Extensive transcriptional variation poses a challenge to thermal stress biomarker development for endangered corals

Running head: Coral thermal stress biomarker challenges

John Everett Parkinson\textsuperscript{1,2,*}, Erich Bartels\textsuperscript{3}, Meghann K. Devlin-Durante\textsuperscript{1}, Caitlin Lusting\textsuperscript{4}, Ken Nedimyer\textsuperscript{5}, Stephanie Schopmeyer\textsuperscript{6}, Diego Lirman\textsuperscript{6}, Todd C. LaJeunesse\textsuperscript{1}, Iliana B. Baums\textsuperscript{1}

\textsuperscript{1}Department of Biology, Pennsylvania State University, State College, Pennsylvania, USA; \textsuperscript{2}Department of Integrative Biology, Oregon State University, Corvallis, Oregon, USA; \textsuperscript{3}Center for Coral Reef Research, Mote Marine Laboratory, Summerland Key, Florida, USA; \textsuperscript{4}The Nature Conservancy, Florida Keys Office, Summerland Key, Florida, USA; \textsuperscript{5}Coral Restoration Foundation, Tavernier, Florida, USA; \textsuperscript{6}Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, USA

Keywords: conservation genetics, coral reef, gene expression, restoration, symbiosis, Symbiodinium

*Author for Correspondence
e-mail: parkinjo@oregonstate.edu, +1 (845) 258-0803

Article Type: Original Article
Abstract

As climate changes, sea surface temperature anomalies that negatively impact coral reef organisms continue to increase in frequency and intensity. Yet, despite widespread coral mortality, genetic diversity remains high even in those coral species listed as threatened. While this is good news in many ways it presents a challenge for the development of biomarkers that can identify resilient or vulnerable genotypes. Taking advantage of three coral restoration nurseries in Florida that serve as long-term common garden experiments, we exposed over thirty genetically distinct Acropora cervicornis colonies to hot and cold temperature shocks seasonally and measured pooled gene expression responses using RNAseq. Targeting a subset of twenty genes, we designed a high-throughput qPCR array to quantify expression in all individuals separately under each treatment with the goal of identifying thermal stress biomarkers. We observed extensive transcriptional variation in the population, suggesting abundant raw material is available for adaptation via natural selection. However, this high variation made it difficult to correlate gene expression changes with colony performance metrics such as growth, mortality, and bleaching susceptibility. Nevertheless, we identified several promising biomarkers for acute thermal stress that may improve coral restoration and climate change mitigation efforts in the future.
Introduction

Colonies of the branching staghorn coral, *Acropora cervicornis*, once formed dense thickets along shallow reef zones throughout the entire Caribbean region. Due to anthropogenic impacts, disease, and temperature-induced bleaching events, the abundance of this prevalent species has declined by more than 80% in recent decades (Bruckner 2002; Dudgeon et al. 2010). *A. cervicornis* is currently listed as threatened under the U.S. Endangered Species Act. In an effort to conserve and restore declining populations, coordinated coral propagation and reef restoration efforts have been developed throughout the Caribbean (reviewed by Young et al. 2012). Most programs propagate corals within in-water nurseries that serve as common gardens, where genetically diverse colonies are reared in close proximity for extended periods before being planted back onto degraded reefs. These nurseries function as active restoration tools as well as genetic repositories that protect diversity during stress events (Schopmeyer et al. 2012).

A guiding principle of coral conservation is to enhance the genetic diversity of dwindling populations and thus improve resilience by preserving varied stress responses among individuals (Baums 2008). This principle was followed when designing the Caribbean restoration programs by selecting donor colonies that were distinct at neutral microsatellite markers (Baums et al. 2005), translating into high genomic variation within and between sites (Drury et al. 2016; Drury et al. 2017b). Additionally, functional variation in the nurseries is high (Lirman et al. 2014; Lohr & Patterson 2017), meaning many *A. cervicornis* genotypes perform differently in the same environment or under the same stress (Lirman et al. 2011a).
Climate change is expected to increase the magnitude and frequency of hot and cold sea surface temperature anomalies in their respective seasons (Easterling et al. 2000), especially in high-latitude marginal regions such as the Florida Keys, USA (Boesch et al. 2000). Both hot and cold events are known to drive coral bleaching (symbiont loss) and mortality (Jokiel & Coles 1990; Saxby et al. 2003; LaJeunesse et al. 2007). However, hot and cold events impact different physiological mechanisms (Roth et al. 2013), and tolerance tradeoffs may exist. For example, coral mortality was rampant during the extreme 2010 Florida Keys cold water event, which disproportionately affected inshore colonies previously resilient to summer hot water bleaching (Kemp et al. 2011; Lirman et al. 2011b; Kemp et al. 2016). Such patterns suggest that hot and cold stress require unique physiological and molecular responses.

