Local and systemic gene expression responses to a white syndrome-like disease in a reef building coral, *Acropora hyacinthus*.

R. M. Wright[*$], G. V. Aglyamova[†], E. Meyer[#] and M. V. Matz[†]

*Institute for Cell and Molecular Biology, The University of Texas at Austin

†Department of Integrative Biology, The University of Texas at Austin

Department of Zoology, Oregon State University

$ corresponding author: Rachel M Wright, rachelwright8@gmail.com

Email addresses:

RMW: rachelwright8@gmail.com

GVA: aglyamova@austin.utexas.edu

EM: eli.meyer@science.oregonstate.edu

MVM: matz@utexas.edu

Key words: Innate immunity, network analysis, gene expression, coral disease

Abstract

Background

Corals are capable of launching diverse immune defenses at the site of direct contact with pathogens, but the molecular mechanisms of this activity and the colony-wide effects of such stressors remain poorly understood. Here we compared gene expression profiles in eight healthy *Acropora hyacinthus* colonies against eight colonies exhibiting white syndrome-like symptoms, all collected from a natural reef environment near Palau. Two types of tissues were sampled from diseased corals: visibly affected and apparently healthy tissues.

Results

Tag-based RNA-Seq followed by weighted gene co-expression network analysis identified groups of co-regulated differentially expressed genes between all disease states (diseased, ahead of the lesion, and healthy). Most of the differentially expressed genes were found between tissues at the lesions and asymptomatic (healthy and ahead of the lesion) tissues. These genes were related to innate immunity, oxidative stress responses, lipid metabolism, and calcification. Network analysis also revealed groups of genes regulated specifically in the tissues from diseased colonies that were not yet showing obvious symptoms of disease, indicating a systemic response to infection.

Conclusions

These observations suggest that tissues ahead of the lesion of disease progression exist in a transitional state between health and lesion appearance. Alternatively, these gene expression profiles capture physiological differences between colonies with varying disease susceptibilities.

Background

Increasing rates of disease have contributed greatly to global coral population declines over the last few decades [1, 2]. The broadly defined "white syndrome" in Indo-Pacific regions, characterized in the field by tissue loss resulting in exposure of the coral skeleton, has been attributed to *Vibrio* spp. [3], a genus of bacteria involved in several coral diseases [4–8]. Other reports find no evidence of pathogenic bacteria in diseased corals [9] and instead blame stress-triggered programmed cell death for the manifestation of symptoms [10]. These conflicting conclusions, drawn mostly from culturing assays and histological observations, are further confounded by insufficient knowledge of the cnidarian immune response.

Corals, like all invertebrates, rely entirely on innate immunity for protection from invading pathogens. Features of innate immunity in corals include physical barriers [11], molecular pattern recognition [12], secretion of antimicrobial macromolecules [13], and cellular responses (*e.g.*, phagocytosis) [14–16]. Recent efforts to characterize those features of immunity using various cnidarian genome and transcriptome sequence databases have identified putative components of coral stress management and immune response pathways by homology with better-studied organisms. The *Acropora digitifera* genome project revealed striking differences in innate immunity complexity in corals compared to a closely related cnidarian, *Nematostella vectensis* [17]. Whereas the *N. vectensis* genome encodes only a single Toll/Toll-like receptor (TLR), the *A. digitifera* genome included at least four TLRs, along with other related immune signaling molecules. Miller *et al.* reported the presence of TLR signaling components, including adaptor proteins that link that cascade with other signaling events, in the expressed

sequence tag library of another acroporid coral, *A. millepora* [12]. Together these elements suggest an ability of corals to respond to pathogen-associated molecular patterns (PAMPs) via TLR recognition and integrate that signal to cellular responses such as inflammation and apoptosis. Toll/TLR signaling can activate NF-κB transcription factor that, upon nuclear localization, up-regulates transcription of immune response genes. In corals, the identities of those response genes and the roles they play remain unclear. Some suggested immune response genes include lectins, complement c3 and apextrin, proteins involved in non-self recognition, aggregation and cell lysis respectively [18]. Metabolism and calcification genes demonstrated differential expression in addition to immunity genes in *A. millepora* challenged with bacterial and viral immunogens, providing a more comprehensive picture of cellular events during an acute infection [19]. Global RNA-sequencing of *A. cervicornis* displaying signs of White Band Disease (WBD) revealed that disease significantly affected the expression of genes involved in immune processes and apoptosis [20]. The up-regulation of phagocytic cell surface receptors and reactive oxygen species (ROS) producing enzymes suggested that the phagocytosis and degradation of damaged cells drives the WBD response in corals. These coral sequencing projects and experimental immune challenges have provided conclusive evidence that corals are capable of launching defensive responses upon direct contact with pathogens. A coral's ability to communicate the recognition of that pathogen along the colony, however, is less understood. Coral polyps utilize a gastrovascular system lined with flagellated gastrodermal cells to transport organic products and zooxanthellae within the colony [21]. These channels are used to allocate energetic resources to areas that need them most, such as fast-growing branch tips [22–

24] and wounded regions [25]. Radiolabeled carbon accumulation experiments have shown that corals preferentially direct energetic resources *towards* physically damaged regions [25] but *away* from disease-induced lesions [26]. These findings suggest that healthy coral tissues might possess means to detect and respond to an advancing disease lesion, but it is still unclear what the physiological consequences of this action might be. Here we examine the gene expression profiles of *A. hyacinthus* affected with white syndrome-like symptoms (Figure 1A) to determine the molecular consequences of the disease. White syndrome advances along a colony in a way such that a distinct lesion forms between affected and unaffected tissues (Figure 1B). Tissues ahead of the lesion are presumably healthy, while tissues at the lesion boundary are actively sloughing cells in response to infection. We compared gene expression profiles among three disease states: (1) affected and (2) apparently healthy tissues from diseased colonies and (3) tissues from completely unaffected colonies. Comparing healthy regions of diseased colonies to completely disease-free individuals provided an opportunity to look for expression patterns that might indicate disease susceptibility and/or a colony-wide systemic effect of infection. We used tag-based RNA-Seq, which allows systems-level investigations, to probe all cellular responses to infection. Gene co-expression network analysis revealed groups of co-regulated genes with shared cellular functions correlating to different disease states. We compiled this information to provide a system-wide assessment of the systemic molecular response to chronic disease in corals.

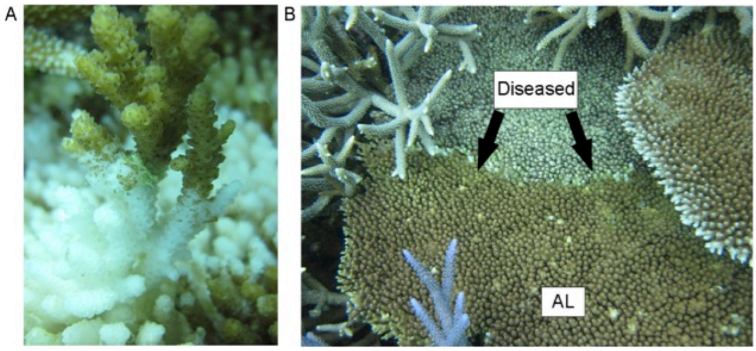

Figure 1: Directional lesion progress along *A. hyacinthus* colonies affected with white syndrome-like symptoms (A). Coral fragments were sampled at the lesion boundary (diseased, "D") or ahead of the lesion ("AL") (B). Tissues were also sampled from completely healthy individuals ("H", not pictured). Photo credit: Carly Kenkel.

Results

Differential Gene Expression between Disease States

Sequencing yielded an average of 6,367,219 reads per sample. An average of 19.5% of these remained after filtering and of these, an average 31.45% mapped to the transcriptome. A total of 44,701 isogroups (clusters of contigs representing the same gene, from here on referred to as "genes") were detected. Reads were converted to unique transcript counts by removing PCR duplicates, yielding an average of 78,325 counts per sample. Technical replicates were summed for a final average of 156,650 unique transcript counts per sample. A generalized linear model with contrasts between all three tissues detected differentially expressed genes between disease states (Additional File 1). The disease-healthy contrast yielded 646 DEGs passing a Benjamini-Hochberg FDR cutoff of 10% and 2459 DEGs with an unadjusted p-value of less than 0.05. The disease-AL contrast yielded 333 DEGs passing an FDR cutoff of 10% and 2054 DEGs with an unadjusted p-value of less than 0.05. No DEGs passed the 10% FDR cutoff for the healthy-AL contrast, but 777 DEGs had an unadjusted p-value of less than 0.05. Between all contrasts, a total of 757 genes passed the FDR cutoff of 10%.

Principal coordinate analysis of the variance-stabilized data for all genes revealed expression differences mainly between disease (D) and the other two health states (Figure 2A). PCoA using the 3827 isogroups with an unadjusted p-value of less than 0.05 for any contrast revealed more differences between disease states, but a significant overlap in expression of healthy and AL tissues remains (Figure 2B).

