8-1-2014

Microbial Community Assembly found with Sponge Orange Band Disease in Xestospongia muta (Giant Barrel Sponge)

Rebecca Mulheron
Nova Southeastern University, rm1361@nova.edu

NOVA SOUTHEASTERN UNIVERSITY OCEANOGRAPHIC CENTER

Microbial Community Assembly found with Sponge Orange Band Disease in *Xestospongia muta* (Giant Barrel Sponge).

By

Rebecca Mulheron

Submitted to the Faculty of Nova Southeastern University Oceanographic Center in partial fulfillment of the requirements for the degree of Master of Science with a specialty in:

Biological Sciences

Nova Southeastern University

Date: August 2014

Thesis of Rebecca Mulheron

Submitted in Partial Fulfillment of the Requirements for the Degree of

Masters of Science:

Biological Sciences

Nova Southeastern University

Oceanographic Center

Approved:

Thesis Committee

Major Professor: _____

Dr. Jose Lopez, Ph.D.

Committee Member: _____

Dr. Aurelien Tartar, Ph.D.

Committee Member: _____

Dr. David Gilliam, Ph.D.

TABLE OF CONTENTS

AKNOWLEGEMENTS

I thank my advisor Jose Lopez for the guidance he has given me throughout my thesis research and my committee members for their support throughout my thesis writing. I also thank Nova Southeastern for the President's Grant that provided financial support for my thesis research. I also want to acknowledge and thank the divers who helped collect initial samples in 2012: Emily Smith, Dawn Formica, Alexandra Campbell, Shira Anteby and Ed Tichenor. I thank Rebecca Vega Thurber and Jesse Zaneveld for their early assistance and consultation with SOB analysis.

ABSTRACT

The giant barrel sponge, *Xestospongia muta* is an iconic and essential species of the coral reefs in South Florida. The sponge has primary roles providing ecosystem services and creating unique habitats for diverse microbial communities. On April 27, 2012 an outbreak of Sponge Orange Band Disease (SOB) was detected off the coast of South Florida. The disease begins with sponge bleaching, followed by mesohyl or "mesohyl" necrosis and often total mesohyl disintegration. Sampling from two diseased populations at Boynton Beach and Fort Lauderdale, FL took place on May 11th and May 29th, 2012. Each of the nine diseased sponges from Boynton Beach and the five diseased sponges from Fort Lauderdale had three separate mesophyl samples collected to examine the effects of disease progression on the microbial community. These included healthy mesohyl from a diseased sponge (HoD), the boundary layer which captured the advancing line of diseased mesohyl (BL) and diseased mesohyl from a diseased sponge (D). Mesohyl from three sponges with no visible signs of SOB disease were also collected from each sampling location to use for healthy controls (HC). Sequencing of the V4 region of the 16S rRNA gene was performed on all of these samples via the "454" pyrosequencing on a Titanium GS FLX platform. The microbial communities associated with the diseased samples revealed a microbiome shift that followed the progression of Sponge Orange Band Disease (SOB) and was dominated by Bacteroidetes, Protebacteria and Chloroflexi. No singular or group of microbes were solely found within the infected mesohyl of *Xestospongia muta* from both sampling site populations; therefore there is no unequivocal candidate as a definite microbial causative SOB agent. But there were bacteria associated with disease progression that included Armatimonadetes, Caldithrix, Chlorobi, Fibrobacteres, Fusobacteria, GN02, KSB3, OP1, OP2, OP8, Planctomycetes, SR1, TM6, Tenericutes, Verrucomicrobia, WPS-2 and ZB3. Verrucomicrobia and Plantomycetes increased significantly within the D and the BL populations, which was consistent within all the diseased sponges. This study provides a deep sequencing profile of microbial communities within *Xestospongia muta* affected with SOB Disease and provides a new insight into the sponge healthy microbiome.

Keywords: Sponge disease, 16S rRNA gene amplification; bacterial community

List of Figures

List of Tables

List of Abbreviations as Encountered

1.) SOB = Sponge Orange Band Disease
2.) ARBS = *Aplysina* Red Band Syndrome
3.) HoD = Healthy mesohyl from a diseased sponge.
4.) BL = Boundary layer, which captured the advancing line of diseased mesohyl.
5.) D = Diseased mesohyl from a diseased sponge.
6.) HC = Completely healthy sponges were used as healthy controls for *Xestospongia muta*.
7.) *X. muta* = *Xestospongia muta* or Giant Barrel Sponge
8.) NGS = Next generation DNA sequencing
9.) SRA = Sequence Read Archive
10.) DNA = Deoxyribonucleic acid
11.) rRNA = ribosomal ribonucleic acid
12.) RNA= Ribonucleic acid
13.) QIIME = Quantitative Insights into Microbial Ecology
14.) OTUs = Operational taxonomic units
15.) PCR = Polymerase chain reaction
16.) NOAA = National Oceanic and Atmospheric Administration
17.) PCoA plots = Principle Coordinate Analysis Plots
18.) BLAST = Basic Local Alignment Search Tool
19.) NCBI = National Center for Biotechnology
20.) PCB = Polychlorinated biphenyl
21.) PAHs = Polycyclic aromatic hydrocarbons

INTRODUCTION

Marine Diseases

Currently there is an increase in the incidence of marine diseases, which is unmatched in known history (Lafferty et al, 2004; McMurry et al, 2008). The most well-known marine diseases are those effecting scleractinian corals, which since the 1970's have increased in occurrence, prevalence and severity (Harvell et al, 2007; Green et al, 2000). The sea urchin *Diadema antillarum* is another well-known species whose population was basically eliminated from reef ecosystems in 1982 by a species-specific marine disease, which caused benthic community shifts resulting in reduced stony coral cover, coral recruitment, crustose coralline algae and clionid sponges (Liddell and Ohlhorst, 1986; Huges, 1994). Reduced coral cover has impacted reef ecosystems across the Caribbean by diminishing habitat structure that provides shelter and protection for a multitude of marine organisms.

Disease outbreaks in sponge populations are less publicized but have also increased in occurrence during the recent decades and have caused reductions in sponge biodiversity in marine ecosystems. Sponge diseases have been observed and documented across the globe, but the Caribbean in particular has been deemed a disease hot spot; as seen with coral disease (Harvell et al, 1999; Green, 2000). *Aplysina* Red Band Syndrome (ARBS) was found within the Bahamas affecting the sponge *Aplysina caulifomis*. A rust colored lesion forms and advances along branches leaving necrotic mesohyl behind (Olson et al, 2006). Evaluation of diseased mesohyl has shown that a Cyanobacteria was responsible for the red color, yet it is unknown whether this *Cyanobacteria* is the etiologic agent or a secondary opportunistic pathogen (Olson et al, 2006). Another sponge disease was seen in the mangroves of Jamaica where an increase in Cyanobacteria concentrations within the sponge mesohyl of *Geodia papyracea* became too toxic, causing the sponge to create spongin barriers that encapsulated high Cyanobacteria concentrations in an effort to reduce abundance (Rutzler, 1988). Sponge disease epidemics can critically reduce the populations of these slow growing organisms, which are crucial components in reef structure and function. Determining causes and reducing disease outbreaks in sponges are critical for reef survival and vitality.

The increase of marine diseases as a whole reduces reef health, resilience, diversity, and fecundity, while also endangering the services and resources that reef ecosystems supply. It is unclear whether changing environmental conditions such as, anthropogenic pollution, nutrient runoff, and invasive species are to blame for the increase in marine disease. However one aspect of global climate change; elevated seawater temperatures has been linked to disease frequency and severity in sponges. Increases in pathogen distributions, invasions, virulence, and a reduction in sponge immunity have amplified disease regularity (Sutherland et al, 2004; Lopez-Legentil et al, 2008).

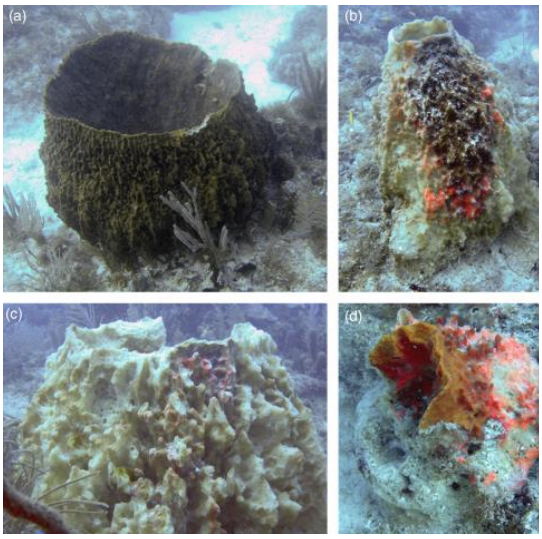


Figure 1: *Xestospongia muta* at different stages of Sponge Orange Band Disease progression: (a) Healthy sponge, (b) Sponge bleaching and tissue regression, (c) Almost complete tissue loss from disease, (d) illustrates the presence of the "orange band" that the disease is named after. Image was taken from Angermeier et al, 2011.

Sponge Orange Band Disease

Two types of events affecting sponges have been recorded in the Caribbean. The first is cyclic bleaching which is a seasonal event where the sponge loses its coloration and in most cases the sponge completely recovers (Angermeier et al, 2011). The other and main focus of this research is Sponge Orange Band Disease (SOB). SOB is described as bleaching that spreads across the sponge from lesions. The bleached mesohyl then starts to die and disintegrate, leaving an exposed skeleton behind (Angermeier et al, 2011). Bleaching progression can take a matter of weeks or progress rapidly, and as the bleaching spreads it is accompanied by the characteristic orange transition band; as its name implies (Cowart et al, 2006). SOB has been seen within the Caribbean and most often occurs in early spring to late summer. The disease has been observed to be more widespread and aggressive in August with high water temperatures often associated with coral bleaching (McMurray et al, 2011).

Xestospongia muta (Giant Barrel Sponge) has been observed to undergo SOB bleaching events, and the progression of the disease can be seen in Figure 1. *X. muta* harbors photosynthetic endosymbionts in its surface layers (pinacoderm) similar to coral, but instead of zooxanthellae they contain Cyanobacteria of the genus *Synechococcus*. The Cyanobateria gives the sponge its reddish brown surface coloration. The Cyanobacteria are dispersed during cyclic bleaching, which gives the sponge its "bleached" coloration and Cyanobacteria return later as the sponge recovers. During SOB the mesohyl dies and crumbles off (as seen in Figure 1, b-d), providing no option to the sponge for regeneration. No cause or pathogen has been identified for the disease to date. The rapid nature of the disease and increasing observations of new cases in new areas shows that SOB is spreading. Bleaching was first documented in Puerto Rico in 1987, then again in Curacao, Nicaragua in 1997 where SOB was described as sponge wasting disease (Vicente, 1990; Nagelkerken et al, 2000). During 2000-2005 SOB was also found in Cozumel Mexico, where it decimated *X. muta* populations (Gammill and Fenner, 2005). SOB was first reported in U.S. waters at Key Largo on Conch Reef and Dixie Shoals Reef during 2007 and then again at Key Largo in Conch Reef during 2009 (Angermeier et al, 2011). On April 27, 2012 SOB was again reported in U.S. waters, off South Florida. The full extent of the epidemic is unknown since sponge monitoring projects are few and far between. This uncertainty provides an increased urgency and demand to better understand the disease for sponge conservation and ecosystem management.

Xestospongia muta

X. muta is found in reefs throughout the Caribbean, it has key roles in providing ecosystem services that are used by the reef system as a whole (Diaz et al, 2001). The sponge is a fundamental part of the coral reef community, and occupies greater than nine percent of reef substrate habitat in the Florida Keys (McMurry et al, 2008). The abundance of *X. muta* increases ecosystem biodiversity, and provides a food source for a variety of species (Bell, 2008). The sponge structure itself creates an un-replicable habitat and protection for invertebrates, microbial, and macrofaunal communities (Bell, 2008; McMurry et al, 2008). Sponges further contribute to reef ecosystems by actively completing the nitrogen cycle that converts ammonia to nitrate; providing nutrients for

organisms to use in the "nutrient poor" coral reef system (Bell, 2008). *X. muta* has specifically been found to posess a sponge derived *nif*H gene that codes for a specific iron protein component of the nitrogenase enzyme, which is responsible for nitrogen fixation (Fiore et al, 2013B). The sponge also facilitates benthic and pelagic coupling through filtering seawater, which transfers carbon from the water column to higher trophic levels (Webster, 2007; Bell, 2008). It has been estimated that 1 kilogram or 2.2 pounds of sponge can filter up to 24,000 liters of seawater per day (Vogel et al, 1977). *X. muta* has also been called "the redwood of the reef" because of the massive size they can obtain, and because of their long life spans (McMurry et al, 2008). The sponge has been known to grow to over 1 meter in height and diameter (McMurry et al, 2008). It has also been found to be as old as 2,300 years, which is around the same age as the actual redwood trees found in California (McMurry et al, 2008). The role that sponges play in reef ecosystems has often been overlooked but recent research has identified that they are valuable, functional members of reef communities and are necessary for reef resilience and health. Reefs themselves comprise a nursery habitat for declining fish populations, reduce storm surges and help control Global Climate Change by fixing $700x\ 10^{12}$g of gross carbon dioxide a year (Veron et al, 2009; Aronson et al, 2006).

Microbial Diversity

Over time molecular phylogenetic studies have identified that the diversity of life is due to microbial communities (Pace, 1997). Their mass numbers, high diversity and large geographic expanse makes the community as a whole difficult to understand. There population size has been estimated to be greater than 10^{30} microbial cells, which surpasses the amount of known stars in our universe by nine times (Knight et al, 2012). Not all microbial organisms can be studied in laboratory pure-culture conditions, which further restricted the known diversity and characterization (Pace, 1997). Overall little is known about the microbial world; yet they are responsible for the function and services of our biosphere (Pace, 1997).

Sponges are well known as hosts and habitats for microbial communities, as their mesophyl contains diverse and abundant microbial populations (Webster and Blackall, 2009). Sponge microbial counterparts (bacteria, archaea, fungi and microalgae) have been

estimated to make up > 50% of total sponge biomass (Santavy et al, 1990; Fieseler et al, 2004) *Xestospongia muta* specifically has been classified as a bacteriosponge, because of the high microbial abundance found within its mesohyl that can be 2-4 orders of magnitude greater than surrounding seawater concentrations (Hentschel et al, 2005).These microbial inhabitants can provide services to the sponge such as nitrogen fixation, nitrification, sulfate reduction, dehalogenation, methane oxidation, photosynthesis, stabilization of the skeleton, processing metabolic waste and secondary metabolite production (Hentschel et al, 2002; Taylor et al, 2007b). For example Cyanobacteria can provide novel metabolic pathways to the sponge in exchange for products from the metabolism and the nutrient rich environment that the sponge itself provides (Freeman and Thacker, 2005; Taylor et al, 2007b; Flatt et al, 2005; Webster and Blackall, 2009). Shade studies estimate that Cyanobacteria symbionts can provide up to 50% of the sponges needed energy intake (Freeman and Thacker, 2005; Wilkinson 1983; Cheshire and Wilkinson 1991). Therefore microbial shifts can greatly influence the sponge health and functional capabilities (Webster et al, 2012). These microbe-metazoan associations have been called the most ancient mutualistic relationships, since vertical transfer has created bacterial speciation within sponge species (Montalvo and Hill, 2011; Freeman and Thacker, 2005; Schmitt et al, 2008; Sharp et al, 2007; Collin et al, 2010; Taylor et al, 2007b). A sponge specific bacterial phylum *Poribacteria* has been found; providing a tangible example about how little is known about microbial diversity in sponges and within ecosystems (Fieseler, et al, 2004; Taylor et al, 2007b; Lafi et al, 2009).

While sponges utilize the compounds that their inhabitants produce, they also have evolved over time to make their own secondary metabolites and bioactive compounds. These products have very specific antibacterial, antiviral, anti-inflammatory and anti-fouling properties that aid in sponge survival, defense and competition for space (Webster, 2007; Bell, 2008). These natural products and secondary metabolites produced by the sponge have also been used for biomedical applications, such as producing antibiotics and anti-tumor medicines (Webster, 2007). For instance seventy five percent of antitumor natural products have been derived from sponges (Koopermans et al, 2009). *X. muta* specifically is known to harbor a high density of symbiotic bacteria and can produce acetylenic acids, which is a compound used by sponges because of its antimicrobial and

cytotoxic properties (Patil, et al, 1992; Morinaka et al, 2007; Richelle-Maurer et al, 2003). Acetylenic acid has been derived from *X. muta* and has been shown to inhibit HIV proteases ability to replicate; providing new techniques to fight a very deadly human autoimmune disease (Patil et al, 1992). Another natural product derived from *X. muta* is mutafurans A-G (seven chiral brominated ene-yne 2,5-disubstituded tetrahydrofurans), which have been found to be an antifungal against a pathogenic fungus (*Cryptococcus neoformans*) associated with HIV patients (Figure 2) (Morinaka et al, 2007). *X. muta* has also been found to produce the sterols: mutasterol and xestostanol, which are structurally unique from terrestrial sterols and are believed to play a role in cell membrane stabilization (Kokke et al, 1979; Niang et al, 1980). The differences between marine and terrestrial sterols provide the possibility of understanding the phylogenic significance of sterols and their role in membrane and dietary pathways (Niang et al, 1980). Understanding cellular pathways will enhance medical understanding of related disorders and help advance biomedical applications.

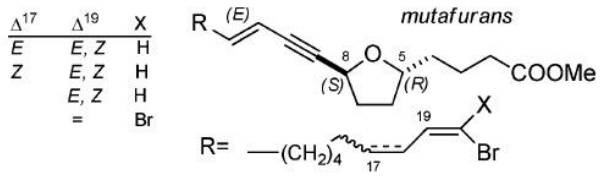


Figure 2: mutafurans compound derived from *X. muta.* Image taken from Morinaka et al. 2007

Therefore sponges are an ideal host to study microbial diversity for novel discovery which could advance biotechnology, ecosystem bioremediation, disease biomarkers and environmental characterization (Knight et al, 2012). *X. muta* itself has shown that the pharmaceutical opportunities that sponges possess are applicable and waiting to be discovered. Understanding microbial diversity will provide knowledge of the factors that can drive ecosystem functions and provide the services essential for life. Understanding how SOB Disease is affecting *X. muta* will improve conservation of its ecological role in reef ecosystems and maintain its availability for potential biomedical applications.

16S rRNA

The paradigm shifting approach to study microbial communities was developed by Carl Woese, who pioneered the use of the ribosomal RNA (rRNA); in particular the small subunit (16S) (Woese, 1987; Rappe and Giovannoni, 2003). This method allowed

microbial communities to be sampled from the environment, instead of laboratory pure cultures. Without the limitations that culturing inflicts, >99% of environmental microbial communities can be revealed; included rare or un-culturable microbial communities that otherwise would have been lost (Hentschel et al, 2002).

16S genes are highly conserved within bacteria because they are part of an essential organelle present within all bacteria, which makes them ideal for studying phylogenetic analysis and microbial diversity. The 16S rRNA taxonomic marker has also been shown to produce high-resolution snapshots of microbial community assemblages and provide insight to the functional relationships between microbes and their environment (Larson et al, 2012; Caporaso et al, 2011).

The 16S rRNA region for this research will be amplified using primers 515F and 806R (Caparaso et al, 2010). These primers target the hypervariable V4 region (Figure 3). This region was selected due to its higher accuracy in taxonomic classification stemming from a lower sequencing error rate (Wang et al, 2007). Regions V3-V5 (including V4) have also been shown to produce the lowest chimera rates compared to other 16S regions (Haas et al, 2010). No hypervariable region to date has received universal acceptance as the optimal region over another (Schloss et al, 2011).

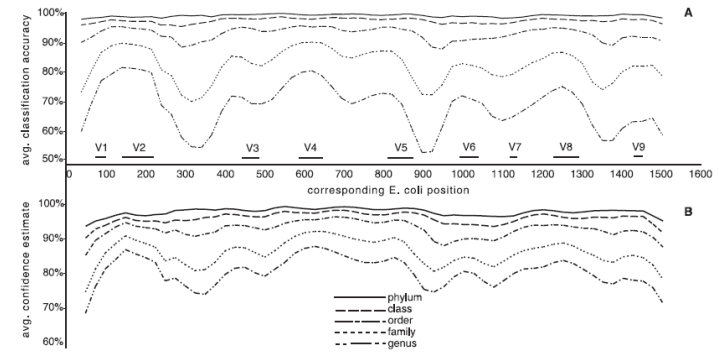


Figure 3: (A) Classification accuracy rates, the black lines on the x axis display the range of the hypervariable regions. Region V4 has one of highest accuracy rates, when compare to other regions. (B) Average bootstrap confidence estimates for each segment, which also shows V4 having one of the best confidence estimates. Image taken from Wang et al, 2007.

Next Generation Sequencing

Next generation DNA sequencing (NGS) has advanced the scientific communities ability to analyze microbial communities taken directly from their environment. This was previously restricted to microbes that were cultivable in the laboratory, as new technologies only require a gene sequence (16S rRNA) instead of a functioning cell from culture (Pace,

1997). Next generation sequencing has further reduced sequencing error by removing the vector-based cloning step, which greatly improves accurately (Mardis et al, 2008).

The di-deoxyribonucleotide (also known as Sanger) sequencing was invented in 1977. It detects 500-600 bases from the maximum 96 reactions and produces sequences of ~500bp (Mardis et al, 2011). Although used for many years, Sanger sequencing for 16S rRNA genes has insufficient coverage needed to define and compare complete environmental microbial communities (Bartram et al, 2011).

Advances in sequencing technology have created the ability to produce millions of sequences compared to the previous methods, by several orders of magnitude and at a lower cost (Mardis et al, 2008). This allowed an in-depth look at microbial diversity and community assemblages never before studied. The "454" Titanium GS FLX system creates single stranded DNA libraries which are bound by adapters to sequencing beads. The DNA-capture beads are within a water/ oil emulsion and are coated with complementary primers which hybridize when mixed with DNA templates. The emulsion PCR reaction creates millions of oligomers from the specified DNA library (Mardis et al, 2011). The

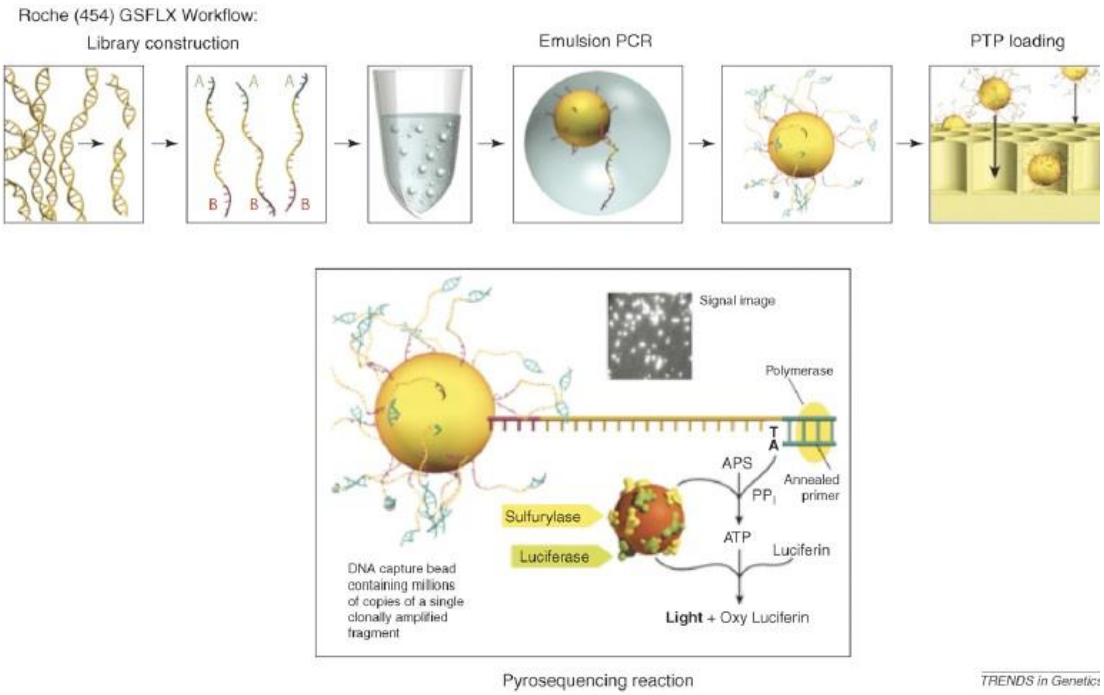


Figure 4: 454 Pyro-sequencing overview. Starts with library construction and ligation with primers to produce DNA fragments. DNA capture beads and DNA fragments are combined and emulsion PCR amplifies sequences. Beads are placed on the picotiter plate (PTP). The bottom image displays the reaction within the sequencing machine, which polymerases complimentary nucleotides and emits the corresponding light intensity. Image taken from Mardis et al. 2008.

beads are then loaded onto a picotiter plate where they adhere to separate wells. The sequencing machine simultaneously identifies the specific nucleotide order in the amplified templates by "pyrosequencing". Polymerization occurs as a complementary nucleotide is added to the well, this triggers pyro-phosphates conversion into ATP, which is then utilized by luciferase to produce light (Mardis et al, 2008; Mardis et al, 2011). The intensity of the light produced is directly correlated to the number of nucleotides present within the template. A standard run takes no longer then a day and can produce 1 to 1.5 million single-end reads, with an average length of 500 bases (Figure 4) (Logares et al, 2012).

The advantages of using 454 pyrosequencing is the use of PCR oil emulsion, which keeps the DNA templates constrained within the oil and therefore avoids the loss of genomic sequences (Metzker et al, 2010). It also produces average read lengths of 700bp which are larger than other high throughput sequencing systems (Illumina Hiseq and SOLiD 4 platforms). 454 pyrosequencing also includes faster run times and a lower substitution error rate. The weaknesses of using 454 pyrosequencing are the cost per sequence read, which often restricts data sets and causes a failure in reaching diversity saturation for a given environment (Gloor et al, 2010; Bartram et al, 2011). The detector which interprets light intensity also has a saturation point, which can cause indel errors with homopolymers (Mardis et al, 2008). Pyrosequencing also has a much larger overall per-read error rate then both the Sanger method and Illumina Hiseq (Gloor et al, 2010).

Bioinformatics

The field of bioinformatics encompasses how the biological sequence (DNA, RNA or protein) data is stored, retrieved, organized and analyzed; specifically utilizing computational software tools. There are a variety of very specific software programs and databases that can analyze the sequences derived from NGS platforms targeting the 16S rRNA region. These tools help interpret data (millions of bits of biochemical compounds) and can process hundreds of communities simultaneously allowing patterns within those communities to be seen. (Jansson, 2013; Caporaso et al, 2010).

QIIME (Quantitative Insights Into Microbial Ecology) is an open source software suite which evaluates raw DNA sequence reads by clustering them into operational taxonomic units (OTUs), measuring diversity by creating visualizations to better represent and explain the datasets (Caporaso et al, 2010). QIIME specializes on analyzing microbial

communities from NGS and was created to analyze and annotate raw 16S rRNA sequence data. Computational steps can be taken in QIIME to avoid artificial diversity, as errors can be created during pyrosequencing and included molecular chimeras. PCR amplification bias can also create greater diversity within the data-set then what really exists. These can be removed through "de-noising" which removes these sequencing errors and "chimera slayer" which removes chimeric sequences (Reed and Knight, 2010).

HYPOTHESES and OBJECTIVES

Hypothesis:

1.) H1: Sponge Orange Band Disease is produced by a bacterial causative agent.
2.) H2: The microbial communities in healthy mesohyl will have a distinct profile compared with diseased mesohyl and thus reveal probable causative agents of SOB.
3.) H0: A non-microbial agent is causing SOB.

Objectives:

1.) To characterize the microbial community (bacteria and archaea) present in both healthy and Sponge Orange Band Disease diseased *Xestospongia muta* (Giant Barrel Sponge) archived mesohyl using 16S RNA and high throughput DNA sequencing methods.

2.) Compare the two microbial communities with current bioinformatics software and programs.

3.) Analyze "environmental" meta-data from NOAA meteorological data for possible correlations with microbial community shifts.

Overall this project will provide the science background necessary for future experiments to identify the causative agent for SOB by verifying Microbial/Koch's Postulate.

MATERIALS and METHODS

Sample Collection and Preservation:

Diseased *X. muta* samples were collected on two separate occasions and at two different sites under Florida Fish and Wildlife Conservation Commission fishing licenses and Special Activity License-12-1372-37a. The first collection was on May 11th 2012 at Gulfstream Reef, Boynton Beach FL (N26º 30.595, W80º 02.019). Nine diseased sponges and three apparent healthy neighboring *X. muta* sponges were sampled. The second collection was on May 29th 2012 in Fort Lauderdale, FL (N26º 10.090, W80º 04.600). Five diseased sponges and three apparent healthy neighboring *X. muta* sponges were sampled from the second site. Diseased sponge specimens had approximately three inches of mesohyl sampled from active diseased mesohyl (bleached/ completely dead), healthy mesohyl present on diseased sponges (retaining normal brown pigmentation) and the BL mesohyl so that the receding mesohyl boundary was captured. Samples from each sponge were then partitioned into the following categories: "Healthy on Diseased" (HoD); completely "Healthy" control sponge (HC); sponges with "Boundary Layer" (BL) of healthy and diseased tissue; or a completely diseased sponge with disintegrated mesohyl (D) (Please see list of abbreviations). Overall 20 sponge individuals yielding 48 distinct sponge samples were collected for library analysis, while two soil samples and one water

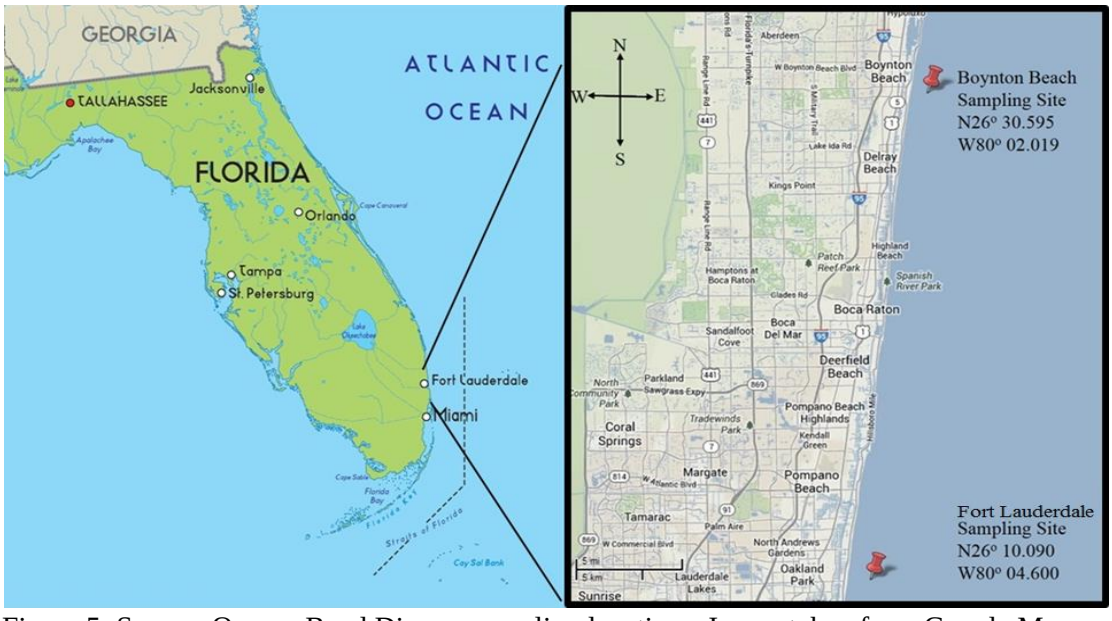


Figure 5: Sponge Orange Band Disease sampling locations. Image taken from Google Maps and http://www.ezilon.com/maps/united-states/florida-geographical-maps.html

sample were ultimately analyzed. Location of collection sites can be seen in Appendix Table 1 and Figure 5, which shows the sampling sites location on the South Florida coast line and the relative location of the SOB outbreak. Samples were preserved by storing them in Trizol (Invitrogen) for future RNA work and at -80^{o}C until further processing. The metadata for both sampling locations can be seen in Appendix Table 1.

Genomic DNA Extraction:

To extract DNA from collected sponge samples, the mesohyl was first conditioned using the Squeeze Method (Lopez Molecular Genetics lab manual; unpublished). This entailed extracting a 3cm^3 piece of sponge from the samples collected, which was then placed within a sterile Petri dish. For the BL mesohyl samples the advancing line of necrotic mesohyl was characterized by a band of brown mesohyl. This characteristic feature was intentionally included in all BL mesohyl samples excised for DNA extraction. The sponge was then soaked with 1mL of L-Buffer, which started lysing cells and allowed the DNA to be more accessible. L- Buffer is comprised of 0.01 M Tris-Cl (pH 7.6), 0.02 M NaCl, and 0.1 M EDTA (pH 8.0) (Sambrook and Russell, 2006). The sponge was then minced as much as possible with a sterile scalpel and another 1mL of L-Buffer was added. The minced sponge was then squeezed using the scalpel blade against the side of the petri dish to remove as much L-Buffer as possible. The extracted L-Buffer was collected in sterile tubes and kept at 4^{o}F until DNA extraction. For DNA extraction the tubes are then centrifuged at 9 rpm for 5 minutes at room temperature. The supernatant was discarded and the remaining pellet was centrifuge again at the previous conditions. The remaining supernatant was removed without disturbing the pellet.

