall et al. (2013) and highlights the pivotal roles that bacteria have played in animal evolution.

In the marine environment, as with certain terrestrial environments (e.g., leached soils), close associations between animals and prokaryotes have often revolved around increasing access to nutrient supplies for one or both partners, or to some extent, provision of defense mechanisms, usually for the animal host. As reviewed in Chapter 1, there are many examples of symbioses where there is exchange of nitrogen, a limiting nutrient in much of the marine environment, between symbiont and host. Symbioses based on the exchange of carbon are also common in the marine environment, and the most conspicuous example is that of corals and their dinoflagellate symbionts, which provide photoautotrophically fixed carbon to the coral host.
Coral reefs in particular, are ecosystems with symbiotic associations at the core of its structure, function, and resilience. Sponges are important members of coral reef ecosystems, providing vital ecosystem services such as consolidation of reefs, erosion prevention, filtering large quantities of seawater and they provide habitat and food for many invertebrates and fishes (Diaz and Rützler 2001; Ribiero et al. 2003), and coupling water column processes to benthic habitats (Lesser 2006; Southwell 2008b). That many sponges harbor a high density of diverse prokaryotic symbionts is likely a major factor behind the fact that sponges have survived for so long relatively unchanged (Wood 1995). Furthermore, these prokaryotic symbionts are also the main reason that sponges are able to effectively influence the biogeochemical cycling within their habitat.

Recent research into the complex prokaryotic community of sponges has led to major advances in our understanding of the distribution and stability of prokaryotic sponge symbionts. In this study I have utilized the Caribbean giant barrel sponge, Xestospongia muta, to address key questions on the specificity, composition, and function of the sponge prokaryotic symbiont community. X. muta is abundant on Caribbean coral reefs, and on one coral reef it is known to be increasing in abundance (McMurray et al. 2008). Xestospongia muta is also known to harbor a dense community of prokaryotic symbionts, making it a good model for further investigating these key questions.

My first goal was to characterize the composition and relative abundance of the X. muta prokaryotic community. Furthermore, I compared the prokaryotic community composition between three distinct locations within the Caribbean basin, to examine the specificity of the community exposed to different environmental factors. I observed a highly diverse prokaryotic community including representatives from domains of
Bacteria and Archaea. There was also high diversity within many of the bacterial and archaeal phyla that were recovered. While the taxonomic composition of the prokaryotic community was largely similar between locations, some distinct differences could be observed based on phylogenetic and ecological multivariate statistics. The main underlying factor of the significance of these tests is likely differences in abundance of key taxa between locations. Differences observed here might have important implications for understanding the acquisition and maintenance involved in the sponge-prokaryote symbiosis. Another important finding was that major functional groups of bacteria and archaea that are known to be involved in carbon, nitrogen, and sulfur cycling were recovered from *X. muta*, although, their corresponding functional role in the sponge could not be evaluated based on ribosomal genes alone.

As nitrogen is an important limiting nutrient on coral reefs, and as sponges are known for serving as a major source of dissolved inorganic nitrogen (DIN) to coral reefs, I wanted to further investigate the influence of *X. muta* on coral reef nitrogen biogeochemistry. *Xestospongia muta* populations were observed to generate high amounts of DIN to the water column, often several fold higher than that generated by nitrification in coral reef sediments or microbial mats (Capone et al. 1992; Bonin and Michotey 2006) and similar to a previous study of *X. muta* form Conch Reef, Florida Keys (Southwell et al. 2008 a, b). What was unique from previous studies, however, was that some *X. muta* individuals were found to be a net sink for DIN, indicating that anaerobic nitrogen transformations were also likely occurring. Because *X. muta* is an abundant and prominent member on coral reefs and is capable of filtering hundreds of
liters of seawater per hour (see Chapter 3), the release or utilization of DIN could have a significant impact on the availability of DIN on these coral reefs.

