Development Of A Biomarker System For SCTLD In Four Species Of Caribbean Corals

Florida Department of Environmental Protection
Office of Resilience and Coastal Protection

Development Of A Biomarker System For SCTLD In Four Species Of Caribbean Corals

Final Report

Prepared By:
Nikki Traylor-Knowles, Ph.D.

University of Miami
Rosenstiel School of Marine and Atmospheric Sciences
4600 Rickenbacker Causeway
Miami, FL 33149

December 31, 2020

This report should be cited as follows:
Nikki Traylor-Knowles, 2020. Development Of A Biomarker System For SCTLD In Four Species Of Caribbean Corals. Florida DEP. Miami, FL. 1-17 pages.

Executive Summary

Since 2014, Florida's Coral Reefs have been dying at an unprecedented rate due to stony coral tissue loss disease (SCTLD). Biomarker development is a valuable tool for identifying genetic markers for the presence of a disease and baseline health. In the case of SCTLD, it is not known why certain species of coral are more susceptible to the disease than others, and it is also not known if there are genetic markers indicative of SCTLD. In this report, we described the transcriptomic outcomes of three different SCTLD transmission experiments performed at the Smithsonian and Mote Marine Labs between 2019 and 2020. Overall, we found that *Montastrea cavernosa* and *Orbicella faveolata* had a similar differential gene expression response to SCTLD, with controls grouping together, and diseased samples spreading out based on genotype. We also found that *O. faveolata* had 1128 differentially expressed genes (DEGs), whereas *M. cavernosa* had 61 DEGs. Lastly, we found that there were many unique and shared genes between *O. faveolata* and *M. cavernosa* that may be important biomarkers for SCTLD. We recommend that further studies be done on this transcriptomics of corals with SCLTD and in particular focus on the role of apoptosis in SCTLD progression.

Acknowledgements

Table of Contents (Created using the Microsoft Word Table of Contents tool)

Table of Contents

List of Figures

List of Tables

List of Acronyms

SCTLD: Stony Coral Tissue Loss Disease

cDNA: Complementary deoxyribonucleic acid

RNA: Ribonucleic acid

PC: Principal Component

WGCNA: Weighted gene coexpression network analysis

DEGs: Differentially expressed genes

ECM: extracellular matrix

MITO: mitochondrial

1. BACKGROUND

1.1. Overview

Florida's Coral Reef is currently experiencing a multi-year disease-related mortality event that has resulted in massive die-offs in multiple coral species. Approximately 21 species of coral, including both Endangered Species Act-listed and the primary reef-building species, have displayed tissue loss lesions that often result in whole colony mortality. First observed near Virginia Key in 2014, the disease has since spread to the northernmost extent of Florida's Coral Reef, and south to the Marquesas in the Lower Florida Keys. The best available information indicates that the disease outbreak is continuing to spread southwest and throughout the Caribbean.

Biomarker development is a valuable tool for identifying genetic markers for disease susceptibility (Traylor-Knowles and Palumbi, 2014). In the case of stony coral tissue loss disease (SCTLD), it is not known why certain species of coral are more susceptible to the disease than others, and it is also not known if there are genetic markers that could indicate the presence of SCTLD before the physical signs.

The development of a biomarker system for SCTLD would provide a very valuable tool for resource managers for the following reasons:
1. Early Intervention: Development of a biomarker would allow for testing of the disease presence in corals before the coral is showing signs of the disease.
2. Genetic diversity: Genetic information gathered from this study can be used in predictive models to identify the corals which are at most risk genetically for developing the disease—thus help with prioritizing which ones to protect or collect.
3. Pathogen identification: Information on the pathogen type can be understood by the coral host response and can enrich current microbiome efforts.

As part of the recent Coral Disease Workshop held in July 2019 the use of transcriptomics and immunology were identified as major areas that needed to be addressed in collaboration with the current research projects. By identifying the coral's genetic signature in response to SCTLD, we will be able to better target a treatment. For example, if specific viral recognition genes are highly upregulated in SCTLD corals, then it would give us evidence that viruses may be involved. Alternatively, if pathogen recognition receptors are differentially expressed, then we may be more confident that a bacterial pathogen is involved. Therefore, this would bolster current research efforts to identify the cause of this disease, while also developing an important tool for early detection of SCTLD.

1.2. Project Goals and Objectives:

The overall goals of this project are: 1) to identify core coral biomarkers for SCTLD, 2) to examine the immune system of SCTLD exposed and apparently healthy corals, and 3) to identify species specific biomarkers in response to SCTLD. The outcomes of this project will be incorporated into an on-going coral disease response effort which seeks to

improve understanding about the scale and severity of the Florida's Coral Reef coral disease outbreak, identify primary and secondary causes, identify management actions to remediate disease impacts, restore affected resources and ultimately prevent future outbreaks. To achieve the project goals, Dr. Traylor-Knowles collaborated with Dr. Val Paul (The Smithsonian) and Dr. Erinn Muller (Mote Marine Labs) to sample corals that are currently being used for disease transmission projects. These included healthy corals that were collected in advance of the disease front compared with healthy corals subjected to SCTLD transmission experiments. By collaborating with these groups there was less redundancy in collections and more comparative information that will be gathered allowing for a holistic approach to mitigate and understand this disease in corals.

2. METHODOLOGY

2.1. Transmission assays

Three separate experiments have been performed, which I will summarize below. First was an experiment run at the Smithsonian in Ft. Pierce, which was a transmission experiment on four different species (*C. natans, O. faveolata, M. meandrites,* and *M. cavernosa*). The healthy corals were collected from Dry Tortugas in January 2020. Second was a transmission experiment run at Mote Marine Labs, on *P. clivosa* Mote nursery fragments. Third, was an experiment previously run by Dr. Erinn Muller and Katie Eaton at Mote Marine Labs to examine the rates of transmission from different parts of an SCTLD infected coral. I will provide more details below on each of these experiments.