For DNA extractions the Power Lyzer Power Soil DNA Isolation Kit (MO BIO) was used and the amendments to the protocol made by the Earth Microbiome Project were followed (www.earthmicrobiome.org). The kit works well with sponge mesohyl because it contains two steps that remove metabolites from the samples. Sponge mesohyl is known for having metabolites, which can infer with PCR tests. DNA from D, BL, HoD, and HC mesohyl from both sites was extracted. The PowerLyzer 24 homogenizer (MO BIO) was be used with 2.8mm ceramic bead tubes (MO BIO) at 4000 rpm for 45 seconds to further cell lyses. Ceramic beads where chosen because within laboratory trials they proved to be

the least DNA shearing, and improved DNA yield. The speed of the homogenizer was optimized due to the fine consistency of the sponge mesohyl, which was similar to clay and silt.

16S rRNA Amplification:

16S rRNA gene fragments were amplified with PCR primers E. coli 9 (5' GAG TTT GAT CCT GGC TCAG 3') and Loop27rc (5' GAC TAC CAG GGT ATC TAA TC 3') under standard PCR conditions to test the PCR yield and quality of extracted DNA (Lopez et al, 1999). Then the 16S rRNA was amplified by PCR using universal bacteria/ archaeal primers 515F (AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT ATG GTA ATT GTG TGC CAG CMG CCC CGG TAA) and 806R (CAA GCA GAA GAC GGC ATA CGA GAT TCC CTT GTC TCC AGT CAG TCA GCC GGA CTA CHV GGG TWT CTA AT) (Caporaso et al, 2012). Different Multiplex Identifiers (MIDs) were included in the reverse primer and are a unique DNA barcode that is essential for library preparation. The 454 pyro-sequencer parallel sequences all samples at once, then each sequence must be quantified individually before they can be analyzed as a whole. The MIDs are attached to all sequence reads produced by the run and will allow the sequences produced to be sorted and the given its associated sample identification.

TaKaRa Ex Taq DNA polymerase (Takara Bio) was used due to its high accuracy which results in less mutation and its ability to work efficiently with DNA that has higher than 65% GC content. Thermocycler conditions for PCR amplification of 16S rRNA are 94^{o}C for three minutes, thirty five cycles of 94^{o}C for forty five seconds, 50^{o}C for one minute, 72^{o}C for one and a half minutes, 72^{o}C for ten minutes and 4^{o}C on hold (Caporaso et al, 2010). The 16S rRNA amplification procedure was consistent with the Earth Microbiome Protocol. Two separate PCR reactions of every sample were run and combined.

The combined amplicon products were then run on a 2% low melting point agarose gel. Bands at the correct size (~300bp) were extracted with a sterile scalpel and purified using the standard protocol from the QIAquick PCR Purification Kit (Qiagen). The purified PCR products were checked with both the Nanodrop 1000 spectrophotometer (Thermo Scientific) and the Agilent 2100 Bioanalyzer G2939A DNA high sensitivity assay. This allowed the amplicon products purity and quality scores to be analyzed. Purity limits were

taken from the Earth Microbiome standards and are 1.80-2.00 for the 260/280 value. Staying within this quality limit ensured that amplicon samples had limited contamination and the results represented the true specimen. The samples were run on the 454 Titanium GS FLX system located at the University of Kentucky.

PCR Error:

PCR amplification was the first step where bias was mostly likely to occur, with the creation of artifacts from point substitutions and chimeras which would skew sequence reads (Acinas et al, 2005; Aird et al, 2011; Patin et al, 2013). PCR polymerases have been known to produce error rates of 1 substitution per 10^5- 10^6 bases (Cline et al, 1996). This bias creates uncertainly and error when estimating microbial diversity and community structures (Acinas et al, 2005; Huse et al, 2010). To reduce PCR bias two separate PCR reactions were ran for every sample and combined. This reduced PCR drift that would cause low frequency error artifacts. Chimeras are incomplete extensions of sequences during amplification, which then continue to multiply in future amplification cycles. Chimeric sequences most often occur between closely related sequences but when they form from multiple parent sequences and different phyla they can be detected as novel organisms; creating false diversity (Hass et al, 2011). The rate of chimerism is estimated as 5- 45% of all diversity present, causing the resulting rare biosphere diversity estimates to be called into question as being a product of sequencing error rather than present community assemblages (Schloss et al, 2011; Haas et al, 2011). To manually reduce chimera creation and further reduce PCR artifacts the amplification cycles were as few as possible, which minimized chimeras and Taq polymerase errors (Acinas et al, 2005; Haas et al, 2011). Precautions made at the PCR amplification step can also reduce GC suppression; such as reducing the ramp speed on the thermocycler. Extending denaturalization of the DNA provided the time necessary to fully separate densely coiled GC rich sections, allowing proper amplification of those areas (Aird et al, 2011). The diverse bacterial assemblages within sponge mesohyl also help avoid PCR bias by providing a multitude of templates for amplification. This avoids over concentration of one template, which causes re-annealing inhibition effects (Acinas et al, 2005; Suzuki and Giovannoni, 1996).

Next Generation Sequencing:

The project utilized the 454 Titanium GS FLX platform provided by a core lab at the University of Kentucky. The sequencing of the 16S rRNA V4 region was performed on 14 different sponges, each with three mesohyl samples from the three different health states (D, BL and HoD) (Appendix Table 1). Sampling multiple sponges from the same site and from different sites was intended to reduce true biological, temporal and technical variation (Knight et al, 2012; Thomas et al, 2012; Logares et al, 2012). To further reduce error three HC sponges that did not visually display SOB from each sampling site was sequenced and used for comparison against diseased sponges.

Sequence Analysis:

The sequences produced were analyzed using the Quantitative Insights into Microbial Ecology v.1.8.0 (QIIME) (Caporaso et al, 2010). During sequencing each nucleotide was given a quality score (Phred Score), the QIIME default quality standards (Q25) was used to evaluate the data (Caporaso et al, 2010). The raw sequences were processed using AmpliconNoise (ampliconnoise.py) which had five simultaneous steps (Quince et al, 2009). The first process filtered the flowgram files into individual sample files, and then the PyroNoise algorithm clustered the flowgrams in order to remove noise, which allowed for accurate OTU and phylogenetic tree construction (Quince et al, 2009). The SeqNoise algorithm preformed sequence clustering to remove single base errors, while the Perseus algorithm identified chimeras, which are amplified sequences or "artifacts" produced during PCR and inflate alpha diversity measures by falsely identifying the sequence read as a rare species (Ley et al, 2008). Finally the Fasta files created by AmpliconNoise were used for downstream analysis. After de-noising the reverse primers were removed from the sequences (truncate_reverse_primers.py). Then OTUs were assigned with UCLUST (Edgar, 2010) which utilized open reference OTU picking with 97% similarity (pick_open_reference_otus,py). The Greengenes (8/13 release) reference database (DeSantis et al, 2006) was used to pick a representative sequence set, which was later aligned (pick_rep_set.py and align_seqs.py). Taxonomic classifications were assigned to the representative sequences using the 8/13 release of Greengenes at a 97% confidence cutoff and UCLUST (assign_taxonomy.py). The sequences were then filtered and the corresponding phylogenetic tree was created (filter_alignment.py and

make_phylogeny.py). A Biom file or OTU table was then created and utilized for downstream data-set analysis (make_otu_table.py). Sequences from this pyrosequencing were submitted to the Sequence Read Archive (SRA).

Statistical Analysis:

Within QIIME 1.8.0, a variety of statistical tests were used to better understand the relationship between the different microbial community assemblages and SOB Disease. Alpha diversity analysis (alpha_diversity.py, collate_alpha.py, make_rarefaction_plots.py) included a variety of statistical tests which measured the species richness found within a health state based on the number of species present and the proportion of species represented within the community. Determining the diversity present in each sample highlights the differences within the microbial community composition of *X. muta* and how the progression of SOB Disease influences the microbial diversity. Alpha diversity analysis usually relies upon evaluating species richness (amount of species) and/ or species evenness (relative abundance). Since all of the health states contain relatively the same core communities, the distinctive variation found between the health states was from the rare microbiota. Therefore when choosing diversity statistics to evaluate species richness it was important to choose formulas that took into account the rare microbial communities. The alpha diversity statistics chosen to evaluate the species richness were Chao1, ACE, Shannon Index of Diversity, and Phylogenetic Distance.

Chao1 is a qualitative measure which takes into account the observed richness, numbers of singletons and the amount doubletons when estimating species richness (Chao, 1984). This statistic assumes that rare species provide the most pertinent information about the number of undetected species, which influences overall diversity (Nicholas and Chao, 2013). The abundance-based coverage estimator (ACE) is also a qualitative measure that uses the number of singletons and doubletons, but also the frequency of other rare species to determine the amount of sample comprehensiveness (Chazdon, 1998; Nicholar and Chao, 2013). The Shannon Index of Diversity is a quantitative measure that differs from the ACE and the Chao1 statistics because it takes into consideration both richness and evenness when it determined the significant differences in abundance. Overall the diversity metric measures entropy within a data set by determining the difficulty in predicting the identity

of the next individual sampled (Shannon and Weaver, 1949). The phylogenetic diversity is also important to consider as it focuses on how much of the phylogenetic tree a sample will cover and measures diversity by agreeing that more diverse communities consist of a larger number of deeper branching lineages (Lozupone et al, 2008). After alpha diversity statistics were determined the script compare_alpha_diversity.py, which was a nonparametric two-sample t-test (Monte Carlo) was utilized to identify significantly different relationships by comparing each pair of groups (health states and location).

Beta Diversity:

Beta diversity (beta_diversity.py) included statistical tests which measured the dissimilarities between two or more communities. This identified how the microbial community composition had shifted between different health states and the influence of sampling location. The resulting similarities and differences comprising the microbial assemblages identified the effect that SOB Disease had on the microbial inhabitants of *X. muta* and helped determine the possibility of a microbial causative agent. The beta diversity statistical tests used included the Jaccard Abundance Index, the Bray-Curtis Dissimilarity Index, the weighted UniFrac and the unweighted UniFrac. These statistics were displayed using Principle Coordinate Analysis plots (PCoA plots) which allowed lineages responsible for divergence in the microbial communities to be identified. This analysis identified if the microbial communities within *X. muta* were similar or significantly different within sponges at the same site and between the two different sampling sites. Hypothetically this would allow a microbial causative agent found at both sites to be compared by relative abundance and phylogeneticlly, which would determine if it was indeed the same microbial agent at both sites.

The Jaccard Abundance Index is an occurrence based measure which takes into account the comparative compositional similarity between two microbial communities. It compares the amount of shared species to the total number of species in the combined communities (Gotelli and Chao, 2013). The Bray-Curtis Dissimilarity Index (Bray and Curtis, 1957) measures the variation within the composition of two communities, in order to expose diversity relationships within the microbial communities.

When relative abundance measurements are they only methods used to analyze beta diversity the community composition analysis is lacking the depth of knowledge provided by phylogenetic measures. Phylogenetic distance measurements are critical to understanding the degree of divergence between the different sequences and provide details about ecological patterns that could be present within the dataset (Lozupone et al, 2005; Lozupone and Knight, 2008; Hamady and Lozupone, 2010). They also determine if microbial communities significantly differ based on the phylogenetic lineages they contain. The unweighted UniFrac statistic is a qualitative measure that depicts the distances between two communities (the amount of evolution) by calculating the amount of branch length within the phylogenetic tree that led to descendants in either group but no both (Lozupone et al, 2007; Lozupone et al, 2008). Therefore for communities to differ there had to be distinctive evolutional lineages which occurred and allowed a community to adapt to the environmental pressures, which in the end allowed them to outcompete and survive other microbes (Lozupone et al, 2005; Lozupone et al, 2008). Overall the unweighted UniFrac statistic measures the presence and absence of microbes and by association identifies factors that restrict taxon abundance within the communities. This determines present lineages and taxa in the microbial communities, which allows comparisons between multiple communities simultaneously

The weighted UniFrac statistic is a quantitative measure that calculates changes within the relative abundance of sequences from the microbial lineages between communities using the abundance measurements of branches that comprise the phylogenetic tree (Lozupone et al, 2008). When changes in the abundance of sequences from each lineage and taxa present are discovered, it highlights the presence and absences of microbial lineages (Lozupone et al, 2007) It is sensitive enough to detect changes in abundance even when the core communities remain consistent throughout all samples (Lozpone et al, 2007). The weighted UniFrac also exposes “transient factors” that affect the presence of microbial lineages within communities (Lozupone et al, 2007). Overall the weighted UniFrac plots show the relative OTU abundance per sample. Providing a 3D representation of how similar samples were to each other based their orientation; as distinct and highly similar samples depict a dense clustering pattern in relation to other samples. After the beta diversity statistics were determined the script beta_signficance.py, which

utilized the UniFrac significance test was used to identify significantly different relationships by comparing each pair of groups (health states and location). The quantitative and qualitative approaches both analyze different factors that influence the microbial communities. Together they provide a more comprehensive look at the microbial community composition, influencing factors and community shifts. These statistical tests allowed accurate analysis of microbial community composition in the populations of *X. muta*.

Jackknife Plots:

To test the robustness of the unweighted and weighted UniFrac PCoA plots jackknife plots (jackknifed_beta_diversity.py) were used, which illustrated the effects of sample depth and evenness (Lozupone et al, 2007; Ley et al, 2008). The vast size of microbial communities makes it challenging to sample communities completely; "jackknifing" utilizes Procustes analysis which repeatedly resamples each sample from a population and creates an ellipse that represents the interquartile range (IQR) for each sample (Lozupone, 2007; Lozupone et al, 2011). A larger IQR shows more variability and possibilities within the sample, which alters how the sample is portrayed within the dataset. While a smaller IQR shows that the results portrayed an accurate representation of the sample and that similar results will be achieved even with resampling.

Distance Boxplots:

Distance boxplots (make_distance_boxplots.py) compare the beta diversity differences between and within the microbial communities. Translating the specific distribution, location and spread which are specific for each health state. The distances present within the separate health states help to convey the nature and relationship of the samples that comprised them.

Rarefaction Plots:

The rarefaction plots (make_rarefaction_plots.py) display diversity saturation and a variety of diversity statistics. This determines whether the samples within the dataset have fully captured existing diversity or if additional sampling is needed to accurately

represent the community. Capturing existing diversity is essential because in order to describe a community it is necessary to know who comprises it first. Comparing microbial communities where diversity has not been fully captured can be problematic since trends and patterns within the data set cannot be linked to environmental conditions or disturbances since those changes may be contributed to a lack of represented diversity. With the lack of an accurate microbial population representation there would also be no way to accurately identify a possible microbial pathogen, even if the causative agent was present.

Rank Abundance:

Rank abundance curves (plot_rank_abundance_graph.py) relate a pattern of the absolute abundance for the species found within the samples. The curves are another way to display the changes in species richness and abundance. The y-axis measures the absolute abundance, while the x-axis represents species rank which directly represents the species richness. The evenness is also correlated to the slope of the graph.

Group Significance:

Group Significance (group_significance.py) compares the OTU frequencies within the different health states and identifies significant statistical differences present within the microbial communities. The non-parametric ANOVA (Kruskal Wallis) compares within treatment variance to between treatment variance, identifying which OTUs differ within the distribution. The OTU table is rarified and filtered to remove OTUs that are not found within at least 10% of samples. This helps to reduce noise and highlight OTUs more commonly associated with at least one health state. The composition of the D health state was evaluated by determining which bacteria were present but also absent within the HC and were statistically significant by having a p-value < 0.05.

Adonis:

Adonis (compare_categories.py) analyzes the differences in composition and relative abundance of different species that comprised different treatments (Anderson, 2001). It is ideal for ecological community comparison because the data does not have to

follow a normal distribution; as species abundance data is often highly aggregated or skewed (Anderson, 2001). Species abundance data can also contain both discrete and continuous values, as species with small means and/ or rare species often contribute asymmetric distributions with many zeros; which therefore require an algorithm that can handle both sets of inputs to fully analyze a community (Anderson, 2001). Overall Adonis describes the strength and significance that a group has in determining variation of distances, and is used instead of ANOSIM because of a more robust statistic. Adonis is typically used for analyzing ecological community data or genetic data, which generally have limited number of samples and a multitude of gene expression data (Zapala and Schork, 2006). The test uses an F-statistic instead of the common p-value. The F-statistic compares the variability within groups versus variability among different groups (Anderson, 2001).

Rational Behind Algorithms Chosen to Process Sequence Data:

The abilities of bioinformatics and the algorithms used to interpret pyro-sequencing data are constantly evolving and improving. To date no consensus in the field has been reached on how to best analyze data to ensure accuracy and avoid artificial microbial diversity. This has caused concerns about overestimating diversity and the need to accurately distinguish artifacts (sequences containing error) from rare environmental samples in order to accurately determine microbial community richness and composition (Patin et al, 2013). Highlighting the need for transparency when presenting bioinformatics research and disclosing the quality control measures taken in this particular study.

AmpliconNoise processing was chosen over standard de-noising because of its ability to also save rare OTUs that would otherwise be discarded; producing higher per sample richness (Bakker et al, 2012). The algorithm also has a stricter OTU read retention standard, which reduces "noise" that is often mistaken for community diversity (Bakker et al, 2012). During 454 pyro-sequencing, errors often occur with homoploymers (single base repeats) since the light intensities do not accurately reflect homoploymer lengths and often result in insertions and deletions (Patin et al, 2013; Kunin et al, 2010; Quince et al, 2009). The AmpliconNoise algorithm takes into account the flowgrams produced by pyro-sequencing and determines if the read is "noise" or a novel sequence by recognizing that

sequence errors are rare and that they would be highly similar to the accurate sequence within the dataset (Quince et al, 2009 and 2011). The Perseus algorithm is more accurate than other chimera removal algorithms because of its high sensitivity, which allows it to accurately remove 99% of the very small pyro-sequencing chimeras (Quince et al, 2011). Final OTU construction is performed by AmpliconNoise using complete linkage and exact pairwise alignments. Accurate OTU construction is of concern as the standard methods of aligning sequences to calculated distances of 3% often overestimates OTU diversity six-fold (Quince et al, 2009). This is because sequence errors during pyro-sequencing create sequence differences which then increase OTU numbers and concomitant artificial diversity within the dataset. As mentioned previously the flowgram clustering step removes the majority of the false OTUs and the remaining are removed during chimera deletion (Quince et al, 2009). AmpliconNoise has been found to reduce OTU diversity from the norm by 40%; allowing accurate OTU assignment that determines microbial diversity even at low sequence differences (Quince et al, 2009).

After denoising open-reference OTU picking was performed, instead of closed-reference OTU picking because it does not remove unknown sequences. Open-reference OTU picking took the sequences and clustered them to the 8/13 Greengenes reference database, those that did not match were clustered de novo instead of being thrown out, which is standard in the closed-reference OTU picking. This difference allows analysis of previously unknown microbial communities.

Even Sampling Depth:

During pyro-sequencing, thousands of sequences are produced with a barcode that corresponds to a specific sample. The amount of sequences per sample varies, for example low sequence counts can be around 300, while high sequence counts can be around 30,000 and higher. This can be problematic as each sample needs to be compared equally to each other, since varying sample sequence depths influence diversity estimates and non-equalized libraries tend to inflate richness measurements (Lundin et al, 2012; Gihring et al, 2012). To produce an unbiased diversity analysis of the microbial communities within a data set, an even sampling depth has to be chosen. The even sampling depth will randomly subsample each sample by a given number and samples with less sequences then the

sampling depth number will be removed from analysis. Recent publications have analyzed and compared data sets in order to estimate the minimum sampling depth needed to accurately measure trends in diversity for microbial communities. Beta diversity has been found to be less sensitive to sequencing depth then alpha diversity and needs a sampling depth of 1,000 sequences to capture 90% of the diversity using the Bray-Curtis distance within a community (Lundin et al, 2010). While alpha diversity, using the Shannon Index which was the least sensitive metric when compared to Choa1 and Pielous needed at least 3,000 sequences per sample to capture 90% of the diversity (Lundin et al, 2010; Gihring et al, 2012). Choosing an even sampling depth can be difficult when there is a wide range of sequences per sample within the dataset. Most often researchers choose the lowest number of reads in order to retain as much sequencing information as possible, which is not always accurate. To determine adequate sample depth for this study the library size range was taken into account. Multiple different sample depths were implemented, choosing the smallest in order retain as much data as possible and a sequencing depth in the middle of the range in order to determine if the results are significantly different. Then the results were compared using ANOISM. This evaluates whether the chosen even sampling depth will bias the diversity metrics within the dataset. This allowed the selection of 1210 as the even sampling depth for this dataset, which produced accurate results.

Identifying a Possible Causative Agent:

To scientifically identify a causative agent a set of specific criteria for microbial pathogen identification was standardized in the 1870's by microbiology pioneer Robert Koch; Koch formulated his four Postulates as the definitive guideline for legitimate disease identification (Koch, 1876). To demonstrate pathogenicity, the pathogenic causative agent must be documented as always being found associated with the disease. The agent must then be isolated from the diseased state and grown in pure culture in the laboratory. The pure culture must then produce the disease in a healthy organism when introduced and the pathogen must be able to be re-isolated from the newly diseased organism (Boyd and Hoerl, 1981; Richardson, 1998). These postulates cannot be specifically adapted to microbial communities, because 90-99% of microbes cannot be cultured in laboratory environments (Amann et al, 1995). Thus, identifying a particular microbial gene needed by the

microorganism to produce pathogenesis and possibly identifying the microbial pathogen's ability can substitute for Koch's postulates and are thus known as "molecular" postulates (Falkow, 2004). The present research does not intend to verify a causative agent through completion of Koch's or "molecular" postulates.

RESULTS

Biom Summary:

Fifty one 16S rRNA amplicon libraries were sequenced, which were derived from 14 individual *X. muta* samples. Overall these samples provided 255,605 raw reads, which were later reduced to 228,315 quality checked sequences after de-noising and chimera removal. The sequence read dataset had a minimum sequence read count of 200 per sample, a maximum sequence read count of 14,597 per sample and a mean sequence read count of 4,477 per sample (Table 1). Due to the wide range of sequences counts per sample for the dataset an even sampling depth of 1210 was used to reduce the statistical difference caused by varying sampling depth. Choosing an even sampling depth that was larger than some of the sequences counts per sample meant that samples with less sequences would not be included in the analysis when the even sampling depth was utilized. The total amount of OTUs which were created once taxonomy was assigned to the sequence reads for all 51

Biom Summary for Sequence Count for *Xestospongia muta* Dataset					
Sample ID	**Sequence Count**	**Sample ID**	**Sequence Count**	**Sample ID**	**Sequence Count**
11.V.12.1.Water	200	29.V.12.1.8D	2565	11.V.12.1.4H	6190
11.V.12.1.5H	609	29.V.12.1.7H	3064	136H	7298
11.V.12.1.Soil	979	11.V.12.1.7I	3079	121H	7363
29.V.12.1.1H	1210	29.V.12.1.5I	3284	11.V.12.1.5I	7753
29.V.12.1.7D	1373	11.V.12.1.143D.I	3298	11.V.12.1.4I	8027
29.V.12.1.5D	1389	11.V.12.1.6H	3327	11.V.12.1.120D.I	8131
29.V.12.1.6H	1391	29.V.12.1.7I	3555	29.V.12.1.3H	8352
11.V.12.1.5D	1541	11.V.12.1.8I	3621	11.V.12.1.6D	9008
29.V.12.1.6D	1541	11.V.12.1.142D.D	3683	137H	10275
29.V.12.1.4H	1576	29.V.12.1.4I	4152	11.V.12.1.142D.I	10505
11.V.12.1.120D.H	1596	11.V.12.1.6I	4169	11.V.12.1.8D	10653
29.V.12.1.8H	1795	29.V.12.1.6I	4387	11.V.12.1.142D.H	12046
29.V.12.1.2H	1804	11.V.12.1.8H	4431	29.V.12.1.Soil	14597
11.V.12.1.7D	2023	11.V.12.1.9I	4724	Healthy Control	36302
11.V.12.1.7H	2221	11.V.12.1.9D	4971	Healthy on Diseased	47898
29.V.12.1.5H	2221	11.V.12.1.9H	4997	Boundary Layer	71094
11.V.12.1.120D.D	2383	11.V.12.1.4D	5016	Diseased	57245
29.V.12.1.8I	2409	29.V.12.1.4D	5119	Soil	15576
11.V.12.1.143D.H	2434	11.V.12.1.143D.D	5980	Water	200

Table 1: Sequence counts per sample for the dataset.

samples was 26,166. The maximum OTUs per sample was 2908, the minimum OTUs per sample was 67 and the mean OTUs per sample was 513 (Table 2).The D state had the greatest OTU counts, while BL had the greatest amount of sequence counts per health state. The BL population may therefore include more unknown bacteria then the other health states. The HC had the least amount of sequence counts and OTU counts compared the other health states.

Biom Summary for OTU Count for *Xestospongia muta* Dataset					
Sample ID	**OTU Count**	**Sample ID**	**OTU Count**	**Sample ID**	**OTU Count**
11.V.12.1.Water	67	29.V.12.1.6D	311	29.V.12.1.3H	617
11.V.12.1.5H	148	11.V.12.1.9H	314	11.V.12.1.120D.I	674
11.V.12.1.7H	163	136H	319	11.V.12.1.9I	686
11.V.12.1.143D.H	227	121H	321	29.V.12.1.4D	699
29.V.12.1.8D	228	11.V.12.1.5D	325	11.V.12.1.142D.D	705
11.V.12.1.8I	247	11.V.12.1.6I	341	11.V.12.1.142D.I	777
11.V.12.1.7I	250	29.V.12.1.8H	348	11.V.12.1.142D.H	817
11.V.12.1.6H	250	11.V.12.1.143D.I	360	11.V.12.1.4D	1011
11.V.12.1.120D.H	258	11.V.12.1.4I	379	11.V.12.1.9D	1053
29.V.12.1.7D	269	29.V.12.1.5I	381	11.V.12.1.6D	1093
11.V.12.1.8H	271	29.V.12.1.7H	402	11.V.12.1.143D.D	1284
29.V.12.1.5D	277	29.V.12.1.4I	411	11.V.12.1.8D	1301
29.V.12.1.4H	278	29.V.12.1.7I	416	29.V.12.1.Soil	2908
11.V.12.1.4H	290	29.V.12.1.5H	440	Healthy Control	2160
29.V.12.1.8I	291	29.V.12.1.6I	464	Healthy on Diseased	4500
29.V.12.1.1H	293	11.V.12.1.5I	514	Boundary Layer	6191
29.V.12.1.6H	294	11.V.12.1.7D	579	Diseased	9725
29.V.12.1.2H	307	11.V.12.1.120D.D	590	Soil	3216
11.V.12.1.Soil	308	137H	610	Water	67

Table 2: OTU counts per sample for the dataset.

Alpha Diversity:

The rarefaction plots for the Chao1 statistic compared the differences in species richness between the health states and the effect of the different sampling locations on the health states (Figure 6). When the species richness between the different health states was compared, the D state had the greatest amount of species richness (Figure 6, Graph A). The diversity of the whole D population was compared against the two subpopulations from the sampling locations, the D Boynton population was overall more diverse then the rest of the health states (Figure 6, Graph C). The D Broward population did not have comparable diversity levels and was significantly less diverse; it was surpassed in diversity by the HoD Broward population. The order of species richness from greatest to least for the different health states were D Boynton > HoD Broward > D Broward > HC Broward >

BL Broward > BL Boynton > HoD Boynton > HC Boynton. When the diversity between the health states were compared the results helped to highlight the differences that also existed between the HoD and the HC subpopulations from the separate sampling locations; the Broward subpopulations (HC and HoD) were noticeably more diverse then the Boynton subpopulations (HC and HoD). To find significant differences between the species richness values for each health state (Chao1 statistic) a nonparametric two sample t-test (Monte Carlo) was utilized (Figure 7). The boxplots created from the two sample t- test displayed the differences in diversity for all the samples which comprised each health state. The D boxplot illustrated that the samples which comprised the population had more variety in the diversity levels; since the box was positioned higher and did not overlap with the other boxplots (Figure 7, Boxplot A). The BL boxplot displayed the exact opposite of the D boxplot with a short, condensed box which showed that species richness within the samples that comprised the community were more similar than any other health state. When the effect of sampling location on diversity was evaluated the samples that comprised the Boynton sampling site also displayed a greater range of diversity compared to the Broward sampling site (Figure 7, Boxplot B). The statistics produced from the two sample t-test

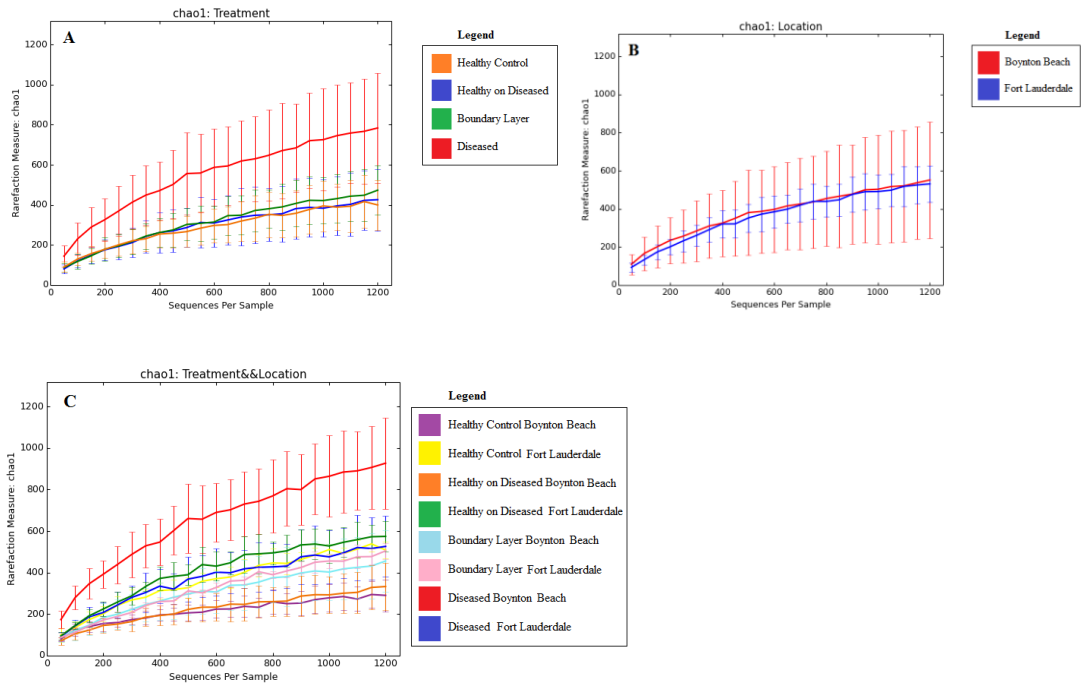
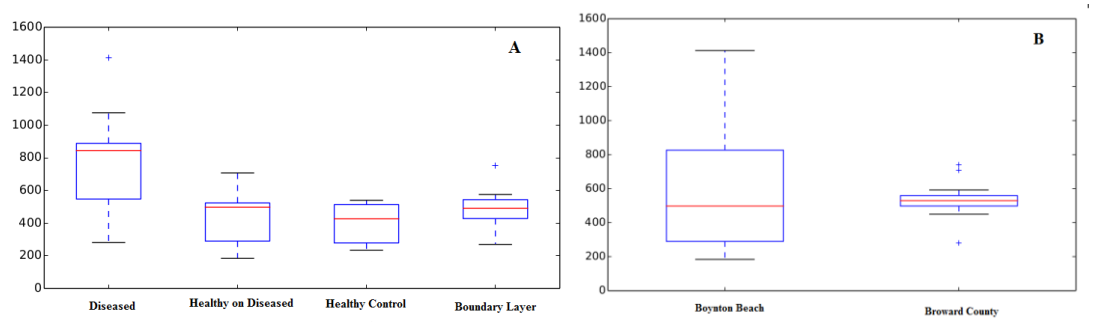


Figure 6: Rarefaction plots displaying Chao1 statistic. Graph A: Displays Chao 1 by separate health states, Graph B: Displays Chao 1 by location and Graph C: Displays Chao 1 within the health states separated by location. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

highlighted the presence of significant differences between the species richness for the D health state when compared to either the BL or the HoD state (P-value < 0.05) (Figure 7). This supported the presence of a relationship between the progression of Sponge Orange Band Disease and the altered species richness found within diseased sponge mesohyl. No significant differences were found between the Boynton Beach and Fort Lauderdale sampling locations (p-value > 0.005) (Figure 7).