The health of a coral colony is also tied to the identity and physiological qualities of its dinoflagellate endosymbionts (Sampayo et al. 2008). Mature A. cervicornis colonies are typically found to associate with just one Symbiodinium species (S. ‘fitti’ = ITS2 type A3Caribbean) at depths above ten meters (Thornhill et al. 2006). Moreover, most of the population of symbiont cells within a colony comprises one clonal cell line (strain), similar to symbiont populations observed in colonies of a related host species, A. palmata (Baums et al. 2014; Thornhill et al. 2017). However, low-abundance background symbionts from other Symbiodinium “clades” have been detected in many A. cervicornis colonies (Baums et al. 2010; Silverstein et al. 2012). Symbiodinium “clade” identity can affect host transcription in corals (DeSalvo et al. 2010), but it is unknown whether within-species diversity in the dominant symbiont and/or variation in the abundance of background symbionts have similar effects.
Genetic variation that correlates with environmental conditions or thermal tolerance can be developed as coral biomarkers (Lundgren et al. 2013; Jin et al. 2016). Gene expression biomarkers (GEBs), which capture dynamic stress responses, are promising tools for coral restoration (reviewed by Louis et al. 2016). By characterizing the molecular profiles that correspond to stress-tolerance and stress-sensitivity, it is possible to monitor the health of reef organisms, assess acute anthropogenic impacts (e.g. pollution effects), and identify resilient genotypes for propagation in coral nurseries. However, the initial analyses of GEBs have been restricted in the number of host genes investigated, the number of individuals assessed, and the time frame over which colonies have been sampled, all without detailed knowledge of the resident symbionts. Here, we use high-throughput molecular approaches to expand the scope of A. cervicornis GEB development and quantify gene expression variation in a marginal coral population.
Materials and methods

For full methodological details, see Supporting Information Text S1. All raw data, R code, and additional supplements can be accessed in the Pennsylvania State University’s ScholarSphere database [https://doi.org/10.18113/S1RP4R].

Coral nurseries and colony selection

Three in-water coral propagation nurseries in Florida were targeted for this experiment: the University of Miami nursery in Miami (referred to as the Miami nursery; 25°28'24.24"N, 80°07'42.24"W), the Coral Restoration Foundation nursery in Tavernier (Upper Keys nursery; 24°58'55.84"N, 80°26'10.69"W), and the MOTE Marine Lab / Nature Conservancy nursery in Summerland Key (Lower Keys nursery; 24°33'45.68"N, 81°24'00.54"W) (Fig. 1a). These three locations fall within the biogeographic range of one intermixing A. cervicornis population near the northernmost boundary of the species (Baums et al. 2010; Drury et al. 2016). All colonies had been growing in their respective common garden nurseries for at least one year prior to experimentation, minimizing environmental variation. At the Miami nursery, the annual linear tissue extension rates, annual mortality rates among replicate colonies, and relative bleaching sensitivities for each genotype were monitored throughout the experimental period.

As part of the initial collection when the nurseries were established, donor colonies were genotyped at four host-specific microsatellite loci (Baums et al. 2005) to establish each colony’s genet identity. Subsequently, the dominant Symbiodinium ‘fitti’ strain in each colony was genotyped at 13 symbiont-specific loci (Pinzon et al. 2011; Baums et al. 2014). Host genets with anecdotally variable growth rates, mortality rates,
and bleaching susceptibilities were chosen from each nursery (Miami: n = 10; Upper Keys: n = 11; Lower Keys: n = 10). Each collection included seven unique S. ‘fitti’ strains. No host genets or symbiont strains were shared across nurseries (Table S1).

To examine intra-individual variation, three host genets were subsampled three times each from physically separate colonies (three ramets of the same genet), so a total of 37 colonies were included in the study. To examine the effect of genotypic variation in the dominant symbiont, collections at each nursery included at least three genetically distinct colonies each associating with the same S. ‘fitti’ strain (‘monotypic’ group), as well as three colonies each associating with unique S. ‘fitti’ strains (‘diverse’ group).