2.1.1. Smithsonian experiment

This experiment occurred at Smithsonian in Ft. Pierce from February 21- March 9, 2020 in collaboration with Dr. Val Paul and Dr. Blake Ushijima. Four different species of coral *C. natans, O. faveolata, M. meandrites,* and *M. cavernosa,* were exposed to an active lesion of SCTLD or a control with no disease sign. For the transmission experiment: *C. natans* had nine different genotypes exposed, *O. faveolata* had 6 different genotypes exposed, *M. meandrites* had 5 different genotypes exposed, and *M. cavernosa* had 5 different genotypes exposed (Figure 1). Whole tissue and skeleton samples were preserved in RNAlater and brought back to University of Miami for processing. Due to low replication, *M. meandrities* was not processed further.
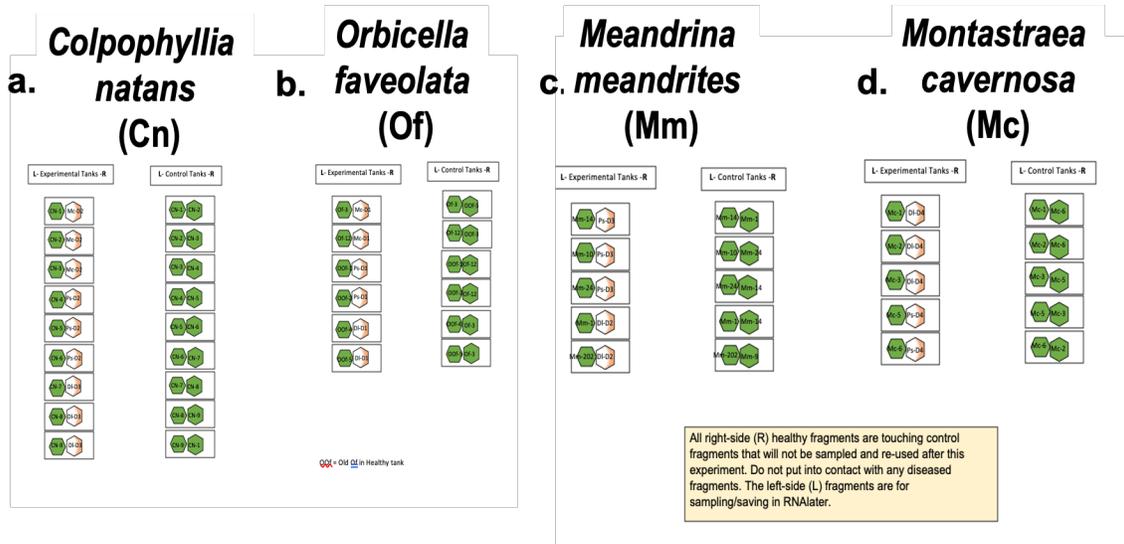
Figure 1: *Experimental design for the Smithsonian experiment.*

2.1.2. Mote Marine Lab experiment

2.1.2.1. *P. clivosa* experiment

This experiment occurred at Mote Marine Lab from late October to early November 2019 in collaboration with Dr. Erinn Muller and Katie Eaton. There were 44 initial samples (taken from each individual *P. clivosa* replicate, as well as the 4 diseased colonies collected), 16 disease samples, and 17 paired control samples, for a total of 77 samples collected.

For this experiment, plugs of 9 genotypes of *P. clivosa* from Mote's nursery were exposed to an active SCTLD lesion and sampled by taking a tissue scrap from the area that transmitted. Each main disease colony had two healthy *P. clivosa* plugs that were in contact with the coral for exposure. The disease lesions used for transmission were from either *M. cavernosa* or *O. faveolata* colonies. Plugs of *P. clivosa* were put into contact with either *M. cavernosa* or *O. faveolata* colonies. Each *P. clivosa* plug for the disease treatment, and the control treatment were sampled using a tissue scrap, and flash frozen. The main diseased colony was also sampled using a tissue scrap. The frozen tissue scraps were then sent to RSMAS on dry ice.

2.1.2.2. *M. cavernosa* and *O. faveolata* experiment

This experiment occurred at Mote Marine Labs collaboration with Dr. Erinn Muller and Katie Eaton in April and July 2019. This experiment was done previously to understand if the parts of lesioned SCTLD colony not showing disease can transmit SCTLD. For our analysis we only processed the samples which were exposed to an active lesion (called D1). Dr. Muller and Katie Eaton generously sent us these samples which were also being examined for microbiome analysis, and microbiome transcriptomics (S. Rosales). The *M. cavernosa* from Mote were 7 healthy colonies collected at one of Mote's inshore

outplant sites off Key West. Four diseased colonies were collected at Looe Key. For the *O. faveolata* samples, 7 healthy *O. faveolata* colonies were collected at the airport coral heads in Key West. Four diseased M. cavernosa colonies were also collected at Wonderland.

2.2. RNA extraction, cDNA library preparation, sequencing

In total, 23 samples of *P. clivosa* (6 genotypes: 13 disease and 10 controls), 38 samples of *O. faveolata* (18 disease and 20 controls), 42 samples of *M. cavernosa* (22 disease and 20 controls), and 13 samples of *C. natans* (7 genotypes: 6 disease and 7 controls), consisting of both controls and transmission samples were prepared for either traditional RNAseq (*P. clivosa, C. natans*) or Tagseq (*O. faveolata* and *M. cavernosa)*. For RNA extraction, total RNA was extracted using the Qiagen RNeasy Minikit which included a recommended 15-minute DNase digestion. The Nanodrop and Qubit fluorometer were used to assess the quantity and quality of total RNA using manufacturer's instructions.

2.2.1. Tagseq library preparation

For the *O. faveolata* and *M. cavernosa* samples, RNA was extracted from each sample and converted into cDNA libraries using the Lexogen Quantseq Library Preparation Strategy. Quantification/quality control steps followed practices as recommended by Lexogen. Samples were sequenced for 50 base pair length reads on a NOVAseq at the University of Miami, Miller School of Medicine, John P. Hussman Institute for Human Genomics. A total of 76 samples passed quality control and were sequenced.

2.2.2. RNAseq library preparation and sequencing

For the *C. natans* and *P. clivosa* samples, total RNA was then converted to complementary DNA (cDNA) libraries using Illumina TruSeq RNA Library poly A-tail selection prep kit following the manufacturer protocol. During cDNA library preparation, Illumina adaptors were randomly assigned to samples to reduce bias between sequencing lanes. cDNA libraries were quantified using a Qubit fluorometer and sent to the Utah Huntsman Cancer Institute High Throughput Genomics Shared Resource Center. cDNA quality control was performed using High Sensitivity D100 Screentape. A total of 36 passed quality control and were sequenced for 150 base pairs, pair-end reads a NOVAseq at the University of Miami, Miller School of Medicine, John P. Hussman Institute for Human Genomics.