Two Sample T-Test (Monte Carlo) using Chao1 for Comparison							
Group1	**Group2**	**Group1 mean**	**Group1 std**	**Group2 mean**	**Group2 std**	**t stat**	**p-value**
Diseased	Healthy Control	36.94	9.29	25.48	3.36	2.79	0.13
Healthy on Diseased	Healthy Control	24.36	3.12	25.48	3.36	-0.67	1.00
Healthy on Diseased	Boundary Layer	24.36	3.12	25.06	3.24	-0.55	1.00
Diseased	Boundary Layer	36.94	9.29	25.06	3.24	4.36	0.006
Healthy on Diseased	Diseased	24.36	3.12	36.94	9.29	-4.47	0.01
Boundary Layer	Healthy Control	25.06	3.24	25.48	3.36	-0.25	1.00
Boynton	Fort Lauderdale	551.16	304.23	530.81	94.99	0.27	0.78

Figure 7: Nonparametric t-tests (Monte Carlo) using Choa1 statistics between the different health states in Boxplot A and between the different sampling locations in Boxplot B. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

The rarefaction analysis using ACE also showed that the diversity levels within the separate D subpopulations differed (Figure 8, Graph A). The diversity between the HoD and the HC subpopulations also contrasted due to location; the Boynton populations were found to be the least diverse, while the Broward populations were closely grouped around the Broward D state (Figure 8, Graph C). The diversity richness trends were similar to the results previously seen using the Chao1 statistic, which supported the diversity estimates produced by the separate health states and showed that they were accurate and reproducible. When the ACE statistics were compared with the nonparametric two sample t-test (Monte Carlo) the results closely mirrored the Chao1 results. The only difference between the t-tests was that ACE found the D health state and the HC state to have significant differences between the levels of diversity, instead of just the D population against HoD and BL

(Figure 9, Graph A). This supports that SOB Disease influenced the changes found in diversity within all the mesohyl samples taken from diseased sponges.

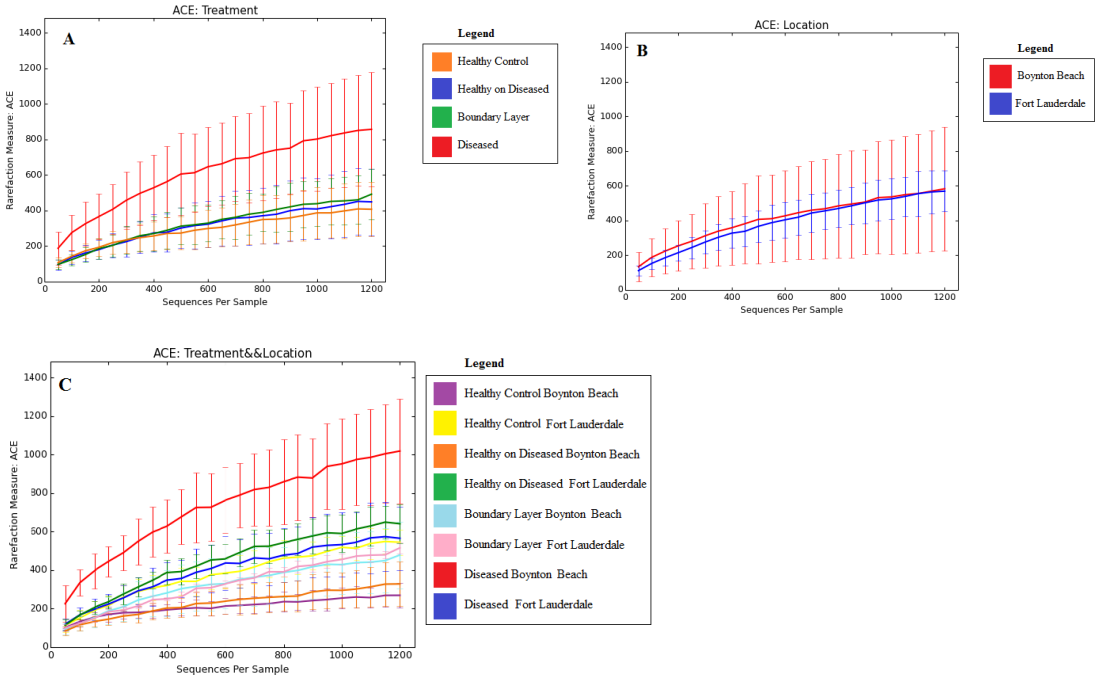
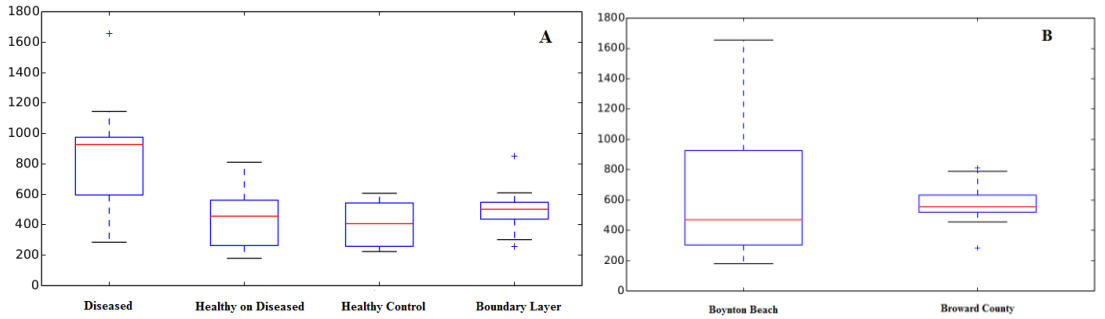


Figure 8: Rarefaction plots displaying the ACE statistic. Graph A: Displays ACE within the separate health states, Graph B: Displays ACE by the separate locations and Graph C: Displays ACE by the health states and by location. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

The rarefaction plots displaying the Shannon Index of Diversity compared the separate health states and the effect of location on the health states (Figure 10). The diversity trends previously seen within Chao 1 and ACE for the health states still held true, and the only difference was the increased evenness between the populations (Figure 10, Graph A). The D population was no longer overwhelmingly diverse when compared to the other health states. The order of diversity from the greatest to least for the Shannon Index of Diversity was: D Boynton > HC Boynton > HoD Broward > HC Broward > D Broward > HoD Boynton > BL Boynton > BL Broward. This order changed the degree diversity between the D populations by putting more distance between them, while making the HoD and HC populations more similar to each other (Figure 10, Graph C). The comparison of the Shannon Index of Diversity using the nonparametric two sample t-test (Monte Carlo) showed more variation between the diversity present within the samples at the Broward sampling site but not enough to make the sampling sites significantly different (Figure 11,

Boxplot B). The statistics for the health states display that the D state was significantly different from both the BL and the HoD state.



Two Sample T-Test (Monte Carlo) using ACE for Comparison Between the Health States							
Group1	**Group2**	**Group1 mean**	**Group1 std**	**Group2 mean**	**Group2 std**	**t stat**	**p-value**
Diseased	Healthy Control	857.06	322.73	407.36	152.00	3.09	0.03
Healthy on Diseased	Healthy Control	448.94	187.49	407.36	152.00	0.45	1.00
Healthy on Diseased	Boundary Layer	448.94	187.49	492.32	140.82	-0.66	1.00
Diseased	Boundary Layer	857.06	322.73	492.32	140.82	3.73	0.01
Healthy on Diseased	Diseased	448.94	187.49	857.06	322.73	-3.83	0.01
Boundary Layer	Healthy Control	492.32	140.82	407.36	152.00	1.15	1.00
Boynton	Fort Lauderdale	583.45	355.64	569.53	116.83	0.16	0.88

Figure 9: Nonparametric t-tests (Monte Carlo) using ACE statistics between the different health states in Boxplot A and between the different sampling locations in Boxplot B. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

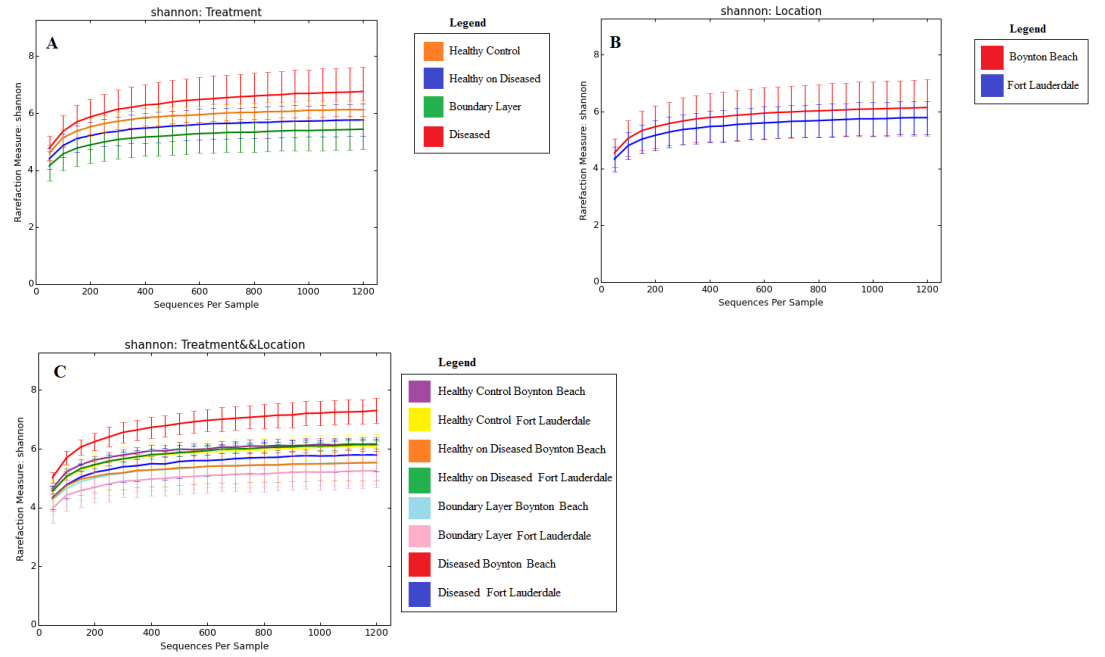
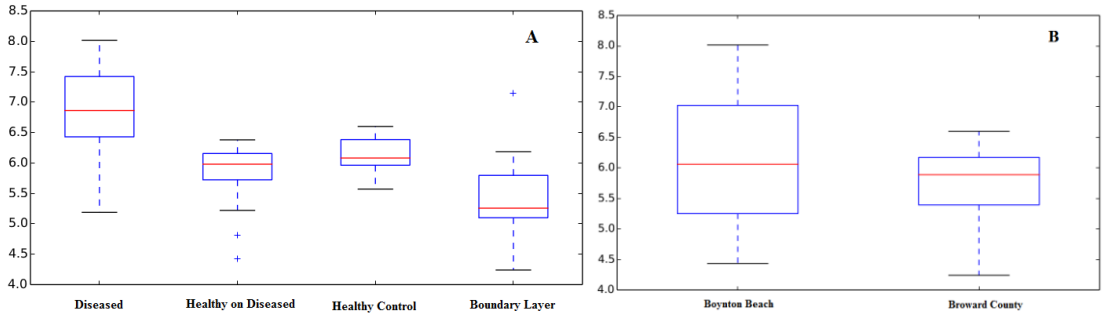


Figure 10: Shannon Diversity Index rarefaction plots. Graph A: between the separate health states, Graph B: by location and Graph C: displays by health states and by location. Plots were created using QIIME 1.8.0 (Caporaso, 2010).



Two Sample T-Test (Monte Carlo) using the Shannon Diversity Index for Comparison Between the Health States							
Group1	**Group2**	**Group1 mean**	**Group1 std**	**Group2 mean**	**Group2 std**	**t stat**	**p-value**
Diseased	Healthy Control	6.76	0.87	6.12	0.34	1.66	0.74
Healthy on Diseased	Healthy Control	5.76	0.57	6.12	0.34	-1.36	1.00
Healthy on Diseased	Boundary Layer	5.76	0.57	5.44	0.70	1.26	1.00
Diseased	Boundary Layer	6.76	0.87	5.44	0.70	4.28	0.01
Healthy on Diseased	Diseased	5.76	0.57	6.76	0.87	-3.39	0.01
Boundary Layer	Healthy Control	5.44	0.70	6.12	0.34	-2.16	0.23
Boynton	Fort Lauderdale	6.15	0.99	5.79	0.58	1.37	0.19

Figure 11: Nonparametric t-tests (Monte Carlo) using the Shannon Index of Diversity statistics between the different health states in Boxplot A and between the different sampling locations in Boxplot B. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

The rarefaction curves for phylogenetic diversity compared the taxon richness found within each health state, location and health state divided by location (Figure 12). The D Boynton population was the most phylogenetically diverse compared the other health states and there was also a distinct difference in how diverse the D states were from each other (Figure 12, Graph C). For instance the diversity within D Broward population was similar with the diversity found in the HC Broward and the HoD Broward. The nonparametric two sample t-test (Monte Carlo) was used to identify significant phylogenetic relationships. The D population overall had more sample variety compared to the other health state populations, which was shown by the elongated boxplot (Figure 13, Boxplot A). Statistically the D state also had significant differences in phylogenetic diversity when compared to both the BL and the HoD state. Which supported the progression of SOB to reductions and changes in phylogenetic diversity found within *X. muta*.

Overall the alpha diversity metrics have shown that D state was more diverse then the other health states and was significantly different from both the BL and HoD states. Significant differences in species richness where also found between sampling locations when the Choa1 and the ACE statistics were used but was not found when using the

Shannon Index of Diversity. When the phylogenetic diversity was compared between the health states the D state was more diverse than the others. The D state was also found phylogenetically significantly different when compared to the BL and the HoD populations.

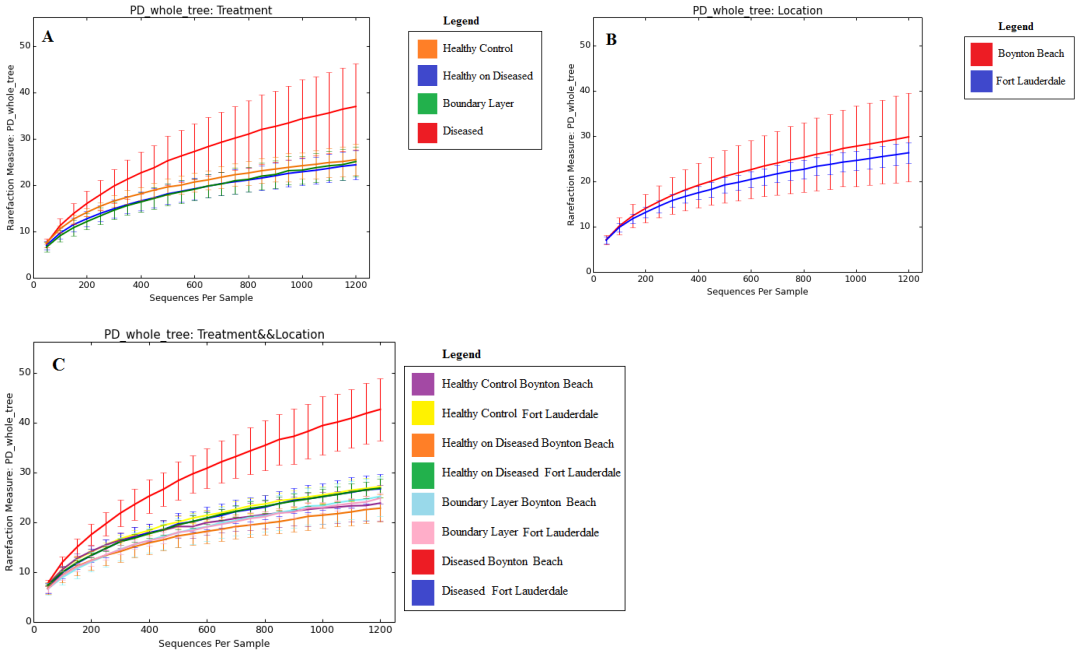
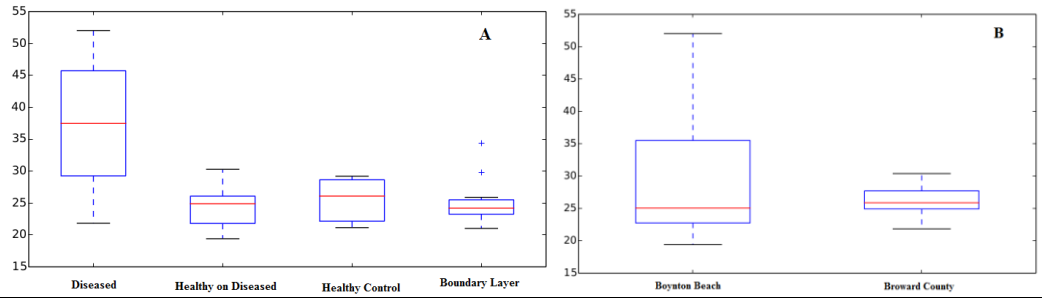


Figure 12: Rarefaction plots displaying phylogenetic diversity. Graph A: Phylogenetic diversity by separate health states, Graph B: Displays Phylogenetic Diversity by location and Graph C: Displays Phylogenetic Diversity by health states and by location. Plots were created using QIIME 1.8.0 (Caporaso, 2010).



Two Sample T-Test (Monte Carlo) using Phylogenetic Diversity Whole Tree Statistics for Comparison Between the Health States

Group1	Group2	Group1 mean	Group1 std	Group2 mean	Group2 std	t stat	p-value
Diseased	Healthy Control	36.94	9.29	25.48	3.36	2.79	0.13
Healthy on Diseased	Healthy Control	24.36	3.12	25.48	3.36	-0.67	1.00
Healthy on Diseased	Boundary Layer	24.36	3.12	25.06	3.24	-0.55	1.00
Diseased	Boundary Layer	36.94	9.29	25.06	3.24	4.36	0.01
Healthy on Diseased	Diseased	24.36	3.12	36.94	9.29	-4.47	0.01
Boundary Layer	Healthy Control	25.06	3.24	25.48	3.36	-0.25	1.00
Boynton	Fort Lauderdale	29.80	9.75	26.29	2.29	1.47	0.15

Figure 13: Nonparametric t-tests (Monte Carlo) using the Phylogenetic Diversity Whole Tree statistics between the different health states in Boxplot A and between the different sampling locations in Boxplot B. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

Beta Diversity:

The Jaccard Abundance Index PCoA plots displayed the presence of microbial community compositional similarities by the grouping and overlapping of multiple health states (Figure 14, Graph A). The plots also presented a visible distribution of samples and the absence of densely clustered individuals, which showed the presence of community variation within the health states. Dissimilarity within the samples that comprised a health state increased as SOB Disease progressed. This was best displayed in the BL and the D populations where the individuals from those populations become more dispersed and distributed (Figure 14, Graph D and E). The diseased population specifically contained seven samples which closely clustered together and away from the rest of the dataset. These seven outlier samples were all from the Boynton Beach sampling site (11.V.12.1.120D.D, 11.V.12.1.9D, 11.V.12.1.8D, 11.V.12.1.7D, 11.V.12.1.6D, 11.V.12.1.5D, and 11.V.12.1.4D), but this grouping did

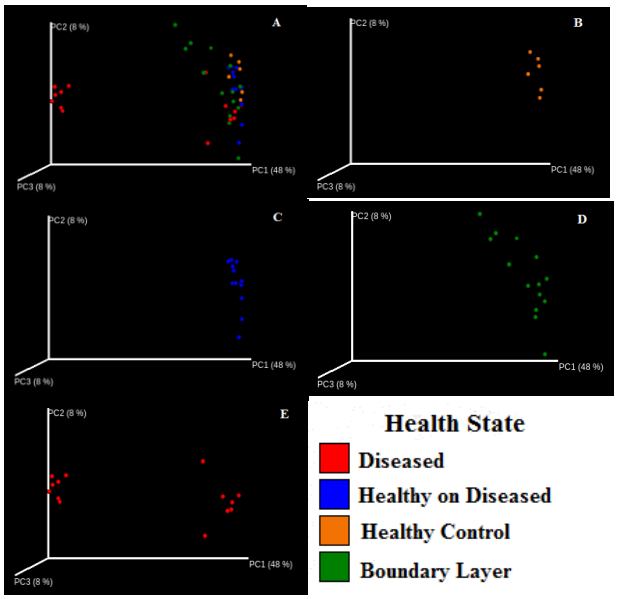


Figure 14: Jaccard Abundance plots separated by health state. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

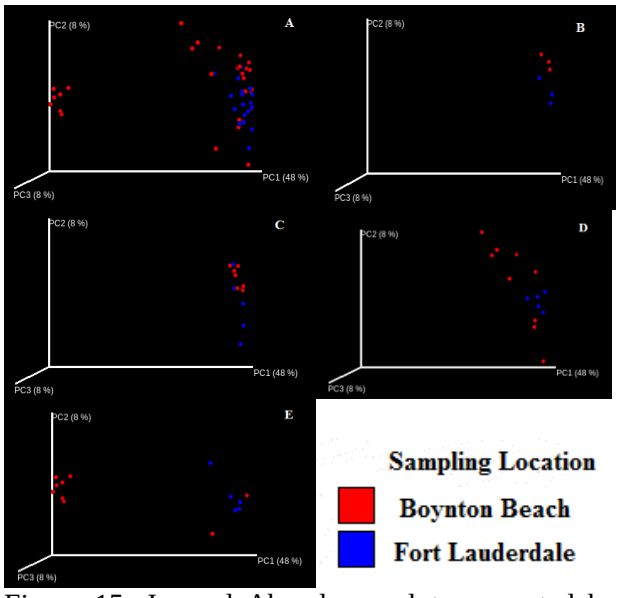


Figure 15: Jaccard Abundance plots separated by sampling location. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

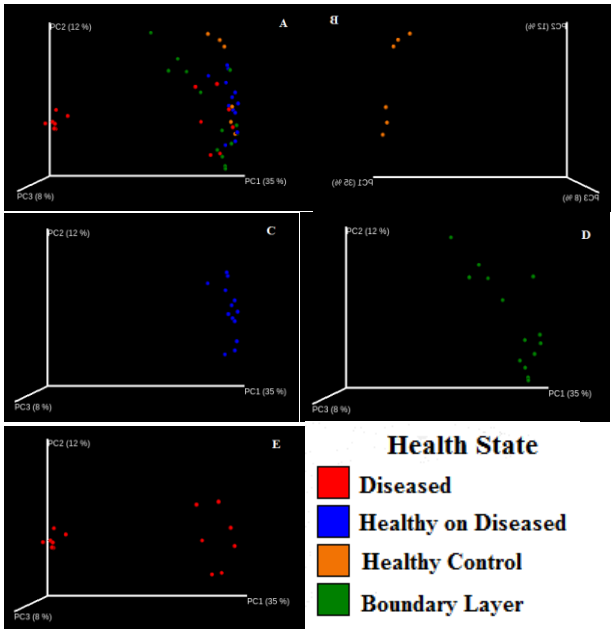


Figure 16: Bray Curtis Dissimilarity plots separated by health state. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME

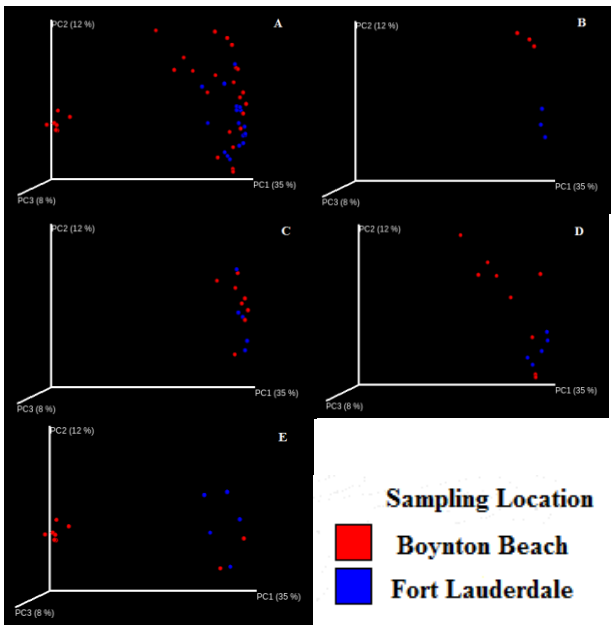


Figure 17: Bray Curtis dissimilarity plots separated by sampling location. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

not include every diseased sample taken from Boynton site; with the two remaining individuals clustered with the D samples from Fort Lauderdale (Figure 15, Graph E). The influence of location on community composition was not a defining factor for most health states depicted in the PCoA plots. There was minimal clustering within the HC population and within the Fort Lauderdale subpopulation of the BL (Figure 15, Graph A and D). The other sampling site subpopulations had no clear clustering due to sampling location.

The Bray-Curtis PCoA plots displayed the presence of more dissimilarity between health states than the Jaccard Abundance Index; as the samples were displayed in a more dispersed pattern from each other (Figure 16, Graph A). The HoD state (Graph C) was the only health state to retain a similar distribution pattern seen previously in the Jaccard Abundance analysis. The HC (Graph B) and the B layer communities (Graph D) had more

variation with the samples that comprised the two communities then previously seen. The HC state also showed a greater division in similarity within subpopulations based on location, which supported the concept that sampling location can influence the microbial community (Figure 17, Graph B).

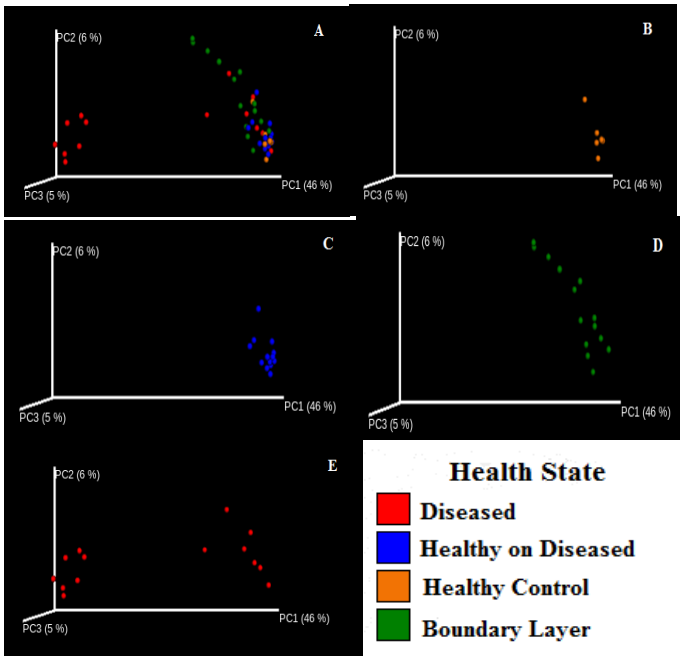
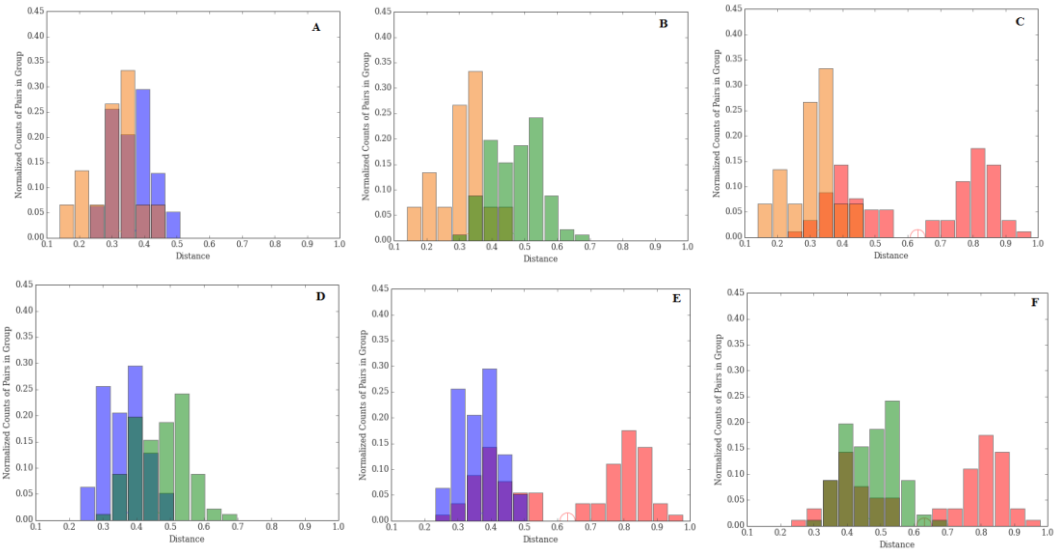


Figure 18: Unweighted UniFrac PCoA plots separated by health state. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

Within the unweighted UniFrac PCoA plots both the HC and the HoD populations were closely clustered to samples within their own health state (Figure 18, Graphs B and C). This showed that the samples were phylogenetically more similar to members of the same health state than other populations. The D and BL populations were comparatively more dispersed, which displayed that the samples comprising the populations were more phylogenetically diverse (Figure 18, Graphs D and E). There were still seven D outlier samples which clustered together away from the majority of the samples, which showed that those samples were phylogenetically distinct from the rest of the D samples and the entire dataset To determine significant relationships within the dataset and obtain further statistical support for the beta diversity estimates the UniFrac significance test was used, which compared the group distances between the health states. When the separate health states were compared to each other they all produced a p-value < 0.05, which showed that there was significant difference within the phylogenetic lineages that comprised the microbial composition of the different health states. When considering the degree of difference, the p-value produced during comparisons for the most significant difference was between the HC and the BL populations. While the HC and the HoD had the least

significant difference with the highest p-value which revealed that those two health states were the most similar (Figure 19).



Group Distances for Health State Pairwise Comparisions									
Category_1a	**Category_1b**	**Avg**	**Category_2a**	**Category_2b**	**Avg**	**t**	**p**	**p_greater**	**p_less**
Healthy on Diseased	Healthy on Diseased	0.37	Healthy Control	Healthy Control	0.32	3.00	3.50E-03	0	1
Healthy on Diseased	Healthy on Diseased	0.37	Boundary Layer	Boundary Layer	0.47	-10.14	3.95E-19	1	0
Healthy on Diseased	Healthy on Diseased	0.37	Diseased	Diseased	0.63	-10.80	5.86E-21	1	0
Boundary Layer	Boundary Layer	0.47	Healthy Control	Healthy Control	0.32	7.57	1.60E-11	0	1
Diseased	Diseased	0.63	Healthy Control	Healthy Control	0.32	5.77	8.31E-08	0	1
Diseased	Diseased	0.63	Boundary Layer	Boundary Layer	0.47	6.69	2.78E-10	0	1

Figure 19: UniFrac Significance Test performed on unweighted UniFrac beta diversity measure for the separate health states. Each health state was assigned a color, healthy control state is orange, the healthy on diseased state is blue, the boundary layer state is green and the diseased state is red. The different histograms assess the significance between two groups. Graph A displays the healthy control in orange, the healthy on diseased in blue and the purple is the overlap between the health states. Graph B displays the healthy control in orange the boundary layer in green and the darker green is the overlap between the health states. Graph C displays the healthy control in orange, the diseased in red and the darker orange is the overlap between the health states. Graph D displays the healthy on diseased in blue, the boundary layer in green and the teal is the overlap between the health states. Graph E displays the healthy on diseased in blue, the diseased in red and the purple is the overlap between the health states. Graph F displays the boundary layer in green, the diseased in red and the dark green is the overlap between the health states. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

Comparing the health state subpopulations by sampling location allowed for an evaluation of whether location phylogenetically influenced the microbial communities (Figure 20). The health state subpopulations displayed within the PCoA plots were overall not densely clustered by location. The intermixing of samples from both subpopulations meant that the health states shared phylogenetic similarity between microbial communities regardless of location (Figure 20, Graph C). This relationship was often seen when

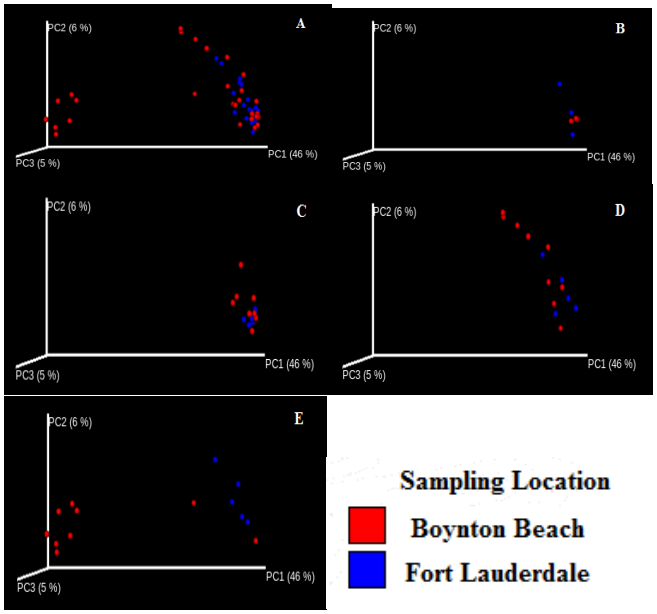
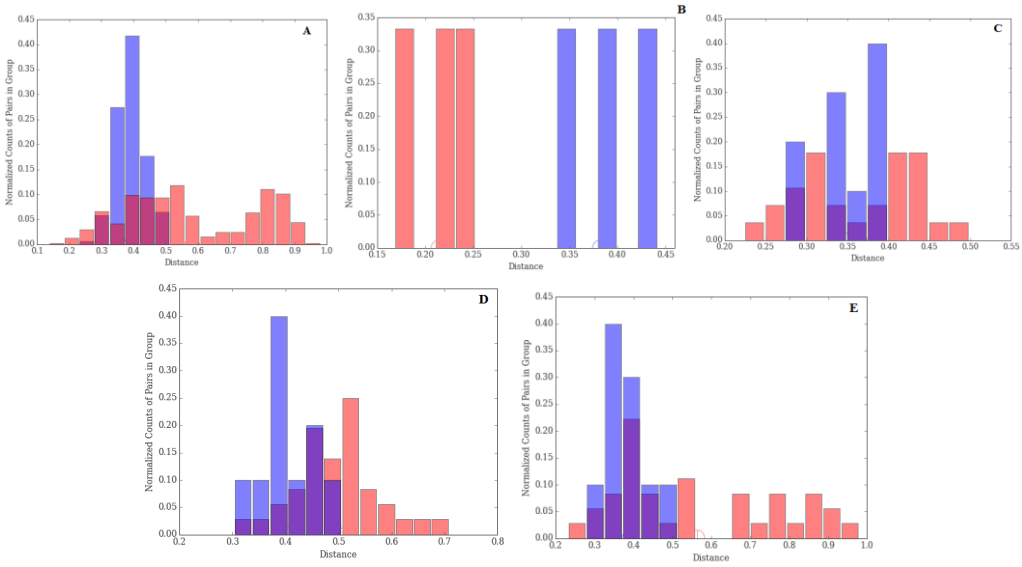


Figure 20: Unweighted UniFrac PCoA plots separated by sampling location. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

environments were similar between populations and few habitat exclusive adaptions were needed to survive at either location (Lozupone et al, 2005). The UniFrac significance test for sampling location showed that the HoD population was not significantly different due to location with a p-value > 0.05 (Figure 21, Graph C). The HC population was significantly different due to location with a p-value < 0.05 (Figure 21, Graph B). The BL and D populations were also significantly different by location with p-values < 0.05 (Figure 21, Graph D and E). Having significant differences between all of the health states from diseased sponges (HoD, BL and D) because of location supports the idea that microbial lineages that comprise the communities are phylogenetically influenced by the environment they live in.