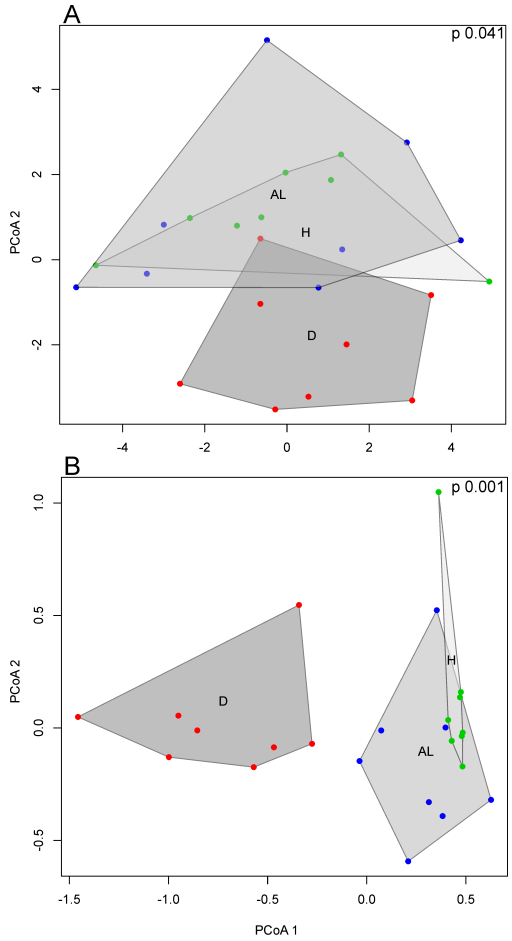

Figure 2: Principal coordinate analysis of variance stabilized data clusters samples by disease state. Samples by the presence of disease symptoms (D vs. AL and H) when all isogroups are included in the PCoA (A). Differences between disease states become more evident when PCoA is performed on DEGs (unadjusted p-value < 0.05) only (B).

Gene Ontology (GO) Enrichment

Functional enrichments between all three contrasts allow a general examination of the molecular functions and biological processes being differentially regulated between disease states. The enriched groups of both the disease-healthy and disease-AL contrasts were largely identical (Figures 3 and 4).

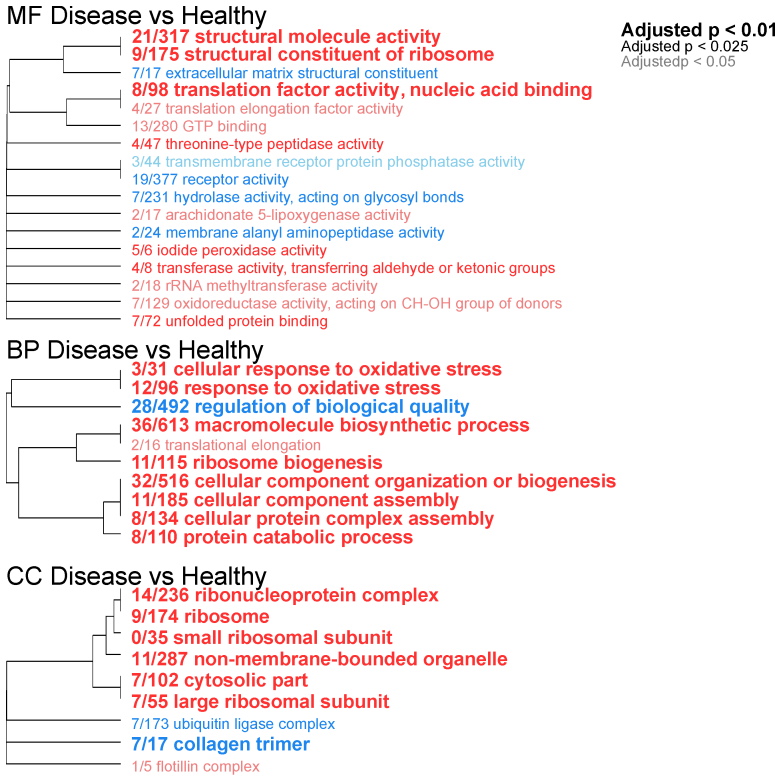

Figure 3: GO enrichment by molecular function (MF), biological process (BP), and cellular component (CC) of genes by p-value generated for the disease-healthy contrast. The color of the text indicates the direction of expression change between disease and healthy tissues (red = up-regulated in disease, blue = down regulated in disease). The size of the text indicates the significance of the term as indicated by the inset key. The fraction preceding the GO term indicates the number of genes annotated with the term that pass an unadjusted p-value threshold of 0.05. The trees indicate sharing of genes among GO categories (the categories with no branch length between them are subsets of each other).

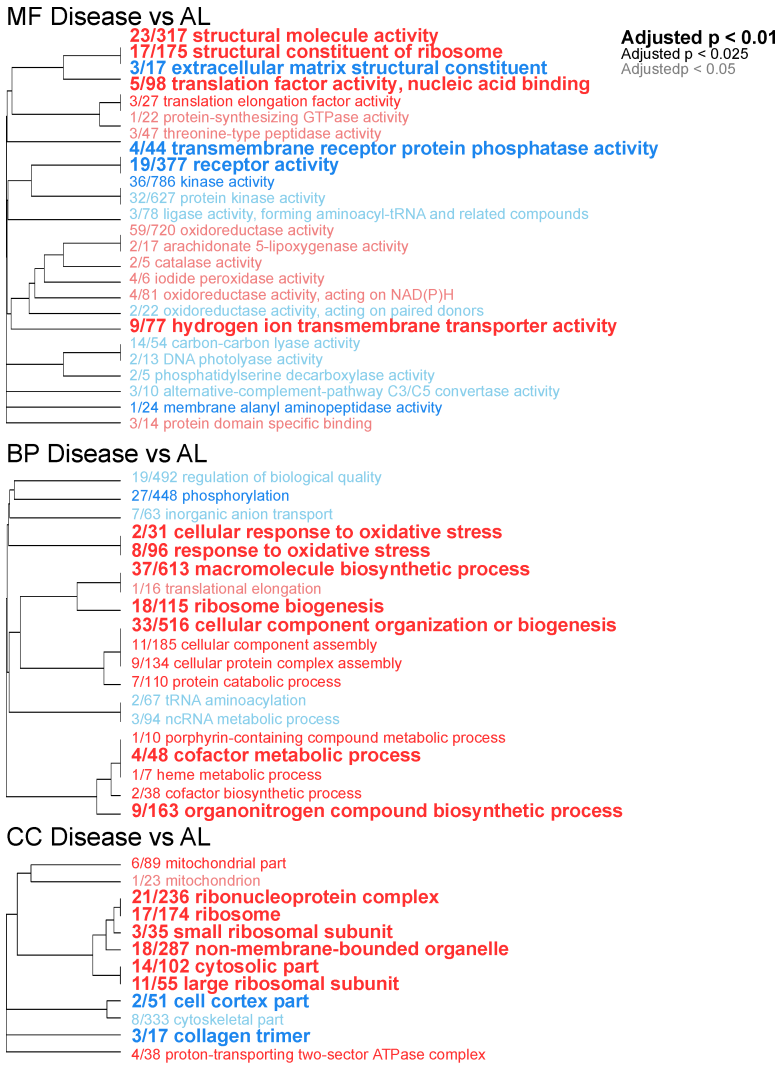

Figure 4: GO enrichment by molecular function (MF), biological process (BP), and cellular component (CC) of genes by p-value generated for the disease-AL contrast. The color and font representations are the same as described in Figure 2.

Ribosomal subunits, oxidative stress responses, and translation factor activity were up regulated in diseased tissues compared to both AL and healthy tissues. Likewise, receptor activity, regulation of biological quality, and extracellular matrix components (collagens) were down regulated in diseased tissues compared to both healthier states. Dissimilarities in enriched functions between the diseased-healthy and diseased-AL contrasts were few, reflecting the fact that healthy and AL states were quite similar. However, there were some differences in functional enrichment. Unfolded protein

binding was enriched for the comparison between healthy and diseased tissues (Figure 3), while tRNA metabolism was enriched only between diseased and AL tissues (Figure 4). To further investigate these differences between health and AL tissues we performed GO enrichment analysis on genes based on the p-value generated for the healthy-AL contrast. However, no enriched GO terms were identified with an adjusted p-value of less than 0.1.

Gene Expression Analysis by Contrast

Hierarchically clustered gene expression heatmaps were constructed to show the relative expression patterns of the top most significant annotated DEGs for each contrast (Figures 5-7). Complete lists of all annotated DEGs with log fold changes for each contrast can be found in Additional Files 2-4.