2.3. Sequence analysis

Using FastOC ready quality was assessed (Brown, Pirrung, & McCue, 2017). Using Trimmomatic, low quality reads were trimmed (Bolger, Lohse, & Usadel, 2014). These reads were then aligned to reference genomes or reference transcriptomes using STAR

(Dobin et al., 2013). Sequences that did not align to reference genomes or transcriptomes were excluded from analysis. Aligned coral sequence reads were then quantified using Salmon (Patro, Duggal, Love, Irizarry, & Kingsford, 2017) before being read into R (v3.6.1) and RStudio (v1.2.1335) using tximport (Team, 2020). Pre-filtered counts were then used for differential gene expression analysis and co-expression analysis. Using the variance stabilizing transformation (VST) function in DeSeq2, sample counts were transformed and then plotted using principal component analysis (PCA) and plotted using ggplot2 (Wickham, 2016). Weighted gene coexpression network analysis (WGCNA) was used to identify groups of coexpressed transcripts that correlated to control, disease, or disease progression (Langfelder & Horvath, 2008). Clustering was performed using the Ward method in WGCNA. A single signed network was built with manual network constructions. Key parameters included soft power = 12, minimum module, size = 40, deep split = 2, merged cut height = 0.40, minimum verbose = 3, and cutHeight = 0.997. The eigengene values of each module were correlated to treatment: disease progression, control, or disease.

3. RESULTS

3.1. *M. cavernosa* and *O. faveolata* sequencing outcome

Forty-two samples from *M. cavernosa* were successfully sequenced. The average read depth was 4,830,551 reads with a total of 117,433,794 combined reads for all of the samples. 45.58% of the transcript had a unique alignment to the *M. cavernosa* genome (Matz, 2018). After filtering, 17 *M. cavernosa* samples passed quality control, changing the final average read count to 7,636,834 reads, and 45.51% unique alignment to *M. cavernosa* genome. For *O. faveolata,* 29 samples were successfully sequenced out of the 38 libraries that were prepared. The average read depth was 9,746,697 reads with a total of 399,614,583 combined reads for all of the samples. Of these reads, 44.83% aligned to the *O. faveolata* genome (Prada et al., 2016). After filtering, 26 samples passed quality control, changing the average read depth to 10,531,729 reads per sample and 42.41% with a unique alignment to the *O. faveolata* genome.

3.2. *C. natans* and *P. clivosa* sequencing outcome

Of the 36 samples initially processed for RNAseq, only 29 were deemed suitable for sequencing. Of the 29 that were sequenced, 13 additional libraries failed to generate more than 1 million reads. The 16 samples that did generate reads (7 *C. natans* and 9 *P. clivosa*) ranged from 6 million to 52 million reads per sample, which is double the 25 million reads per sample target and is suggestive of adapter barcode overlap. In total, 216,063,897 reads were generated for *C. natans*, and 141,101,165 for *P. clivosa.*

Additionally, while further analysis is being attempted, there were poor rates of transcript quantification using the *P. strigosa* reference transcriptome from Ocampo et al. 2015 for the *C. natans* samples (~15%) and only marginally better for *P. clivosa* samples (~20 - 40%). As a result of these issues and due to technical difficulties with installation of the Trinity software to perform *de novo* transcriptome assembly, analysis of these samples

was paused to focus on the other species. For the rest of this report, we will focus on the *O. faveolata* and *M. cavernosa*.

3.3. Principal component analysis of disease and control samples

3.3.1. Principal component analysis of disease and control samples for *O. faveolata* from Mote

Principal component analysis of disease and control *O. faveolata* from the Mote Marine Laboratories revealed that disease samples are different in their gene expression from control samples. Overall, 45% of the variance is explained by the differences between controls and disease on principal component (PC) axis 1, whereas 11% of the variance is found on the PC2 axis (Figure 2). Controls group together, whereas the disease samples are scattered (Figure 2). One exception to this is that there are four disease samples which appear to group with the control samples. While we do not know why this is the case, it may have to do with what samples were used to transmit the disease. Future studies will need to follow up on these individual genotypes to better understand their gene expression.
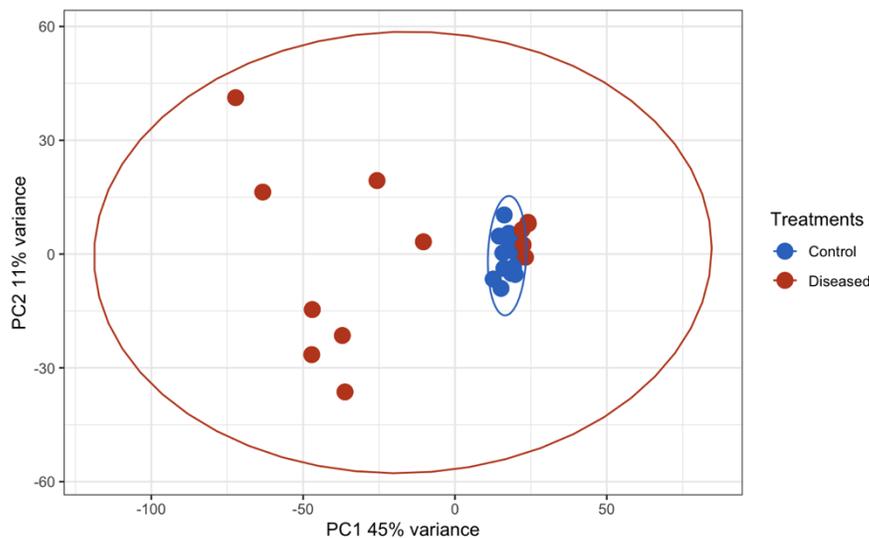
Figure 2: Principal component analysis of control and diseased O. faveolata *reveals the effects of treatment on gene expression. Color indicates the treatment (red =diseased, blue= control).*