There was one significant difference in the flux of DIN between locations: the volume-normalized flux of $\text{NO}_x^-$, which was higher for LSI sponges than for FL sponges. Additionally, the fluxes of DIN were often observed to contain a significant contribution of $\text{NH}_4^+$ as opposed to mostly $\text{NO}_3^-$, which had not been observed in this sponge before. Both of these findings may have important implications for understanding nutrient cycling on different coral reefs and predicting how future environmental changes may differentially impact these reefs. While, further research into the underlying mechanisms of the variability in fluxes of DIN observed in this study are needed, we do know there are prokaryotes present that could perform anaerobic and aerobic nitrogen transformations, and that the prokaryotic community is capable of taking up inorganic nitrogen compounds such as $\text{NH}_4$ and $\text{NO}_3^-$, which could then be used in assimilatory and/or dissimilatory processes.

To obtain a more detailed view of the function and activity of the prokaryotic community associated with $X. \text{muta}$; the last major component to this project was to analyze the metatranscriptome of the sponge. Based on the metatranscriptomic analyses, a diverse active community of prokaryotes was recovered, indicating that at least most major groups of prokaryotes present in the sponge are also active. Genes that were recovered in DNA sequencing such as $\text{nifH}$ (involved in nitrogen fixation) and archaeal $\text{amoA}$ (ammonia oxidation), were not recovered from the metatranscriptome. This may be a result of the genes not being expressed at the time of sampling (mid-morning, when oxygen levels are likely to be high from photosynthesis), or that they were not recovered
in the assembly. However, transcripts involved in denitrification, anaerobic ammonium oxidation and sulfate reduction pathways were recovered, providing further support for complex nitrogen cycling, including anaerobic transformations of nitrogen and other compounds, in *X. muta*.

Also recovered from the metatranscriptome was a suite of transcripts involved in a variety of processes that may represent pathways for host-microbe interactions. Extracellular proteins, capsule synthesis, and secondary metabolites all may provide mechanisms for communication between prokaryotes and the sponge to determine such information as: whether the prokaryote is pathogenic or not, whether useful nutrients may be present (for either member), or to “negotiate” for specific nutrients or other useful compounds such as bioactive metabolites. Further research into the metatranscriptomic data will likely yield more insight into host-prokaryote interactions, potential transfer of various compounds between symbiont and host, and into nutrient cycling within the prokaryotic community and the sponge.

Results from this project support the notion that sponges and their prokaryotic symbionts have maintained a close and stable relationship over geographic distances, and as other studies have shown there is support for a core sponge-specific community of prokaryotes (Schmitt et al. 2011). There is also evidence for region specific sponge clades of bacteria as described by Schmitt et al. (2011), and as shown in the current study, there may also be subtle but distinct differences in prokaryotic community composition on smaller geographic scales such as within the Caribbean.

This work has also provided insight into the microbial ecology of the abundant and ecologically important sponge *Xestospongia muta*. Results presented here indicate that *X.*
*muta* hosts a diverse and active prokaryotic community that is relatively stable over geographic distances, although environmental factors may influence the relative abundances of certain symbiotic taxa. The differences observed in the symbiotic prokaryotic community provide an avenue for further investigation of the importance of vertical and horizontal transmission in the acquisition and maintenance of the symbiotic community. The stability of the prokaryotic community of *X. muta* suggests a close relationship where the prokaryotic community has coevolved with the sponge. Coevolution of sponge prokaryotic communities has been investigated in other studies specifically (Erpenbeck et al. 2002; Thacker and Starnes 2003; Fan et al. 2012), and in regards to *X. muta*, would suggest that the diverse prokaryotes present in the sponge perform a variety of functions that are important to the health and functioning of the sponge and the coral reef.

*Xestospongia muta* has been shown to be the primary contributor to DIN on Caribbean coral reefs (Southwell et al. 2008 b), largely due to its symbiotic prokaryotic community. Fluxes of DIN from *X. muta* in the current study were more complex than previously shown and indicated that both aerobic and anaerobic nitrogen transformations by the prokaryotic community were occurring. Results from the current study also indicate that complex nitrogen and sulfur cycling occur in the sponge, and it is likely that these processes not only influence each other (e.g., Purubsky et al. 2009), but other nutrient related processes in the sponge. These results have important implications as the contribution of DIN and other important nutrients to the coral reef environment from *X. muta* may be even more extensive than previously shown, and may be different for different locations. The differences in fluxes observed here could differentially influence
nutrient cycling on the different reefs, and could influence how future environmental changes impact these reefs.