Diseased vs. Healthy

Genes found to be up regulated (Benjamini-Hochberg FDR < 0.01) in diseased tissues compared to healthy corals include key members of the oxidative stress response in corals (*e.g.*, catalases and peroxidases) and pentose phosphate metabolism (transketolase, transaldolase, and 6-phosphogluconate dehydrogenase). Both proteinases (astacin and cathepsin L) and protease inhibitors (alpha-macroglobulin and serine proteinase inhibitor Ku-type) were up regulated in diseased tissues. C-type lectin, a carbohydrate-binding protein, and malate synthase, a key enzyme of the glyoxylate cycle, were also up regulated in symptomatic tissues.

Down regulated genes (Benjamini-Hochberg FDR < 0.01) include those encoding extracellular matrix constituents (collagens, heparin sulfate proteoglycans) and carbonic anhydrase, a key enzyme in coral skeletal deposition. Red fluorescent protein was also down regulated in diseased tissues, a hallmark of the coral stress response [27–31].

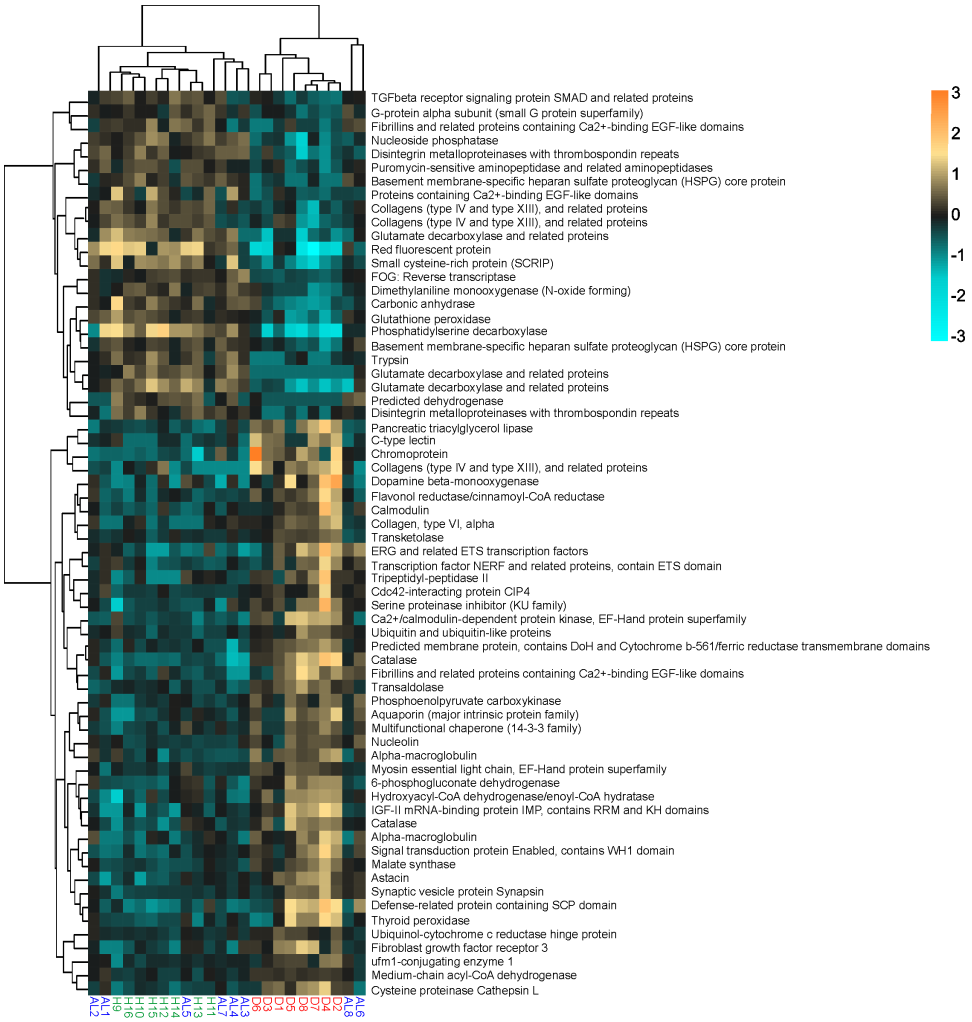

Figure 5: Gene expression heatmaps for annotated DEGs (adjusted p-value < 0.01) for the disease-healthy contrast. Rows are genes, columns are samples ordered as in the bottom panel: ahead-of-lesion (AL), healthy (H), and diseased (D). The color scale is in $\log_2$ (fold change relative to the gene's mean). The trees are hierarchical clustering of genes based on Pearson's correlation of their expression across samples.

Diseased vs. AL

The expression differences between diseased and AL tissues within a colony paralleled the expression differences between diseased tissues and healthy corals. At the same significance threshold (Benjamini-Hochberg FDR < 0.01), almost the exact same genes were identified (Figure 6). Catalases, peroxidase, alpha-macroglobulin, cathepsin L, and malate synthase were up regulated in diseased tissues compared to tissues ahead of the lesion. Likewise, fluorescent proteins (red and green/cyan) and extracellular matrix components were down regulated in diseased compared to AL tissues.

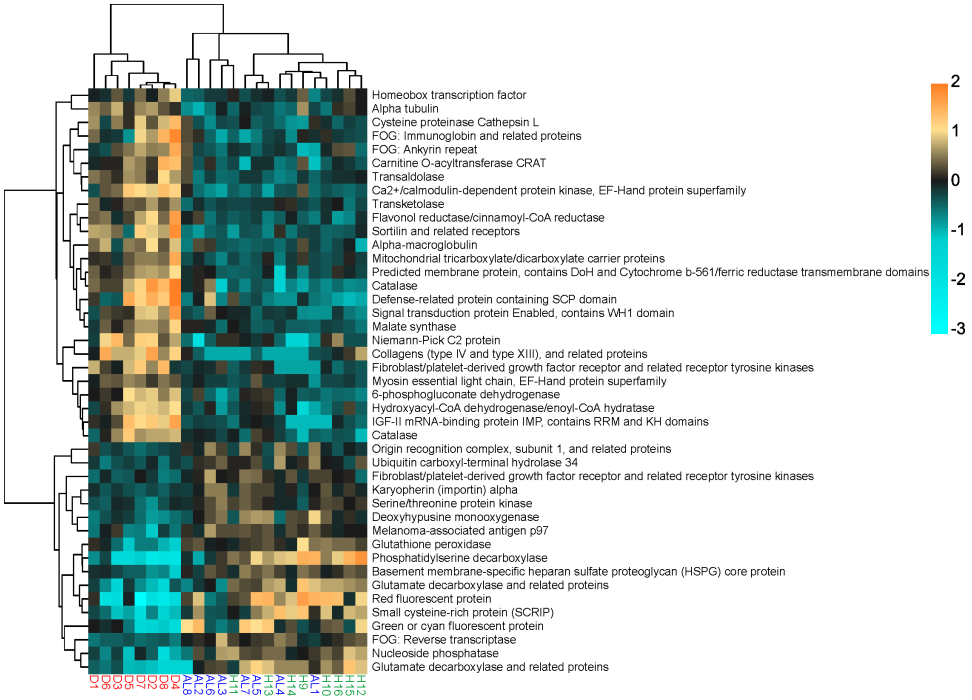

Figure 6: Gene expression heatmaps for annotated DEGs (adjusted p-value < 0.01) for the disease-AL contrast. Rows are genes, columns are samples ordered as in the bottom panel: ahead-of-lesion (AL), healthy (H), and diseased (D). The color scale is in $\log_2$ (fold change relative to the gene's mean). The trees are hierarchical clustering of genes based on Pearson's correlation of their expression across samples.

AL vs. Healthy

Although no differentially expressed genes passed the Benjamini-Hochberg 10% FDR when comparing the AL tissues and healthy corals, at a more lenient unadjusted p-value threshold of 0.01 we see that AL tissues appear to down-regulate hypoxia-inducible factor prolyl 4-hydroxylase and up-regulate a caspase, metalloendopeptidase and transcription factor AP-1. The absence of DEGs passing the FDR cutoff and paucity of DEGs at a more lenient threshold reaffirm the finding that gene expression is broadly similar between healthy corals and corals ahead of progressing disease lesions. However, slight signatures of differential gene expression suggest that there might still be some systemic physiological effect of nearby disease symptoms along a coral colony, but DESeq2 analysis alone does not provide enough power to detect it.