3.3.2. Principal component analysis of disease and control samples for *M. cavernosa* from Mote and Smithsonian

Principal component analysis of disease and control *M. cavernosa* from Mote and Smithsonian revealed again that disease and controls are unique in their disease expression profiles. However, the Smithsonian samples, both control and disease were closely clustered with the control samples of Mote (Figure 3a) however, when analyzed separated, both the Mote samples and the Smithsonian samples showed similar patterns to what was observed in *O. faveolata*, where controls were closely grouped and the disease samples were more scattered (Figure 3b, 3c). For this, the 26% of the variance on PC1 was explained by the treatment for *M. cavernosa* from Mote, whereas 82% of the variance on PC1 was explained by the treatment for *M. cavernosa* from Smithsonian (Figure 3b, 3c).
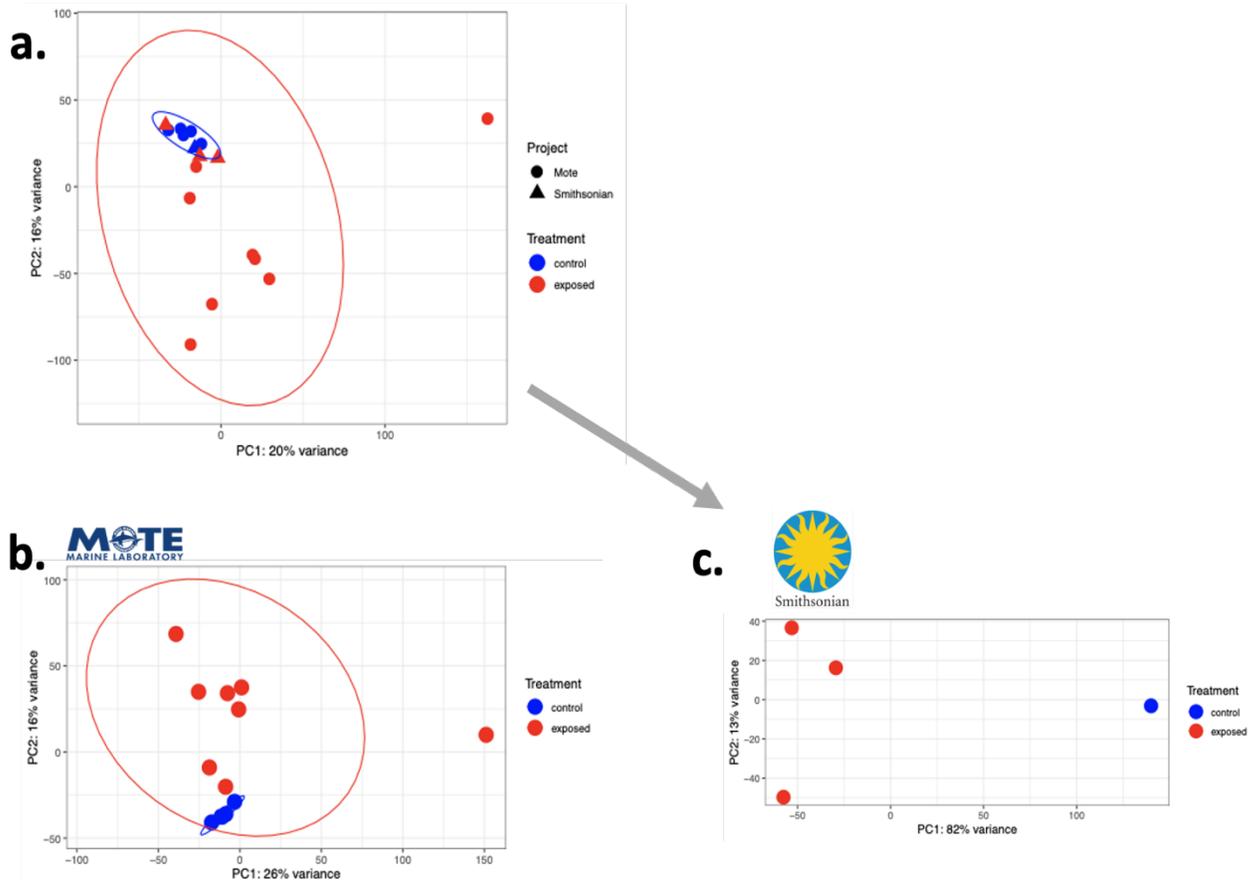
Figure 3: Principal component analysis of disease and control M. cavernosa *samples reveals the effects of treatment on gene expression. a) Overall Principal component analysis of control and diseased* M. cavernosa *reveals the effects of treatment on gene expression. Color indicates the treatment (red =diseased, blue= control, circle=Mote, triangle = Smithsonian). b) separating out the Mote and Smithsonian samples, the mote samples show a clear separation of controls (blue) and diseased (red). c) Smithsonian samples were separated by control and disease.*

3.4. Differential Gene Expression analysis

Differentially expressed genes (DEGs) were identified using a false-discovery-rate adjusted p-value cut off of ($p_{adj} < 0.05$) (Figure 4). In the overall comparison between disease and controls, *O. faveolata* had 1128 DEGs, whereas *M. cavernosa* had 61 DEGs that met the p-value threshold (Figure 4).
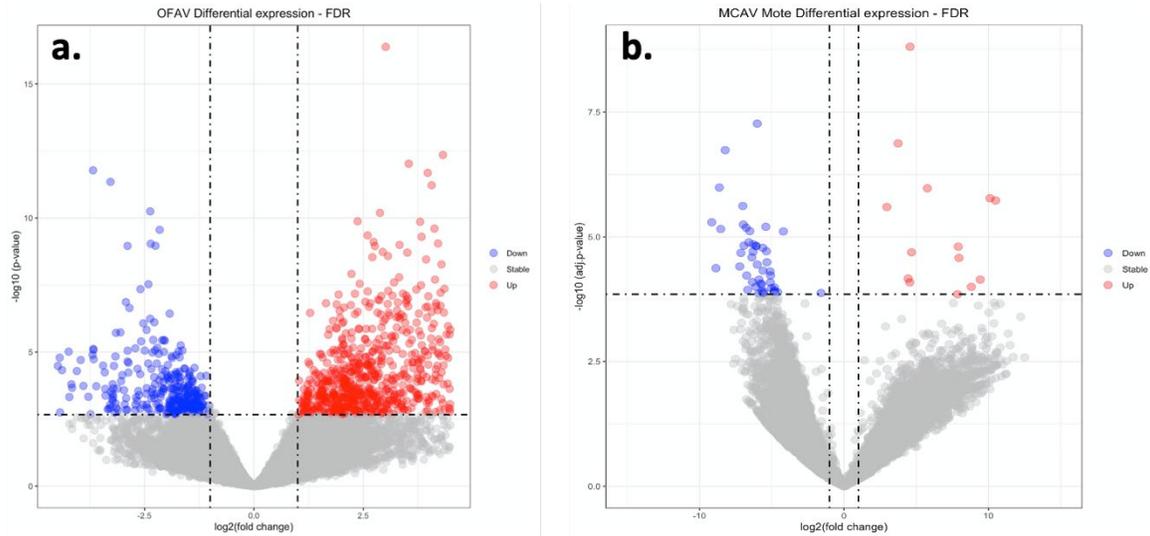
Figure 4: *Volcano plot depicting differential gene expression (DEGs) between the control and disease treatment. Log$_2$ fold-change (LFC) is on the x-axis, and the significance value (-log$_{10}$ p_{adj}) is on the y-axis. Significant DEGs ($p_{adj} < 0.05$) are colored red or blue, and all other non-DEGs are dark grey a) O. faveolata has a total of 1128 DEGs . b) M. cavernosa has a total 61 DEGs.*