A recent hypothesis in regards to sponge mediated nutrient cycling on coral reefs and specifically including *X. muta* due to its prevalence, is that the excess nutrients from sponges (particularly if they increase in abundance (e.g., McMurray et al. 2008)) will help fuel a phase shift on many tropical coral reefs to sponge and algal only reefs (Bell et al. 2013). While this is one possible scenario given the expected significant environmental changes resulting from global climate change, it should be noted that factors underlying such a phase shift are complex and can vary by location (Hughes et al. 2010). Complex interactions, particularly involving predation, herbivory, and competition, make reefs subject to multiple stable states, and in many cases are the primary drivers for shifting from a coral dominated reef to one dominated by macroalgae (Sutherland 1974, Knowlton 1992). Chronic stressors such as pollution and overfishing also underly such phase shifts, but are typically considered slow drivers of change and are more likely to be one factor contributing to changes in coral reef cover as opposed to the main driver (Hughes et al. 2010). The fact that global changes overlay many of these complex interactions and chronic factors further complicates the process (Hughes et al. 2010), making it difficult to ascribe one factor as a cause of a coral reef phase shift. Furthermore, in regards to DIN from sponges, the effects of these fluxes on coral reef ecosystems are poorly constrained. For example, we do not know the fate of DIN after it leaves *X. muta*: does the composition and amount of DIN have local influences on the planktonic community? How much of the DIN is utilized by the planktonic community and what portion is recycled back into the benthic habitat (e.g., corals and sponges)
versus cycled within the plankton or exported out of the ecosystem? In future work it will be important to further investigate the factors influencing prokaryotic mediated nutrient cycling in *X. muta* and the extent to which this impacts the structure and function of the surrounding coral reef.

It is also now widely accepted that prokaryotes, whether they are mutualistic, commensal, or pathogenic to the sponge, influence their host physiology and ecology, and often influence the biogeochemistry and ecology of their surrounding habitat. Because of this close and often necessary relationship, it will also be important to investigate how these relationships may be altered in the face of major environmental changes such as pollution, warmer waters and ocean acidification.

Future goals that I think are important to address in regards to the response of the holobiont (sponge plus prokaryotes) to environmental change include determining the response of the prokaryotic community to environmental stressors, and more specifically to multiple stressors, and determine the effect that this may have on the host sponge. One recent study observed a shift of the composition of prokaryotic symbionts of the sponge *Rhopaloeides odorabile*, from diverse sponge-specific clades to general seawater-like clades and lineages known to be present in coral and sponge disease studies following exposure to elevated seawater temperature (Webster et al. 2008). Another recent study examined nutrient and temperatures stresses on *R. odorabile*, however, and found little change in the prokaryotic community in response to both stressors (Simister et al. 2012). There is still more work to be done to better understand how the holobiont will respond to not only increasing temperatures, but also to concomitant increase in CO₂ concentrations and increases or sustained exposure to pollution. As we uncover the physiological
responses to these environmental stressors, and if there are in fact detrimental effects from such stressors, we may be able to provide more evidence for enacting policy changes, or enable changes in the behavior of nearby human populations to prevent further degradation of these unique ecosystems that are coral reefs.

Additionally, there are several areas for further investigation based on work presented here that would shed light on the distribution, function, and resilience to environmental change, of sponge prokaryotic communities. First is to expand the geographic comparisons of the prokaryotic community composition (based on 16S rRNA genes) to include a depth range of ~10 – 100 m at multiple locations. There is a growing realization of the importance of mesophotic reefs as a refuge and population source for their shallow water counterparts, thereby increasing resilience of many coral reef species (Lesser et al. 2009). One major goal would be to determine if the prokaryotic community of \textit{X. muta} changes over such a large depth gradient and if this any changes are observed in more than one location. One recent study (Olson and Gao 2013) has examined a similar depth gradient for three sponge species from reefs near LC and it will be interesting to see if similar results, of a core prokaryotic community over depth, will hold for multiple locations. Second is to examine potential shifts in the prokaryotic community of \textit{X. muta} in response to elevated temperature and/or elevated CO$_2$ concentrations. Samples from such a factorial experiment (n=4 treatments: control, high temperature, high CO$_2$, and high temperature + high CO$_2$) are in hand and are currently being prepped for analysis.