**Temperature experiments**

A single branch (~9 cm) from each colony at each nursery was clipped underwater via SCUBA. Corals were sampled within two hours of solar noon and the growing tips were removed to minimize effects of diel cycle and branch position on gene expression (Hemond & Vollmer 2015). The remaining branch was divided into three equal fragments, which were exposed for one hour to one of three treatments: extreme hot (35 °C), extreme cold (10 °C), or ambient (season-dependent: 24-28 °C). Treatments were conducted shipboard immediately after collection. Temperatures were maintained in insulated water buckets using temperature regulators connected to aquarium heaters and/or ice as needed. This set-up was designed to be relatively inexpensive and simple for restoration workers to repeat, and based on an instant-read thermometer maintained target temperatures for the one hour treatment duration. After the temperature exposure, fragments were preserved in RNALater.
The experiment was repeated four times over a 12-month period: June 2011 (Summer 1), September 2011 (Summer 2), February 2012 (Winter 1), and May 2012 (Winter 2; Fig. 1b). The Summer collections were scheduled to capture gene expression before and after the Summer thermal maximum. Due to logistical issues, it was impossible to sample around the Winter thermal minimum in the same way. Instead, the Winter collections began immediately after the Winter thermal minimum and spanned the same time interval as the Summer collections.

Our treatment temperatures were extreme, as bleaching and mortality are typically observed at <16°C and >31°C in this species, and it is unlikely that any coral would experience such dramatic instantaneous temperature changes naturally. An alternative approach would have been to bring the coral fragments to a laboratory to acclimatize them using more ecologically-relevant temperatures. However, the reestablishment of an aquarium-based common garden would negate the power provided by the field-based nursery (namely, the myriad factors—both known and unknown—that influence survival in the wild and cannot be replicated in an aquarium). The immediate short and extreme stress treatment was thus the best approach to accentuate transcriptomic response differences among genotypes.

RNAseq experiment

A total of 444 *Acropora cervicornis* experimental fragments were collected from the Florida coral nurseries (37 colonies x 3 temperature treatments x 4 seasons). First, an RNAseq experiment was carried out on pooled samples to identify important stress response genes that could later be assayed in all samples using high-throughput qPCR.
(HTqPCR). The subset included seven host genets each from two of the three nurseries (Miami and Lower Keys), two of the four seasonal collections (Summer 1 and Winter 1), and all three temperature treatments (ambient, hot, and cold), resulting in 12 libraries.

Total RNA extraction, library preparation, Illumina sequencing, read processing, functional annotation, and differential expression analyses were carried out as by Parkinson et al. (2016) with minor modifications (Text S1), resulting in ~37 million high quality reads per library. Reads were mapped to the *A. cervicornis* transcriptome of Libro et al. (2013) and separated into host and symbiont components bioinformatically using additional coral and *Symbiodinium* genomic resources. The host gene set (n = 22,772) was analyzed separately from the symbiont gene set (n = 21,094). Transcripts with unassigned or ambiguous origin (n = 21,669) were excluded from further analysis. All genes were modeled in the R package EdgeR (Robinson et al. 2010) and assessed for differential expression with respect to temperature treatment (season ignored) and season (temperatures analyzed separately) using nurseries as replicates. Results were visualized through principal component analysis and Venn diagrams.

**HTqPCR array design**

The RNAseq experiment yielded the targets for the HTqPCR array. The array was developed to assess expression separately for each individual coral, in contrast to the pooled design of the RNAseq experiment. The TaqMan OpenArray platform (Thermo Fisher Scientific, Waltham, MA) included 28 targets in duplicate per plate. Based on the RNAseq experiment, it appeared that certain host genes responded in the same direction under both hot and cold stress (allied pattern), others responded in opposite directions.
(opposing pattern), still others responded to only one stress (uncoupled pattern), and some varied across time (seasonal pattern).

We chose 21 functionally relevant host genes from each expression category along with three low-variance host control genes and four clade-specific *Symbiodinium* ribosomal genes for the HTqPCR array. The symbiont genes were used to track abundances of Clades A, B, C, and D, but no further *Symbiodinium* genes were included due to the low incidence of differential expression. Host genes were chosen based on meeting a majority of several selection criteria (Text S1). Ideally they were annotated, differentially expressed, functionally enriched, intron-spanning, universal (similar expression across nurseries), and part of a temperature coexpression module. Complete gene IDs, array order, annotation information, response categories, and RNAseq expression patterns for each target are presented in Fig. 2.