Within *O. faveolata*, there were many different genes that were differentially expressed. The top ten highest expressed genes included orexin receptor type 1-like, sphingolipid delta(4)-desaturase/C4-monooxygenase DES2-like, centromere protein F-like, trifunctional purine biosynthetic protein adenosine-3-like, insulin-degrading enzyme-like, galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 3-like, mismatch repair endonuclease PMS2-like, CCR4-NOT transcription complex subunit 2-like isoform X2, monoacylglycerol lipase ABHD2-like isoform X2, and LOW QUALITY PROTEIN: tyrosine-protein kinase-like (Table 1, Figure 5a). Additionally, there were many immune and apoptosis/autophagy genes that were differentially expressed including peroxidasin homolog, TRAF5-like, NACHT, LRR and PYD domains-containing protein 3-like, ubiquitin carboxyl-terminal hydrolase CYLD-like, 14-3-3 protein 2-like, ETS-related transcription factor, Elf-1-like, TRAF 5-like, TNFRSF27-like, Elf-4-like, programmed cell death protein 4-like, apoptosis regulator BAX-like, protein C-ets-2-like isoform X3, UPF0577 protein, and KIAA1324-like homolog isoform X (Table 1, Figure 5b).

Table 1: *Top differentially expressed genes in* O. faveolata. *LFC = Log fold change.*

| LFC	Name	Function
2	apoptosis regulator BAX-like	apoptosis
2	protein C-ets-2-like isoform X3	apoptosis
-2	tumor protein p53-inducible nuclear protein 2-like	autophagy
2	UPF0577 protein KIAA1324-like homolog isoform X1	autophagy
-2	tumor protein p53-inducible nuclear protein 2-like	autophagy
-4	alpha-dioxygenase 2-like	catalyze
6	monoacylglycerol lipase ABHD2-like isoform X2	catalyze
9	trifunctional purine biosynthetic protein adenosine-3-like	catalyze
-2	hepatocyte nuclear factor 4-gamma-like isoform X1	Cell division
9	centromere protein F-like	cell division
8	insulin-degrading enzyme-like	degradation
6	LOW QUALITY PROTEIN: mismatch repair endonuclease PMS2-like	DNA repair
-5	angiopoietin-related protein 7-like, partial	ECM
-2	collagen alpha-1(I) chain-like	ECM
-2	solute carrier family 26 member 6-like	homeostasis
9	sphingolipid delta(4)-desaturase/C4-monooxygenase DES2-like	hypoxia
6	peroxidasin homolog	immunity
3	TRAF5-like, partial	immunity
3	NACHT, LRR and PYD domains-containing protein 3-like	immunity
3	ubiquitin carboxyl-terminal hydrolase CYLD-like	immunity
2	14-3-3 protein 2-like	immunity
2	ETS-related transcription factor Elf-1-like	immunity
2	TRAF 5-like	immunity
2	TNFRSF27-like	immunity
2	Elf-4-like	immunity
1	programmed cell death protein 4-like	immunity
-1	UDP-glucose 4-epimerase-like	metabolism
-1	putative glutamate synthase	metabolism
7	galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase 3-like	metal binding
6	CCR4-NOT transcription complex subunit 2-like isoform X2	mRNA degradation
-3	uncharacterized protein LOC110069624	N/A
-2	5-hydroxytryptamine receptor 1-like	neuro-related
-2	sn1-specific diacylglycerol lipase beta-like	neuro-related
6	LOW QUALITY PROTEIN: tyrosine-protein kinase-like otk	neuro-related
10	orexin receptor type 1-like	neuro-related
-3	organic cation transporter protein-like	other
-3	caveolin-3-like	other
-2	amiloride-sensitive sodium channel subunit alpha-like	other
-2	reversion-inducing cysteine-rich protein with Kazal motifs-like	other
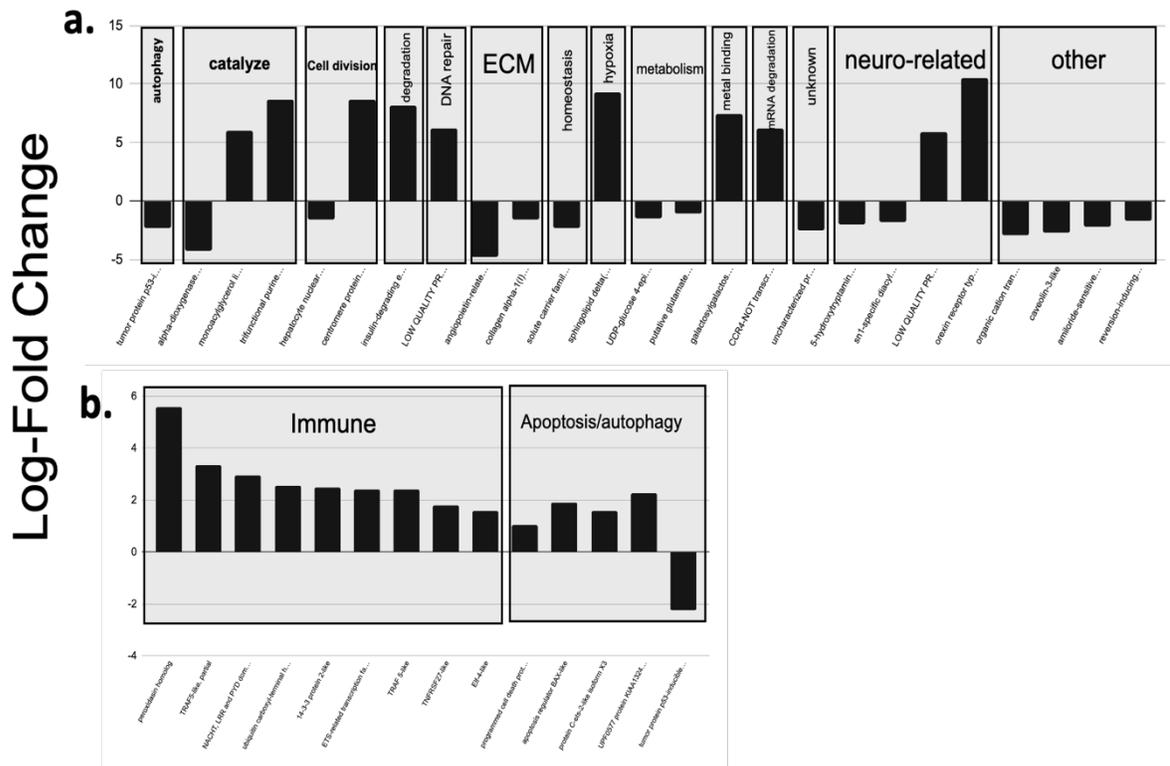
Figure 5: Differentially expressed genes for O. faveolata *exposed to disease. a) Summary of top DEGs. These include genes involved in autophagy, catalysis, cell division, degradation, DNA repair, extracellular matrix (ECM), homeostasis, hypoxia, metabolism, metal binding, mRNA degradation, neuro-related transcripts. b) DEGs related to immunity, apoptosis and autophagy.*