The jackknife PCoA plots for the BL population had larger and more frequent IQRs around the samples that comprised the population which revealed that there was more variability within the dataset compared to the other health states (Figure 22, Graph D). The presence of large IQRs within the BL population may be due the fact that the mesohyl sample naturally included both HoD and D communities, which increased the amount of diversity present within the samples. The rest of the health states had relatively small IQR which helped to support the validity of the results for the presence and absence of microbial lineages (Figure 22).



Group Distances for Location Pairwise Comparisons										
Category_1a	**Category_1b**	**Avg**	**Category_2a**	**Category_2b**	**Avg**	**t**	**p**	**p_greater**	**p_less**	**Health State**
Boynton	Boynton	0.58	Fort Lauderdale	Fort Lauderdale	0.39	11.44	2.33E-27	0	1	All
Boynton	Boynton	0.21	Fort Lauderdale	Fort Lauderdale	0.38	-5.98	0	0.99	0	Healthy Control
Boynton	Boynton	0.37	Fort Lauderdale	Fort Lauderdale	0.35	0.63	0.54	0.25	0.75	Healthy
Boynton	Boynton	0.50	Fort Lauderdale	Fort Lauderdale	0.41	3.49	0	0	1	Boundary Layer
Boynton	Boynton	0.56	Fort Lauderdale	Fort Lauderdale	0.39	2.60	0.01	0	1.00	Diseased

Figure 21: Unifrac Significance Test on unweighted Unifrac beta diversity measures for sampling location. Each sampling location was assigned a color, Boynton Beach is red, Fort Lauderdale is blue and any overlap between the communities is in purple. Graph A compares all the samples from Boynton Beach in red and all the samples from Fort Lauderdale in blue. Graph B compares the healthy control samples from Boynton Beach in red to the healthy control samples from Fort Lauderdale in blue. Graph C compares the healthy on diseased samples from Boynton Beach in red to the healthy on diseased samples from Fort Lauderdale in blue. Graph D compares the boundary layer samples from Boynton Beach in red to the boundary layer samples from Fort Lauderdale in blue. Graph E compares the diseased samples from Boynton Beach in red to the diseased samples from Fort Lauderdale in blue. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

The weighted UniFrac PCoA plots displayed wide spread sample intermixing between the majority of the different health states, which showed that each health state had relatively similar microbial relative abundances (Figure 23, Graph A). Some congruency within the microbial compositions was expected as three health states (HoD, BL and D) were all samples from the same fourteen diseased sponges. Sampling the same diseased individual at different disease stages was done in order to possibly show the disease progression within an individual. Each health state therefore would share to some extent the same core microbiome of that individual and only differ by bacteria unique to the disease progression. Even the HC samples which were taken from non-diseased sponges at both sampling sites were spatially intermixed with the D samples on the PCoA plots (Figure 23). *X. muta* is known to pass on beneficial bacteria to its offspring, so it was

expected that the species would have similar microbial communities even in populations that are separated by location (Montalvo and Hill, 2011). While the majority of samples within the separate health states were not widely segregated from each other there was a clustering of seven D samples from the Boynton Beach sampling location (11.V.12.1.120D.D, 11.V.12.1.9D, 11.V.12.1.8D, 11.V.12.1.7D, 11.V.12.1.6D, 11.V.12.1.5D, and 11.V.12.1.4D) (Figure 23, Graph E). Two more D samples from Boynton Beach and the entire Broward sampling population may not have clustered to the outliers due to the timing of the sampling and the natural progression from initial infection to secondary and opportunistic microbes taking advantage of an already compromised host. To better understand the relationships between the health states and location, the two sample t-test

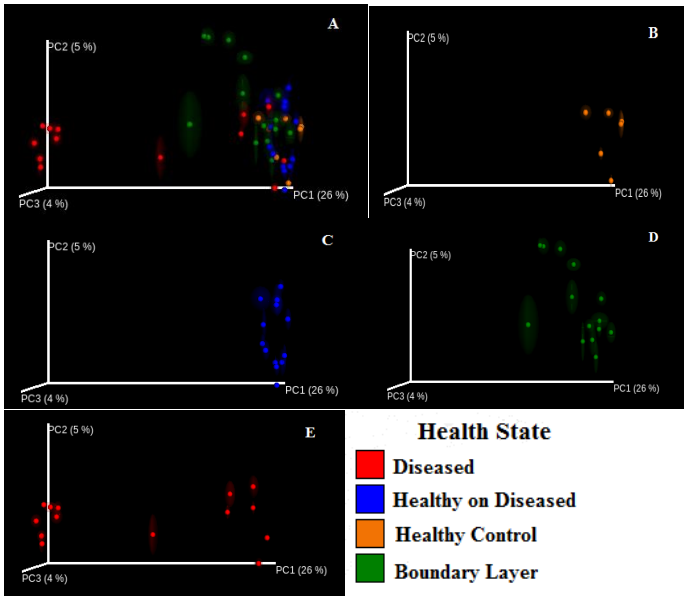


Figure 22: Jackknife PCoA plots for unweighted UniFrac results. Graph A includes all the health states, Graph B displays only the Healthy Control population, Graph C includes only the healthy on diseased population, Graph D includes only the boundary layer population, and Graph E includes only the diseased population. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

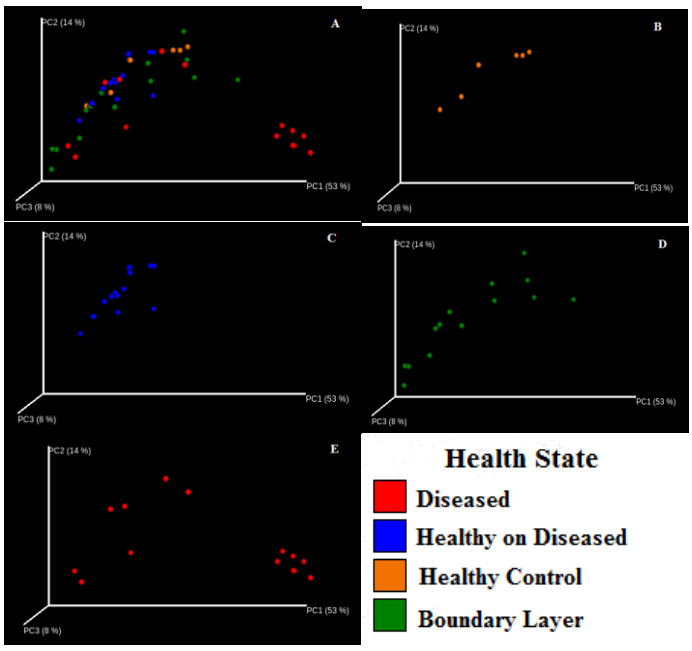
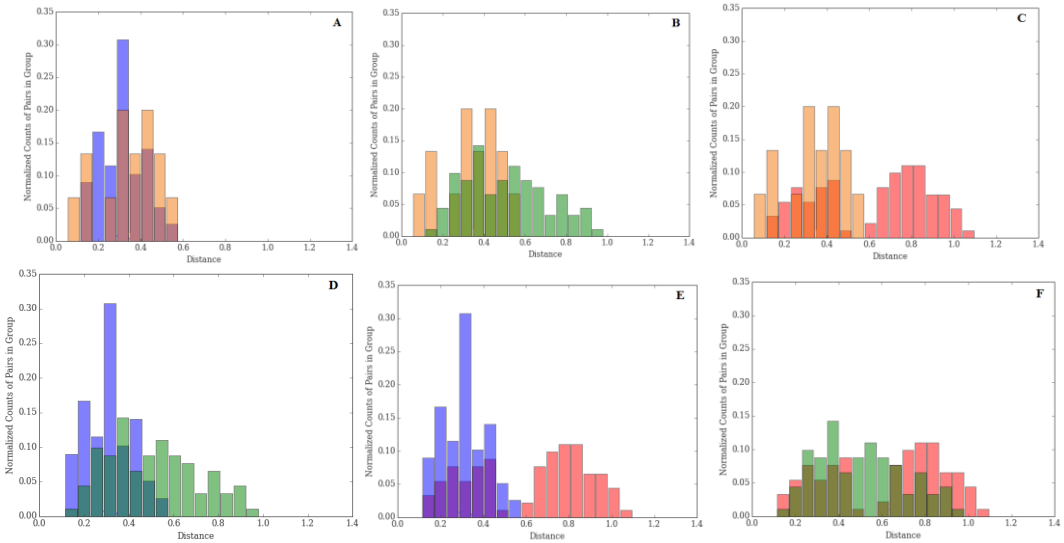


Figure 23: Weighted UniFrac PCoA plots separated by health state. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

(Monte Carlo) was used to determine significant differences. The taxonomic relative abundance differences between the HC and the HoD were not significant (p-value > 0.05) [Figure 24, Graph A]. Therefore the HC and the HoD population had similar microbial community lineages. When the other health states were compared, the relationships between the taxon relative abundances proved to be significantly different. This was expected since both the BL and the D populations showed an increase in diversity and a reduced similarity of their microbial lineages comparatively (Figure 24).



Group Distances for Health State Pairwise Comparisons									
Category_1a	**Category_1b**	**Avg**	**Category_2a**	**Category_2b**	**Avg**	**t**	**p**	**p_greater**	**p_less**
Healthy on Diseased	Healthy on Diseased	0.31	Healthy Control	Healthy Control	0.35	-1.15	0.25	0.89	0.11
Healthy on Diseased	Healthy on Diseased	0.31	Boundary Layer	Boundary Layer	0.51	-7.75	8.51E-13	1	0
Healthy on Diseased	Healthy on Diseased	0.31	Diseased	Diseased	0.62	-9.61	1.12E-17	1	0
Boundary Layer	Boundary Layer	0.51	Healthy Control	Healthy Control	0.35	2.98	3.59E-03	0	1
Diseased	Diseased	0.62	Healthy Control	Healthy Control	0.35	3.88	1.82E-04	0	1
Diseased	Diseased	0.62	Boundary Layer	Boundary Layer	0.51	3.23	1.49E-03	0.01	0.99

Figure 24: UniFrac Significance Test on the weighted UniFrac beta diversity measures. Each health state was assigned a color, healthy control state is orange, the healthy on diseased state is blue, the boundary layer state is green and the diseased state is red. The different histograms assess the significance between two groups. Graph A displays the healthy control in orange, the healthy on diseased in blue and the purple is the overlap between the health states. Graph B displays the healthy control in orange the boundary layer in green and the darker green is the overlap between the health states. Graph C displays the healthy control in orange, the diseased in red and the darker orange is the overlap between the health states. Graph D displays the healthy on diseased in blue, the boundary layer in green and the teal is the overlap between the health states. Graph E displays the healthy in blue, the diseased in red and the purple is the overlap between the health states. Graph F displays the boundary layer in green, the diseased in red and the dark green is the overlap between the health states. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

When comparing the effects of sampling location on the health states, the PCoA plots displayed the HC subpopulations clustered accordingly by location, which

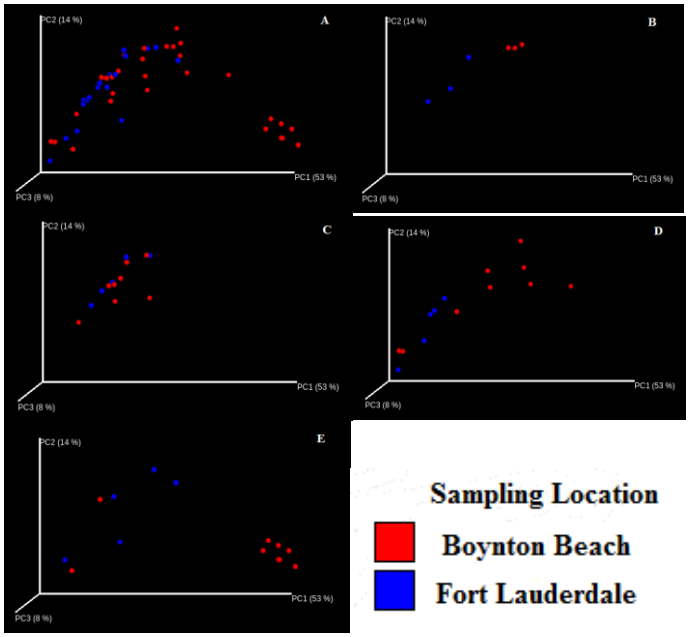


Figure 25: Weighted Unifrac PCoA plots separated by sampling location. Plot A includes all health states, plot B displays the healthy control community, plot C displays the healthy on diseased community, plot D is the boundary layer community and plot E is the diseased community. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

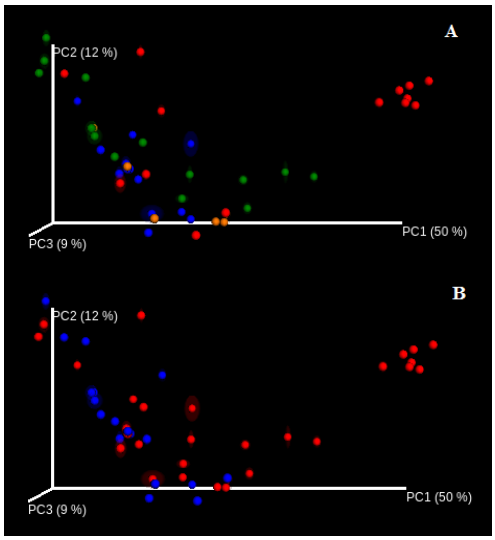
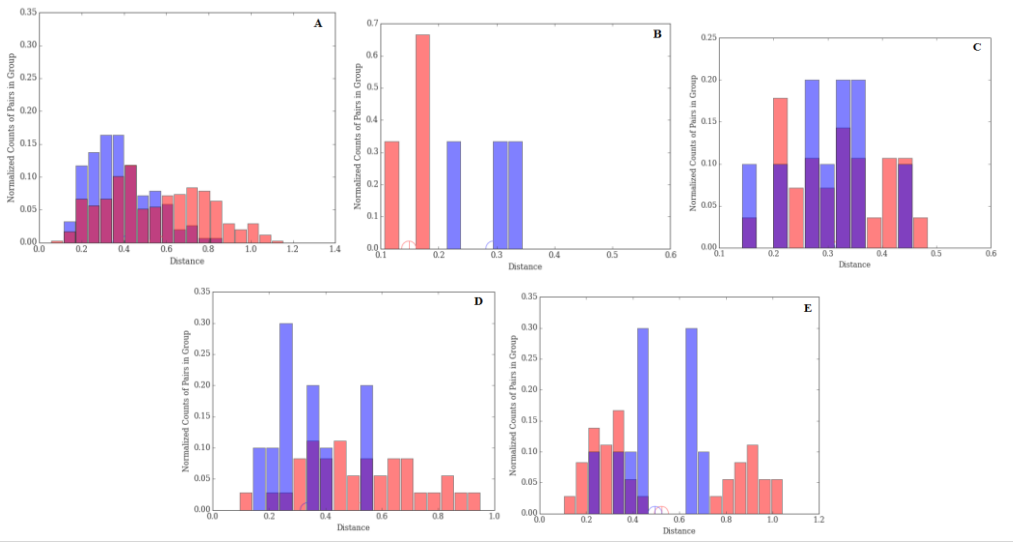


Figure 27: Jackknife PCoA plots for weighted Unifrac results. Graph A displays all samples defined by health state (red = Diseased, green = Boundary Layer, blue = Healthy on Diseased and yellow = Healthy Control), while Graph B displays all the samples by sampling location (red = Boynton Beach and blue = Fort Lauderdale). Plots were created using QIIME 1.8.0 (Caporaso, 2010).

demonstrated that the two subpopulation differed in microbial composition (Figure 25, Graph B). The two sample t-test (Monte Carlo) confirmed that the HC subpopulations were significantly different then compared by sampling location, with a p-value < 0.05 (Figure 26, Graph B). The D population was not significantly different with a p-value > 0.05, while the BL population was significantly different with a p-value < 0.05 (Figure 26, Graph D and E). The jackknife PCoA plots for the weighted UniFrac statistics showed small to non-exist IQRs within the dataset (Figure 27). This helped to validate the relative abundance relationships visually and statistically present between the different health states.

Overall beta diversity was examined by multiple statistics. The PCoA plots created from the Jaccard Abundance Index depicted a level of congruency for shared species between the health states as samples for each state were intermixed together rather than segregated. Seven outlier samples from the D dataset were

present from the Boynton Beach sampling site and were the only samples to routinely cluster away within the PCoA plots from the rest of the samples. The Bray Curtis Index established the presence of more dissimilarity between samples and found that the HC subpopulations differed due to location. Phylogenetic analysis of the microbial lineages completed by the unweighted UniFrac revealed that all the health states were significantly different from each other. The influence of location was also significant for the HC, BL and the D health states. The relative abundance changes due the presence and absence of branches within the phylogenetic tree (weighted UniFrac) showed that all the health states were significantly different when compared to each other, except the HC against HoD. The HC and the BL microbial communities were the only subpopulations, which were significantly influenced by location.



Group Distances for Location Pairwise Comparisons

Category_1a	Category_1b	Avg	Category_2a	Category_2b	Avg	t	p	p_greater	p_less	Health State
Boynton	Boynton	0.56	Fort Lauderdale	Fort Lauderdale	0.38	8.55	1.21E-16	0	1	All
Boynton	Boynton	0.15	Fort Lauderdale	Fort Lauderdale	0.29	-3.79	0.02	0.94	0.05	Healthy Control
Boynton	Boynton	0.32	Fort Lauderdale	Fort Lauderdale	0.30	0.59	0.56	0.28	0.72	Healthy on Diseased
Boynton	Boynton	0.53	Fort Lauderdale	Fort Lauderdale	0.34	2.89	0.01	0.00	1.00	Boundary Layer
Boynton	Boynton	0.52	Fort Lauderdale	Fort Lauderdale	0.50	0.27	0.79	0.41	0.59	Diseased

Figure 26: UniFrac Significance Test on the weighted UniFrac beta diversity measures for sampling location. Each sampling location was assigned a color, Boynton Beach is red, Fort Lauderdale is blue and any overlap between the communities is in purple. Graph A compares all the samples from Boynton Beach in red and all the samples from Fort Lauderdale in blue. Graph B compares the healthy control samples from Fort Lauderdale in red to the healthy control samples from Fort Lauderdale in blue. Graph C compares the healthy on diseased samples from Boynton Beach in red to the healthy on diseased samples from Fort Lauderdale in blue. Graph D compares the boundary layer samples from Boynton Beach in red to the boundary layer samples from Fort Lauderdale in blue. Graph E compares the diseased samples from Boynton Beach in red to the diseased samples from Fort Lauderdale in blue. Plots and statistics were created using QIIME 1.8.0 (Caporaso, 2010).

Distance Boxplots:

The distances present within the separate health states helped to convey the nature and relationship of the samples that comprised them (Figure 28). The HC and the HoD boxplots were comparatively shorter than the others, which suggested similarity within the microbial composition of the separate treatments. While the D and BL boxplots were proportionally taller, which advocated that the samples comprising the separate health states had more variation. The D microbial composition was also significantly different from the other treatments since the plot was

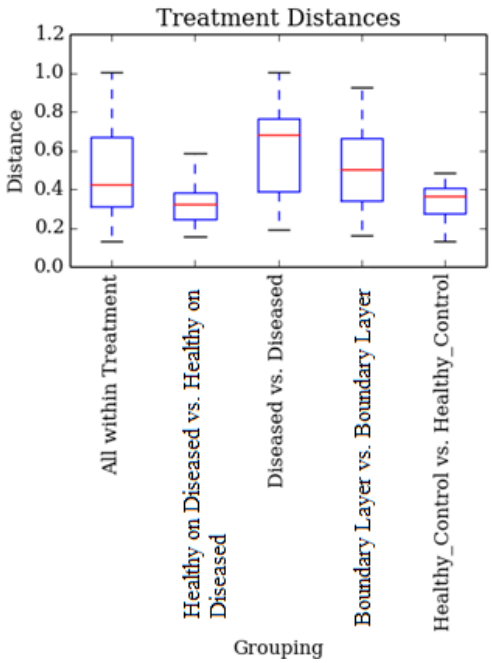


Figure 28: Boxplot displaying distances within the treatments. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

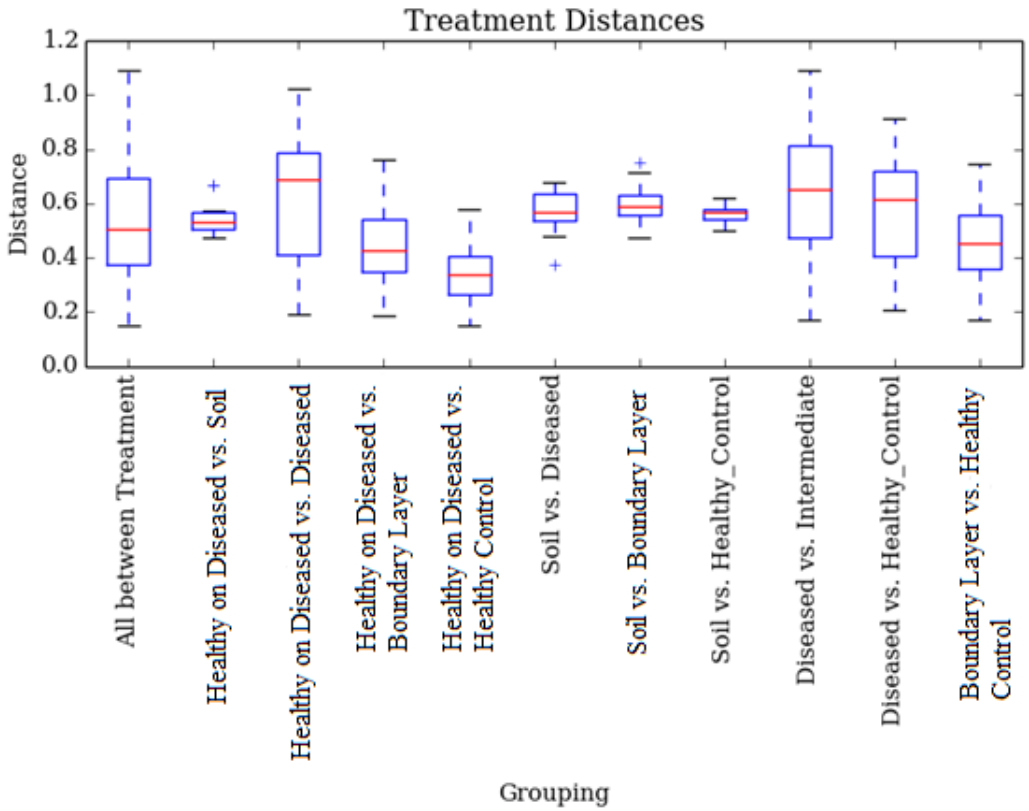


Figure 29: Boxplot displaying distances between treatment groups. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

positioned higher on the graph in relation to the other boxplots (Figure 28). The degree of difference found in the D community was illustrated by visually comparing how much of the box overlapped with the other boxes. The D boxplot overlapped the most with the BL health state, which exposed that those microbial community compositions were the most similar. The D boxplot did not overlap with the HoD state and only very slightly with the HC, which suggested that the disease progression of SOB had altered the normal or expected microbial community for *X. muta*. The inter-quartile range (median) was also closer in proximity to the upper quartile line in the D population which displayed that the D samples had significant differences within their microbial compositions comparatively (Figure 28).

The between treatment boxplots directly compared the differences in microbial community composition of the different health states to each other (Figure 29). When the D health state was compared to any of the other health states (HC, HoD and B layer) the

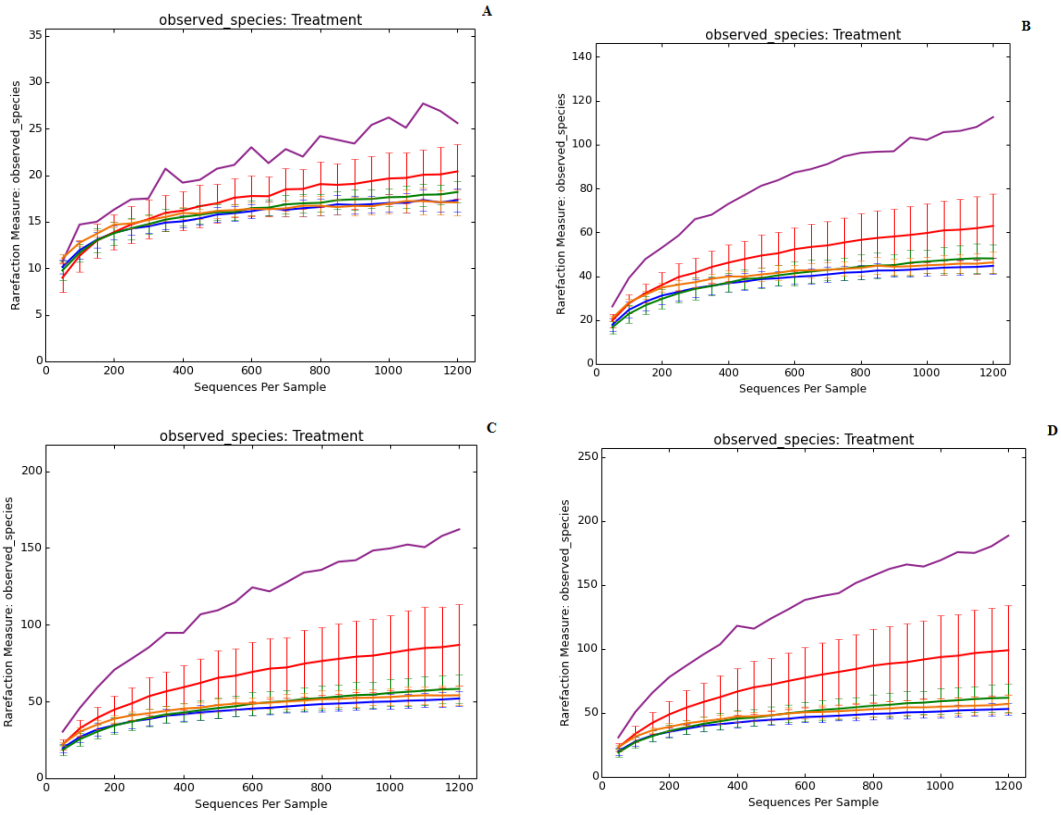


Figure 30: Rarefaction plots displaying Observed Species for different taxonomic levels (A= Phylum, B = Order, C= Family, D = Genus). The red line = Diseased, blue = Healthy on Diseased, yellow = Healthy Control, green = Boundary layer and purple = Soil. Plots were created using QIIME 1.8.0 (Caporaso, 2010).

box plot was relatively larger which suggested that the D microbial community was significantly different and had a unique microbial composition. The boxplots also showed that the microbial community found within the soil sample taken from Fort Lauderdale was highly similar with the microbial composition of all the separate healthy states (Figure 29). This suggested and revealed the possibility that soil may have contributed to the microbial community composition of *X. muta*.

Rarefaction Plots:

The rarefaction plots displayed the observed species for each health state at different taxonomic levels (Figure 30). Analyzing diversity saturation at multiple taxonomic levels allowed the full extent of taxonomic diversity at the different genetic distances within the database to be understood. The rarefaction plots at the different genetic distances displayed that diversity was fully captured for each health state, which was characterized by the curves plateau after the incline. The plots also showed that the D state had the greatest diversity, followed by the BL, HC and the HoD state.

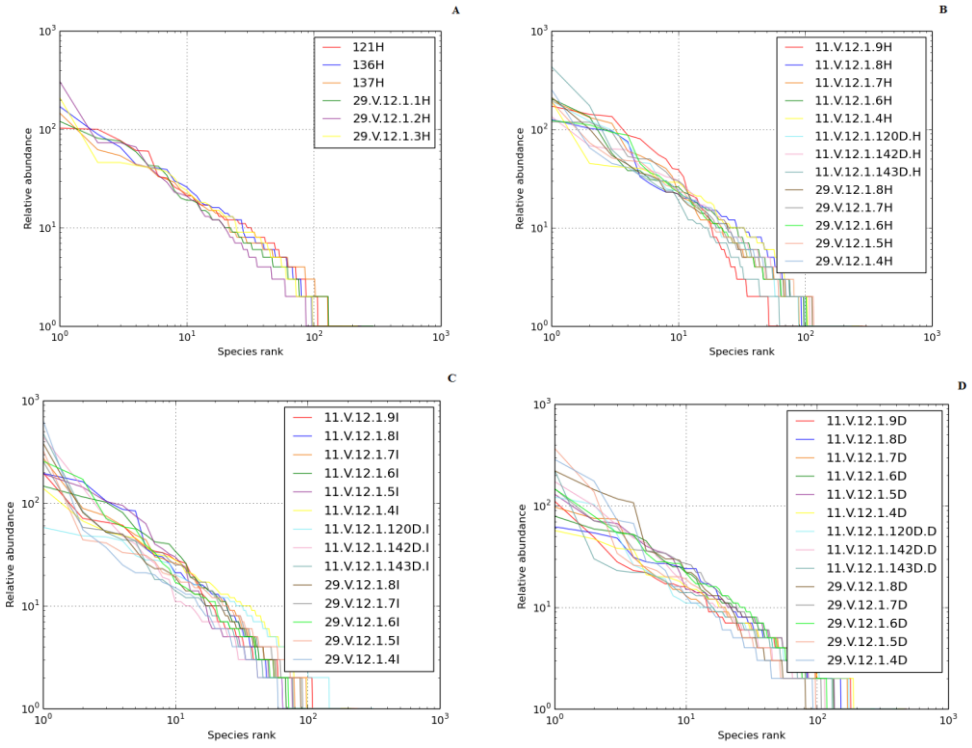


Figure 31: Rank abundance plots by health state (A = Healthy Control Samples, B = Healthy on Diseased Samples, C = Boundary Layer, and D = Diseased). Plots were created using QIIME 1.8.0 (Caporaso, 2010).