**HTqPCR array experiment**

After choosing genes of interest, all 28 targets were printed in duplicate on 10 TaqMan OpenArray custom plates, accommodating the 444 samples (24,864 unique qPCR reactions). Total RNA from each sample (100 ng) was treated with DNase I to remove gDNA, converted to cDNA with a High-Capacity cDNA Reverse Transcription kit, and pre-amplified using TaqMan Custom PreAmp Pools (all kits from Thermo Fisher Scientific). qPCRs were performed on a QuantStudio 12K Flex Real-Time PCR System at the Pennsylvania State University Genomics Core Facility.

Raw cycle threshold (C_T) values for each qPCR reaction were processed in R with the package MCMC.qpcr (Matz et al. 2013). After accounting for differences in
amplification efficiencies and removing outliers, these values were fit to the “classic”
linear mixed model, which uses Markov Chain Monte Carlo simulations for maximum
likelihood analysis. The model specifically tested the fixed effects of temperature, season,
and their interaction on each gene, normalized to the control genes and incorporating a
random effect of host genotype. One gene target (A17_Gnat3) had very low efficiency,
multiple cases of non-amplification, and highly variable expression values; therefore, it
was dropped from the model and further consideration. A separate time series analysis
was performed using the R package maSigPro (Conesa et al. 2006) on the averaged,
normalized relative expression values generated with MCMC.qpcr.

Visualizing host gene expression variation

To visualize gene expression variation among host genets, DataAssist (Thermo
Fisher Scientific) was used to remove outlier C_T values, calculate mean C_T values based
on the technical replicates for each sample, and normalize to the average expression of
the three endogenous control genes (dC_T). The values were then imported into the R
statistical environment, where expression relative to ambient temperature control samples
(ddC_T) was calculated for each gene separately and plotted on a log_2 scale.

To examine the effect of dominant S. ‘fitti’ genotypic variation on host expression
variation, ‘monotypic’ vs. ‘diverse’ groups were compared. For each gene at each nursery
and season, the variance in stress response values (ddC_T for hot or cold) across the three
‘monotypic’ or three ‘diverse’ colonies were calculated. Genewise differences in mean
variance between groups were assessed via t-test (α = 0.05).
Background symbiont analysis

Special consideration was given to the quantification of background symbiont types. The PreAmp Pools included primers for all targets except the Clade A Symbiodinium ribosomal gene, as this target already yielded very high signal because of the overwhelming numerical dominance of S. ‘fitti’ in all samples. Pre-amplification would have potentially negatively influenced the efficiency of other reactions. Thus, the Clade A values were not directly comparable to those of Clades B, C, and D, and were excluded from the analyses below.

Correlation analysis

To explore correlations among expression profiles (both raw expression levels and fold-change responses relative to ambient), background symbiont abundances (Symbiodinium Clades B, C, and D), and physiological metrics (growth rate, mortality, and bleaching), all relevant data from the Miami nursery were combined in a single matrix. Growth was measured as annual linear extension averaged over all ramets of a genet, mortality was measured as annual percent mortality among ramets of a genet, and bleaching was a qualitative measure of susceptibility during a bleaching event in 2014 (categorized as 1 = no bleaching, 2 = bleached during summer and recovered, and 3 = bleached and died). Pearson correlation coefficients with a Holm multiple comparison correction were calculated pairwise ($\alpha = 0.05$). Calculations were performed in R, and correlation heatmaps were generated in Gene-E (Gould 2015).
Results

Global expression patterns (RNAseq)

The RNAseq experiment on a subset of samples and seasons showed temperature-based differential expression was more apparent in the coral host than the algal symbiont (e.g. 949 host differentially expressed genes (DEGs) vs. 28 symbiont DEGs in the heat shock vs. ambient treatment; Fig. 1c). Few changes in gene expression for both hosts and symbionts were observed between Summer 1 and Winter 1. Given the low proportion of dynamic gene expression in Symbiodinium, we subsequently focused only on coral host genes. There were more host genes responsive to heat stress than cold-stress (e.g. 949 hot DEGs vs. 237 cold DEGs). As expected, a Gene Ontology enrichment analysis revealed many differentially expressed genes were components of stress response processes (Table S2). In general, heat stress mostly resulted in gene upregulation, whereas cold stress mostly resulted in gene downregulation (Fig. 1e).