Within *M. cavernosa* the top ten differentially expressed genes including Galaxin , Swi5-dependent recombination DNA repair protein 1 homolog, Aggrecan core protein, A disintegrin and metalloproteinase with thrombospondin motifs 16, Dolichyldiphosphatase, Hemicentin-1, WD repeat-containing protein 36, ATP synthase subunit alpha-chloroplastic, probable methylcrotonoyl-CoA carboxylase beta chain-mitochondrial, and a sodium-and chloride-dependent GABA transporter 1 (Table 2, Figure 6a). Additionally, there were several differentially expressed genes that were in notable categories including calcification, extracellular matrix, and immunity that were significantly differentially expressed (Table 2, Figure 6b).

Table 2: Top differentially expressed genes in M. cavernosa. LFC= Log fold change

| LFC	Name	Function
12	Galaxin	calcification
9	Polycystic kidney disease protein 1-like 2	calcification
23	Swi5-dependent recombination DNA repair protein 1 homolog	DNA break repair
-7	Collagen triple helix repeat-containing protein 1	ECM
-6	Short-chain collagen C4	ECM
-6	Collagen alpha chain	ECM
-5	Collagen alpha-2(I) chain	ECM
10	Aggrecan core protein	ECM
10	A disintegrin and metalloproteinase with thrombospondin motifs 16	ECM
13	Dolichyldiphosphatase	glycosylation
11	Hemicentin-1	immunity
-6	Macrophage-stimulating protein receptor	immunity
25	WD repeat-containing protein 36	immunity
10	ATP synthase subunit alpha, chloroplastic	metabolic
23	Probable methylcrotonoyl-CoA carboxylase beta chain, mitochondrial	mitochondrial
9	Sodium- and chloride-dependent GABA transporter 1	neuro-related
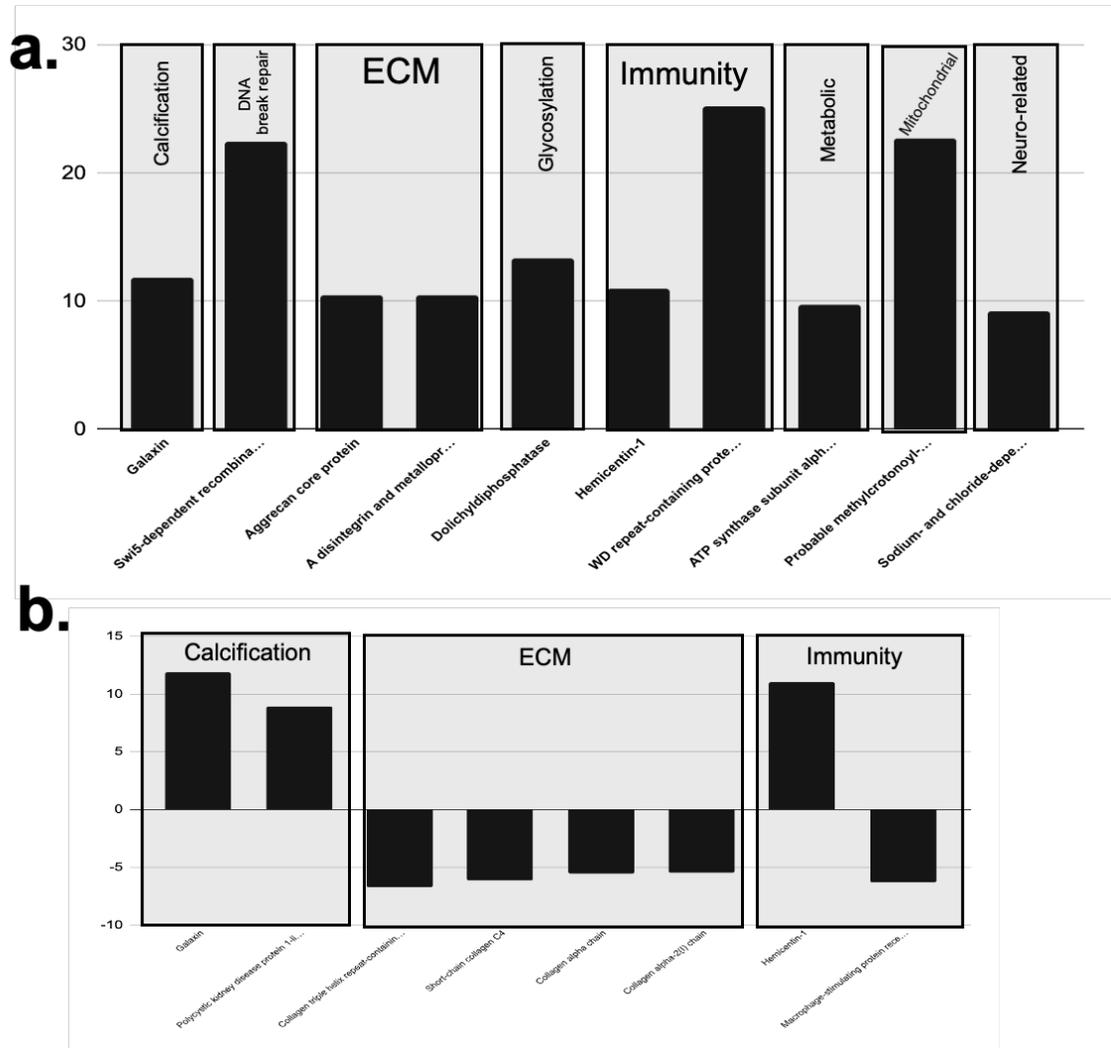
Figure 6: *Differentially expressed genes for* M. cavernosa *exposed to disease. a) Summary of top 10 DEGs. These include genes involved in Calcification, DNA break-repair, ECM, Glycosylation, Immunity, Metabolic, Mitochondrial and Neuro-related transcripts. b) Differential gene expression in response to disease included transcripts involved in calcification, ECM, and immunity.*