Rank Abundance:

All the rank abundance curves were similar for each health state, with a general clustering of samples around 10^{25} for relative abundance and 10^{2} for species rank. (Figure 31). When the differences were compared the D samples had the greatest amount of variation within relative abundance values and also had the largest species rank value (Figure 31, Graph D). The variation present within the relative abundance axis for D samples was likely associated with the stage of SOB progression; as seen with most diseases the community will start to reduce as the disease progresses and the host becomes less able to support a large community. The stage of SOB progression can also influence the variation and explain some of the high values found in the species ranks. Since a compromised host will attract opportunistic bacteria which will boost the species richness only to decline as the habitat become unsuitable. The BL population had the greatest observed relative abundance in comparison to the other health states (Figure 31, Graph C). The advancing line of diseased mesohyl represented the mixing of both HoD and D communities, which included any bacteria responsible for the physical spread or progression of SOB. The HC and the HoD states had relatively similar distributions for the relative abundance and species rank values (Figure 31, Graph A and B). This showed that the two microbial communities were highly similar and the results coincided with the data from both alpha and beta diversity (Figure 31, Graph A and B). All the rank abundance curves had comparable slopes which translated to similar evenness between the health states. This supported the idea that within each health state the microbial community assemblages have core microbiomes which were found to fluctuate with diversity and abundance but always had a common level of congruity.

Taxon Summary at the Phylum Level:

A taxon summary listing all of the bacteria found within a dataset is overwhelming large; to better notice changes within one disease state versus another it is easier to first view the microbial community at the phylum level. Within the HC samples the phyla which dominated the population were Crenarcheota (12.5%), Chloroflexi (19.2%) and Proteobacteria (21.7%) (Table 3). Since samples were collected from two separate locations it was also important to compare the separate populations in order to take into account site variation. Phyla which were found present within all samples from both

sampling sites were: Crenarchaeota, Acidobacteria, Actinobacteria, AncK6, Bacteroidetes, Chloroflexi, Cyanobacteria, Gemmatimonadetes, Nitrospirae, PAUC34f, Poribacteria, Proteobacteria, SBR1093, Spirochaetes, and Thermi (Figure 32 and Table 3). Phyla which were present only in the Boynton Beach population include: Lentishaerae, NKB19, and OD1. Phyla only found within the Fort Lauderdale population include: Euryarcheota, *H*-178, and SAR406. The majority of the variation between the microbial communities of the separate populations was within the taxon relative abundances measurements. This was most evident in Actinobacteria and Cyanobacteria which were found in greater abundances

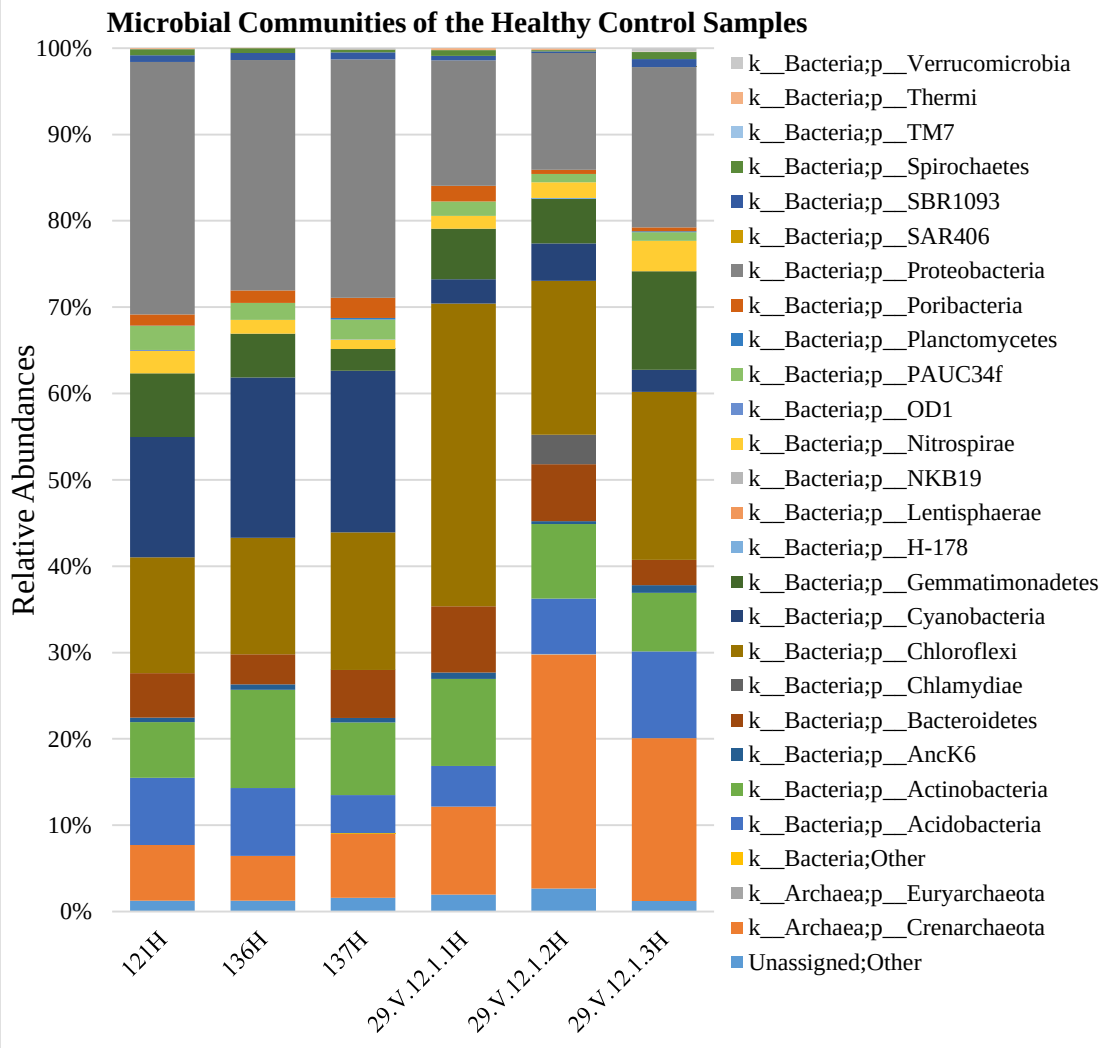


Figure 32: Taxon summary at the phylum level for healthy tissue samples from healthy sponges (Healthy Control). Samples 121H, 136H and 137H are from the Boynton Beach sampling location. Samples 29.V.12.1.1H, 29.V.12.1.2H and 29.V.12.1.3H are samples from the Fort Lauderdale sampling location. Plot was created using data retrieved from QIIME 1.8.0 (Caporaso, 2010).

in the Boynton County sampling site, while Choroflexi and Crenarcheota were found in greater abundance in the Fort Lauderdale sampling site (Figure 32).

The HoD samples were also dominantly compromised of Crenarchaeota (19.3%), Chloroflexi (19.1%) and Proteobacteria (18.9%) (Table 3). The HC and the HoD communities were similar but there was a noticeable increase in the abundance of Crenarchaeota by 6.8% within the HoD state. Phyla which were present in both HoD subpopulations were: Crenarchaeota, Acidobacteria, Actinobacteria, AncK6, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Nitrospirae, PAUC34f, Poribacteria, Proteobacteria, and Spirochaetes (Figure 33 and Table 3). Cyanobacteria, SBR1093 and Thermi were also very common, as they were found in all samples except for one. Phyla only present within the Boynton population include: Euryarchaeota, Firmicutes, SAR406 and WS3. The phyla only present within the Fort Lauderdale population were Lentisphaerse.

Healthy Control Taxon	Percentage	Observations in Boynton Population	Observations in Broward Population	Healthy on Diseased Taxon	Percentage	Observations in Boynton Population	Observations in Broward Population
Unassigned	1.667	3	3	Unassigned	2.632	9	5
k__Archaea;p__Crenarchaeota	12.538	3	3	k__Archaea;p__Crenarchaeota	19.308	9	5
k__Archaea;p__Euryarchaeota	0.011	0	2	k__Archaea;p__Euryarchaeota	0.002	1	0
k__Bacteria	0.003	1	0	k__Bacteria;p__Acidobacteria	4.582	9	5
k__Bacteria;p__Acidobacteria	6.875	3	3	k__Bacteria;p__Actinobacteria	8.273	9	5
k__Bacteria;p__Actinobacteria	8.625	3	3	k__Bacteria;p__AncK6	0.343	9	5
k__Bacteria;p__AncK6	0.596	3	3	k__Bacteria;p__Bacteroidetes	5.067	9	5
k__Bacteria;p__Bacteroidetes	5.242	3	3	k__Bacteria;p__Chlamydiae	0.007	1	1
k__Bacteria;p__Chlamydiae	0.565	1	1	k__Bacteria;p__Chloroflexi	19.053	9	5
k__Bacteria;p__Chloroflexi	19.193	3	3	k__Bacteria;p__Cyanobacteria	9.356	9	4
k__Bacteria;p__Cyanobacteria	10.155	3	3	k__Bacteria;p__Firmicutes	0.099	3	0
k__Bacteria;p__Gemmatimonadete	6.234	3	3	k__Bacteria;p__Gemmatimonadetes	5.447	9	5
k__Bacteria;p__H-178	0.018	0	1	k__Bacteria;p__Lentisphaerae	0.015	0	1
k__Bacteria;p__Lentisphaerae	0.002	1	0	k__Bacteria;p__NKB19	0.001	1	1
k__Bacteria;p__NKB19	0.002	1	0	k__Bacteria;p__Nitrospirae	3.437	9	5
k__Bacteria;p__Nitrospirae	2.010	3	3	k__Bacteria;p__OD1	0.027	4	2
k__Bacteria;p__OD1	0.021	2	0	k__Bacteria;p__PAUC34f	0.818	9	5
k__Bacteria;p__PAUC34f	1.773	3	3	k__Bacteria;p__Planctomycetes	0.044	8	1
k__Bacteria;p__Planctomycetes	0.059	2	1	k__Bacteria;p__Poribacteria	0.828	9	5
k__Bacteria;p__Poribacteria	1.314	3	3	k__Bacteria;p__Proteobacteria	18.857	9	5
k__Bacteria;p__Proteobacteria	21.693	3	3	k__Bacteria;p__SAR406	0.014	3	0
k__Bacteria;p__SAR406	0.002	0	1	k__Bacteria;p__SBR1093	0.736	8	5
k__Bacteria;p__SBR1093	0.657	3	3	k__Bacteria;p__Spirochaetes	0.748	9	5
k__Bacteria;p__Spirochaetes	0.536	3	3	k__Bacteria;p__TM7	0.042	6	2
k__Bacteria;p__TM7	0.008	2	1	k__Bacteria;p__Thermi	0.214	9	4
k__Bacteria;p__Thermi	0.102	3	3	k__Bacteria;p__Verrucomicrobia	0.043	6	1
k__Bacteria;p__Verrucomicrobia	0.101	1	1	k__Bacteria;p__WS3	0.004	1	0
	100				100		

Table 3: Microbial community assemblages of Healthy Control and Healthy on Diseased dataset at the phylum level. Any highlighted phyla are found in both populations at the separate sampling locations, while yellow highlighted phyla indicate the dominate phyla within the health state. Observations in the separate sampling locations convey the number of samples a particular taxa was found in. The number of sponges sampled for the Healthy Control population was three from each sampling site. The number of sponges sampled for the Healthy on Diseased population was nine from Boynton Beach and five from Fort Lauderdale. Tables were created using data evaluated and retrieved from QIIME 1.8.0 (Caporaso, 2010).

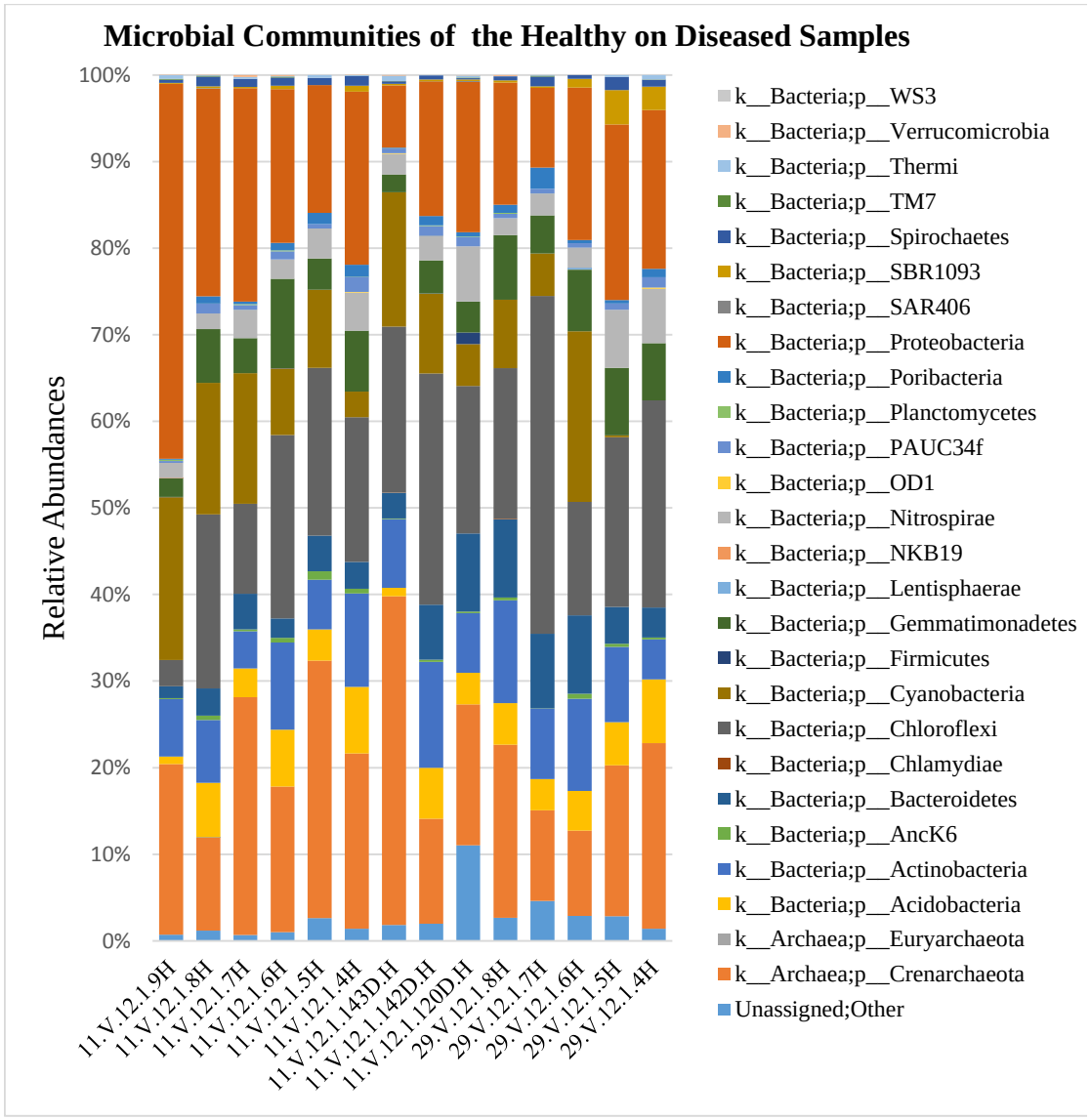


Figure 33: Taxon summary at the phylum level for healthy samples from diseased sponges. Samples 11.V.12.1.9H, 11.V.12.1.8H, 11.V.12.1.7H, 11.V.12.1.6H, 11.V.12.1.5H, 11.V.12.1.4H, 11.V.12.1.143DH, 11.V.12.1.142D.H, 11.V.12.1.120D.H were samples from the Boynton Beach sampling location. Samples 29.V.12.1.8H, 29.V.12.1.7H, 29.V.12.1.6H, 29.V.12.1.5H and 29.V.12.1.4H were samples from the Fort Lauderdale sampling location. Plot was created using data retrieved from QIIME 1.8.0 (Caporaso, 2010).

Within the BL samples (advancing line of diseased mesohyl) the phyla which were most abundant within the community were Crenarchaeota (21.7%), Chloroflexi (14.3%), and Proteobacteria (23.2%) (Table 4). The Phylum Verrucomicrobia increased in abundance to 4.7%; the highest abundance compared to the other health states. Phyla present within all samples from both populations were: Crenarchaeota, Acidobacteria,

Actinobacteria, AncK6, Bacteroidetes, Chloroflexi, Cyanobacteria, Gemmatimonadetes, Nitrospirae, PAUC34f, Planctomycetes, Poribacteria, Proteobacteria, SBR1093, Spirochaetes, Thermi, and Verrucomicrobia (Figure 34 and Table 4). The Boynton Beach population included phyla: Euryachaeota, Firmicutes, Fusobacteria, NKB19, SAR406, Tenericutes, WPS-2 and ZB3 which were not seen within the Fort Lauderdale population. The presence of noticeable site variation within the relative abundances of Cyanobacteria, Proteobacteria and Verrucomicrobia within Boynton Beach (Figure 34). The Fort Lauderdale sampling site had increased relative abundances of Crearchaeota and Chloroflexi (Figure 34). The BL population also included bacteria which were not

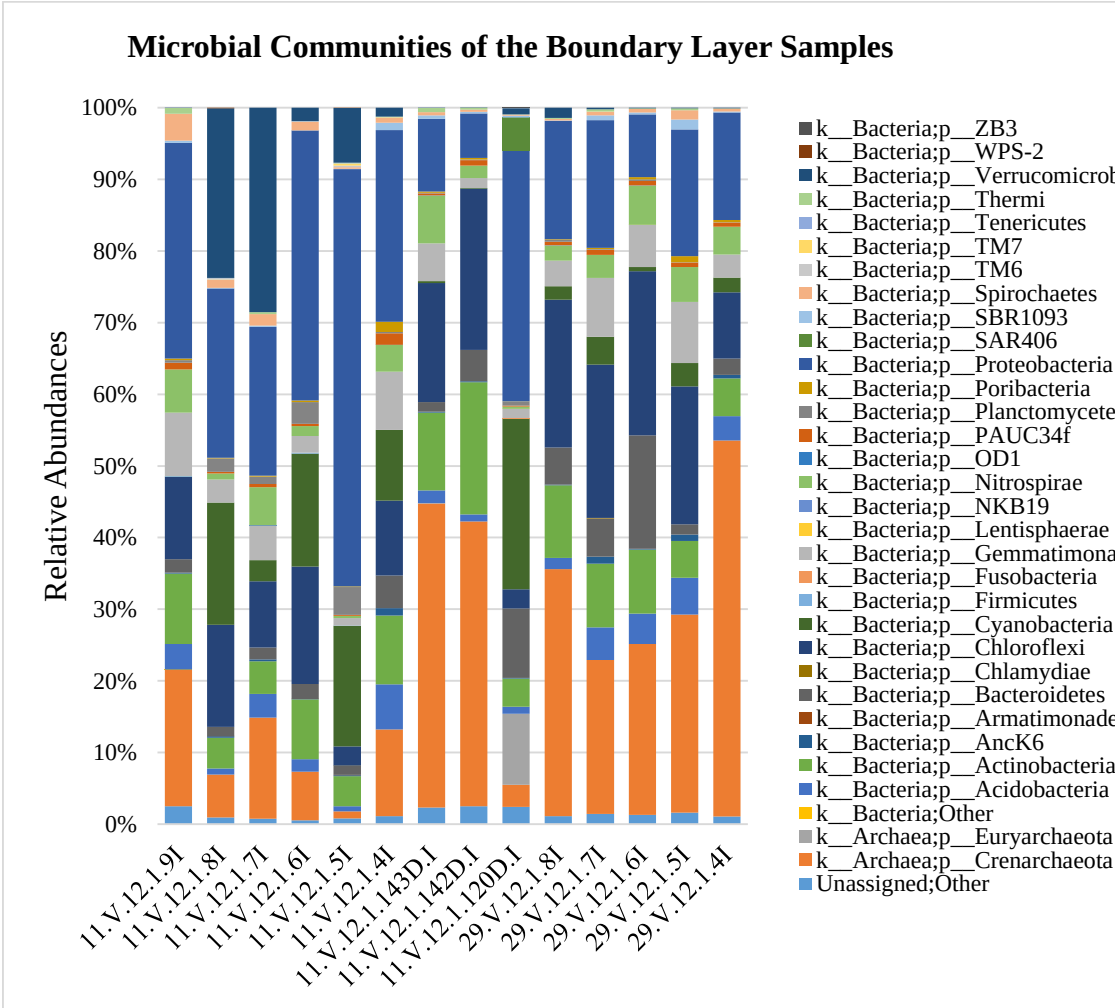


Figure 34: Taxon summary at the phylum level for the Boundary Layer samples. Samples 11.V.12.1.9I, 11.V.12.1.8I, 11.V.12.1.7I, 11.V.12.1.6I, 11.V.12.1.5I, 11.V.12.1.4I, 11.V.12.1.143D.I, 11.V.12.1.142D.I, 11.V.12.1.120D.I were samples from the Boynton Beach sampling location. Samples 29.V.12.1.8I, 29.V.12.1.7I, 29.V.12.1.6I, 29.V.12.1.5I and 29.V.12.1.4I were samples from the Fort Lauderdale sampling location. Plot was created using data retrieved from QIIME 1.8.0 (Caporaso, 2010).

previously found within the HC or the HoD populations, such as Armatimonadetes, Fusobacteria, TM6, Tenericutes, WPS-*2* and ZB3 (Figure 35).

Boundary Layer Taxon	Percentage	Observations in Boynton Population	Observations in Broward Population
Unassigned	1.449	9	5
k__Archaea;p__Crenarchaeota	21.739	9	5
k__Archaea;p__Euryarchaeota	0.710	2	0
k__Bacteria	0.003	1	0
k__Bacteria;p__Acidobacteria	2.795	9	5
k__Bacteria;p__Actinobacteria	8.030	9	5
k__Bacteria;p__AncK6	0.343	9	5
k__Bacteria;p__Armatimonadetes	0.011	3	1
k__Bacteria;p__Bacteroidetes	4.144	9	5
k__Bacteria;p__Chlamydiae	0.006	1	1
k__Bacteria;p__Chloroflexi	14.276	9	5
k__Bacteria;p__Cyanobacteria	7.021	9	5
k__Bacteria;p__Firmicutes	0.023	4	0
k__Bacteria;p__Fusobacteria	0.003	1	0
k__Bacteria;p__Gemmatimonadetes	4.692	9	5
k__Bacteria;p__Lentisphaerae	0.003	1	1
k__Bacteria;p__NKB19	0.002	1	0
k__Bacteria;p__Nitrospirae	3.277	9	5
k__Bacteria;p__OD1	0.009	2	2
k__Bacteria;p__PAUC34f	0.562	9	5
k__Bacteria;p__Planctomycetes	0.844	9	5
k__Bacteria;p__Poribacteria	0.311	9	5
k__Bacteria;p__Proteobacteria	23.167	9	5
k__Bacteria;p__SAR406	0.333	1	0
k__Bacteria;p__SBR1093	0.363	9	5
k__Bacteria;p__Spirochaetes	0.895	9	5
k__Bacteria;p__TM6	0.024	3	0
k__Bacteria;p__TM7	0.035	3	1
k__Bacteria;p__Tenericutes	0.008	3	0
k__Bacteria;p__Thermi	0.220	9	5
k__Bacteria;p__Verrucomicrobia	4.691	9	5
k__Bacteria;p__WPS-2	0.009	2	0
k__Bacteria;p__ZB3	0.004	1	0
	100		

Diseased Taxon	Percentage	Observations in Boynton Population	Observations in Broward Population
Unassigned	3.996	9	5
k__Archaea;p__Crenarchaeota	9.624	9	5
k__Archaea;p__Euryarchaeota	0.008	2	1
k__Bacteria	0.010	2	1
k__Bacteria;p__Acidobacteria	2.676	9	5
k__Bacteria;p__Actinobacteria	9.526	9	5
k__Bacteria;p__AncK6	0.200	4	5
k__Bacteria;p__Armatimonadetes	0.388	7	1
k__Bacteria;p__Bacteroidetes	11.634	9	5
k__Bacteria;p__Caldithrix	0.035	4	0
k__Bacteria;p__Chlamydiae	0.329	8	1
k__Bacteria;p__Chlorobi	0.048	6	0
k__Bacteria;p__Chloroflexi	10.331	9	5
k__Bacteria;p__Cyanobacteria	5.676	9	5
k__Bacteria;p__Fibrobacteres	0.001	1	0
k__Bacteria;p__Firmicutes	0.173	5	2
k__Bacteria;p__Fusobacteria	0.036	7	0
k__Bacteria;p__GN02	0.081	7	0
k__Bacteria;p__Gemmatimonadetes	2.443	8	5
k__Bacteria;p__H-178	0.053	5	0
k__Bacteria;p__KSB3	0.001	1	0
k__Bacteria;p__Lentisphaerae	0.040	3	1
k__Bacteria;p__NKB19	0.012	3	0
k__Bacteria;p__Nitrospirae	1.330	5	5
k__Bacteria;p__OD1	0.014	4	1
k__Bacteria;p__OP1	0.009	1	0
k__Bacteria;p__OP3	0.012	2	0
k__Bacteria;p__OP8	0.002	1	0
k__Bacteria;p__PAUC34f	0.527	6	5
k__Bacteria;p__Planctomycetes	4.448	8	4
k__Bacteria;p__Poribacteria	0.393	6	5
k__Bacteria;p__Proteobacteria	31.980	9	5
k__Bacteria;p__SBR1093	0.264	9	5
k__Bacteria;p__SR1	0.012	4	0
k__Bacteria;p__Spirochaetes	0.466	9	5
k__Bacteria;p__TM6	0.482	8	0
k__Bacteria;p__TM7	0.011	4	1
k__Bacteria;p__Tenericutes	0.014	3	1
k__Bacteria;p__Thermi	0.129	5	5
k__Bacteria;p__Verrucomicrobia	2.498	9	5
k__Bacteria;p__WPS-2	0.059	7	0
k__Bacteria;p__WS3	0.024	5	1
k__Bacteria;p__ZB3	0.001	1	0
k__Bacteria;p__[Caldithrix]	0.004	1	0
	100		

Table 4: Microbial community assemblages of the Boundary Layer and the Diseased datasets at the phylum level. Any highlighted phyla are found in both populations at the separate sampling locations, while yellow highlighted phyla indicate the dominate phyla within the health state. Observations in the separate sampling locations convey the number of samples a particular taxa was found in. The number of sponges sampled for the Boundary Layer and Diseased populations was nine from Boynton Beach and five from Fort Lauderdale. Tables were created using data evaluated and retrieved from QIIME 1.8.0 (Caporaso, 2010).

The D samples most abundant phyla were Bacteroidetes (11.6%), Chloroflexi (10.3%), and Proteobacteria (32 %) (Table 4). Noticeable community shifts within the D population included the increase in abundance of Bacteroidetes by 6.4%, Planctomycetes by 4.4% and Proteobacteria by 10.3% when compared to the HC community. Overall there were decreases in the relative abundances of most taxon that previously were associated with all mesohyl samples. Phyla which were present in all D samples from both sampling

sites were: Crenarchaeota, Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Proteobacteria, SBR1093, Spirochaetes and Verrucomicrobia (Figure 36 and Table 4). Not all phyla were found in every sample from both populations, but bacteria which were abundant and common members of the D community include: Gemmatimonadetes, Planctomycetes, Lentisphaerae, Nitrospirae, PAUC34f, Poribacteria, Thermi, Firmicutes and AncK6. Phyla absent within the Fort Lauderdale population

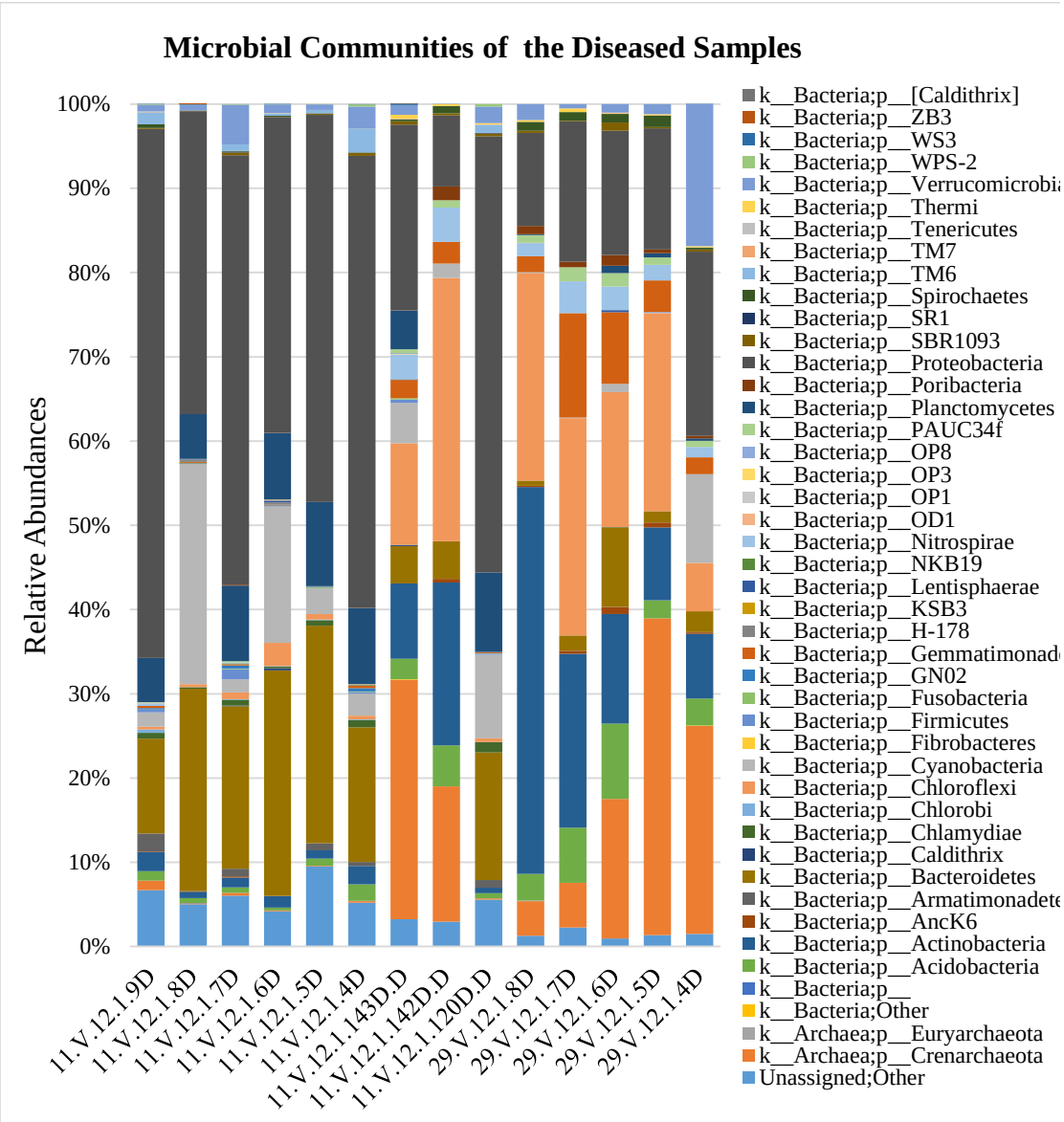
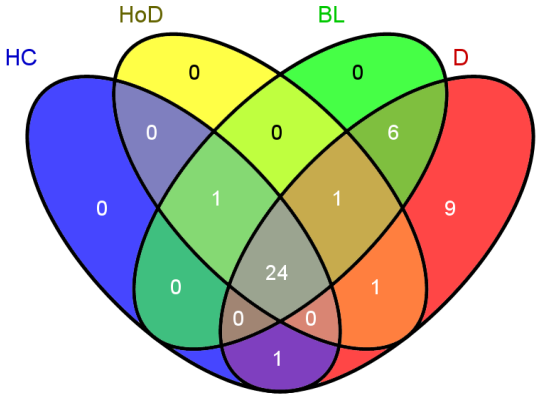


Figure 36: Taxon summary at the phylum level for the Diseased samples. Samples 11.V.12.1.9D, 11.V.12.1.8D, 11.V.12.1.7D, 11.V.12.1.6D, 11.V.12.1.5D, 11.V.12.1.4D, 11.V.12.1.143D.D, 11.V.12.1.142D.D, 11.V.12.1.120D.D were samples from the Boynton Beach sampling location. Samples 29.V.12.1.8D, 29.V.12.1.7D, 29.V.12.1.6D, 29.V.12.1.5D and 29.V.12.1.4D were samples from the Fort Lauderdale sampling location. Plot was created using data retrieved from QIIME 1.8.0 (Caporaso, 2010).