Samples with similar host expression patterns were grouped visually through principal component analysis (Fig. 1d,f). The first principal component accounted for 54% of expression variation and was largely correlated with temperature ($r^2 = 0.63; p = 0.014$). The second principal component accounted for 25% of variation and was correlated with nursery/location ($r^2 = 0.75; p = 0.002$). Season did not correlate with the ordination ($r^2 = 0.14; p = 0.498$). Many of the stress-response genes appeared to show allied or opposing patterns (Fig. 1e). Of the allied patterns, 2 genes were upregulated in both hot and cold, while 56 genes were both downregulated in hot and cold. Of the opposing patterns, 10 genes were upregulated in hot and downregulated in cold, while no genes were downregulated in hot and upregulated in cold.
Targeted expression patterns (HTqPCR)

The HTqPCR experiment on a subset of 20 host genes showed the average expression of each gene under each treatment generally matched the pattern expected based on the RNAseq results. All allied genes responded to hot and cold stress in the same direction, all opposing genes responded in opposite directions, and all control genes diverged very little from zero (Fig. 2; Table S3a). According to the Bayesian model, most of these trends were statistically significant (or not significant, in the case of controls), and ultimately 19 of the 23 genes analyzed met significance expectations under both hot and cold conditions.

According to the separate time series analysis (Figs. 3, S1), nine of the 20 genes had expression profiles that varied with season under ambient conditions, while those nine plus an additional three genes varied seasonally in terms of their responses to hot and cold (Table S3b), largely in agreement with the results of the Bayesian model. In the time series model, none of the control genes were differentially expressed at any point or under any treatment. Interestingly, neither were the two genes that were expected to be seasonal based on RNAseq data (A20_EXD1 and A21_ANKRD44). Representative expression time series are presented in Fig. 3, while plots for all genes can be found in Fig. S1.

Genotypic expression variation

Expression patterns were highly variable when each host genotype was considered separately. This is evident in the wide range of individual expression values for a given
gene at each time point in Fig. 3. An expanded example is provided in Fig. 4 for target
A12_hsp-16.2. While the hot treatment drove upregulation of hsp-16.2 in all colonies, the
extent of variation ranged from a log₂ fold change (LFC) of ~1 to ~10 (or 2-fold to 1,024-fold). Notably, this range was observed among ramets of the same genet (U11 in Winter 1). If considered instead in terms of dominant symbiont genotype, results were similarly variable, with both high and low expression among colonies sharing S. ‘fitti’ strains. Ultimately no genes featured expression patterns that could be explained easily by host or dominant symbiont identity.

Nor was host expression related to dominant symbiont genotypic diversity. When comparing groups of colonies with identical host diversity but varying intraspecific symbiont diversity, no genes showed significant differences in host heat stress response variances. For the cold stress response, only one gene was significant (A05_Drip), but in the opposite direction than might be expected (the ‘monotypic’ variance was greater than the ‘diverse’ variance). An example is given in Fig. 6a for target A15_wnt4.

Correlations of host expression with physiological data and background symbiont diversity

Host gene relative expression levels ($dC_T$) under ambient, hot, and cold temperatures were generally unrelated to colony growth rates, mortality, or bleaching categories (Fig. 5), with typically low Pearson correlation coefficients ($|\rho| < 0.3$). Nevertheless, some correlations were quite strong and significant ($|\rho| > 0.7$; Fig. 5a), but patterns were inconsistent across seasons. Results were similar for stress responses ($ddC_T$) and for nonparametric Spearman rank correlations (data not shown).
Background *Symbiodinium* abundance did not have an obvious effect on ambient host gene expression (Fig. 6b), although four heat stress genes showed strong positive correlations with Clade D abundance during Summer 1 ($\rho > 0.7$). The relative abundance of different *Symbiodinium* did not predict host colony performance metrics at the Miami nursery consistently (Fig. 6c), with the exception of a negative correlation between bleaching frequency and background symbiont abundance during Winter 2. While background Clade B, C, and D symbionts were detected in most colonies at most time points, the time series analysis revealed no seasonal influence on their relative abundances.
Discussion

Extensive expression variation among and within coral colonies

We observed a high degree of gene expression variation among Acropora cervicornis colonies, similar to other acroporid corals (e.g. Granados-Cifuentes et al. 2013; Parkinson et al. 2015). Host genotype, nursery of origin, stress type, and time point all contributed to the diverse expression patterns (Figs. 1-4). Remarkably, the expression of a given gene in physically separate colonies of the same clone could vary up to 1,024-fold despite shared proximity, environmental history, and genetic makeup. Minor differences in handling, experimental treatment, circadian cycle, or microhabitat may have contributed to these patterns, though we explicitly aimed to control these factors in the experimental design.