When comparing both the *O. faveolata* gene expression to the *M. cavernosa* gene expression there were five transcripts that were shared between the two species. These genes included cAMP-responsive element-binding protein-like 2, Collagen alpha chain, fibroblast growth factor receptor 1-like, Mitochondrial dicarboxylate carrier, Solute carrier family 26 member 6 (Table 3, Figure 7). All shared genes have the same directionality of expression for both *O. faveolata* and *M. cavernosa*.
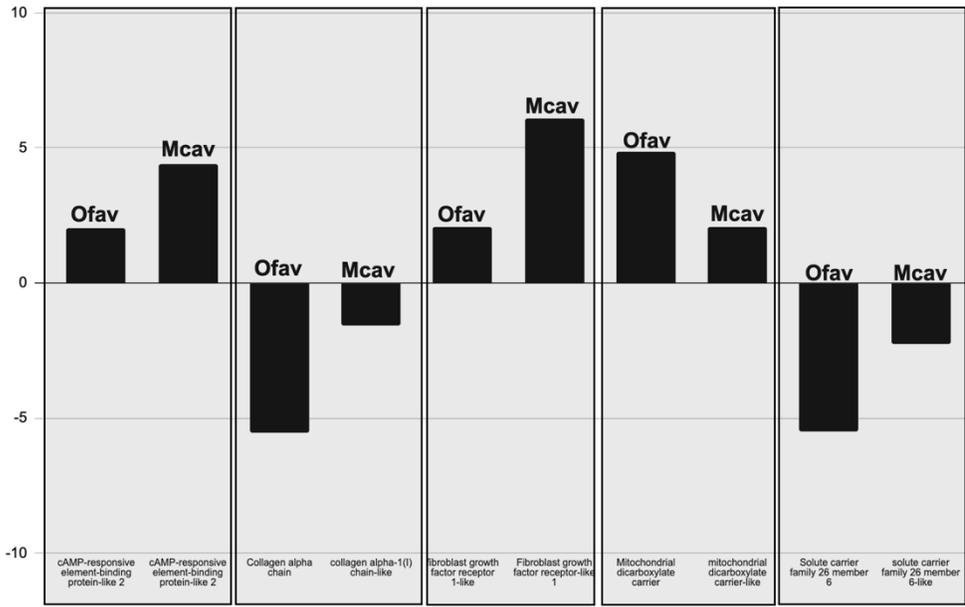
Figure 7: *Differentially expressed genes shared by* O. faveolata *and* M. cavernosa *exposed to disease. These transcripts include cAMP responsive element binding protein like 2, collagen alpha chain, fibroblast growth factor receptor 1-like, mitochondrial dicarboxylate carrier, and solute carrier family 26 member 6.*

Table 3: *Differentially expressed genes that are shared by* O. faveolata *and* M. cavernosa*. LFC = Log Fold Change.*

Coral species	LFC	Name	Function
O. faveolata	2	cAMP-responsive element-binding protein-like 2	cell cycle division
M. cavernosa	4	cAMP-responsive element-binding protein-like 2	cell cycle division
O. faveolata	-2	collagen alpha-1(I) chain-like	ECM
M. cavernosa	-6	Collagen alpha chain	ECM
O. faveolata	2	fibroblast growth factor receptor 1-like	cell surface receptor
M. cavernosa	6	Fibroblast growth factor receptor-like 1	cell surface receptor
O. faveolata	2	mitochondrial dicarboxylate carrier-like	mitochondrial
M. cavernosa	5	Mitochondrial dicarboxylate carrier	mitochondrial
O. faveolata	-2	solute carrier family 26 member 6-like	homeostasis
M. cavernosa	-6	Solute carrier family 26-member 6	homeostasis