included: Caldithrix, Chlorobi, Fibrobacteres, Fusobacteria, GN02, H-178, KSB3, NKB19, OP1, OP3, OP8, SR1, TM6, WPS-2, and ZB3 (Table 4). Chlamydiae was also rarely found within the Fort Lauderdale sampling population (one sample), while it was abundant within the Boynton Beach population; as it was found within eight samples out of nine. Phyla which were exclusive to the D population included: Caldithrix, Chlorobi, Fibrobacteres, GN02, KSB3, OP1, OP3, OP8, SR1, and WSB (Figure 35). The microbial composition relative abundance differences due to sampling location were also present and included reductions in Actinobacteria, Crenarchaeota and Chloroflexi from the Boynton Beach sampling site, while Proeobacteria and Bacteroidetes were more abundant when compared to the Fort Lauderdale sampling site (Figure 36). It was unclear why the Boynton Beach



Phyla Exclusive to the Health States						
HC, HoD, BL and D	**HC, HoD and BL**	**HoD, BL and D**	**HC and D**	**HoD and D**	**BL and D**	**D**
Unassigned	SAR406	Firmicutes	H-178	WS3	Armatimonadetes	Caldithrix
Crenarchaeota					Fusobacteria	Chlorobi
Euryarchaeota					TM6	Fibrobacteres
Acidobacteria					Tenericutes	GN02
Actinobacteria					WPS-2	KSB3
AncK6					ZB3	OP1
Bacteroidetes						OP3
Chlamydiae						OP8
Chloroflexi						SR1
Cyanobacteria						
Gemmatimonadetes						
Lentisphaerae						
NKB19						
Nitrospirae						
OD1						
PAUC34f						
Planctomycetes						
Poribacteria						
Proteobacteria						
SBR1093						
Spirochaetes						
TM7						
Thermi						
Verrucomicrobia						

Figure: 35 Venn Diagram of Phyla found within the Health States. (Olivers, 2007)

population has obvious differences in diversity but it may be contributed to the amount of samples taken from the sampling site or the two week gap between sampling. Having double the number of samples from one population will certainly increase the amount of diversity found at the Boynton Beach sampling location. A difference of two weeks

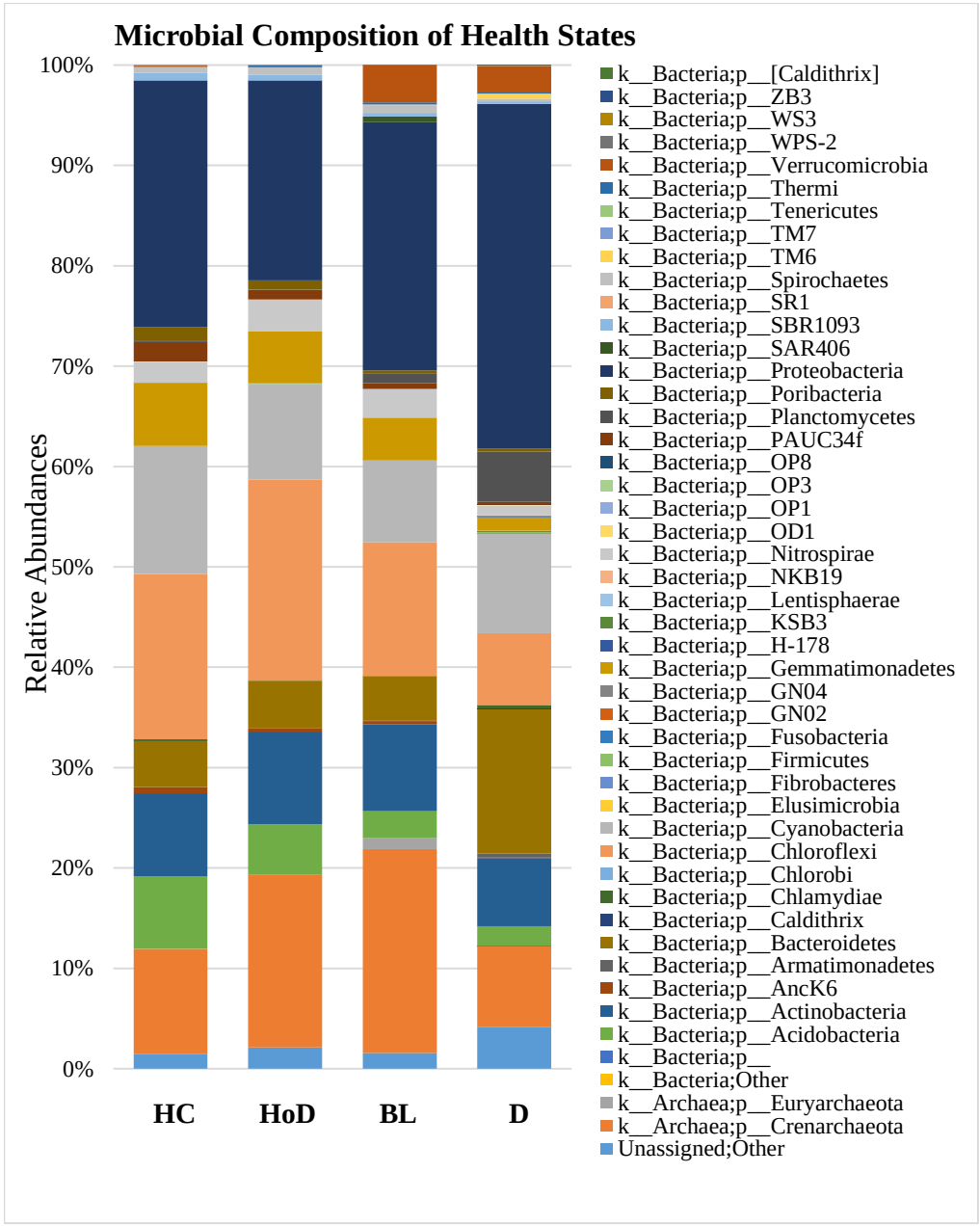


Figure 37: Microbial composition of health states at the phylum level. This figure shows that even with the presence of Sponge Orange Band Disease there exists a group of core microbiome which are present in every health state. These bacteria may be fundamentally necessary for *Xestospongia muta* survival. Plot was created using data retrieved from QIIME 1.8.0 (Caporaso, 2010).

between sampling could have influenced what microbes were found within the sponge mesohyl and what could have survived within the diseased mesohyl.

Overall as SOB Disease progressed it caused microbial community shifts within *X. muta* (Figure 37). Phyla which were found exclusively within the BL and the D health states included: Armatimonadetes, Caldithrix, Chlorobi, Fibrobacteres, Fusobacteria, GN02, KSB3, OP1, OP3, OP8, SR1, TM6, Tenericutes, WPS-2, and ZB3 (Table 5 and Figure 35). Phyla which became more abundant within the BL and D populations included Planctmycetes and Verrucomicrobia. Whether these phyla were opportunistic bacteria or secondary infections is to be determined, but their presence within actively diseased sponge mesohyl helps to understand the effect SOB has on the microbial community.

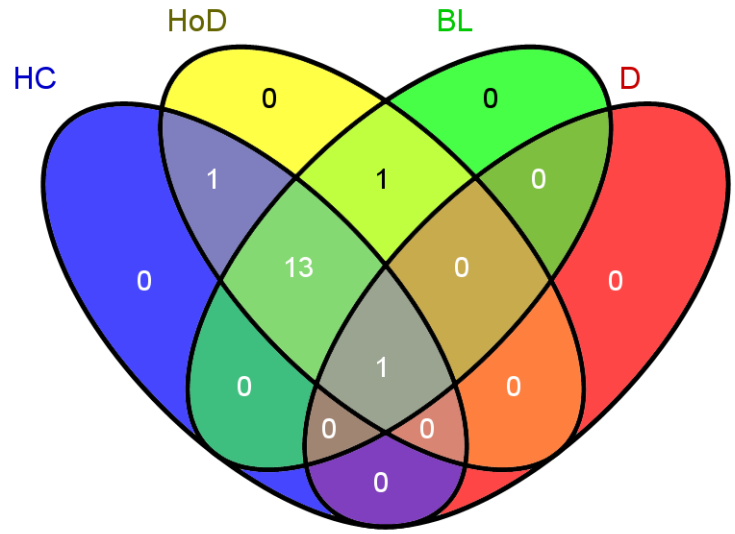
Comparison of Taxon at Phylum Level Between Health States	
Taxon	**Found Within/ Shared Between**
Unassigned	All
k__Archaea;p__Crenarchaeota	All
k__Archaea;p__Euryarchaeota	All
k__Bacteria	Healthy Control & Boundary Layer & Diseased
k__Bacteria;p__Acidobacteria	All
k__Bacteria;p__Actinobacteria	All
k__Bacteria;p__AncK6	All
k__Bacteria;p__Armatimonadetes	Boundary Layer & Diseased
k__Bacteria;p__Bacteroidetes	All
k__Bacteria;p__Caldithrix	Diseased
k__Bacteria;p__Chlamydiae	All
k__Bacteria;p__Chlorobi	Diseased
k__Bacteria;p__Chloroflexi	All
k__Bacteria;p__Cyanobacteria	All
k__Bacteria;p__Fibrobacteres	Diseased
k__Bacteria;p__Firmicutes	Healthy on Diseased & Boundary Layer & Diseased
k__Bacteria;p__Fusobacteria	Boundary Layer & Diseased
k__Bacteria;p__GN02	Diseased
k__Bacteria;p__Gemmatimonadetes	All
k__Bacteria;p__H-178	Healthy Control & Diseased
k__Bacteria;p__KSB3	Diseased
k__Bacteria;p__Lentisphaerae	All
k__Bacteria;p__NKB19	All
k__Bacteria;p__Nitrospirae	All
k__Bacteria;p__OD1	All
k__Bacteria;p__OP1	Diseased
k__Bacteria;p__OP3	Diseased
k__Bacteria;p__OP8	Diseased
k__Bacteria;p__PAUC34f	All
k__Bacteria;p__Planctomycetes	All
k__Bacteria;p__Poribacteria	All
k__Bacteria;p__Proteobacteria	All
k__Bacteria;p__SAR406	All
k__Bacteria;p__SBR1093	All
k__Bacteria;p__SR1	Diseased
k__Bacteria;p__Spirochaetes	All
k__Bacteria;p__TM6	Boundary Layer & Diseased
k__Bacteria;p__TM7	All
k__Bacteria;p__Tenericutes	Boundary Layer & Diseased
k__Bacteria;p__Thermi	All
k__Bacteria;p__Verrucomicrobia	All
k__Bacteria;p__WPS-2	Boundary Layer & Diseased
k__Bacteria;p__WS3	Healthy on Diseased & Diseased
k__Bacteria;p__ZB3	Boundary Layer & Diseased
k__Bacteria;p__[Caldithrix]	Diseased

Table 5: Microbial community comparison between health states at the phylum level. Allowing microbial community shifts due the progression of the SOB to be recognized. Table was created using data retrieved from QIIME 1.8.0 (Caporaso, 2010).

Core Microbiome:

The taxon summary at the phylum level provided evidence of a core microbiome that was present within the HC, HoD and BL health states. The relative abundance of the

microbes that comprised the core composition was found to decrease as SOB progressed. The core microbiome was determined and defined as an OTU that was present in at least 70% of samples within a health state. Phyla which constituted the core microbiome (found within the HC, HoD and BL health states) were Crenarchaeota, Acidobacteria, Actinobacteria, AncK6, Bacteroidetes, Chloroflexi, Cyanobacteria, Gemmatimonadetes, Nitrospirae, PAUC34f, Proteobacteria, SBR1093, Spirochaetes and Unassigned (Figure 38). A noticeable shift was present within the core microbiomes which corresponded with the progression of SOB.



Phyla Exclusive to the Core Microbiomes by Health State			
HC, HoD, BL and D	**HC, HoD and BL**	**HC and HoD**	**HoD and BL**
Proteobacteria	Crenarchaeota	Poribacteria	Thermi
	Acidobacteria		
	Actinobacteria		
	AncK6		
	Bacteroidetes		
	Chloroflexi		
	Cyanobacteria		
	Gemmatimonadetes		
	Nitrospirae		
	PAUC34f		
	SBR1093		
	Spirochaetes		
	Unassigned		

Figure 38: Core Microbiome Venn Diagram at the Phylum Level. (Olivers, 2007)

The HC core microbiome had the largest abundance of OTUs (total of 90 OTUs) compared the other core microbiome assemblages and the dataset total OTU count of 228,315. The community was mainly comprised of 30 OTUs of Proteobacteria, 21 OTUs of Chloroflexi, and 8 OTUs of Acidobacteria. At lower taxonomic levels the most abundant bacteria based on OTU counts within the HC population was SAR202 (Choloroflexi), then Rhodospirillaceae (Proteobacteria) and the third was a five way tie between A4b (Chloroflexi), Caldilineaceae (Chloroflexi), Gemm-2 (Gemmatimonadetes), Rhodobacteraceae (Proteobacteria) and HTCC2089 (Proteobacteria) (Appendix Table 2).

The HoD core microbiome had reduced microbial abundance when compared to the HC core microbiome with a total of 65 OTUs. Dominant phyla included Proteobacteria with 22 OTUs, Chloroflexi with 12 OTUs and Actinobacteria with 6 OTUs. At lower taxonomic levels the most abundant bacteria within the community composition was a three way tie between A4b (Chloroflexi), Gemm-2 (Gemmatimonadetes), and Rhodospirllaceae (Proteobacteria) (Appendix Table 3). The second most abundant bacteria belonged to Caldilineaceae (Chloroflexi) and thirdly Rhodobacteraceae (Proteobacteria). The microbial composition between the HC and HoD core communities were highly similar with the main deviation found in the absence of SAR202 (Chloroflexi) (Appendix Table 3). SAR202 was the most abundant bacteria in the HC with a total of 9 separate OTUs. Then the sponge exclusive phylum Poribacteria was only present within the core microbiome of the HC and the HoD states and absent with the BL and D states (Figure 38) (Fieseler, et al, 2004; Taylor et al, 2007b; Lafi et al, 2009). There was also the presence of the phyla Thermi which previously was not present within the HC.

The BL core microbiome community abundance was notably decreased with a total of 54 OTUs when compared to the HC and HoD core microbiome. The four main phyla continued to decrease in abundance as SOB progressed with 15 OTUs for Proteobacteria, 9 OTUs for Chloroflexi, and 7 OTUs for Actinobacteria. The most prevalent bacteria within the community was a tie between Rhodobacterales (Proteobacteria) and Gemm-2 (Gemmatimonadetes) (Appendix Table 4). The second most abundant bacteria was Rhodospirllaceae (Proteobacteria) and then Caldilineaceae (Chloroflexi) (Appendix Table 4). The transition from the HoD to the BL health state was marked by the continued presence of phyla Thermi (Figure 38).

The D core microbiome community radically different than the previous health states because it was devoid of a core microbiome and there was only two OTUs which were found in 70% of samples that comprised the D population (Figure 38). Both of these OTUs were from the phyla Proteobacteria and at lower taxonomic classifications both were identified as Rhodobacteraceae (Appendix Table 5).

Taxon Summary at the Genus Level or Lowest known Taxonomic Classification:

Comparing the microbial community at deeper taxonomic levels provided greater insight in community shifts due SOB progression and difference due to sampling location. The microbial communities at the genus level or furthest known taxonomic classification provided a summary of the microbial community assemblages within *X. muta* and a community comparison between the separate health states. To recognize the dominant bacteria within each health state the three largest groups by percentage were compared. The most abundant bacteria that comprised the microbial community of the HC samples were Nitrosopumilus (11.3%), which oxidizes ammonia to nitrite and therefore can reduce metabolic waste produced by the sponge that would otherwise cause damaging health effects if left to accumulate (Hardoim et al, 2014; Könneke et al, 2005; Steger et al, 2008). There has also been evidence that ammonia oxidizing Archaea are included in some Demosponge larvae, which links ammonia detoxification as a highly important process for overall sponge health and explains their abundance within sponge microbial communities (Steger et al, 2008). The second most abundant bacteria was Prochlorococcus (9.8%), which are very abundant within marine environments and are the smallest known marine phytoplankton (Partensky et al, 1999). Finally Caldilineaceae (6.3%) belongs to the phylum Chloroflexi, which produce energy from light and are described as filamentous green non-sulfur bacteria (Yamada et al, 2006). A full list of all taxon found within the HC population can be found at Appendix Table 6A-E.

A similar microbial community assemblage, including the most abundant bacteria was also found within the HoD samples. The most abundant bacteria was Nitrosopumilus which increased to 17.4%, then Calilineaceae which increased to 8.3% and *A4b* which was another strain of Chloroflexi at 5.4% (Appendix Table 7A-E). The HC and HoD states were highly similar to each other in terms of bacterium present and the relative abundances. This

suggests that the advancing band of diseased mesohyl which was captured within the BL samples spreads SOB to minimally affected mesohyl.

The three most abundant bacterium in the BL samples were Nitrosopumilus at 20.3%, Rhodobacteraceae at 9% and Caldilineaceae at 7.2% (Appendix Table 8A-H). Rhodobacteraceae completes both carbon and sulfur cycling, as well as reducing trace metals in marine environments (Dang et al, 2008). *Rhodobacteraceae bacterium* extracted from sponge *Aaptos-aaptos* was found to contain an antibacterial metabolite, which was effective against pathogenic *Vibrio eltor*, *Bacillus subtilis* (gut commensal found associated with food poisoning) and *Staphylococcus aureus* ("Staph" infections) (Muriasih et al, 2014). The bacteria that comprised this family were also known to quickly exploit available nutrients and were primary surface colonizers within coastal waters (Dang et al, 2008). SOB Disease is known to cause sponge mesohyl to degrade and crumble away, which creates a new food source in a nutrient poor environment. Opportunistic bacteria would naturally exploit this energy source and colonize the mesohyl, while the microbial communities originally associated with *X. muta* would shift as environmental conditions change. The BL health state was the first to have noticeable fluctuations in the microbial community inhabitants. Bacteria found within the BL population which were previously absent from the microbial communities of HC and HoD include: Armatimonadetes which was classified down to the order Fibriimonadales, Fusobacteria, TM6 which was classified down to the class SJA-*4*, Tenericutes which was classified down to the order Mycoplasmatales, WPS-2 and ZB3 (Appendix Figure 2). Another change in the microbial composition of the BL samples was the increased percentages of the phylum Verrucomicrobia. Previously Verrucomicrobia was not found within all sampled individuals and was not present in high percentages within the HC (0.10%) and HoD (0.04%) communities but within the BL samples it comprised 4.7% of the community assemblage (Appendix Table 6A and 7A).

The D taxonomic summary at the genus level had noticeably more bacterial strains than the other health states, but the bacteria inhabitants present were in smaller quantities. The three bacteria that were most abundant within the community were Rhodobacteraceace with 11%, Nitrosopumilus with 8.8% and Flavobacteriaceae with 5.2% (Appendix Table 9A-L). Flavobacteriaceae OTUs were only found within the D subpopulation from

Boynton Beach. In general few bacteria were found in high quantities from both sampling locations, which advocated that the microbial communities from the D samples were more influenced by sampling location then the other health states. Genera that were abundant within the D population and found at both sampling locations were Luteolibacter, Rubritlea and Verrucomicrobium from the phylum Verrucomicrobia (Appendix Table 9K-4L). Luteolibacter specifically was only found associated with BL and D health states. The phylum Planctomycetes was also more abundant within the D population where it compromised 4.44% of the community. This was comparatively higher than within the HC health state (0.059%), the HoD state (0.044%) and the BL state (0.844%) [Appendix Tables 6-9]. Other pathogenic bacterium (Clostridia, Fusobacteria, Rickettsiales, Burkholderiaceae, Legionellales, Pseudomonadales, Vibrio, and Leptospiraceae) were associated with the dataset as whole but were more frequent within the samples from the Boynton Beach sampling location and rarely associated with the samples from the Broward Beach sampling site. Bacteria which were exclusive to only the D population included: Caldithrix;g_LCP-26 and g_Caldithrix, Chlorobi;c_OPB56, Fibrobacteres;o_Ucp1540, GN02, KSB3, OP1;c_OPB14, OP3;c_PBS-25 and f_kpj58rc, OP8;o_HMMVPog-54, SR1 and Tenericutes;f_Anaeroplasmataceae (Appendix Figure 2). Along with the new bacteria not seen previously in the HC or HoD health states were the shifts in the microbial communities taxon based on the lowest taxonomic classifications within the BL and D populations. This provided never seen before genus classifications within the BL and D health states for phyla already present within the HC and HoD populations. The numerous microbial shifts at the lower taxonomic level for the BL and D populations suggested that the sponge phyla may have been adapting the habitat pressures and alterations caused by SOB Disease. For a full list of all the microbial shift within the separate health states see Appendix Table 10.

Non-parametric ANOVA (Group Significance):

Bacteria capable of causing disease were of particular interest as well as the microbial community compositional shift from HoD to D (Arnone et al, 2007; Bourne et al, 2009; Harvell et al, 2007; Negandhi et al, 2010). A non-parametric ANOVA (Kruskal Wallis) was used to identify the OTUs within the dataset which were significantly different

within the health states (Appendix Tables 11A-11D). Bacteria of interest which were found to be significantly different within the D population were (ordered by significance): Xanthomonadales, Chlamydiales, Leginellales, Verrucomicrobiales and Vibrionales.

To examine the potentially pathogenic bacteria found significant within the D population from the non-parametic ANOVA, sequences of interest were uploaded to the Basic Local Alignment Search Tool (BLAST), while using the National Center for Biotechnology Information (NCBI) microbial database and using a representative genome of the lowest taxonomic classification already provided by the open reference OTU picking with UCLUST. OTU RO13519 was identified as Legionellales and when BLAST searched the sequence closely matched *Coxiella burnetii* with an E-value of 1e-90 and a percent identity of 90% (Altschul, 1997; Appendix Table 11A). *Coxiella burnetii* is the causative agent of Q fever, which spreads through inhalation of the organism from infected secretions and is known to be one of the most infectious organisms according to the Center of Disease Control (Tigertt et al, 1961). The other Legionellales OTUs (269030) when BLAST searched the sequence matched *Coxiella burnetii*, with an E-value of 2e-121 and a percent identity of 97% (Altschul, 1997; Appendix Table 11A). The OTU (339560) was previously identified as Vibrio, when the sequence was BLAST searched it matched *Vibrio anguillarum* with an E-value of 1e-116 and a percent identity of 96% (Altschul, 1997; Appendix Table 11B). *Vibrio anguillarum* is a pathogen found in fish (salmon, eel, mullet, tilapia and catfish) and needs direct contact to spread (Frerichs and Roberts, 1989). The second Vibrio OTU (2670730) was found to evenly match both *Vibrio harveyi* and *Vibrio parahaemolyticus* with an E value of 5e-130 and a percent identity of 99% (Altschul, 1997; Appendix Table 11D). *Vibrio harveyi* is a microbe that lives within the gut of animals and has been found linked to multiple diseases in marine organisms (coral, oysters, prawns, lobsters, snook, barramundi, turbot, milkfish and seahorses)(Austin et al, 2006; Owens et al, 2006). *Vibrio parahaemolyticus* causes gastrointestinal illness in humans which is transmitted through fecal contact and ingestion of contaminated seafood (Ryan et al, 2004). When Verrucomicrobiales and Xanthomonadales OTUs were BLAST against the NCBI database, they did not match to any known bacteria. Therefore the OTUs my represent new strains of the bacteria. A lower taxonomic classification of Xanthomonadales;

Xanthomonas are a plant necrogenic pathogen which causes plant cell death and are associated with cruciferous vegetables and citrus trees (Van et al, 2002).

Permutational Multivariate Analysis of Variance Using Distance Matrices (Adonis):

Adonis was performed on the Bray-Curtis, weighted UniFrac and unweighted UniFrac statistics. The Bray-Curtis statistic exposed that there was significant differences within the microbial composition of the separate health states and the microbial composition between sampling locations (Table 6). Differences in the microbial composition due to location was expanded upon further by evaluating each health state individually by location to see if differences were more abundant in one health state versus the others. The only health state to show significant difference due to location was the D health state (Table 6). Then each health state was compared to the HC which showed that both the BL and D communities were significantly different while the HoD was not. The BL and D populations were then compared against each other to see how drastically SOB Disease progression changed the microbial community. Both were found to be significantly different from each other with an F-value < 0.05 (Table 6). This linked the progression of SOB with changing microbial communities of *X. muta*, and supports that they continued to

Permutational Multivariate Analysis of Variance Using Distance Matrices (Adonis)								
Bray-Curtis Distance Matrix			Weighted Distance Matirx			Unweighted Distance Matrix		
Samples Selected	Compared By	Significance	Samples Selected	Compared By	Significance	Samples Selected	Compared By	Significance
Entire Dataset	Location	0.003**	Entire Dataset	Location	0.004**	Entire Dataset	Location	0.009**
Healthy Control	Location	0.094	Healthy Control	Location	0.117	Healthy Control	Location	0.186
Healthy on Diseased	Location	0.24	Healthy on Diseased	Location	0.273	Healthy on Diseased	Location	0.055
Boundary Layer	Location	0.052	Boundary Layer	Location	0.049*	Boundary Layer	Location	0.1
Diseased	Location	0.017*	Diseased	Location	0.028*	Diseased	Location	0.006**
Entire Dataset	Treatment	0.001***	Entire Dataset	Treatment	0.003**	Entire Dataset	Treatment	0.001***
Healthy Control & Healthy on Diseased	Treatment	0.057	Healthy Control & Healthy on Diseased	Treatment	0.216	Healthy Control & Healthy on Diseased	Treatment	0.507
Healthy Control & Boundary Layer	Treatment	0.028*	Healthy Control & Boundary Layer	Treatment	0.192	Healthy Control & Boundary Layer	Treatment	0.031*
Healthy Control & Diseased	Treatment	0.008**	Healthy Control & Diseased	Treatment	0.045*	Healthy Control & Diseased	Treatment	0.01**
Healthy on Diseased & Boundary Layer	Treatment	0.06	Healthy on Diseased & Boundary Layer	Treatment	0.381	Healthy on Diseased & Boundary Layer	Treatment	0.013*
Healthy on Diseased & Diseased	Treatment	0.001***	Healthy on Diseased & Diseased	Treatment	0.001***	Healthy on Diseased & Diseased	Treatment	0.001***
Boundary Layer & Diseased	Treatment	0.003**	Boundary Layer & Diseased	Treatment	0.013*	Boundary Layer & Diseased	Treatment	0.001***
Significance codes: 0***, 0.001**, 0.01*, 0.05 '.', 0.1' ', 1								

Table 6: Permutational multivariate analysis of variance using distance matrices (Adonis) for Bray-Curtis, weighted Unifrac and unweighted Unifrac. Table was created using data analyzed by QIIME 1.8.0 (Caporaso, 2010).

change over time. Sampling time in relation to disease progression was likely a factor that affected which bacteria were present within the diseased microbial community.

The unweighted UniFrac statistic showed significant differences in the relative abundance of microbes when comparing the dataset by health states and location (Table 6). Comparing the separate health states individually by sampling location only showed that only the D health state was significant different. When comparing the health states to each other both the BL and the D where significantly different from all other health states.

When Adonis was performed on the weighted UniFrac statistic it also showed that significant differences existed within the presence and absence of microbes when the health states and the sampling location influences where compared. When each health state was compared both the BL and the D health states had microbial communities that were significantly different due to the influence of sampling location (Table 6). This supports the existence of a correlation between sampling location and the phylogenetic diversity found within the microbial community assemblages within the BL and D populations. When comparing the health states to each other the D population was found to be significantly different from all other health states. The significant difference within the D population supported a direct connection between mesohyl destruction and the change in phylogenetic diversity present within the mesohyl.

Taxon Summary of Environmental Samples:

Environmental samples were taken to assess the microbial communities within the surrounding marine water and soil. One marine water sample was taken at the Boynton Beach sampling location and it should be noted that the sample when sequenced and re-sequenced only produced 200 sequence reads. Due to the low concentration of sequence reads associated with the sample caution should be taken when using it to make extrapolations. Phyla present within the water sample were: Euryarchaeota, Actinobacteria,

Water Taxon	Percentage
k__Archaea;p__Euryarchaeota	6
k__Bacteria;p__Actinobacteria	6.5
k__Bacteria;p__Bacteroidetes	15
k__Bacteria;p__Cyanobacteria	35
k__Bacteria;p__Planctomycetes	2.5
k__Bacteria;p__Proteobacteria	34.5
k__Bacteria;p__Verrucomicrobia	0.5

Table 7: Taxonomic summary for the Boynton Beach water sample at the phylum level. Table was created using data analyzed by QIIME 1.8.0 (Caporaso, 2010).

Bacteroidetes, Cyanobacteria, Planctomycetes, Proteobacteria and Verrucomicrobia (Table 7). The most abundant phyla within the water column was Cyanobacteria which compromised 35% of the community, then Proteobacteria with 34.5 % and Bacteroidetes with 15%. The presence of Verrucomicrobia within the water column could support Scheuermayer et al, 2006 claim that the sponge acquired Verrucomicrobia from the surrounding environment and that the phylum is not a sponge symbiont. A taxonomic summary at the genera level or lowest taxonomic level possible can be found in Table 8.

Water Taxon	Percentage
k__Archaea;p__Euryarchaeota;c__Thermoplasmata;o__E2;f__Marine group II;g__	6
k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__OCS155;g__	5.5
k__Bacteria;p__Actinobacteria;c__Acidimicrobiia;o__Acidimicrobiales;f__wb1_P06;g__	1
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__;f__;g__	0.5
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__;g__	0.5
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae;Other	0.5
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__Flavobacteriaceae;g__	13.5
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Cryptophyta;f__;g__	1
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Haptophyceae;f__;g__	3.5
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Stramenopiles;f__;g__	5
k__Bacteria;p__Cyanobacteria;c__Synechococcophycideae;o__Synechococcales;f__Synechococcaceae;g__Prochlorococcus	25.5
k__Bacteria;p__Planctomycetes;c__Phycisphaerae;o__Phycisphaerales;f__;g__	2.5
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Hyphomicrobiaceae;g__	1
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodobacterales;f__Rhodobacteraceae;g__	5.5
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhodospirillales;f__Rhodospirillaceae;g__	4.5
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__;g__	0.5
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__Pelagibacteraceae;g__	4.5
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__mitochondria;g__	0.5
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Myxococcales;f__;g__	0.5
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Sva0853;f__S25_1238;g__	0.5
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Alteromonadales;f__OM60;g__	2
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__HTCC2188;f__;g__	0.5
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Oceanospirillales;f__Halomonadaceae;g__Candidatus Portiera	14
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Vibrionales;f__Vibrionaceae;g__Vibrio	0.5
k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__MSBL3	0.5

Table 8: Taxonomic summary for the Boynton Beach water sample at the genus or lowest taxonomic level known. Table was created using data analyzed by QIIME 1.8.0 (Caporaso, 2010).

Soil samples were taken at both the Boynton Beach and the Fort Lauderdale sampling locations. The quantity of OTUs produced by the two soil samples were far apart. The Boynton Beach soil sample produced 979 OTUs and the Fort Lauderdale soil sample produced 14,597 OTUs. The phyla present within the soil samples were far greater than the water sample. Phyla included Crenarchaeota, Euryarchaeota, Acidobacteria, Actinobacteria, Bacteroidetes, Caldithrix, Chlamydiae, Chlorobi, Chloroflexi, Cyanobacteria, Elusimicrobia, Fibrobacteres, Fusobacteria, GN02, GN04, Gemmatimonadetes, KSB3, Lentisphaerae, NKB19, Nitrospirae, OD1, OP1, OP3, OP8, PAUC34f, Planctomycetes, Poribacteria, Proteobacteria, SBR1093, Spirochaetes, TM6,

Tenericutes, Thermi, Verrucomicrobia, WPS-2, and WS3 (Table 9). The phyla which were previously exclusive to the D state (Caldithrix, Chlorobi, Fibrobacteres, GN02, KSB3, OP1, OP3, OP8 and WS3) were also found within the soil samples. Phyla Elusimicrobia and GN04 were solely found within the soil samples. The most abundant phyla within the soil samples were Proteobacteria which compromised 40.6% of the community, then Bacteroidetes with 13.3% and Euryarchaeota with 11.1%. Genera found almost exclusively within the soil samples from Broward include Luteolibacter, MSBL3, Persicirhabdus, Rubritalea and Verrucomicrobium from the phylum Verrucomicrobia. Persicirhabdus was the only genera from the phylum Verrucomicrobia subclass 4 which was found in the Boynton Beach soil sample. Since the Verrucomicrobia genera present in the BL and the D states was not found in both soil samples it remain unclear whether the soil at the sampling locations could have been sources for Verrucomicrobia found in *Xestospongia muta*. A taxonomic summary at the genera level or lowest taxonomic level possible can be found in Appendix Table 12A-J.