Alternatively, epigenetic modifications such as differential methylation of genes among ramets of the same genet could explain large expression variation associated with a single coral genotype (Putnam et al. 2016). In silico searches of the draft A. cervicornis genome (Baums et al., unpubl) recovered multiple homologs for genes that comprise the core methylation machinery. Environmental ‘memory’ within individual colonies, likely driven in part by epigenetic modification, may persist for at least ten years in corals (Brown et al. 2015). Therefore, branches of the same genet originally sourced from different parts of the donor colony may exhibit very different expression profiles reflecting past microhabitats (e.g. shaded vs. unshaded) prior to entering the nursery. Such modifications add an extra layer of complexity when developing biomarkers.
Inconsistent correlations between gene expression and performance metrics

A primary goal of this study was to identify gene expression biomarkers for holobiont performance. Despite finding some of the strongest correlations between gene expression and growth rate, mortality, and bleaching yet reported in corals, none of the candidate biomarkers were consistently predictive of colony performance (Fig. 5). The best candidate genes produced significant correlations with performance only in some of the seasons or with only some of the performance metrics, such that a strong predictor in Summer became a weak predictor in Winter, for example. Likewise, Bay and Palumbi (2017) recently identified two gene coexpression modules (large groups of genes that share similar expression patterns across treatments in an experiment) that significantly correlated with colony survival and/or growth in a reciprocal transplant study. One module showed consistent expression across multiple experiments (Rose et al. 2016), but the other did not, illustrating that high within-species variability in gene expression responses to stress are a general feature in corals.

Development of gene expression markers for colony performance requires that good colony performance indicators are available. Performance measurements, let alone true lifetime fitness estimations, remain challenging in corals (Edmunds 2017). Performance in the nursery was measured on the scale of a year (annual growth rates and mortality) or multiple years (bleaching). However, this study represented a narrow snapshot at four time points, and gene expression proved to be highly variable. These temporal incongruities likely made detection of correlations more challenging. Recently, calcification rate has been identified as a phenotype with strong genetic influence in A. cervicornis (Kuffner et al. 2017). Further development of high throughput quantitative
performance measures indicative of lifetime colony fitness is a high priority for coral research.

Candidate biomarkers

Predicting performance is only one goal of developing gene expression biomarkers; another is to detect stress. Although few genes correlated with colony performance metrics across different seasons, many responded consistently to temperature changes. Four genes were particularly well suited to detecting heat stress in *A. cervicornis* (A11_hsp-16.2, A12_ZFAND2B, A13_RTKN and A14_GADD45A; Figs. 3c, S1). All four are known members of stress pathways, coding for heat shock, zinc finger, rhotekin, and damage-inducible proteins, respectively. They were upregulated under hot conditions by nearly all genotypes regardless of nursery and varied only slightly by season. They were also members of the same gene coexpression module. The first principal component of all genes in the cluster (the eigengene) featured a high and statistically significant correlation with temperature ($\rho = +0.57$; $p = 0.05$) but not season or nursery/location.

Four genes were consistently downregulated in response to both hot and cold stress (A04_ALKBH1, A05_Drip, A18_sesn1, and A19_TIR1; Figs. 2a, S1). These genes encode alkylated DNA repair, aquaporin, sestrin, and toll-like receptor proteins, respectively. Given their similar expression patterns regardless of stress type, these genes may be useful for identifying general stresses such as pollution, disease, and other acute impacts. These were all members of a separate gene coexpression module whose
eigengene correlated with nursery/location ($\rho = -0.66; p = 0.02$) but not temperature or season, suggesting they may be locally acclimatized/adapted (see below).