3.5. Weighted Gene Co-Expression Analysis

A total of three modules were identified as having significant correlation values to either control, disease, or disease outcome. Out of these three only one module was significantly positively correlated with disease treatment. That module, 'Black' module had 242 total genes (Figure 8). Within the 'Black' module, enrichment for apoptosis and calcium binding were identified, indicating that this process may be important for the transcriptional response to disease.
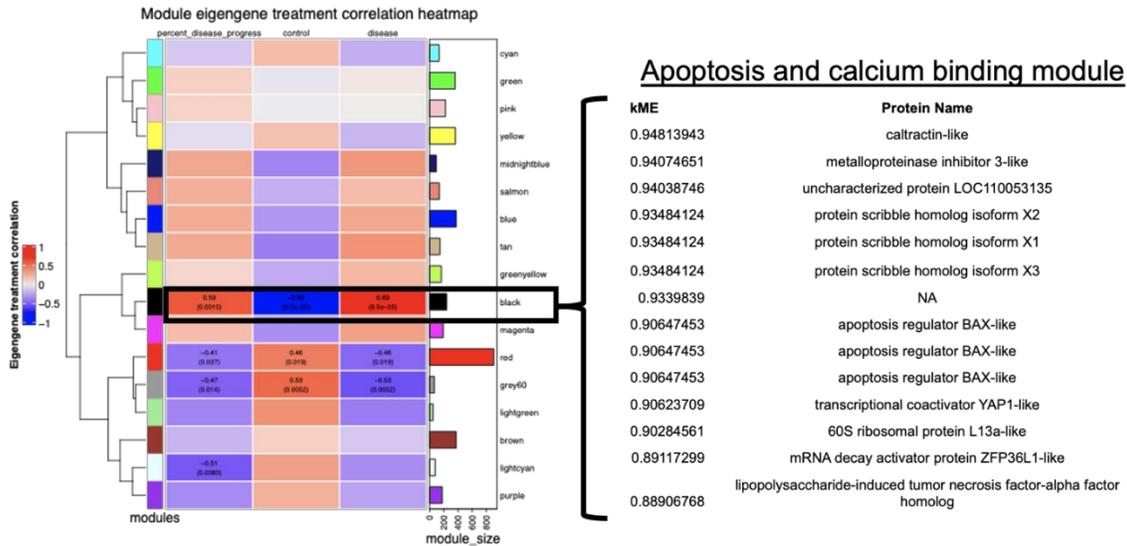
Figure 8: *Co-expression heatmap showing one module that is significantly correlated to disease ("black" module) in* O. faveolata. *Heatmap fill shows positive (red) to negative correlation (blue). The top number in each cell shows the correlation strength and the bottom number shows module significance to percent disease progression, control and disease. The 'Black' module was the only module positively correlated with disease outcome. That module highly enriched for genes involved in apoptosis and calcium binding. This may indicate that disease outcomes are correlated with apoptosis events.*

4. PRELIMINARY CONCLUSIONS

The motivation of this study was to 1) identify if the corals affected by SCTLD were immunocompromised and 2) to identify potential biomarkers that may be developed for monitoring or identifying potentially susceptible corals. Based on our preliminary findings we hypothesize that the *O. faveolata* colonies that we were examining were not immunocompromised. They were expressing immune genes, and in fact, it appears that more of an apoptosis response was occurring (Figure 5). This is in line previous histological evidence. For *M. cavernosa*, the findings are not as straightforward. Because of the low read depth, we didn't get as many differentially expressed genes, so our findings are only very preliminary. We do see that calcification, and immune genes are differentially expressed in *M. cavernosa*, so we hypothesize that they are not immunocompromised, as well (Figure 6). However, we did not see a signal of apoptosis. When comparing both of the gene sets, we did see that there were genes that were shared by both *O. faveolata* and *M. cavernosa* (Table 3, Figure 7). These five genes had the same directionality for both species, and we proposed that these could be good initial candidates for cross-species SCTLD biomarkers.

Another interesting thing to note in our dataset was that the expression of immune genes related to bacteria response were not identified. These genes (e.g., toll-like receptors and lectins) were not found to be differentially expressed in any of our datasets. While we cannot make a definitive conclusion because this is preliminary, we do want to note this, because it appears to point to the fact that whatever is causing the SCTLD reaction in corals is not showing a typical immune reaction that is indicative of an immune response. For example, in previous studies our group has found that putative white band disease transmission in *A. palmata* induces a response that is indicative of a typical innate immune response (Young et al., 2020). Further sequencing of both *M. cavernosa* and *O. faveolata* will hopefully help in our understanding of the SCTLD reaction, and lead to more definitive conclusions.

5. RECOMMENDATIONS

Recommendation 1: *Continued efforts towards in-depth 'omics initiatives to understand the coral host and the symbiont response to SCTLD.* 'Omics initiatives, such as this project are valuable for us to understand the mechanism of the disease, and lead us to developing better tools for diagnosing, and treating affected corals. Much like what is done in human medicine, 'omics tools can be used to profile the health state of corals and promote the development of potential therapeutics.

Recommendation 2: *Continued investment in the development of biomarkers for SCTLD.* Biomarkers are a great tool for diagnosing the health state of any organism. In corals, this type of work is particularly critical because most of the time when we see a disease already developing, the coral is very advanced in their disease. By developing biomarkers as a foundational health tool in corals, we will be better able to protect our reefs for the future.

Recommendation 3: *Advancing coordinated efforts using 'omics tools and histology/pathology tools to do a fine scale time series on multi-species.* Our results really do show that there is a need to coordinate between the different disease working groups, in particular the 'omics and histology working groups. The preliminary data that we have generated shows that apoptosis may be an important mechanism involved in SCTLD progression, supporting what was previously reported by the histology working group. However, based on these findings, and the fact that *M. cavernosa* didn't show the same apoptotic signal, we believe that doing a fine scale time series (every 6-12 hours post transmission), that samples for both sequencing and histology will be extremely valuable.

Recommendation 4: *Expansion of sequencing to isolated symbiont cells versus the coral host tissue in response to SCTLD.* Because we didn't find the typical immune response that would be indicative of a bacteria response, we believe that focusing on the role of symbiont in SCTLD will be very valuable. Thus, using flow cytometry and fluorescence activated cell sorting coupled with high throughput sequencing will be a valuable technique to follow in the future. Using these techniques, we will be able to tease apart the different components of the coral holobiome and examine what their role is in the progression of SCTLD.

6. REFERENCES

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*, *33*(19), 3137–3139. https://doi.org/10.1093/bioinformatics/btx373

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Matz, M. (2018). *Montastraea cavernosa* annotated genome.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

Prada, C., Hanna, B., Budd, A. F., Woodley, C. M., Schmutz, J., Grimwood, J., … Medina, M. (2016). Empty Niches after Extinctions Increase Population Sizes of Modern Corals. *Current Biology*, *26*(23), 3190–3194. https://doi.org/10.1016/j.cub.2016.09.039

Team, Rs. (2020). No Title. *RStudio: Integrated Development for R.* Boston, MA.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag. Retrieved from https://ggplot2.tidyverse.org

Young, B., Serrano, X. M., Rosales, S., Miller, M. W., Williams, D., & Traylor-Knowles, N. (2020). Innate immune gene expression in Acropora palmata is consistent despite variance in yearly disease events. *PloS One*, *15*(10), e0228514.