Soil Taxon	Percentage
Unassigned;Other	2.253
k__Archaea;p__Crenarchaeota	11.153
k__Archaea;p__Euryarchaeota	0.021
k__Bacteria;Other	0.010
k__Bacteria;p__	0.027
k__Bacteria;p__Acidobacteria	2.012
k__Bacteria;p__Actinobacteria	3.463
k__Bacteria;p__Bacteroidetes	13.304
k__Bacteria;p__Caldithrix	0.082
k__Bacteria;p__Chlamydiae	0.075
k__Bacteria;p__Chlorobi	0.058
k__Bacteria;p__Chloroflexi	0.702
k__Bacteria;p__Cyanobacteria	8.931
k__Bacteria;p__Elusimicrobia	0.007
k__Bacteria;p__Fibrobacteres	0.010
k__Bacteria;p__Firmicutes	4.091
k__Bacteria;p__Fusobacteria	0.024
k__Bacteria;p__GN02	0.007
k__Bacteria;p__GN04	0.041
k__Bacteria;p__Gemmatimonadetes	0.677
k__Bacteria;p__KSB3	0.027
k__Bacteria;p__Lentisphaerae	1.458
k__Bacteria;p__NKB19	0.031
k__Bacteria;p__Nitrospirae	0.759
k__Bacteria;p__OD1	0.003
k__Bacteria;p__OP1	0.010
k__Bacteria;p__OP3	0.055
k__Bacteria;p__OP8	0.027
k__Bacteria;p__PAUC34f	0.150
k__Bacteria;p__Planctomycetes	5.947
k__Bacteria;p__Poribacteria	0.003
k__Bacteria;p__Proteobacteria	40.645
k__Bacteria;p__SBR1093	0.455
k__Bacteria;p__Spirochaetes	1.707
k__Bacteria;p__TM6	0.021
k__Bacteria;p__Tenericutes	0.014
k__Bacteria;p__Thermi	0.003
k__Bacteria;p__Verrucomicrobia	1.568
k__Bacteria;p__WPS-2	0.003
k__Bacteria;p__WS3	0.147
k__Bacteria;p__[Caldithrix]	0.017

Table 9: Taxonomic summary for the Boynton Beach and Fort Lauderdale soil samples at the phylum level. Table was created using data analyzed by QIIME 1.8.0 (Caporaso, 2010).

DISCUSSION

Changes Present within the Microbial Communities

The Chao1, ACE and Shannon Index of Diversity statistics all illustrated that the D state had the greatest species richness of all the health states, which was also supported by the D state having the largest OTU counts. The two sample t-test from the Chao1 and the Shannon Index of Diversity statistics showed that the increase in species richness found in the D state was significantly different from the BL and the HoD health states. While the two sample t-test for ACE showed that the D state was also significantly different then the HC states. Overall the alpha diversity statistics that took into account species richness showed that the D state was more diverse when compared to the other health states. This supported the notion that microbial community present in *X. muta* was affected by SOB, which altered the microbiome away from the community seen in the HC and similarities within the HoD. Whether the community shift was a directly caused by the colonization of the causative agents or due to the deteriorating mesohyl which spurred opportunistic bacteria colonization is unknown and further research is needed to understand this shift.

Beta diversity statistics (Jaccard Abundance Index, Bray Curtis Dissimilarity and Weighed UniFrac PCoA plots) all displayed intermixing between the samples that comprised the health states. This supported a level similarity within the microbial communities between the different health states. Some similarities between the microbial communities were expected since diseased sponges had three separate samples taken from the same sponge, therefore similarities between those microbial communities are expected. *X. muta* as a species is also known to pass on beneficial bacteria to the offspring; so as a species they should also share similar microbial community (Montalvo and Hill, 2011; Freeman and Thacker, 2005; Schmitt et al, 2008; Sharp et al, 2007; Collin et al, 2010; Taylor et al, 2007b). The UniFrac significance test for the weighted UniFrac showed that when the microbial communities of the HC and HoD were compared they were not significantly different but when the BL or the D state was compared to any other health state; even each they were significantly different. The variation present between health states could have been caused by the new microbial inhabitants which could adapt to the "new disturbed sponge mesohyl environment". These opportunistic microbes would most likely come from the surrounding environment which would make these new inhabitants

more location specific. This further links SOB with not only altering the microbial community but changing it to such a degree that the communities were significantly different.

Evaluating the core microbiome helped to better understand how SOB Disease affected the microbial composition of *X. muta*. The shift of losing SAR202, the general reduction in core microbiome OTUs, and the addition of phyla Thermi within the HoD core microbiome suggests that while the mesohyl on the diseased sponge may visually look healthy but is biologically being affected by SOB Disease. The loss of sponge specific phyla Poribacteria within the BL and complete loss of a structured core microbiome within the D health state futher emphasises that SOB Disease had shifted the microbiome away from what is naturally found within *X. muta* (Fieseler, et al, 2004; Taylor et al, 2007b; Lafi et al, 2009).

.

Comparing the Microbial Community of Xestospongia muta with the Assessments of Previous Publications:

Comparing previous research which characterized the microbial community of *X. muta* will provide a better understanding of the true microbial community present that was not influenced by the presence of disease. The HC samples for the study were taken within the same sampling location as the D sponges. To the best of our ability healthy sponges were visually assessed and sampled, but since the sampling location was not revisited or the HC sponges tagged for reassessment there remains reasonable possibility that the HC sponges could have been in the very early stages of SOB infection. Therefore comparing previous microbial community assessments will provide a better understanding of healthy *X. muta* microbial populations and allow a better assessment of the microbial shifts caused by SOB Disease.

The microbial community of *X. muta* was previously assessed by Montalvo et al, 2011. The study found the following phyla present within the sponges: Chloroflexi (42%), Acidobacteria (12%), Actinobacteria (12%), Cyanobacteria (9%), DeltaProteobacteria (8%), Bacteroidetes, Deferribacteres (3%), Nitrospira, GammaProteobacteria, Gemmatimonadetes, unclassified bacteria, Proteobacteria, BetaProteobacteria, AlphaProteobacteria, TM7, Deinococcus-Thermus, Planctomycetes, and TM6 (Montalvo et al, 2011). All of these phyla were found within the HC samples but

Chloroflexi comprised 19% of our community assessment and did not dominate the microbial population, which was reported in Montalvo et al, 2011. The assessment by Montalvo et al, 2011 also failed to included phyla Crenarchaeota (third most abundant phylum within this study), Euryarchaeota, AnK6, Chlamydiae, H-178, Lentisphaerae, NKB19, OD1, PAUC34f, Poribacteria (recognized as a sponge specific phylum), SAR406, SBR1093 and Verrucomicrobia. Verrucomicrobia specifically were reported in another *Xestospongia sp.* sampled but were not found within the *X. muta* samples (Montalvo et al, 2011). The study only sampled three sponges for the analysis, more sponges may have been needed to fully capture the rare community present within *X. muta;* like Verrucomicobia.

Another assessment of the microbial community of *X. muta* was completed by Fiore et al, 2013. Within this study a total of 17 phyla were found: Acidobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Deferribacteres, Deinococcus-Thermus, Firmicutes, Gemmatimonadetes, Nitrospirae, Planctomycetes, Proteobacteria, Spirochetes, Verrucomicrobia, Crenarchaeota, Poribacteria, TM7 and SBR1093 (Fiore et al, 2013). This community composition had more similarities with the more abundant bacterium within our HC but overall was still missing some of the rare microbial communities found within our dataset. These phyla include Actinobacteria, AncK6, Chlamydiae, H-178, Lentisphaerae, NKB19, OD1, PAUC43f, SAR406, and Spirochaetes. This study was able to find the presence of Verrucomicrobia within all three separate sampling locations. The relative abundance of Verrucomicrobia within the Fiore samples was equally minor between all three sample locations (Cayman Islands, Bahamas, Key Largo), which corresponds to our measurements within the HC and the HoD communities. The Fiore study also found that the three separate sampling locations influenced the microbial communities found in *X. muta*. The effect of sampling location was also evident with the Boynton Beach and the Fort Lauderdale sampling locations of this study. Proteobacteria was the dominate phylum within the communities of the Fiore et al, 2013 study which does not match our assessment or Montalvo et al, 2011.

Overall the dominate phyla found within these studies have been found within *X. muta* off of South Florida. There were differences within the rare microbial communities between the studies. One phyla that both previous studies found but was

absent within our analysis was Deferribacteres. The phyla present within our study but absent within the others may be the results of seasonal shifts, location specific bacteria that was abundant in the water column at the time of sampling and the effects of SOB Disease on the microbial communities.

Bacteria Associated with Sponge Orange Band Disease:

The microbial assemblages present within the different health states highlighted the presence of pathogenic bacteria that was either exclusive to the BL and the D states or increased in overall relative abundance that deemed it significant. Phyla which were only found within the microbial communities of the BL and the D samples were associated with SOB Disease (Armatimonadetes, Caldithrix, Chlorobi, Fibrobacteres, Fusobacteria, GN02, KSB3, OP1, OP2, OP8, SR1, TM6, Tenericutes, WPS-2 and ZB3). To better understand these bacteria the possible roles and functions that each could provide to the sponge community a brief synopsis for each has been provided. Armatimonadetes (also known as OP10) consists for 4 subgroups, the first is commonly derived from mesophilic terrestrial ecosystems (soils, lakes, manure run-off), while subgroups 2-4 are affiliated with thermophilic environments (Tamaki et al, 2011). The class Fimbriimonadia was found in both the BL and the D populations, but it was found in a greater quantities within the D state (seven samples compared to four in the BL). Fimbriimonadia belonged to subgroup 4 and in previous studies has been found within soils, freshwater and activated sludge from wastewater treatment plants (Im et al, 2012). It has abilities as an anaerobic digester, anammox reactor, aerobic bioreactor and microaerobic bioreactor (Im et al, 2012). The known functions of Armatimonadetes suggests that it may be using the diseased mesohyl as an energy source.

Caldithrix are thermophilic and anaerobic bacteria which have been associated with marine sponges (Hardoim et al, 2014; Simister et al, 2012). The phylum contains bacteria such as *Caldithrix abyssi* which are nitrate-reducing bacteria, found from a hydrothermal vents (Miroshnichenko, 2003). The bacteria was only found within three samples from the D population.

Chlorobi are green sulfur bacteria which are mostly phototrophic and found in a variety of environments from aquatic sediments to hot springs (Ross, 2012). The lower taxonomic classification of OPB56 is an unclassified sulfur bacteria which was found

within six Yellowstone National Park spring mesothermal mats (Ross, 2012). Chlorobi were only found within five samples from the D population.

Fibrobacteres are commonly found within the gastrointestinal tract of animals and are widely distributed. The bacteria was only found within one sample from the D population.

Fusobacteria was found in one sample from the BL population and thirteen samples from the D population (Appendix Table 4E). The phylum includes anaerobic bacteria which have been associated with coral, human and animal diseases (Bennett et al, 1993; Schmitt et al, 2012; Thurber et al, 2009). Within the D population Fusobacteria was classified to the genus level as *Cetobacterium* and *Propionigenium*. *Cetobacterium* have been found within the intestinal track of marine mammals, fish, and humans (Tsuchiya et al, 2008). The genus *Propionigenium* contains anaerobic bacteria found in anoxic marine sediment, which can ferment both succinate and carbohydrates (Schink, 2006). These bacteria were more abundant within the D population where it was found in ten samples compared to one sample within the BL population.

Candidate phylum KSB3 is a filamentous bacteria which was a bulking agent found in the anaerobic sludge within organic wastewater discharged from a sugar manufacturing factory (Yamada, 2007).

TM6 was further classified to the class level as SJA-4, which was found in three samples from the BL population and eight samples from the D population. The candidate phylum TM6 was first discovered in a biofilm found within a sink drain of a hospital in 1996 and is believed to live in treated water systems globally (McLean et al, 2013). TM6 has previously been found associated with marine sponges (Siegl et al, 2010). The bacteria is believed to have a symbiotic relationship with an unknown host due to the loss of biosynthetic pathways in its genome. Facultative or obligate symbionts that are possible hosts for TM6 include Chlamydiae, GammaProteobacteria (Legionella and Francisella), AlphaProteobacteria (*Rickettsia*), Spirochaetales (Borrelia) and Planctomycetaceae (McLean et al, 2013). All of these phyla have been found within *Xestospongia muta*, which may explain its presence within the sponge mesohyl but TM6 was first seen within the BL population and was not present within the HC or the HoD populations. SJA-4 have been found in fluidized bed reactors used to treat polluted soil, sediments and aquifers

contaminated with chlorobenzenes (von Wintzingerode et al, 1999). *SJA*-4 was found related to the genus Dehalobacter which are anaerobic bacteria known to grown by reductive dechlorination of chlorinated ethane (von Wintzingerode et al, 1999). Therefore *SJA*-4 may have a role in anaerobic transformation or degradation of chlorinated organic compounds within polluted environments but more research is needed to truly define their role in the environment (von Wintzingerode et al, 1999). Nine OTUs were found within the D population and three were found in the BL population.

Tenericutes was further classified to Mycoplasmatales, which is a microorganism that does not possess a true cell wall and include pathogenic and saprophytic species (Ryan et al, 2004). Mycoplasmatales contains three genera: *Mycoplasma*, *Hepatoplasma* and *Ureaplasma*. Diseases caused my *Mycoplasma* species are spread through contact with infected fluids and can infect plants (corn stunt and white leaf of sugarcane), animals (contagious bovine, caprine pleuropneumonia, enzootic pneumonia, and chronic respiratory disease) and humans (atypical pneumonia and pelvic inflammatory diseases) (Maramorosch et al, 1970; Nicholas et al, 2009). *Mycoplasmatales* was found in three samples from the BL population and three samples from the D population, but two more samples from the D population were found to contain Tenericutes classified to Anaeroplamataceae (family level).

WPS-2 was first found in a single cloned sequence from a microbial community of highly polychlorinated biphenyl (PCB) polluted soil (Nogales et al, 2001). The most similar 16S rRNA sequences at the time were sequence SJA-5 and SJA-22 which are related to previously found SJA-4 sequence (Nogales et al, 2001). WPS-2 was later found within the rare phyla associated with anaerobic digesters used at treatment plants to combat wastewater sludge (Riviere et al, 2009). The bacteria were found within two samples form the BL population and three samples within the D population.

OP1 was found within the anoxic bulk soil of flooded rice microcosms (Derakshani, 2001). The bacteria was found in one sample from both the BL and D populations.

ZB3 is a candidate division proposed from the microbial community analysis of the sulfur rich Zodletone spring through cloning and sequencing 16S rRNA genes (Elshahed et al, 2003). The candidate division includes two OTUs collected from Zodletone spring (ZB15 and ZB21) and a clone from the oxic-anoxic boundary layer from the hyper-saline

Mono Lake (Elshahed et al, 2003). Past research suggests that ZB3 may have a role in sulfur cycling. ZB3 was found within one sample from both the BL and the D population.

The phyla which were exclusive to the BL and the D populations were usually always more abundant within the D state. The D population also had a greater diversity and abundance of new phyla (Appendix Figure 2). This further provided evidence that the progression of SOB Disease not only physically altered the sponge mesohyl environment but also the microbial microbiome.

One phylum which was found to change drastically in abundance between the healthy (HC and HoD) and infected mesohyl (BL and D) was Verrucomicrobia, which associated its increase with the SOB progression. The proportional percentage that Verrucomicrobia comprised within the HC samples was 0.10%, then it diminished to 0.04% within the HoD samples, it reached its maximum of 4.69% within the BL samples and still presented elevated abundances within the D community with 2.50% (Table 1A and Table 2A). Since the BL state included the line of advancing disease it was possible that the bacteria could be concentrated there and aid in infecting HoD sponge mesohyl. Verrucomicrobia OTUs 574883, 560276 and 113182 were deemed significantly different (P value < 0.05) by a nonparametric ANOVA and found exclusively within the D and/ or BL states but missing in the HC state (Appendix Table 6A-D). Verrucomicrobia are found within a wide range of environments such as soil, fresh water, marine environments, and even the intestinal tract of humans (Dubourg et al, 2013; Freitas et al, 2012; Yoon, 2011). The phylum has intracellular compartments; a trait it shares with Planctomycetes and Chlamydiales (Lee et al, 2009). The different bacterial strains that comprised the phylum utilize sources of carbon and methane compounds for energy. Genera present within the *X. muta* dataset include: *Rubritalea*, *Luteolibacter*, *Persicirhabdus* and *Verrucomicrobium*. *Rubritalea marina* gen. nov., was previously isolated from Mediterranean sponge *Axinella polypoides* (Scheuermayer et al, 2006). The 16S rRNA gene sequence analysis for the strain and other Verrucomicrobial collected from the *Axinella polypides* showed that the sequences were more closely related to seawater 16S rRNA gene sequences from the collection site, rather than the sponges themselves (Scheuermayer et al, 2006). This revealed that the bacteria was introduced to the sponge through seawater filtering versus being a sponge symbiont. *Rubritalea squalenifaciens* was another strain isolated from

marine sponge *Halichondria okdai* and was found to produce diapolycopenedioic acid xylosyl esters, which have the ability to act as a potent antioxidant (Shindo et al, 2008). *Persicirhabdus* was more frequently found within sediment and has the ability to reduced nitrate to nitrite (Freitas et al, 2012; Yoon et al, 2008). *Luteolibacter* has been found within marine ecosystems, driftwood, red algae, tundra soil and on the surface of medical leeches; but little was known about its functional properties (Glaeser et al, 2012; Jiang et al, 2012; Yoon et al, 2008). Even with the large distribution and abundance of Verrucomicrobia little is known about the functions and interactions of the phylum; especially within the marine sponge environment (Freita et al, 2012; Rappe and Giovannoni, 2003).

Planctomycetes was another phylum which increased dramatically in overall abundance, from the HC population (0.059%) to the D population (4.44%) [Appendix Table 6-9]. The bacteria was found in seven individuals from Boynton Beach and two samples from Fort Lauderdale. The specific bacterial strain Pirellulaceae at the family level accounted for 2.772% of the Planctomycetes phylum. Planctomycetes OTUs 332383, 4355495, 240923, 2220974, 4479622, 607278, 4437176, New.CleanUp.ReferenceOTU10739, 338967, 2658746, New.CleanUp.ReferenceOTU10537, New.CleanUp.ReferenceOTU2880, 311201 and 314251 were deemed significantly different (P value < 0.05) by a nonparametric ANOVA and found exclusively within the D and/ or BL states but missing in the HC state (Appendix Table 11A-D) This phylum in previous research was also found in small quantities within Healthy *Xestospongia muta*, which highlighted the unusual abundance seen in D mesohyl (Montalvo et al, 2011; Fiore et al, 2013). Planctomycetes are known to degrade polymers and anammox (Anaerobic Ammonium Oxidation) which utilizes fixed nitrogen (nitrite) and coverts it to dinitrogen gas (Freitas et al, 2012). Overall the D health state contained an abundance of bacterium not seen within the HC samples (Appendix Table 9A-L).

Overall the bacteria associated with SOB Disease provided a direct picture of how the microbiome of *X. muta* shifted with infection and disease progression. It is unclear if the bacteria exclusive to the BL and the D populations are opportunistic bacteria utilizing an available energy source, secondary infections or are simply present within the environment. More research is needed to understand the relationships between the

microbial inhabitants and there host, but the bacteria associated with SOB provide a first look at the microbial community of diseased *X. muta*.

Annual Meteorological Data for South Florida 2012

Climate patterns and extreme weather events influence environmental conditions which then affect the organisms that inhabit them. This is more evident in terrestrial environments where habitat variation is visually apparent. But the habitat conditions in marine environments similarly control species ranges, survival, fecundity and dispersal (Bruno et al, 2007). To fully take into consideration all factors that were effecting *X. muta* and its microbial community during the SOB Disease outbreak and during research sampling, the meteorological data from NOAA was analyzed for the year of 2012. The month of May was of special interest as the disease was first seen at the end of April 2012 and was confirmed by the first week of May by Reef Rescue. It was also in the month of May that both Boynton Beach and Fort Lauderdale samples were collected.

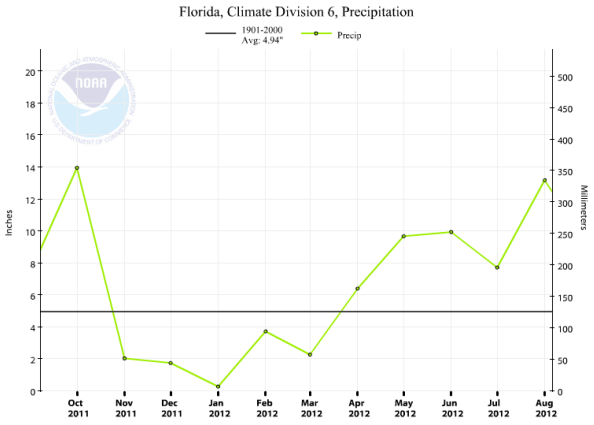


Figure 39: Monthly precipitation for the South Florida coastline. This graph show the steep rise in precipitation in March -April and its continued increase into May 2012. Image taken from NOAA, National Climatic Data Center: http://www.ncdc.noaa.gov/cag/time-series

Meteorological conditions were measured at the NOAA station in Fort Lauderdale Beach, FL USA (26.1408°, -80.1061°). A comparison of monthly precipitation rates in the South Florida Coastline region showed below average precipitation during November 2011 with a monthly total of 2 inches versus the monthly average of 4.95 inches (Figure 39). From November 2011 till the middle of March 2012 below average precipitation continued until April when monthly precipitation reached 6 inches. The graph in Figure 39 illustrates a steep increase in precipitation which started in March and continued through May. Precipitation records showed that the 10th of May had the highest daily precipitation of the whole month; which was the day before the first SOB Disease sample collection in

Boynton Beach took place (Table 10). The records also showed a large precipitation event on the 21st of April which the precipitation level was not surpassed until July and was around the same time that SOB was first noticed. The precipitation rates in Florida for the month of May when compared against the continental United States also displayed above average precipitation for the South Florida coastline (Figure 40). Overall precipitation rates in South Florida a month before SOB detection and during sampling were elevated and stayed above average until the month of October 2012.

Precipitation for 2012 (inches)			
Month	Total Precipitation	Greatest Daily Precipitation	Day of Greatest Precipitation
January	0.23	0.19	30
Febuary	3.09	1.23	7
March	2.39	1.02	8
April	9.21	2.50	21
May	**9.06**	**1.55**	**10**
June	10.49	2.22	1
July	7.63	3.47	17
August	10.56	5.6	27
September	6.73	1.48	23
October	7.46	3.32	2
November	0.34	0.15	13
Annual Precipitation	67.19		
Mean	**6.11**		

Table 10: Annual precipitation summary for 2012 in inches. Data taken from NOAA's Annual Climatological Summary 2012 for station (COOP:083168): Fort Lauderdale Beach, FL US.

The average monthly temperature records for the South Florida Coastline in 2012 showed that temperatures began to rise above average in April when SOB was first noticed and continued to rise until their peak in August 2012 (Figure 41). Sea surface temperatures also increased in temperature between the first sample collection on May 11th and the second sample collection on May 29th (Figure 42). Temperatures in South Florida for the month of May were also elevated when compared to the Continental United States (Figure 43). Previous research which compared the temperature thresholds of microbial symbionts for marine sponge *Rhopaloeides odorabile* found that dramatic shifts in the microbial communities were not present until sponges were exposed to temperature between 31 and 33°C (Webster et al,

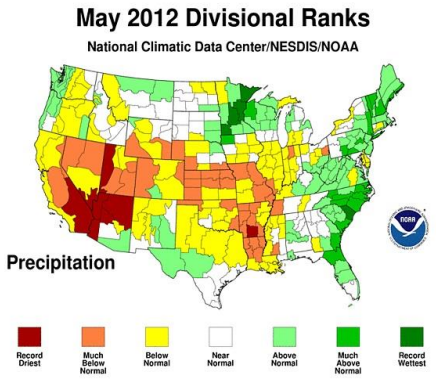


Figure 40: U.S. divisional precipitation ranks for May 2012, which takes into account historical climate analyses from 1895. Data taken from NOAA, National Climatic Data Center: http://www.ncdc.noaa.gov/temp-and-precip/maps

2008). A water temperature of 33ºC was uninhabitable, so that sponges ejected cellular material and displayed less than 10% surface necrosis after 24 hours (Webster et al, 2008). At 27ºC which was the closest to the sea surface temperatures during SOB sampling no bacterial shifts were evident (Webster et al, 2008). This helps to verify that the microbial community shift that were found within the BL and D health states were in response to SOB and not high temperatures. In another study that examined the expression of the *hsp70* "the stress indicator gene" in response to increased water temperatures found that *X. muta* exposed to temperatures at 30ºC for 1.5 hours demonstrated higher *hsp70* expression then control sponges and demonstrated mortality in less than 15 hours (Lopez-Legentil et al, 2008). Since the sea surface temperatures around the time period of SOB sample collection were above 27ºC it safe to assume that *X. muta* was suffering some degree stress due the high sea surface temperatures during SOB Disease outbreak.

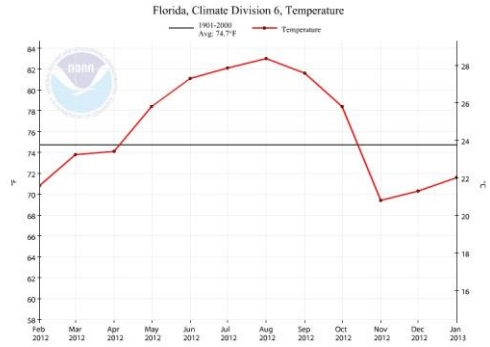


Figure 41: A.) The South Florida's coastline temperature per month for 2012. This shows that temperatures started to rise above the average in April when disease was first noticed and continued to rise till August 2012. Image taken from NOAA, National Climatic Data Center: http://www.ncdc.noaa.gov/cag/time-series

To examine the likely hood that weather events influenced the emergence of SOB Disease, other SOB outbreaks in Florida had the corresponding weather conditions compared. In the Angermeier et al, 2011 publication SOB was found at two reefs (Conch

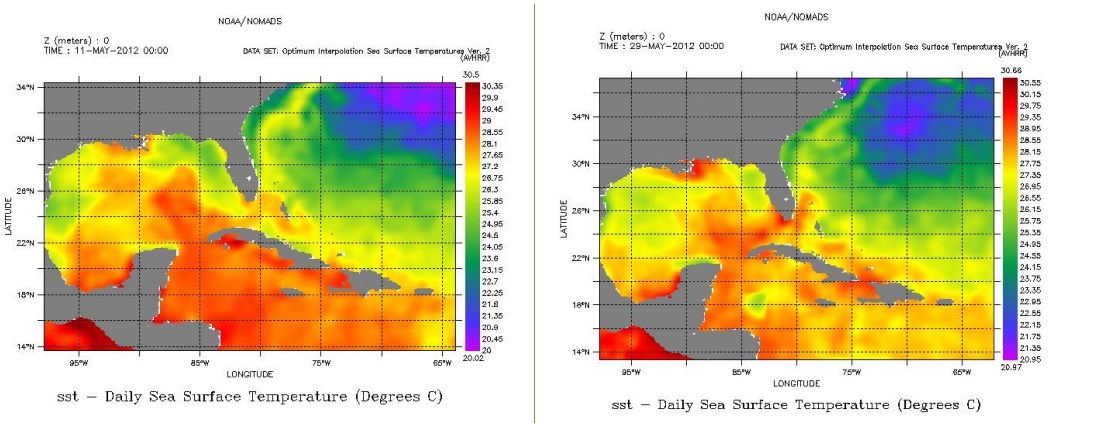


Figure 42: Sea surface temperature during sample collection. Data was retrieve from NOAA's Optimum Interpolation Sea Surface Temperature Ver. 2 (AVHRR).

Reef and Dixie Shoals Reefs) in Key Largo during 2007 and 2009. The first sampling occurred November 11th 2007 on Conch Reef (24º56'863''N/ 80º27'230''W), and then again on December 8th 2007 at Dixie Shoals Reef (25º04'316''N/ 80º19'074''W) (Angermeier et al, 2011). The two months preceding the SOB sampling; October and September had the highest total precipitation of the entire year (Table 11 and Figure 44). Sampling also occurred during September 11th and the 17th at Conch Reef, 2009. The 2009 year started out with low monthly precipitation from January to May, which had an overall monthly mean of 0.86 inches for those five months. The precipitation greatly increased in June and July, with September (the month during sampling) having the highest total monthly precipitation for 2009. August had lower precipitation when expected but had the second largest daily precipitation event of the year at 2.95 inches (Figure 44). Sea surface temperatures were also elevated months prior to the disease outbreak in 2007 and throughout the fall to September of 2009. Figure 45 displays the sea surface temperatures for all three years: 2007, 2009 and 2012. This showed that elevated temperatures were present prior and during both the 2009 and 2012 outbreaks. While in 2007 the sea surface temperatures where last elevated in September with a continuous decline throughout the rest of the year, including sampling in November and December.

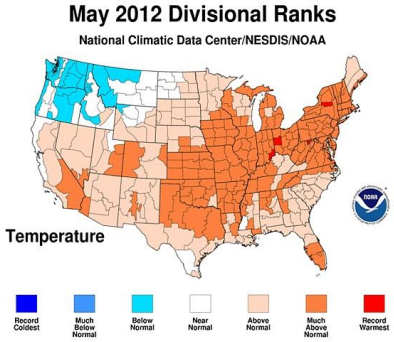


Figure 43: Temperature divisional ranks during May 2012 for the continental United States, which takes into account historical climate analyses from 1895. Data taken from NOAA, National Climatic Data Center: http://www.ncdc.noaa.gov/temp-and-precip/maps

Precipitation for Key Largo 2007 & 2009 (inches)

Month	2007	Greatest Daily Precipitation 2007	2009	Greatest Daily Precipitation 2009
January	1.09	0.65	0.30	0.28
Febuary	4.83	3.28	1.62	1.60
March	0.83	0.42	0.53	0.33
April	3.18	1.84	0.01	0.01
May	2.22	0.80	1.85	1.25
June	7.01	2.10	**5.76**	1.50
July	4.29	1.70	**6.95**	1.37
August	4.09	1.60	3.50	**2.50**
September	**10.90**	**3.30**	7.57	**2.95**
October	**11.05**	**3.70**	1.03	0.40
November	0.59	0.30	1.23	0.74
December	2.57	1.50	7.59	2.00
Annual Precipitation	52.65		37.94	
Mean	**4.38**		**3.16**	

Table 11: Annual precipitation records from NOAA's Key Largo station: City:US120020, during 2007 and 2009. Data was retrieved from NOAA, National Climatic Data Center.

When comparing the weather associated with SOB Disease outbreaks in Florida waters during 2007, 2009 and 2012 each outbreak was proceeded by elevated sea surface temperatures and a large or extended period of precipitation. Raised sea surface temperatures have already been proven to cause stress in *X. muta* (Lopez-Legentil et al, 2008). Similarity elevated sea surface temperatures have previously been linked to disease outbreaks in corals because stress diminishes the immunity of an organism; increasing susceptibility when exposed to pathogens (Bruno et al, 2007; Lafferty et al, 2003). It is unknown if stress lowers the immunity of a sponge and what if any immune response sponges process, but the increase in sponge diseases that corresponds with elevated sea surface temperatures are a coincidence that should not be overlooked. The other pattern of large and/or prolonged precipitation events just before the emergence of SOB could explain why the disease does not occur seasonally (can occur in early spring and late fall through winter) or constantly reoccurs within an area. South Florida is a highly developed area dominated with impervious surfaces, which contribute to runoff during rainfall. Runoff can influence water quality by introducing sediment, pollution, toxic chemicals, microbial

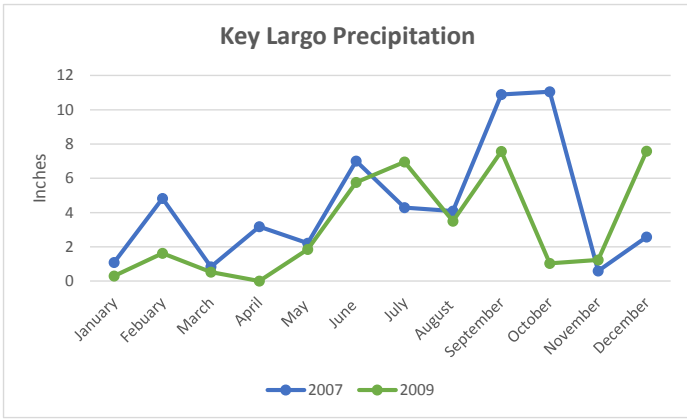


Figure 44: Monthly precipitation comparison of 2007 and 2009 retrieved from NOAA's Key Largo station: City:US120020 records. Data taken from NOAA, National Climatic Data Center: http://www.ncdc.noaa.gov/temp-and-precip/maps

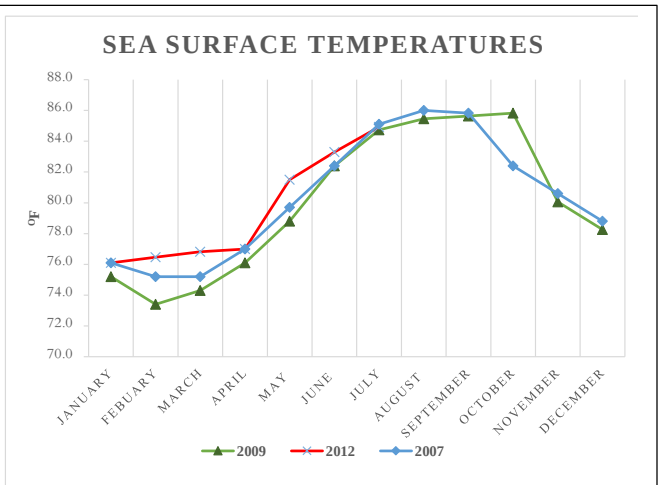


Figure 45: Sea surface temperature comparison estimates using data retrieved from NOAA's Extended Reconstruction Sea Surface Temperatures (ERSST 3b), http://nomads.ncdc.noaa.gov/las/getUI.do

pathogens and nutrients into the marine environment (Jiang et al, 2001; Mitsova et al, 2011). Previous research has found polycyclic aromatic hydrocarbons (PAHs) within the sediments of canals that drain into Biscayne Bay, FL (Mitsova et al, 2011). A direct correlation between the magnitude PAHs pollution run off and prolonged periods of drought was found during 2010-2011 (Mitsova, 2011). Precipitation events after extended periods of drought therefore have the ability to introduce immune compromised *X. muta* to causative agents. SOB Disease is likely influenced by a variety of factors that produce the physical response of mesohyl narcosis. But the precipitation patterns that occur before the outbreak of SOB for the three years the disease has been found in U.S. waters provides evidence of an environmental condition that coincides with SOB Disease emergence.