TNF receptor-associated factor 3 (A07_TRAF3) stood out as a potentially informative marker for seasonal heat acclimatization. This gene inhibits NF-κB activation (Yamamoto et al. 1998), which is a key signaling component of the immune response and stress-induced apoptosis in humans (Gilmore & Wolenski 2012) as well as corals (Davy et al. 2012; Zhou et al. 2017). In this study, the A. cervicornis TRAF3 homologue was typically downregulated during hot and cold exposure, likely activating NF-κB and a subsequent stress response (Fig. 3d). However, TRAF3 was upregulated under heat shock during Summer 2 after the summer thermal maximum, reflecting a possible reduction in stress. This pattern suggests that the corals may have been ‘primed’ by earlier heat exposure, making them more capable of dealing with heat stress in the late summer than any other period during the year.

Local adaptation

By design, our HTqPCR analysis focused on a subset of genes that appeared to have similar expression patterns across nurseries (at least according to the initial RNAseq results). The goal was to identify universal biomarkers for A. cervicornis, and so we tried to filter out genes with nursery-specific patterns. Nevertheless, there was a large nursery signal in the RNAseq data set, as well as among some of the HTqPCR genes (Fig. 1d, Fig. 4). This suggests differential acclimatization among corals in different nurseries. It might also indicate a degree of local adaptation across the latitudinal gradient within Florida.
Local adaptation despite gene flow has been inferred in other Caribbean species (Polato et al. 2010; Kenkel et al. 2013), and A. cervicornis genotype, environment, and their interaction greatly impact colony growth, survivorship, and tolerance in the Florida Keys (Drury et al. 2017a). Reciprocal transplant studies have demonstrated differential acclimatization as well as local adaptation with respect to coral gene expression, indicating transcriptional plasticity itself may be an adaptive trait serving as a genomic basis for resilience to climate change (Barshis et al. 2013; Palumbi et al. 2014; Kenkel & Matz 2016). A. cervicornis is an ideal system to further explore these ideas through transcriptional comparisons between the marginal Florida population and more central Caribbean populations. Additional work will be required to test the feasibility of using locally-adapted genes as expression biomarkers of colony performance.

Opposing vs. allied gene expression responses

One previous study on tropical reef-building corals quantified levels of a thermal stress response protein in the same individuals during hot and cold stress, finding a strong increase in protein abundance under both treatments after six hours (Seveso et al. 2016). Such allied molecular responses are expected to reflect the ability of an organism to respond to multiple stressors using a common core stress response. In contrast, opposing responses suggest the organism requires unique mechanisms to handle different stressors. In the context of threatened corals, a greater ratio of allied to opposing thermal stress response genes in a given colony could indicate greater resilience to future sea surface temperature fluctuations caused by climate change. Only three genes of interest truly featured opposing patterns. Though some allied genes also turned out to be false
positives, the low proportion of opposing genes (n = 10) relative to the high number of allied genes (n = 58) from the RNAseq experiment indicates that it is far more common for coral stress genes to act in the same direction than to be opposed under different thermal shocks (Fig. 1e), and these allied responses tend to be consistent across seasons (e.g. Fig. 3a).

Even though hot and cold stresses generally affect different molecular pathways, certain elements of the response appear to be conserved in *A. cervicornis*. It is also more common for coral stress genes to be downregulated together (n = 56) than to be upregulated together (n = 2) during alternate stress events (Fig. 1e). Therefore, this *A. cervicornis* population does not appear to be constrained to intermediate expression of stress response genes despite exposure to different thermal extremes during summer and winter. Rather, alternate stressors may reinforce the expression patterns of genes shared in both responses. It is possible that any Florida host genotypes showing opposing gene expression patterns with hot and cold stress have died out during the past few decades or that such genotypes do not exist in this species. It thus would be interesting to investigate *A. cervicornis* gene expression patterns in more central locations in the Caribbean with narrower temperature ranges.

*No clear influence of dominant or background symbiont diversity on host gene expression and performance*

Given that selection can act at the level of the coral holobiont (Iglesias-Prieto & Trench 1997; LaJeunesse *et al.* 2010; Parkinson & Baums 2014), we posited that maintaining genetic diversity of both hosts and symbionts—not just the host—should be
an important goal in the conservation of endangered corals. We therefore tested whether
dominant or background symbiont composition influenced host expression or holobiont
phenotypes. However, we found no strong evidence for such an effect.