Third is to expand the gene specific analysis (\textit{nifH} and \textit{amoA} presented here) to include genes involved in denitrification (i.e., \textit{nir, nor, nos} genes), and to include RT-
PCR analysis of all isolated genes and qPCR of certain genes. The qPCR experiment that I would like to do is to examine the expression of nifH, and potentially other isolated genes, in sponge samples that were collected in the morning and samples collected in the evening. There is evidence for diel variation in expression of nifH and other nitrogen specific genes (see Chapter 1), and this may provide some insight into the observed variability of DIN fluxes from X. muta, and into the physiological state of specific symbionts on a temporal scale.

The fourth goal is to further mine the metatranscriptomic data for insight into important metabolic processes and interactions between the sponge and its symbiotic prokaryotic community. Ideally, I would also sequence additional replicate sponges from each location to allow for statistical comparisons in gene expression between locations. However, even with the current dataset many hypotheses can be generated and examined. One example is to hypothesize that the active portion of the prokaryotic community will be the same between different sponges. To examine this I will be able to use the ribosomal portion of the reads (data not shown) to compare the taxonomic composition of prokaryotes between all three sponges. I will also be able to use the taxonomic assignments of the putative mRNA reads, however, these first need to be separated out by sample. Another example is to hypothesize that there are multiple pathways between prokaryotic symbionts and the sponge that are homologous to those found in the human microbiome. One experiment would be to compare metabolic and secondary metabolite pathways present in these two datasets.

As our knowledge has advanced in regards to the sponge holobiont, we have found a more complex host than previously though, and a highly specific and complex
prokaryotic community. The work presented here provides additional insight into the relationship between sponges and prokaryotes, and implications for how this relationship influences the surrounding environment. As sequencing technology advances, we will likely be able to sequence – and annotate – more of the host and the prokaryotic community; however, hypothesis driven experimental manipulations will also be necessary to complement sequencing methods that will allow us to better understand the evolutionary development and complex interactions between sponges and their associated prokaryotic community.
LIST OF REFERENCES


Bergquist, P. R. 1978. Sponges, University of California Press, Berkeley and Los Angeles, California, USA.


in symbiotic associations of heterotrophic dinoflagellates with cyanobacteria in the western Bay of Bengal. Symbiosis 42: 51-58.


Smithsonian Institution Press, Washington D.C.


candidate phylum symbiotically associated with marine sponges. ISME J. 1–10.


Tanaka, Y., T. Miyajima, I. Koike, Y. Hayashibara, and H. Ogawa. 2006. Translocation
and conservation of organic nitrogen within coral-zooxanthella symbiotic system of
_Acropora pulchra_, as demonstrated by dual isotope-labeling techniques. J. Exp. Mar.

Tao, F., dell, L. Daoji, Y. Lihua, G. Lei, and Z. Lihua. 2006. Effects of irradiance and
phosphate on growth of nanophytoplankton and picophytoplankton. Acta Ecologica
Sinica _26_: 2783–2790.

Biogeography of bacteria associated with the marine sponge _Cymbastela concentrica_.

Biol. R. _71_: 1–53.

Thacker, R. W. 2005. Impacts of Shading on Sponge-Cyanobacteria Symbioses: A
Comparison between Host-Specific and Generalist Associations. Integr. Comp. Biol.
_45_: 369–376.


Thakur, N. L., S. Perović-Ottstadt, R. Batel, M. Korzhev, B. Diehl-Seifert, I. M. Muller,

Thiel, T., J. C. Meeks, J. Elhai, M. Potts, F. Larimer, J. Lamerdin, P. Predki, and R.
B. Layzell, J. K. Bessey, and W. Newton [Eds.]), pp 88-94. CABI publishing, New
York, NY, USA.

Mid-chain branched alkanoic acids from “living fossil” demosponges: a link to

Heidelberg, S. Egan, P. D. Steinberg, and S. Kjelleberg. 2010. Functional genomic
signatures of sponge bacteria reveal unique and shared features of symbiosis. ISME J.
_3_: 74.

Thompson, J. E. 1985. Exudation of biologically-active metabolites in the sponge