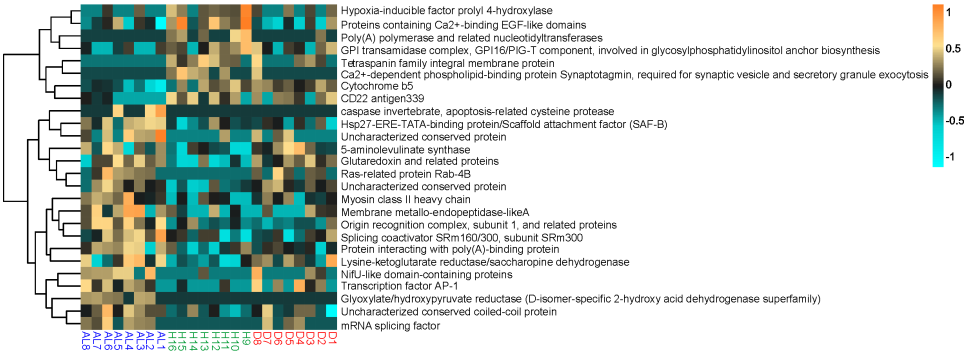

Figure 7: Gene expression heatmaps for annotated DEGs (unadjusted p-value < 0.01) for the AL-healthy contrast. Rows are genes, columns are samples ordered as in the bottom panel: ahead-of-lesion (AL), healthy (H), and diseased (D). The color scale is in $\log_2$ (fold change relative to the gene's mean). The trees are hierarchical clustering of genes based on Pearson's correlation of their expression across samples.

Correlation Between Gene Network Modules and Disease States

A total of 6737 DEGs with unadjusted p-value < 0.1 were input into WGCNA for network analysis. One sample (diseased individual "4") was identified as an outlier and removed from subsequent analysis. Twelve unique modules, denoted by arbitrarily assigned colors, remained after merging highly correlated modules (Additional File 5). Of these twelve modules, eight were highly correlated to a single coral individual and one (grey) is reserved to contain genes that do not fall into any co-expression module. The remaining three modules were highly correlated with the disease states. Since we assembled these modules using a signed network, the sign of the correlation is equivalent to the direction of expression change with respect to the trait. For example, a module that is significantly negatively correlated to diseased corals contains genes that are down regulated in that state.

The dark green module (1155 genes) strongly correlated with diseased and healthy coral. The genes within this module are up regulated in diseased tissues and down regulated in healthy tissues. Similarly, the turquoise module (669 genes) also correlated with diseased and healthy corals but with the reverse sign: these genes are up regulated in healthy tissues and down regulated in diseased tissues. Notably, one module was identified (green, 661 genes) that was significantly up-regulated in AL and (to a lesser extent) down regulated in diseased tissues, while remaining unchanged in healthy tissues.

Within Module Gene Expression Analysis

Hierarchically clustered gene expression heatmaps were constructed to show the relative expression patterns of the genes within each module that best represent the module and show significant correlations to the disease state, based on module membership and gene significance values (Additional File 6 and Figure 8).

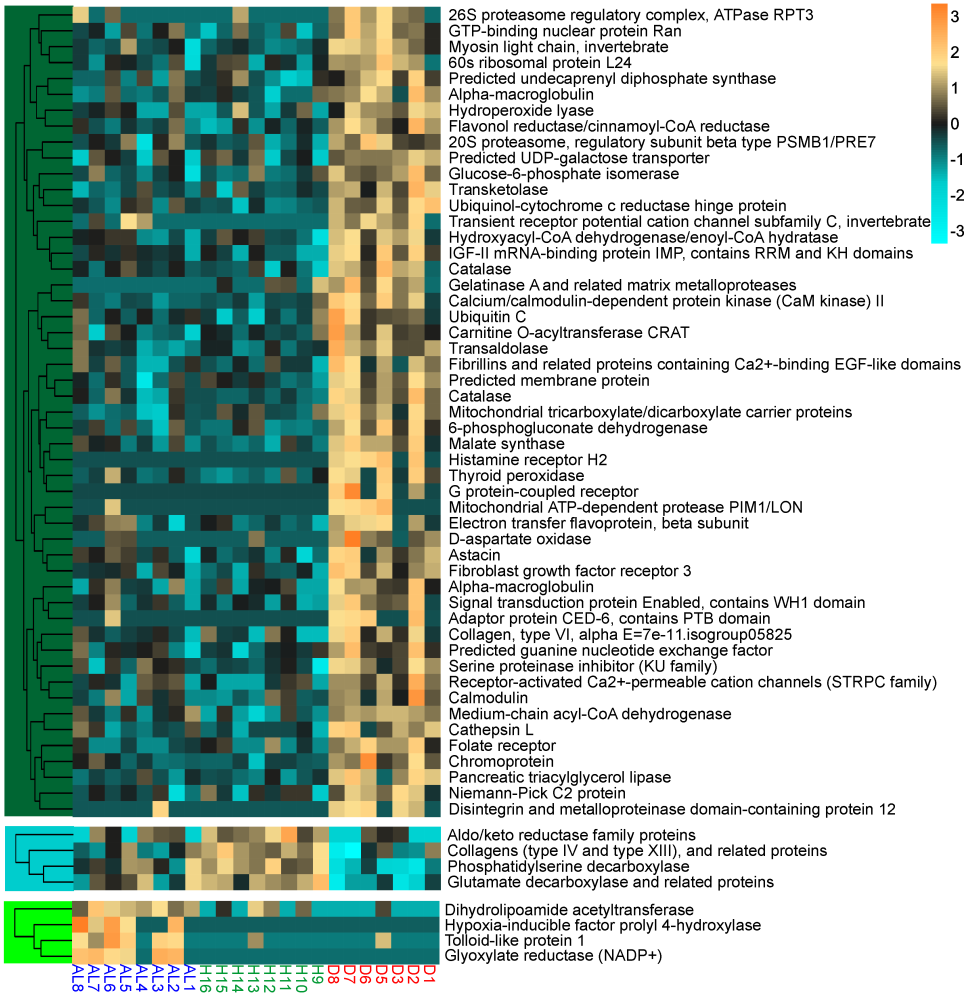

Figure 8: Gene expression heatmaps of annotated DEGs with a module membership and gene significance score greater than 0.6 within their respective modules. Rows are genes, columns are samples ordered as in the bottom panel: ahead-of-lesion (AL), healthy (H), and diseased (D). The color scale is in $\log_2$ (fold change relative to the gene's mean). The trees are hierarchical clustering of genes based on Pearson's correlation of their expression across samples. The color block of the trees indicates the module to which these genes belong (dark green, turquoise, and green).

Discussion

The major pattern of variation in gene expression was between asymptomatic (healthy and AL) and diseased (D) corals (Figure 2). Contrasts between diseased corals with both healthy and ahead-of-the-lesion tissues produced many differentially regulated genes. The differentially regulated genes between diseased and asymptomatic tissues represent elements of an immune response and a general stress response. Contrasts between healthy and AL gene expression generated far fewer differences, suggesting that these tissues are physiologically similar. However, gene co-expression network analysis and subsequent correlation to disease states revealed expression patterns distinct to each of the three states.

Diseased tissues up regulate immune response elements

Innate immunity provides immediate protection against non-self and responds to physical injury. Three general steps are involved in an innate immune response: detection, defense activation, and effector responses to neutralize the threat. Tissues sampled from the lesion of disease progression in corals experiencing white syndrome-like symptoms have enhanced expression of genes involved in each of these three immune response phases. C-type lectins act as pattern recognition receptors to activate pathogen elimination through phagocytosis in invertebrates [32]. Cnidarian genomes encode *c-type lectin* genes with highly variable substrate regions, leading to hypotheses that these proteins recognize a large variety of pathogens [33]. In *A. millepora*, mannose-binding C-type lectins have been shown to respond immediately following an immune challenge (only 45 minutes after lipopolysaccharide injection in [34]), but show no significant response at later time points [35]. The up regulation of C-type lectins in tissues at the

lesion may suggest that these tissues have very recently become infected. The second phase of an immune response prepares targets for elimination via antimicrobial peptide synthesis and immune cell activation. While this experiment did not discover any differentially regulated antimicrobial peptides, we do detect the activation of immune activating proteins C4, alpha-macroglobulin and CD109 in diseased tissues. The lectin pathway of immune activation is triggered by lectins binding a pathogen-associated molecule and results in the activation complement component factor C4 and C3 [36]. These complement factors, along with alpha-macroglobulin and CD109, tag pathogens and secreted proteases for elimination. In the final phase of an innate immune response, foreign organisms are engulfed and destroyed by phagocytic immune cells. Lysosomes within these phagocytic cells contain proteins capable of degrading engulfed material via the production of reactive oxygen species (ROS) or proteolytic enzymes, such as cathepsins. The up regulation of cathepsin L in diseased tissue may be a consequence of such phagocytic activity.