CONCLUSIONS

Evaluating the microbial community assemblages of healthy and SOB Diseased *Xestospongia muta* identified the microbiome and the presence of a microbial community shift which followed SOB progression. As SOB Disease progressed the core phyla (Crenarchaeota, Chloroflexi and Cyanobacteria) which had previously dominated the HC and HoD states decreased in abundances, while Bacteroidetes, Protebacteria and Planctomycetes became the most plentiful phyla within the D population. With the changes in community proportions came the presence of new colonizing bacterium first noticed in the BL state but were more profuse in the D state. There was no singular or group of microbes that were solely found within the infected mesohyl of both sample site populations of *X. muta*. Therefore there is still no unequivocal pathogen as the microbial causative agent, but there were bacteria which appear associated with SOB Disease progression. The BL and the D microbial communities included phyla Armatimonadetes, Caldithrix, Chlorobi, Fibrobacteres, Fusobacteria, GN02, KSB3, OP1, OP2, OP8, SR1, TM6, Tenericutes, WPS-2 and ZB3 which were not previously present within the HC or HoD populations. These phyla therefore represent the bacteria associated with SOB Disease. The relative abundances of phyla Verrucomicrobia and Planctomycetes increased dramatically within the BL and the D populations, which also associated them with SOB Disease. Future research is needed to not only better understand the role and functions of these phyla within sponge communities but to understand their association with SOB Disease. We cannot rule out the possibility that undetected viral or eukaryotic pathogens are responsible or also contributed to SOB Disease.

The effect that weather patterns had on the emergence or progression of Sponge Band Disease remains unclear. The reoccurring pattern of long periods of low precipitation followed by large precipitation events that coincide with elevated sea surface temperatures should be considered and examined in future outbreaks.

REFERENCES

Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL and Polz MF. (2004). Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**: 551-554.

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V and Polz M. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Environ Microbiol* **71**: 8966- 8969.

Aird D, Ross MG, Chen W, Danielsson M, Fennell T, Russ C, Jaffe DB, Nusbaum C and Gnirke A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol* **12**: 1-14.

Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M and Rohani P. (2006). Seasonality and the dynamics of infectious diseases. *Ecol Letters* **9**: 467-484.

Amann RL, Ludwig W, Scheidler KH. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *FEMS Microbiol Rev* **59**:143-69.

Anderson MJ (2001). A new method for non-parametric multivariate analysis of variance. *Austral ecology* **26**: 32-46.

Angermeier H, Kamke J, Abdelmohsen UR, Krohne G, Pawlik JR, Lindquist NL, Hentschel U. (2011). The pathology of sponge orange band disease affecting the Caribbean barrel sponge *Xestospongia muta*. *FEMS Microbiol Ecol* **75**: 218-230.

Aronson RB and PrechtWF. (2006). Conservation, precaution, and Caribbean reefs. *Coral Reefs* **25**: 441-450.

Arnone R and Walling J. (2007). Waterborne pathogens in urban watersheds. *Journal of Water and Health* **5**: 149-162.

Austin B and Zhang XH. (2006). Vibrio harveyi: a significant pathogen of marine vertebrates and invertebrates. *Letters* in *Appl Microbiol* **43**: 119-124.

Bakker MG, Tu ZJ, Bradeen JM and Kinkel LL. (2012). Implications of pyrosequencing error correction for biological data interpretation. *PloS One* **7**: e44357.

Bartram AK, Lynch MD, Stearns JC, Moreno-Hagelsieb G and Neufeld JD. (2011). Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl and Environ Microbiol* **77**: 3846-3852.

Bell JJ. (2008). The functional roles of marine sponges. *Estuarine, Coastal and Shelf Sci* **79**: 341-353.

Bennett KW and Eley A. (1993). Fusobacteria: new taxonomy and related diseases. *J of Med Microbiol* **39**: 246-254.

Bibby, K., Viau, E., and Peccia, J. (2010). Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids. *Water research* **44**: 4252-4260.

Bourne DG, Garren M, Work TM, Rosenberg E, Smith GW and Harvell CD. (2009). Microbial disease and the coral holobiont. *Trends in Microbiol* **17**: 554-562.

Boyd R and Hoerl B. (1991). The Gram-positive cocci: Group A streptococci. Basic Medical Microbiology. Boyd R, Hoerl B (Eds). Fourth Edition. London, Little, Brown, and Company, 393-418.

Bruno JF, Selig ER, Casey KS, Page CA, Willis BL, Harvell CD et al, (2007). Thermal stress and coral cover as drivers of coral disease outbreaks. *PLoS Biol* **5**: e124.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N et al, (2012). Ultra-high-throughput microbial community analysis on the IIIumina HiSeq and MiSeq platforms. *The ISME J* **6**: 1621-1624.

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

Chao A (1984) Non-parametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265–270.

Chazdon RL, Colwell RK, Denslow JS and Guariguata MR. (1998) Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of northeastern Costa Rica. Forest Biodiversity Research, Monitoring and Modeling: Conceptual Background and Old World Case Studies. In: Dallmeier F and Comiskey JA (ed). Parthenon Publishing: Paris, pp. 285–309.

Claesson MJ, Wang Q, O'Sullivan O, Greene-Diniz R, Cole JR, Ross PR, O'Toole PW. (2010). Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Research*. **38**: e200.

Cowart JD, Henkel TP, McMurray SE, Pawlik JR. (2006). Sponge orange band (SOB): a pathogenic-like condition of the giant barrel sponge, *Xestospongia muta*. *Springer-Verlag* **25**: 513.

Dang H, Li T, Chen M, Huang, G. (2008). Cross-ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. *Appl and Environ Microbiol* **74**: 52-60.

Diaz CM, Rutzler, K. (2001). Sponges: essential component of Caribbean coral reefs. *Bulletin of Marine Science*. **69**: 532-546.

Derakshani M, Lukow T, Liesack W. (2001). Novel bacterial lineages at the (sub) division level as detected by signature nucleotide-targeted recovery of 16S rRNA genes from bulk soil and rice roots of flooded rice microcosms. *Appl and Environ Microbiol* **67**: 623-631.

Derrien M, Vaughan EE, Plugge CM, de Vos WM. (2004). Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J of Syst and Evol Microbiol* **54**: 1469-1476.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K. et al, (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl and Environ Microbiol* **72**: 5069-5072.

Dubourg G, Lagier JC, Armougom F, Robert C, Audoly G, Papazian L, Raoult, D. (2013). High-level colonisation of the human gut by Verrucomicrobia following broad-spectrum antibiotic treatment. *Int J of antimicrobial agents* **41**: 149-155.

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

Elshahed MS, Senko JM, Najar FZ, Kenton SM, Roe BA, Dewers TA et al, (2003). Bacterial diversity and sulfur cycling in a mesophilic sulfide-rich spring. *Appl and environ microbiol* **69**: 5609-5621.

Falkow S. (2004). Molecular Koch's postulates applied to bacterial pathogenicity-a personal recollection 15 years later. *Nat Rev Microbiol* **2**: 67-72.

Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. (2010). Forensic identification using skin bacterial communities. *PNAS* **107**: 6477-6481.

Fierer N, Ladau J, Clemente JC, Leff JW, Owens SM, Pollard KS, Knight R, Gilbert JA, McCulley RL. (2013). Reconstructing the Microbial Diversity and Function of Pre-Agricultural Tallgrass Prairie Soils in the United States. *Sci* **342**: 621-624.

Fieseler L, Horn M, Wagner M, Hentschel U. (2004). Discovery of the novel candidate phylum "Poribacteria" in marine sponges. *Appl and Environ Microbiol* **70**: 3724-3732.

Fiore CL, Jarett JK, Lesser MP. (2013). Symbiotic prokaryotic communities from different populations of the giant barrel sponge, *Xestospongia muta*. *Microbiol Open* **2**: 938-952.

Fiore CL, Baker DM, Lesser MP. (2013B). Nitrogen Biogeochemistry in the Caribbean Sponge, *Xestospongia muta*: A Source or Sink of Dissolved Inorganic Nitrogen? *PloS One* **8**: e72961.

Freitas S, Hatosy S, Fuhrman JA, Huse SM, Welch DBM, Sogin ML, Martiny AC. (2012). Global distribution and diversity of marine Verrucomicrobia. *The ISME J* **6**: 1499-1505.

Fox GE, Wisotzkey JD, Jurtshuk P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J of Syst Bacteriology* **42**: 166-170.

Freeman CJ, Thacker RW. (2011). Complex interactions between marine sponges and their symbiotic microbial communities. *Limnol and Oceanogr* **56**: 1577-1586.

Frerichs GN, Roberts RJ, (1989) The bacteriology of teleosts**.** In: Roberts RJ, ed Fish Pathology. London: Baillière Tindall, 289-319.

Fu L, Niu B, Zhu Z, Wu S, Li W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**: 3150-3152.

Fuhrman JA. (2009). Microbial community structure and its functional implications. *Nature* 459: 193-199.

Gammill ER, Fenner D. (2005). Disease threatens Caribbean sponges: report and identification guide. *Reefbase*. Available at http://www.reefbase.org/spongedisease.

Galtsoff PS. (1942). Wasting disease causing mortality of sponges in the West Indies and Gulf of Mexico. *Proc. 8th. Amer. Sci. Congr* **3**: 411-421.

Gilbert JA, Field D, Swift P, Thomas S, Cummings D, et al, (2010) The Taxonomic and Functional Diversity of Microbes at a Temperate Coastal Site: A 'Multi-Omic' Study of Seasonal and Diel Temporal Variation. *PLoS One* **5**: e15545.

Gilbert JA, DuPont CL. (2011). Microbial Metagenomics: Beyond the Genome. *Annu Rev Marine Sci* **3**: 347-371.

Gihring TM, Green SJ, Schadt CW. (2012). Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environ Microbiol* **14**: 285-290.

Glaeser SP, Galatis H, Martin K, Kämpfer P. (2012). Luteolibacter cuticulihirudinis sp. nov., isolated from Hirudo medicinalis. *Antonie van Leeuwenhoek* **102**: 319-324.

Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhe, R, Reid G. (2010). Microbiome profiling by illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One* **5**: e15406.

Green EP, Bruckner AW. (2000). The significance of coral disease epizootiology for coral reef conservation. *Biol Conservation* **96**: 347-361.

Gotelli NJ, Chao A. (2013) Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data. In: Levin S.A. (ed.) Encyclopedia of Biodiversity, second edition, Volume 5, pp. 195-211. Waltham, MA: Academic Press

Hamady M, Knight R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome research*, **19**: 1141-1152.

Handelsman J. (2004). Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and Mol Biol Rev* **68**: 669-685.

Hardoim CC, Costa R. (2014). Temporal dynamics of prokaryotic communities in the marine sponge Sarcotragus spinosulus. *Mol Ecol*.

Harvell CD, Kim K, Burkholder JM, Colwell RR, Epstein PR, Grimes DJ et al, (1999). Emerging marine diseases-climate links and anthropogenic factors. *Science* **285**: 1505-1510.

Harvell D, Aronson R, Baron N, Connell J, Dobson A, Ellner S et al, (2004). The rising tide of ocean diseases: unsolved problems and research priorities. *Frontiers in Ecol and the Environ* **2**: 375-382.

Harvell,D, Jordán-Dahlgren E, Merkel S, Rosenberg E, Raymundo L, Smith G, et al, (2007). Coral disease, environmental drivers, and the balance between coral and microbial associates. *Oceanography* **20**: 172-195.

Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G et al, (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome research* **21**: 494-504.

Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J, Moore BS. (2002). Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl and Environ Microbiol* **68**: 4431-4440.

Hentschel U, Usher KM, Taylor MW. (2006). Marine sponges as microbial fermenters. *FEMS Microbiology Ecol* **55**: 167-177.

Hong HA, Khaneja R, Tam, Nguyen MK, Cazzato A, Tan S, Urdaci M, et al, (2009). "Bacillus subtilis isolated from the human gastrointestinal tract". *Research in Microbiol* 160: 134-43.

Huber JA, Welch DBM, Morrison HG, Huse SM, Neal PR, Butterfield DA, et al, (2007). Microbial population structures in the deep marine biosphere. *Sci* **318**: 97-100.

Hughes TP. (1994). Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. SCIENCE-NEW YORK THEN WASHINGTON-, 1547-1547.

Huse SM, Welch DM, Morrison HG, Sogin ML. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12**: 1889-1898.

Im WT, Hu ZY, Kim KH, Rhee SK, Meng H, Lee ST et al, (2012). Description of Fimbriimonas ginsengisoli gen. nov., sp. nov. within the Fimbriimonadia class nov., of the phylum Armatimonadetes. *Antonie van Leeuwenhoek*, **102**: 307-317.

Jansson JK, Prosser JI. (2013). Microbiology: The life beneath our feet. *Nat* **494**: 40-41.

Jiang S, Noble R, Chu W. (2001). Human adenoviruses and coliphages in urban runoff-impacted coastal waters of Southern California. *Appl and Environ Microbiol* **67**: 179-184.

Jiang F, Li W, Xiao M, Dai J, Kan W, Chen L et al, (2012). Luteolibacter luojiensis sp. nov., isolated from Arctic tundra soil, and emended description of the genus Luteolibacter. *Int J of Syst and Evol Microbiol* **62**: 2259-2263.

Keijser BJF, Zaura E, Huse SM, Van Der Vossen JMBM, Schuren FHJ, Montijn RC et al, (2008). Pyrosequencing analysis of the oral microflora of healthy adults. *J of dental research* **87**: 1016-1020.

Knights D, Costello EK, Knight R. (2011). Supervised classification of human microbiota. *FEMS microbiology reviews* **35**: 343-359.

Knight R, Jansson J, Field D, Fierer N, Desai N, Fuhrman JA et al, (2012). Unlocking the potential of metagenomics through replicated experimental design. *Nat biotechnology* **30**: 513-520.

Koch R. (1876). "Untersuchungen über Bakterien: V. Die Ätiologie der Milzbrand-Krankheit, begründet auf die Entwicklungsgeschichte des *Bacillus anthracis*" (Investigations into bacteria: V. The etiology of anthrax, based on the ontogenesis of *Bacillus anthracis*). *Cohns Beitrage zur Biologie der Pflanzen* (in German) **2**: 277–310.

Kokke WCMC, Tarchini C, Stierle DB, Djerassi C. (1979). Isolation, structure elucidation, and partial synthesis of Xestosterol, a biosynthetically significant sterol from the sponge *Xestosponiga muta*. *J.Org. Chem*. **44**: 3385-3388.

Koopermans M, Martens D, Wijffels RH. (2009). Towards commercial production of sponge medicines. *Marine Drugs* **7**: 787-802.

Könneke M, Bernhard AE, José R, Walker CB, Waterbury JB, Stahl DA. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543-546

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* **12**: 118-123.

Lafferty KD, Holt RD. (2003). How should environmental stress affect the population dynamics of disease? *Ecol Letters* **6**: 654-664.

Lafferty KD, Porter JW, Ford SE. (2004). Are diseases increasing in the ocean? *Annual Review of Ecology, Evol and Syst* 31-54.

Lafi FF, Fuerst JA, Fieseler L, Engels C, Goh WWL, Hentschel U. (2009). Widespread distribution of Poribacteria in Demospongiae. *Appl and Environ Mcrobiol* **75**: 5695-5699.

Larsen PE, Field D, Gilbert JA. (2012). Predicting bacterial community assemblages using an artificial neural network approach. *Nat methods* **9**: 621-625.

Lee OO, Wang Y, Yang J, Lafi FF, Al-Suwailem A. (2010). Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *The ISME J* **5**: 560-664.

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS et al, (2008). Evolution of mammals and their gut microbes. *Science* **320**: 1647-1651.

Liddell WD, Ohlhorst SL. (1986). Changes in benthic community composition following the mass mortality of *Diadema* at Jamaica. *J of Experimental Marine Biol and Ecol* **95**: 271-278.

Ling Z, Kong J, Liu F, Zhu H, Chen X, Wang, Y et al, (2010). Molecular analysis of the diversity of vaginal microbiota associated with bacterial vaginosis. *BMC genomics* **11**: 488.

Li Z, Chen Y, Mu D, Yuan J, Shi Y, Zhang H et al, (2012). Comparison of the two major classes of assembly algorithms: overlap layout consensus and de-bruijn-graph. *Briefings in functional genomics*, **11**: 25-37.

Logares R, Haverkamp TH, Kumar S, Lanzén A, Nederbragt AJ, Quince C et al, (2012). Environmental microbiology through the lens of High-Throughput DNA Sequencing: Synopsis of current platforms and bioinformatics approaches. *J of Microbiol Methods* **91**: 106-113.

Lopez-Legentil S, Erwin PM, Pawlik JR, Song B. (2010). Effects of sponge bleaching on ammonia-oxidizing *archaea*: distribution and relative expression of ammonia monooxygenase genes associated with the barrel sponge *Xestrospongia muta*. *Microb Ecol*. **60**: 561-571.

Lopez-Legentil S, Song B, McMurry SE, Pawlik JR. (2008). Bleaching and stress in coral reef ecosystems: hsp70 expression by the giant barrel sponge *Xestospongia muta*. *Mol Ecol* **17**: 1840-1849.

Lopez J.V, McCarthy PJ, Janda K.E,Willoughby R, Pomponi SA. (1999) Molecular techniques reveal wide phyletic diversity of heterotrophic microbes associated with the sponge genus Discodermia (Porifera:Demospongiae). Proceedings of the 5th International Sponge Symposium. *Memoirs of the Queensland Museum*. **44**: 329-341.

Lozupone C, Hamady M, Knight R. (2006). UniFrac- An online tool for comparing microbial community diversity in a phylogenetic context. *BioMed Central* **7**: 1-14.

Lozupone C, Hamady M, Kelley ST, Knight R. (2007). Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Appl Environ Microbiol* **73**: 1576-1585.

Lozupone CA, Knight R. (2007). Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* **104**: 11436–11440

Lozupone CA, Knight R. (2008). Species divergence and the measurement of microbial diversity. *FEMS Microbiol Rev* **32**: 557–578.

Luo C, Tsementzi D, Kyrpides N, Read T, Konstantinidis KT. (2012). Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. *PloS One* **7**: e30087.

NOAA National Climatic Data Center, State of the Climate (NCDC): Synoptic Discussion for May 2012, published online June 2012, retrieved on October 25, 2013 from http://www.ncdc.noaa.gov/sotc/synoptic/2012/5.

NOAA National Climatic Data Center, State of the Climate (NCDC): El Niño/Southern Oscillation for June 2007, published online July 2007, retrieved on October 31, 2013 from http://www.ncdc.noaa.gov/sotc/enso/2007/6.

Maramorosch K, Granados RR, Hirumi H. (1970). Mycoplasma diseases of plants and insects. *Adv. Virus Res* **16**: 135-193.

Mardis ER. (2008). The impact of next-generation sequencing technology on genetics. *Trends in genetics* **24**: 133.

Mardis ER. (2011). A decade's perspective on DNA sequencing technology. *Nat* **470**: 198-203.

McLean JS, Lombardo MJ, Badger JH, Edlund A, Novotny M, Yee-Greenbaum J et al, (2013). Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. *Pro of the Nat Acad of Sci* **110**: E2390-E2399.

McMurray SE, Blum JE, Leichter JJ, Pawlik JR. 2011. Bleaching of the giant barrel sponge *Xestospongia muta* in the Florida Keys. *Limnol. Oceanogr* **56**: 2243-2250.

McMurray SE, Blum JE, Pawlik JR. (2008). Redwood of the reef: growth and age of the giant barrel sponge *Xestospongia muta* in the Florida Keys. *Mar Biol* **155**: 159-171.

McMurray SE, Henkel TP Pawlik JR. (2009). Demographics of increasing populations of the giant barrel sponge *Xestospongia muta* in the Florida Keys. *Ecol* **91**: 560-570.

Miroshnichenko ML, Kostrikina NA, Chernyh NA, Pimenov NV, Tourova TP, Antipov AN et al, (2003). Caldithrix abyssi gen. nov., sp. nov., a nitrate-reducing, thermophilic, anaerobic bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent, represents a novel bacterial lineage. *Int J of Syst and Evol Microbiol* **53***:* 323-329.

Mitsova D, Vos J, Stafeychuk I, Gardinali P. (2011). Variability in road runoff pollution by polycyclic aromatic hydrocarbons (PAHs) in the urbanized area adjacent to Biscayne Bay, Florida. *J of Environ Protection* **2***:* 1317.

Montalvo NF, Hill RT. (2011). Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts. *Appl and Environ Microbiol* **77**: 7207-7216.

Morinaka BI, Skepper CK, Molinski TF. (2007). Ene-yne tetrahydrofurans from the sponge Xestospongia muta. Exploiting a weak CD effect for assignment of configuration. *Organic Letters* **9**: 1975-1978.

Munn C. (2003). Marine microbiology: Ecology and applications. Garland Science.

Murmasih T, Kosela S, Kardono LBS, Hanafi M, Priyono W. (2014). An Antibacterial Compound Isolated from Sponge-associated bacteria Rhodobacteracea bacterium. *J of Med Sci* **14**: 75-80.

Nagelkerken I, Aerts L, Pors L. (2000). Barrel sponge bows out. *Reef Encounter* **28**:14-15.

Negandhi K, Blackwelder PL, Ereskovsky AV, Lopez JV. (2010). Florida reef sponges harbor coral disease-associated microbes. *Symbiosis* **51***:* 117-129.

Niang L, Sjostrand U, Djerassi C. (1980). Minor and trace sterols in marine invertebrates. 19. Isolation, structure elucidation, and partial synthesis of 24-Methylene-25-ethylcholesterol (mutasterol): First example of sterol side-chain bioanlkylation at position 25. *J. Org. Chem. 103*: 115-119.

Nicholas RAJ, Ayling RD, McAuliffe L. (2009). Vaccines for Mycoplasma diseases in animals and man. *Journal of comparative pathology* **140***:* 85-96.

Nogales B, Moore ER, Llobet-Brossa E, Rossello-Mora R, Amann R, Timmis KN. (2001). Combined use of 16S ribosomal DNA and 16S rRNA to study the bacterial community of polychlorinated biphenyl-polluted soil. *Appl and Environ Microbiol* **67**: 1874-1884.

Oliveros, J.C. (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html.

Olson JB, Gochfeld DJ, Slattery M. (2006). Aplysina red band syndrome: a new threat to Caribbean sponges. *Diseases of aquatic organisms* **71**: 163.

Owens L, Busico-Salcedo N (2006). "Vibrio harveyi: Pretty Problems in Paradise (Chapter 19)". In Thompson, Fabiano; Austin, Brian; Swings, Jean. *The Biol of Vibrios*. ASM Press.

Pace NR. (1997). A molecular view of microbial diversity and the biosphere. *Sci* **276**: 734-740.

Patil AD, Kooke WC, Cochran S, Francis TA, Tomszek T, Westley JW. (1992). Brominated polyacetylenic acids from the marine sponge *Xestospongia muta*: Inhibitors of HIV protease. *J of Nat Prod* **55**: 1170-1177.

Patin NV, Kunin V, Lidström U, Ashby MN. (2013). Effects of OTU Clustering and PCR Artifacts on Microbial Diversity Estimates. *Microbial Ecol* **65:** 709-719.

Pedrós-Alió, C. (2006). Marine microbial diversity: can it be determined? *Trends in microbiol* **14**: 257-263.

Quince C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM et al, (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat methods* **6:** 639-641.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ (2011) Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12**: 38.

Rappé MS, Giovannoni SJ. (2003). The uncultured microbial majority. *Annual Reviews in Microbiol* **57**: 369-394.

Reeder J, Knight R. (2010). Rapid denoising of pyrosequencing amplicon data: exploiting the rank-abundance distribution. *Nat methods* **7***:* 668.

Richardson LL. (1998). Coral diseases: what is really known? *Trends* in *Ecol and Evol* **13**: 438-443.

Richelle-Maurer E, Gomez R, Braekman JC, Van de Vyver G, Van Soest RWM, Devijver C. (2003). Primary cultures from the marine sponge *Xestospongia muta* (Petrosiidae, Haplosclerida). *J of Biotechnology* **100**: 169-176.

Riehle K, Coarfa C, Jackson A, Ma J, Tandon A, Paithankar S et al, (2012). The Genboree Microbiome Toolset and the analysis of 16S rRNA microbial sequences. *BMC Bioinformatics* **13**: S11.

Riitzler K. (1988). Mangrove sponge disease induced by cyanobacterial symbionts: failure of a primitive immune system. *Diseases of aquatic organisms* **5**: 143-149.

Riviere D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach J, et al, (2009). Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *The ISME J* **3:** 700-714.

Ross KA, Feazel LM, Robertson CE, Fathepure BZ, Wright KE, Turk-MacLeod RM et al, (2012). Phototrophic phylotypes dominate mesothermal microbial mats associated with hot springs in Yellowstone National Park. *Micro Ecol* **64**: 162-170.

Ryan KJ, Ray CG. (2004). *Sherris medical microbiology: an introduction to infectious diseases* (pp. xiii-979). New York: McGraw-Hill.

Sambrook J, Russell DW. (2006). The condensed protocols from Molecular cloning: a laboratory manual, Author: Joseph Sambrook, David W. Russell, Publisher.

Santavy DL, Willenz P, Colwell RR. (1990). Phenotypic study of bacteria associated with the Caribbean Sclerosponge, *Ceratoporella nicholsoni*. *Appl. Environ. Microbiol*. **56**: 1750-1762.

Schink, B. (2006). The genus Propionigenium. *The Prokaryotes: Volume 7: Proteobacteria: Delta, Epsilon Subclass*, 955-959.

Schloss P, Handelsman J. (2005). Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. G*enome Biol* **6:** 229.

Schloss PD, Gevers D, Westcott SL. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PloS One* **6**: e27310.

Schmitt S, Tsai P, Bell J, Fromont J, Ilan M, Lindquist N et al, (2012). Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *The ISME J* **6:** 564-576.

Schmitt S, Angermeier H, Schiller R, Lindquist N, Hentschel U. (2008). Molecular microbial diversity survey of sponge reproductive stages and mechanistic insights into vertical transmission of microbial symbionts. *Appl and Environ Microbiol* **74**: 7694-7708.

Schmitt S, Tsai P, Bell J, Fromont J, Ilan M, Lindquist N et al, (2012). Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *The ISME J* **6**: 564-576.

Shannon CE, Weaver W. (1949) The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL.

Siegl A, Kamke J, Hochmuth T, Piel J, Richter M, Liang C et al, (2010). Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *The ISME J* **5***:* 61-70.

Simister RL, Deines P, Botté ES, Webster NS, Taylor MW. (2012). Sponge specific clusters revisited: a comprehensive phylogeny of sponge associated microorganisms. *Environ Microbiol* **14**: 517-524.

Smith DG. (1982). Basic medical microbiology By RF BOYD and BG HOERL. 1981. Little, Brown and Co., Boston. *J of Med Microbiol* **15**: 274-274.

Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR et al, (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Pro of the Nat Acad of Sci* **103**: 12115-12120.

Steger D, Ettinger-Epstein P, Whalan S, Hentschel U, De Nys R, Wagner M et al.(2008). Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges. *Environ Microbiology* **10**: 1087-1094.

Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs". *Nucleic Acids Res* **25**: 3389-3402.

Sun Y, Cai Y, Huse SM, Knight R, Farmerie WG, Wang X et al, (2011). A large-scale benchmark study of existing algorithms for taxonomy-independent microbial community analysis. *Briefings in Bioinformatics* **13**: 107-121.

Sundquist A, Bigdeli S, Jalili R, Druzin M, Waller S, Pullen K et al, (2007). Bacterial flora-typing with targeted, chip-based Pyrosequencing. *BMC Microbiol* **7**: 108.

Sutherland KP, Porter JW, Torres C. (2004). Disease and immunity in Caribbean and Indo-Pacific zooxanthellate corals. *Marine Ecol Progress Series* **266**: 265-272.

Suzuki MT, Giovannoni SJ. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl and Environ Microbiol* **62**: 625-630.

Tamaki H, Tanaka Y, Matsuzawa H, Muramatsu M, Meng XY, Hanada S et al, (2011). Armatimonas rosea gen. nov., sp. nov., of a novel bacterial phylum, Armatimonadetes phyl. nov., formally called the candidate phylum OP10. *Int J of* Syst *and Evol Microbiol* **61**: 1442-1447.

Taylor MW, Radax R, Steger D, Wagner M. (2007). Sponge-associated microorganisms: evolution, ecology, and biotechnological potential, *Microbiol and Mol Biol reviews* **71**: 295-347.

Temperton B, Giovannoni SJ. (2012). Metagenomics: microbial diversity through a scratched lens. Current Opinion in Microbiology.

Thomas T, Gilbert J, Meyer F. (2012). Metagenomics- a guide from sampling to data analysis. *Microbial Informatics and Experimentation.* **2**: 1-12.

Tigertt WD, Benenson AS, Gochenour WS. (1961). Airborne Q fever.*Bacteriological reviews* **25**: 285.

Tsuchiya C, Sakata T, Sugita H. (2008). Novel ecological niche of Cetobacterium somerae, an anaerobic bacterium in the intestinal tracts of freshwater fish. *Letters in appl microbiol* **46**: 43-48.

Van Sluys M., Monteiro-Vitorell CB, Camargo LEA, Menc CFM, Da Silva, ACR et al, (2002). Comparative genomic analysis of plant-associated bacteria. *Annual review of phytopathology* **40**: 169-189.

Veron JEN, Hoegh-Guldberg O, Lenton TM, Lough JM, Obura DO, Pearce-Kelly P et al, (2009). The coral reef crisis: The critical importance of < 350ppm CO< sub> 2</sub>. *Marine Pollution Bulletin* **58**: 1428-1436.

Vicente VP. (1990). Response of sponges with autotrophic endosymbionts during the coral-bleaching episode in Puerto Rico. *Coral Reefs* **8**: 199-202.

Vogel S. (1977). Current-induced flow through living sponges in nature. *Proceedings of the National Academy of Sciences* **74**: 2069-2071.

Von Wintzingerode F, Selent B, Hegemann W, Göbel UB. (1999). Phylogenetic analysis of an anaerobic, trichlorobenzene-transforming microbial consortium. *Appl and Environ Microbiol* **65**: 283-286.

Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl and Environ Microbiol* **73**: 5261-5267.

Webster NS, Blackall LL. (2009). What do we really know about sponge-microbial symbioses? *The ISME J* **3**: 1-3.

Webster NS, Cobb RE, Negri AP. (2008). Temperature thresholds for bacterial symbiosis with a sponge. *The ISME J* **2**: 830-842.

Webster NS. (2007). Sponge disease: a global threat? *Environ Microbiol*. **9**: 1363-1375.

Webster NS, Taylor MW. (2012). Marine sponges and their microbial symbionts: love and other relationships. *Environ Microbiol* **14**: 335-346.

Woese CR. (1987). Bacterial evolution. *Microbiol reviews* **51**: 221.

Wulff JL. (2007). Disease prevalence and population density over time in three common Caribbean coral reef sponge species. *J. Mar. Biol. Ass. U.K.* **87**: 1715-1720.

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H et al, (2006). Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. *Int J of Syst and Evol Microbiol* **56**: 1331-1340.

Yamada T, Yamauchi T, Shiraishi K, Hugenholtz P, Ohashi A, Harada H et al, (2007). Characterization of filamentous bacteria, belonging to candidate phylum KSB3, that are associated with bulking in methanogenic granular sludges. *The ISME J* **1**: 246-255.

Zapala MA, Schork NJ. (2006). Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. *Proceedings of the national academy of sciences* **103***:* 19430-1943.