Glyoxylate cycle and gluconeogenesis transcripts are up regulated in diseased tissues

Transcripts involving lipid and carbohydrate metabolism (triacylglycerol lipase, phosphoenolpyruvate carboxykinase), including glyoxylate cycle metabolism (malate synthase), were up regulated in diseased tissues compared to asymptomatic tissues. The differential regulation of these metabolic genes suggests that diseased corals may be utilizing stored energy reserves more than healthy corals. Fatty acids derived from stored lipids are oxidized by the beta-oxidation pathway and release acetyl-CoA to enter the citric acid cycle. Both carbons in one molecule of acetyl-CoA are consumed during the

decarboxylation steps of the citric acid cycle and energy is released. The glyoxylate cycle is an alternative route through the citric acid cycle that allows organisms to thrive on two-carbon sources by catalyzing the conversion of acetyl-CoA to malate and succinate via a glyoxylate intermediate, bypassing the decarboxylation steps in the citric acid cycle [37]. These four carbon compounds contribute to the energetic requirements of the cell and serve as building blocks for cellular components, fulfilling all the same necessities of the citric acid cycle without the need to replenish oxaloacetate from the diet. One of the most significantly up regulated transcripts in diseased coral tissues was malate synthase, one of the key enzymes of the glyoxylate cycle along with isocitrate lyase. Glyoxylate cycle enzymes are rare throughout the animal kingdom, but bioinformatic analyses suggest they exist in cnidarians [38]. As additional support for a potential role of glyoxylate cycle metabolism in coral stress responses, glyoxylate cycle transcripts were up regulated in *A. palmata* larvae subjected to thermal stress [39]. In higher plants, glyoxylate enzymes are active when the cell is switching from photosynthetic production of sugars to scavenging pathways from stored and structural lipids, as in starvation and/or senescence [40]. In corals, this metabolic shift might indicate a decline in shared energy reserves with zooxanthellae, presumably due to stress-induced symbiont loss.

Oxidative stress response genes are up regulated in diseased tissues

Reactive oxygen species (ROS) are produced as a consequence of fatty acid oxidation. The up regulation of antioxidants that protect the cell from these harmful byproducts corals could be a consequence of increased fatty acid metabolism. This explanation

coincides well with the observed up regulation of lipid metabolism and antioxidant (catalase, peroxidase) transcripts in diseased corals. The production of ROS is also a fundamental element of the innate immune response. While ROS are capable of neutralizing phagocytized pathogens, the harm they cause to the host must be countered if an organism is to withstand its own immune response. Catalases and peroxidases capable of hydrolyzing harmful peroxides provide a mechanism of such self-protection. The up-regulation of oxidative stress response genes is well characterized in corals experiencing thermal stress [41], physical stress [42], and infectious disease [20]. Thus the identification of differentially expressed ROS response genes was anticipated, but it is interesting to note that up regulation of these genes was confined to diseased tissues and was not seen elsewhere in symptomatic colonies. A possible explanation for this finding is that asymptomatic tissues near the lesion are not directly encountering a bacterial pathogen.

Matrix metalloproteinases are up regulated in diseased tissues

Stony corals are subject to many potential sources of physical injury such as predators [43], burrowing bivalves [44] and storms [45, 46]. The tissue regeneration mechanisms employed by corals that have sustained a physical injury are common to wound-healing processes across metazoans [47]. One of these steps involves restructuring of the extracellular matrix to encourage tissue regeneration. Matrix metalloproteinases (MMPs) are a group of enzymes capable of such activities and have been shown to play a direct role in wound repair in Hydra [48]. MMPs also act on pro-inflammatory cytokines to direct inflammation due to wounding and innate immune responses to pathogens

(reviewed in [49]). The up regulation of MMPs in response parasitic protists in a gorgonian coral suggests that these proteins are active in the immune response of cnidarians [50]. Astacin and gelatinase have matrix metalloproteinase activities and were up regulated in affected coral tissues in this study. These activities suggest that these tissues at the disease lesion are launching an active immune defense and/or regenerating lost tissue. Inhibitors of protease activity were also up regulated in diseased coral tissues. One of these proteins, the Kunitz-type serine protease, exhibited antimicrobial activity in Hydra [51]. Another up regulated protease inhibitor, alpha-macroglobulin, is a vital component of the innate immune response that inactivates bacterial secreted proteases, thus compromising their virulence [52]. The up regulation of alpha-macroglobulin in diseased (but not healthy or AL) tissues supports the notion that only diseased tissues are in direct contact with bacterial pathogens. The activation of these protease-inhibitors implies immune function in visibly infected tissues, but not necessarily in regions ahead of the disease lesion.

Calcification genes are down regulated in diseased tissues

Calcification rates in reef-building corals are sensitive to several environmental variables such as light, pH, and temperature [1, 53–55]. While this experiment did not directly measure coral calcification, the identification of DEGs with functions in biomineralization suggests that disease negatively impacts coral skeletal deposition. Both general extracellular matrix structural components and coral-specific calcification functions were differentially regulated in diseased tissues compared to asymptomatic tissues. Coral biomineralization is directed by an extracellular skeletal organic matrix

comprised of secreted glycosylated proteins [56]. These proteins include collagens and negatively charged macromolecules (like chondroitin sulfate proteoglycans) that bind calcium ions to aid in crystal formation [57]. Several collagens and a protein with high similarity to nematogalectin, a collagen family protein that forms a major structural component in Hydra nematocyst tubules [58], were down regulated in diseased tissues. The down regulation of these genes in diseased tissues suggests a weakening of the coral skeletal organic matrix and thus a diminished capacity for biomineral deposition. The potential impact of disease on coral skeletal growth is most clearly revealed by the down regulation of carbonic anhydrase, an enzyme that plays a fundamental role in mediating bicarbonate supplies for calcification in scleractinian corals [59–61].

AL gene expression: a systemic response to disease or factors contributing to disease susceptibility?

Recent studies investigating the colony-wide transcriptomic effects of yellow band disease (YBD) in *Orcibella faveolata* revealed differences between healthy tissues, diseased tissues, and apparently healthy tissues ahead of the yellow band lesion [62]. YBD-infected corals displayed intermediate expression patterns where expression ahead of the lesion responded more similarly to healthy corals in some genes and more similarly to diseased corals in others. Only one DEG was identified between ahead of the lesion and diseased tissues. Conversely, in our study gene expression between asymptomatic (healthy and AL) and diseased corals dominates the observed transcriptomic response. However, principal coordinate analysis and gene co-expression network analysis both demonstrate differences between healthy and AL tissues. Coral tissues ahead of the

lesion of disease progression tend to show more variability in gene expression than both entirely healthy and diseased corals (Figure 1). Expression variation between individuals in the intermediate disease state could signify that these tissues are at different stages in a transition to lesion appearance. This variation may explain the lack of power to detect significantly differentially expressed genes between AL and healthy tissues in DESeq2 analysis. However, gene co-expression network analysis was sensitive enough to identify modules correlated (*post hoc*) with all three coral conditions. The module that correlated specifically with the AL state confirms that these tissues are distinct from both diseased and healthy tissues and may be responding to the adjacent threat. One of the genes that best represent the AL-correlated module encodes hypoxia inducible factor prolyl 4-hydroylase (HIF-P4H). HIF-P4H was up regulated in AL tissues compared to both diseased tissues from the same colonies and healthy individuals. These proteins regulate gene expression in response to cellular oxygen levels and inflammation [63]. HIFs rapidly accumulate in low oxygen environments and induce hypoxia-responsive genes that function in oxidative stress responses, glucose metabolism, extracellular matrix homeostasis, and innate immunity (reviewed in [64]). HIF-P4Hs are responsible for maintaining appropriate levels of HIFs in response to oxygen availability in the cell [65]. In normoxic conditions HIF-P4H modifies the alpha subunit of HIF-1 in an oxygen-dependent reaction, leading to its eventual degradation in the proteasome. A study performed in rat cell culture observed that the expression of prolyl hydroxylases is up regulated in hypoxic environments, possibly as a mechanism to halt hypoxic signaling in freshly re-oxygenated cells [66]. The up regulation of HIF-P4H in apparently healthy tissues ahead of the lesion therefore suggests that these tissues are responding to variable

oxygen availability. HIF-P4H has also been shown to negatively regulate immune responses by modifying the kinase responsible for releasing NF-κB from its inhibitor in an oxygen-dependent reaction [67].

White syndrome is a general term that refers to both active and recovering disease lesions [68]. Since we did not monitor the colonies after sampling it is unknown whether the gene expression response in tissues ahead of the lesion describes tissues that are preparing to defend against or successfully overcoming an infection. The transitional gene expression profile of the AL state could signify either a return to homeostasis or an oncoming immune challenge. Yet one more possible explanation is that the differences in gene expression profiles between AL tissues and healthy corals reflect baseline differences in susceptibility that allowed the healthy corals to resist infection. In other words, the differential regulation of these genes could be either a consequence of disease progression, recovery from the disease, or an indicator of susceptibility.

Conclusions

Our gene expression analysis identified molecular processes specific to corals affected with white syndrome-like symptoms. Information provided by this analysis complements other investigations of coral disease responses to better understand the physiological consequences of infection in both visibly affected tissues as well as apparently healthy tissues adjacent to the lesion of disease progression. An increased knowledge of the systemic molecular effects of coral disease will contribute towards the development of diagnostic tools to predict and manage coral disease outbreaks.

Abbreviations

AL: Ahead-of-the-lesion

DEG: Differentially expressed genes

FDR: False discovery rate

FOG: Fuzzy orthologous groups

GO: Gene ontology

HIF-P4H: Hypoxia-inducible factor prolyl 4-hydroxylase

KEGG: Kyoto encyclopedia of genes and genomes

KOG: EuKaryotic orthologous groups

MMP: Matrix metalloproteinase

NF-κB: Nuclear factor kappa-light-chain-enhancer of activated B cells

PAMP: Pathogen-associated molecular pattern

ROS: Reactive oxygen species

TLR: Toll-like receptor

WBD: White band disease

WGCNA: Weighted gene co-expression network analysis

YBD: Yellow band disease

Competing Interests

The authors declare that they have no competing interests.

Methods

Sampling

Coral fragments from 16 colonies of *A. hyacinthus* were sampled in the spring of 2011 along the eastern coast of Palau (7° 18.738' N, 134° 30.423' E) and immediately stored in RNAlater (Ambion). Eight of these colonies were visibly affected with white syndrome-like symptoms. The remaining eight colonies were completely symptom-free (designated healthy, H). From the eight affected colonies, coral fragments were sampled from both the lesion interface between diseased and healthy tissues (diseased, D) and areas well ahead of the lesion (AL, Figure 1B). AL coral fragments were sampled from approximately midway between the lesion boundary and the edge of the colony. AL tissues are unlikely to be in direct contact with pathogen since previous studies have demonstrated declines in pathogens only ~1 cm in advance of a white syndrome disease lesion [69].

Transcriptome Assembly and Annotation

The *A. hyacinthus* transcriptome has been generated from 5-day aposymbiotic larvae as described previously [70]. It was annotated based on two resources: the proteome of the starlet sea anemone *Nematostella vectensis* [71] and in-depth annotations of the *Acropora digitifera* proteome [72]. Based on manual verifications of a subset of *A. digitifera* annotations, they were pre-filtered to include only protein sequences longer than 60 amino acids with the annotation assigned based on the listed e-value = 1e-20 or better. The GO, KEGG ("Kyoto Encyclopedia of Genes and Genomes"), KOG ("euKaryotic orthologous groups"), and gene name annotations were transferred to an *A. hyacinthus*

contig if the contig matched one or both of these two resources with e-value = 1e-4 or better in blastx [73]. The GO and KOG annotations assigned to genes that were denoted FOG ("fuzzy orthologous group", [74]) in the *N. vectensis* data were removed, since such genes encode proteins with common domains and cannot be functionally annotated based on homology alone. The annotated *A. hyacinthus* transcriptome has been released for unrestricted use prior to this publication, http://www.bio.utexas.edu/research/matz_lab/matzlab/Data.html.

Tag-based RNA-Seq

Duplicate libraries were prepared following [75] and sequenced using Applied Biosystems SOLiD v.3 platform. Read trimming, quality filtering, mapping, and conversion to per-gene counts was performed as described previously [75] with one modification: reads mapping to the same starting coordinate in the reference and aligning with 100% identity along the full length of the shorter read were discarded as potential PCR duplicates. The current step-by-step library preparation protocol as well as bioinformatic pipeline are available at https://sourceforge.net/projects/tag-based-rnaseq/ ; note however that the current version of the tag-based RNA-seq method assumes sequencing on the Illumina HiSeq instrument.

Identification of Differentially Expressed Genes (DEGs)

All statistical analyses were performed using R3.1.1 [76]. DEGs were identified using a generalized linear model implemented by the R package DESeq2 [77]. No outlying samples were detected by the arrayQualityMetrics package [78]. DESeq2 performed

automatic independent filtering to remove lowly abundant transcripts and maximize the rate of DEG discovery post multiple testing correction at an alpha of 0.1. P-values for significance of contrasts between all three disease states were generated based on Wald statistics and were adjusted for multiple testing using the false discovery rate method [79]. The contrasts resulted in tables including adjusted and unadjusted p-values and $\log_2$ fold changes that were used in downstream analyses.

Gene Coexpression Network Analysis

A weighted gene coexpression network analysis (WGCNA, [80]) was used to identify groups of co-regulated genes in an unsupervised way. Genes with an unadjusted p-value < 0.1 for any of the three contrasts as determined by the generalized linear model testing for the effect of disease state were input into WGCNA. A sample network was constructed to identify outlying samples with a standardized connectivity score of less than -2.5 [81]. A signed gene co-expression network was constructed with a soft threshold power of 24. Groups of co-regulated genes (modules) correlated with each other with the Pearson correlation coefficient 0.42 or better were merged. The eigengenes of the resulting modules (the first principal component of the expression matrix corresponding to the genes included in the module) were correlated with disease states (H, AL, or D).

Assessing the Robustness of the Analysis

Low quality of the SOLiD sequencing data resulted in low number of reads mapped, raising concerns about the reliability of the data. In tag-based RNA-seq, unlike standard

RNA-seq, every count represents an observation of an independent transcript. Thus low counts could still provide sufficient quantitative information about transcript abundances. In addition, high level of biological replication (n=8 per group) in our experiment should have compensated somewhat for the low counts within each replicate. To confirm that low counts did not result in inflated false discovery rate, we have simulated a series of count datasets based on the empirical per-gene total counts and coefficients of variation across samples as well as empirical sample size factors, which included no effect of experimental conditions (Additional File 7). Analysis of these simulated datasets recovered nearly identical sample size factors and highly similar dispersion estimates as in real data (Additional File 8). When these simulated datasets were analyzed with DESeq2 using the same models as real data, at most four genes passed the 10% Benjamini-Hochberg false discovery rate (FDR) cutoff for each contrast. Compared to 646 genes passing the FDR 10% cutoff for the disease-healthy comparison and 333 genes passing the same cutoff for the disease-AL comparison in the real dataset this is much less than 10%, indicating that the real data analysis was conservative. The simulation-based p-value cutoff achieving the empirical 10% FDR (Additional File 9A-C) would have yielded 1.05-1.5X more DEGs for these comparisons than the Benjamini-Hochberg procedure. Notably, the Benjamini-Hochberg correction did not yield any DEGs for the healthy-AL comparison, and accordingly, the DEG discovery rate in this comparison was even slightly lower than the simulation-based false discovery rate (Additional File 9C), indicating that for this comparison DESeq2 analysis did not provide sufficient power. To keep the analysis conservative, we chose to report DESeq2-based DEGs discovered using the Benjamini-Hochberg procedure.

WGCNA analysis provides one additional confirmation that the observed gene expression differences are driven by biological factors rather than stochasticity. WGCNA constructs gene co-expression modules from their correlation pattern across samples without using the information about how the samples are distributed among experimental conditions. The fact that *post hoc* the module eigengenes correlate strongly with coral condition (Additional File 5) indicates that the gene expression patterns in the data truly reflect the biological processes related to disease. To verify this, we shuffled the condition designations among samples and indeed observed that the correlations with co-expression modules disappeared (Additional File 10).

Principal Coordinate Analysis

Principal coordinate analysis to visualize clustering of gene expression between disease states was performed using the "adegenet" package [82] using variance stabilized data for all genes and subsequently with only candidate differentially expressed genes (unadjusted p-value < 0.05), based on Manhattan distances which correspond to the sum of absolute log-fold changes across all genes. Effects of the three disease states ("D", "H", and "AL") were calculated using the multivariate analysis of variance function "adonis" of the R package "vegan" [83].

Functional Summaries

Gene ontology enrichment analysis was performed using adaptive clustering of GO categories and Mann-Whitney U tests [84] based on ranking of signed log p-values (GO-

MWU, https://sourceforge.net/projects/go-mwu/). Gene expression heatmaps with hierarchical clustering of expression profiles were created with the pheatmap package in R [85].

Authors Contributions

GVA prepared the libraries for sequencing, EM assembled the transcriptome, RMW conducted the tag-based RNA-seq data analysis. MVM supervised the study, annotated the transcriptome, and performed simulations. MVM and RMW wrote the manuscript.

Acknowledgements

We thank C Kenkel, D Stump, and C Palmer for sample collections and photographs. Funding for this study was provided by National Science Foundation grant DEB-1054766 to MVM.

References

1. Hoegh-Gulberg O: **Climate change, coral bleaching and the future of the world's coral reefs**. *Mar Freshw Resour* 1999:839–866.

2. Porter JW, Dustan P, Jaap WC, Patterson KL, Kosmynin V, Meier OW, Patterson ME, Parsons M: **Patterns of spread of coral disease in the Florida Keys**. *Hydrobiologia* 2001:1–24.

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

4. Kushmaro A, Banin E, Loya Y, Stackebrandt E, Rosenberg E: **Vibrio shiloi sp. nov., the causative agent of bleaching of the coral Oculina patagonica.** *Int J Syst Evol Microbiol* 2001, **51**:1383–1388.

5. Ben-Haim Y: **Vibrio coralliilyticus sp. nov., a temperature-dependent pathogen of the coral Pocillopora damicornis**. *Int J Syst Evol Microbiol* 2003, **53**:309–315.

6. Cervino JM, Thompson FL, Gomez-Gil B, Lorence E a, Goreau TJ, Hayes RL, Winiarski-Cervino KB, Smith GW, Hughen K, Bartels E: **The Vibrio core group induces yellow band disease in Caribbean and Indo-Pacific reef-building corals.** *J Appl Microbiol* 2008, **105**:1658–1671.

7. Vezzulli L, Previati M, Pruzzo C, Marchese A, Bourne DG, Cerrano C: **Vibrio infections triggering mass mortality events in a warming Mediterranean Sea.** *Environ Microbiol* 2010, **12**:2007–2019.

8. Ushijima B, Videau P, Burger A, Shore-Maggio A, Runyon CM, Sudek M, Aeby GS, Callahan SM: **Vibrio coralliilyticus strain OCN008 is an etiological agent of acute Montipora white syndrome.** *Appl Environ Microbiol* 2014, **80**:2102–2109.

9. Roff G, Kvennefors ECE, Fine M, Ortiz J, Davy JE, Hoegh-Guldberg O: **The ecology of "Acroporid white syndrome", a coral disease from the southern Great Barrier Reef.** *PLoS One* 2011, **6**:e26829.

10. Ainsworth TD, Kvennefors EC, Blackall LL, Fine M, Hoegh-Guldberg O: **Disease and cell death in white syndrome of Acroporid corals on the Great Barrier Reef**. *Mar Biol* 2007, **151**:19–29.

11. Brown B, Bythell J: **Perspectives on mucus secretion in reef corals**. *Mar Ecol Prog Ser* 2005, **296**:291–309.

12. Miller DJ, Hemmrich G, Ball EE, Hayward DC, Khalturin K, Funayama N, Agata K, Bosch TCG: **The innate immune repertoire in cnidaria--ancestral complexity and stochastic gene loss.** *Genome Biol* 2007, **8**:R59.

13. Vidal-Dupiol J, Ladrière O, Destoumieux-Garzón D, Sautière P-E, Meistertzheim A-L, Tambutté E, Tambutté S, Duval D, Fouré L, Adjeroud M, Mitta G: **Innate immune responses of a scleractinian coral to vibriosis.** *J Biol Chem* 2011, **286**:22688–22698.

14. Mydlarz LD, Harvell CD: **Peroxidase activity and inducibility in the sea fan coral exposed to a fungal pathogen.** *Comp Biochem Physiol Part A* 2007, **146**:54–62.

15. Mydlarz LD, Holthouse SF, Peters EC, Harvell CD: **Cellular responses in sea fan corals: granular amoebocytes react to pathogen and climate stressors.** *PLoS One* 2008, **3**:e1811.

16. Olano CT, Bigger CH: **Phagocytic activities of the gorgonian coral Swiftia exserta.** *J Invertebr Pathol* 2000, **76**:176–184.

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

18. Puill-Stephan E, Seneca FO, Miller DJ, van Oppen MJH, Willis BL: **Expression of putative immune response genes during early ontogeny in the coral Acropora millepora.** *PLoS One* 2012, **7**:e39099.

19. Weiss Y, Forêt S, Hayward DC, Ainsworth T, King R, Ball EE, Miller DJ: **The acute transcriptional response of the coral Acropora millepora to immune challenge: expression of GiMAP/IAN genes links the innate immune responses of corals with those of mammals and plants.** *BMC Genomics* 2013, **14**:400.

20. Libro S, Kaluziak ST, Vollmer S V: **RNA-seq profiles of immune related genes in the staghorn coral Acropora cervicornis infected with white band disease.** *PLoS One* 2013, **8**:e81821.

21. Gladfelter EH: **Circulation of Fluids in the Gastrovascular System of the Reef Coral Acropora cervicornis**. *Biol Bull* 1983, **165**:619–636.

22. Pearse VB, Muscatine L: **Role of Symbiotic Algae (Zooxanthellae) in Coral Calcification**. *Biol Bull* 1971, **141**:350–363.

23. Taylor DL: **Intra-colonial transport of organic compounds and calcium in some Atlantic reef corals**. In *Proceedings, 3rd Int Coral Reef Symp*; 1977(May):431–436.

24. Fang LS, Chen YWJ, Chen CS: **Why does the white tip of stony coral grow so fast without zooxanthellae?**. *Mar Biol* 1989, **103**:359–363.

25. Oren U, Rinkevich B, Loya Y: **Oriented intra-colonial transport of 14C labeled materials during coral regeneration**. *Mar Ecol Prog Ser* 1997, **161**:117–122.

26. Roff G, Hoegh-Guldberg O, Fine M: **Intra-colonial response to Acroporid "white syndrome" lesions in tabular Acropora spp. (Scleractinia)**. *Coral Reefs* 2006, **25**:255–264.

27. Roth MS, Deheyn DD: **Effects of cold stress and heat stress on coral fluorescence in reef-building corals.** *Sci Rep* 2013, **3**:1421.

28. Smith-Keune C, Dove S: **Gene expression of a green fluorescent protein homolog as a host-specific biomarker of heat stress within a reef-building coral.** *Mar Biotechnol* 2007, **10**:166–180.

29. Rodriguez-Lanetty M, Harii S, Hoegh-Guldberg O: **Early molecular responses of coral larvae to hyperthermal stress.** *Mol Ecol* 2009, **18**:5101–5114.

30. DeSalvo MK, Voolstra CR, Sunagawa S, Schwarz J a, Stillman JH, Coffroth M a, Szmant a M, Medina M: **Differential gene expression during thermal stress and bleaching in the Caribbean coral Montastraea faveolata.** *Mol Ecol* 2008, **17**:3952–3971.

31. Bay LK, Ulstrup KE, Nielsen HB, Jarmer H, Goffard N, Willis BL, Miller DJ, Van Oppen MJH: **Microarray analysis reveals transcriptional plasticity in the reef building coral Acropora millepora.** *Mol Ecol* 2009, **18**:3062–3075.

32. Fujita T: **Evolution of the lectin-complement pathway and its role in innate immunity.** *Nat Rev Immunol* 2002, **2**:346–353.

33. Kvennefors ECE, Leggat W, Hoegh-Guldberg O, Degnan BM, Barnes AC: **An ancient and variable mannose-binding lectin from the coral Acropora millepora binds both pathogens and symbionts.** *Dev Comp Immunol* 2008, **32**:1582–1592.

34. Kvennefors ECE, Leggat W, Kerr CC, Ainsworth TD, Hoegh-Guldberg O, Barnes AC: **Analysis of evolutionarily conserved innate immune components in coral links immunity and symbiosis.** *Dev Comp Immunol* 2010, **34**:1219–1229.

35. Brown T, Bourne D, Rodriguez-Lanetty M: **Transcriptional Activation of c3 and hsp70 as Part of the Immune Response of Acropora millepora to Bacterial Challenges**. *PLoS One* 2013, **8**:e67246.

36. Matsushita M, Fujita T: **Activation of the Classical Complement Pathway by Marmose-binding Protein in Association with a Novel Cls-like Serine Protease**. *J Exp Med* 1992, **176**:1497–1502.

37. Kornberg HL, Krebs HA: **Synthesis of Cell Constituents from C2-Units by a Modified Tricarboxylic Acid Cycle**. *Nature* 1957, **179**:988–991.

38. Kondrashov FA, Koonin E V, Morgunov IG, Finogenova T V, Kondrashova MN: **Evolution of glyoxylate cycle enzymes in Metazoa: evidence of multiple horizontal transfer events and pseudogene formation.** *Biol Direct* 2006, **1**:31.

39. Polato NR, Altman NS, Baums IB: **Variation in the transcriptional response of threatened coral larvae to elevated temperatures.** *Mol Ecol* 2013, **22**:1366–1382.

40. Wanner L, Keller F, Matile P: **Metabolism of radiolabelled galactolipids in senescent barley leaves**. *Plant Sci* 1991, **78**:199–206.

41. DeSalvo M, Sunagawa S, Voolstra C, Medina M: **Transcriptomic responses to heat stress and bleaching in the elkhorn coral Acropora palmata**. *Mar Ecol Prog Ser* 2010, **402**:97–113.

42. Lõhelaid H, Teder T, Tõldsepp K, Ekins M, Samel N: **Up-Regulated Expression of AOS-LOXa and Increased Eicosanoid Synthesis in Response to Coral Wounding**. *PLoS One* 2014, **9**:e89215.

43. Rotjan R, Lewis S: **Impact of coral predators on tropical reefs**. *Mar Ecol Prog Ser* 2008, **367**:73–91.

44. Fang L-S, Shen P: **A living mechanical file: the burrowing mechanism of the coral-boring bivalve Lithophaga nigra**. *Mar Biol* 1988, **97**:349–354.

45. Bythell JC, Bythell M, Gladfelter EH: **Initial results of a long-term coral reef monitoring program : Impact of Hurricane Hugo at Buck Island Reef National Monument, St. Croix, U.S. Virgin Islands**. *J Exp Mar Bio Ecol* 1993, **172**:171–183.

46. Bythell JC, Hillis-Starr ZM, Rogers CS: **Local variability but landscape stability in coral reef communities following repeated hurricane impacts**. *Mar Ecol Prog Ser* 2000, **204**:93–100.

47. Palmer C V, Traylor-Knowles NG, Willis BL, Bythell JC: **Corals use similar immune cells and wound-healing processes as those of higher organisms.** *PLoS One* 2011, **6**:e23992.

48. Bosch TCG: **Why polyps regenerate and we don't: towards a cellular and molecular framework for Hydra regeneration.** *Dev Biol* 2007, **303**:421–433.

49. Parks WC, Wilson CL, López-Boado YS: **Matrix metalloproteinases as modulators of inflammation and innate immunity.** *Nat Rev Immunol* 2004, **4**:617–629.

50. Burge C a, Mouchka ME, Harvell CD, Roberts S: **Immune response of the Caribbean sea fan, Gorgonia ventalina, exposed to an Aplanochytrium parasite as revealed by transcriptome sequencing.** *Front Physiol* 2013, **4**:1–9.

51. Augustin R, Siebert S, Bosch TCG: **Identification of a kazal-type serine protease inhibitor with potent anti-staphylococcal activity as part of Hydra's innate immune system.** *Dev Comp Immunol* 2009, **33**:830–837.

52. Armstrong PB, Quigley JP: **a 2 -macroglobulin : an evolutionarily conserved arm of the innate immune system**. *Dev Comp Immunol* 1999, **23**:375–390.

53. Tentori E, Allemand D: **Light-enhanced calcification and dark decalcification in isolates of the soft coral Cladiella sp. during tissue recovery.** *Biol Bull* 2006, **211**:193–202.

54. Clausen CD, Roth AA: **Effect of Temperature and Temperature Adaptation on Calcification Rate in the Hermatypic Coral Pocillopora damicomis**. *Mar Biol* 1975, **33**:93–100.

55. Kleypas J a.: **Geochemical Consequences of Increased Atmospheric Carbon Dioxide on Coral Reefs**. *Science (80-)* 1999, **284**:118–120.

56. Goffredo S, Vergni P, Reggi M, Caroselli E, Sparla F, Levy O, Dubinsky Z, Falini G: **The skeletal organic matrix from Mediterranean coral Balanophyllia europaea influences calcium carbonate precipitation.** *PLoS One* 2011, **6**:e22338.

57. Weiner S, Addadi L: **Acidic macromolecules of mineralized tissues: the controllers of crystal formation**. *Trends Biol Sci* 1991, **16**:252–256.

58. Hwang JS, Takaku Y, Momose T, Adamczyk P, Özbek S, Ikeo K, Khalturin K, Hemmrich G, Bosch TCG, Holstein TW, David CN, Gojobori T: **Nematogalectin, a nematocyst protein with GlyXY and galectin domains, demonstrates nematocyte-specific alternative splicing in Hydra.** *Proc Natl Acad Sci U S A* 2010, **107**:18539–18544.

59. Moya A, Tambutté S, Bertucci A, Tambutté E, Lotto S, Vullo D, Supuran CT, Allemand D, Zoccola D: **Carbonic anhydrase in the scleractinian coral Stylophora pistillata: characterization, localization, and role in biomineralization.** *J Biol Chem* 2008, **283**:25475–25484.

60. Furla P, Galgani I, Durand I, Allemand D: **Sources and mechanisms of inorganic carbon transport for coral calcification and photosynthesis.** *J Exp Biol* 2000, **203**(Pt 22):3445–3457.

61. Allemand D, Ferrier-pagès C, Furla P, Houlbrèque F, Darwin C, Fitz C, London R, Eds JM: **Biomineralisation in reef-building corals : from molecular mechanisms to environmental control**. *Gen Paleontol* 2004, **3**:453–467.

62. Closek CJ, Sunagawa S, DeSalvo MK, Piceno YM, DeSantis TZ, Brodie EL, Weber MX, Voolstra CR, Andersen GL, Medina M: **Coral transcriptome and bacterial**

community profiles reveal distinct Yellow Band Disease states in Orbicella faveolata. *ISME J* 2014:1–12.

63. Wang GL, Semenza GL: **General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia.** *Proc Natl Acad Sci U S A* 1993, **90**:4304–4308.

64. Myllyharju J, Koivunen P: **Hypoxia-inducible factor prolyl 4-hydroxylases: common and specific roles.** *Biol Chem* 2013, **394**:435–448.

65. Epstein ACR, Gleadle JM, Mcneill LA, Hewitson KS, Rourke JO, Mole DR, Mukherji M, Metzen E, Wilson MI, Dhanda A, Tian Y, Masson N, Hamilton DL, Jaakkola P, Barstead R, Hodgkin J, Maxwell PH, Pugh CW, Schofield CJ, Ratcliffe PJ, Drive R, Ox O: **C . elegans EGL-9 and Mammalian Homologs Define a Family of Dioxygenases that Regulate HIF by Prolyl Hydroxylation**. *Cell* 2001, **107**:43–54.

66. D'Angelo G, Duplan E, Boyer N, Vigne P, Frelin C: **Hypoxia up-regulates prolyl hydroxylase activity: a feedback mechanism that limits HIF-1 responses during reoxygenation.** *J Biol Chem* 2003, **278**:38183–38187.

67. Cummins EP, Berra E, Comerford KM, Ginouves A, Fitzgerald KT, Seeballuck F, Godson C, Nielsen JE, Moynagh P, Pouyssegur J, Taylor CT: **Prolyl hydroxylase-1 negatively regulates IkappaB kinase-beta, giving insight into hypoxia-induced NFkappaB activity.** *Proc Natl Acad Sci U S A* 2006, **103**:18154–18159.

68. Work TM, Aeby GS: **Pathology of tissue loss (white syndrome) in Acropora sp. corals from the Central Pacific.** *J Invertebr Pathol* 2011, **107**:127–131.

69. Sweet MJ, Bythell JC, Nugues MM: **Algae as reservoirs for coral pathogens.** *PLoS One* 2013, **8**:e69717.

70. Meyer E, Aglyamova G V, Wang S, Buchanan-Carter J, Abrego D, Colbourne JK, Willis BL, Matz M V: **Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx**. *BMC Genomics* 2009, **10**.

71. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, Kapitonov V V, Jurka J, Genikhovich G, Grigoriev I V, Lucas SM, Steele RE, Finnerty JR, Technau U, Martindale MQ, Rokhsar DS: **Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization**. *Science (80-)* 2007, **317**:86–94.

72. Dunlap WC, Starcevic A, Baranasic D, Diminic J, Zucko J, Gacesa R, van Oppen MJH, Hranueli D, Cullum J, Long PF: **KEGG orthology-based annotation of the predicted proteome of Acropora digitifera: ZoophyteBase - an open access and searchable database of a coral genome**. *BMC Genomics* 2013, **14**.

73. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res* 1997, **25**:3389–3402.

74. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin E V, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov A V, Vasudevan S, Wolf YI, Yin JJ, Natale DA: **The COG database: an updated version includes eukaryotes.** *BMC Bioinformatics* 2003, **4**:41.

75. Meyer E, Aglyamova G V, Matz M V: **Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure.** *Mol Ecol* 2011, **20**:3599–3616.

76. R Core Team: **R: A Language and Environment for Statistical Computing**. *Vienna R Found Stat Comput* 2014, **ISBN 3-900**.

77. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2 Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2**. *bioRxiv* 2014, **doi: 10.11**.

78. Kauffmann A, Gentleman R, Wolfgang Huber: **arrayQualityMetrics--a bioconductor package for quality assessment of microarray data**. *Bioinformatics* 2009, **25**:415–416.

79. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing**. *J R Stat Soc Ser B* 1995, **57**:289–300.

80. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.** *BMC Bioinformatics* 2008, **9**:559.

81. Horvath S: *Weighted Network Analysis. Applications in Genomics and Systems Biology*. Springer; 2011.

82. Jombart T: **adegenet: an R package for the multivariate analysis of genetic markers**. *Bioinformatics* 2008:doi:10.1093/bioinformatics/btr521.

83. Dixon P: **VEGAN, a package of R functions for community ecology.** *J Veg Sci* 2003, **14**:927–930.

84. Voolstra CR, Sunagawa S, Matz M V, Bayer T, Aranda M, Buschiazzo E, Desalvo MK, Lindquist E, Szmant AM, Coffroth MA, Medina M: **Rapid evolution of coral proteins responsible for interaction with the environment.** *PLoS One* 2011, **6**:e20392.

85. Kolde R: **pheatmap: Pretty Heatmaps**. 2013:R package version 0.7.7.