The dominant S. ‘fitti’ strain was no more likely to predict host expression levels
than host genotype (Fig. 4), and host expression response variance was just as high
among groups of colonies sharing a single S. ‘fitti’ strain as it was among groups of
colonies with multiple strains (Fig. 6a). Moreover, correlations between background
symbiont composition and host expression were mostly weak and variable (Fig. 6b). Of
note, background Clade D abundance, which in the Caribbean is represented
predominately by S. trenchii (= ITS2 type D1a), strongly correlated with the ambient
eexpression levels of four candidate heat shock proteins during Summer 1, when
temperatures in the Florida Keys were rapidly climbing. Additionally, all three
background clade abundances correlated negatively with longer-term bleaching
frequency during Winter 2 (Fig. 6c). At this point it is unclear if these were spurious
associations or meaningful biological interactions, but they warrant further investigation.

It is possible that symbiont composition did not affect colony performance at all,
but it is more likely that host expression is simply a poor metric by which to measure a
symbiont genotype and/or species effect during acute thermal stress. Host gene
expression changes are rapid and dynamic, whereas Symbiodinium appear far less
transcriptionally responsive to temperature shocks (Leggat et al. 2011; Barshis et al.
2014; this study; but see McGinley et. al 2012, Baumgarten et al. 2013, Levin et al.
2016), and a ‘host buffering’ effect may temporarily protect symbiont cells from rapid
changes in the host environment (Parkinson et al. 2015). Given the different time frames
over which coral hosts and their algal symbionts experience and respond to stress at a molecular level, transient shocks appear insufficient to detect genotypic interactions among host-symbiont partners. Longer-term exposures to less extreme temperatures or alternate metrics such as proteomics, metabolomics, or cellular physiological assays should be used to further characterize the phenotypes of these fine-scale interactions and assess the conservation priority of symbiont diversity.

Conclusion

We have improved on previous efforts to identify gene expression biomarkers for corals by developing new high-throughput methods, interrogating a large number of host genes simultaneously, expanding the number of individuals assessed, repeating the experiments at multiple time points throughout a year, and incorporating Symbiodinium diversity. Working with nursery-reared corals reduced the influence of environmental variation and ensured that our results produced useful information tied directly to ongoing restoration projects. Despite finding few consistent correlations between ambient or temperature shock expression variation, holobiont performance, and symbiont diversity, we nevertheless identified several potentially useful thermal stress biomarkers. Although challenging for GEB development, the high levels of standing transcriptional variation observed among individuals in this study suggest restoration nurseries are fulfilling their role as repositories for coral genetic and phenotypic diversity.

Acknowledgements

This research was funded by the NOAA Coral Reef Conservation Grant Program
(NA14NOS4820085 to IBB) and National Science Foundation (NSF DGE-0750756 to JEP; NSF OCE-0928764 to TCL and IBB). Nursery corals were collected under the following permits: FKNMS-2007-041; FKNMS-2009-026; FKNMS-2009-099-A2; US DOI/NPS #BISC-2011-SCI-0019. Special thanks are extended to Deborah Grove, Craig Praul, and Ashley Price of the Penn State Genomics Core Facility for helping implement the RNAseq and HTqPCR methods, to Nick Polato for contributing ideas during project development, to Sebastian Baumgarten for updating the A. cervicornis transcriptome annotation, and to Ford Drury, Sean Griffin, Meaghan Johnson, Jessica Levy, Kayla Ripple, Stephanie Roach, Pedro Rodriguez, and Cory Walter for diving assistance.
References


and their algal symbiont to thermal stress. Plos One 6, e26687.

Levin RA, Beltran VH, Hill R, et al. (2016) Sex, scavengers, and chaperones:
transcriptome secrets of divergent Symbiodinium thermal tolerances. Molecular
Biology and Evolution, msw119.

Libro S, Kaluziak ST, Vollmer SV (2013) RNA-seq profiles of immune related genes in
the Staghorn coral Acropora cervicornis infected with white band disease. Plos
One 8, e81821.

Caribbean staghorn coral Acropora cervicornis: influence of host genotype,
symbiont identity, colony size, and environmental setting. Plos One 9, e107253.

Lirman D, Schopmeyer S, Manzello D, et al. (2011a) Severe 2010 cold-water event
caused unprecedented mortality to corals of the Florida Reef Tract and reversed
previous survivorship patterns. Plos One 6, e23047.

Lirman D, Schopmeyer S, Manzello D, et al. (2011b) Severe 2010 Cold-Water Event
Caused Unprecedented Mortality to Corals of the Florida Reef Tract and

Lohr KE, Patterson JT (2017) Intraspecific variation in phenotype among nursery-reared
staghorn coral Acropora cervicornis (Lamarck, 1816). Journal of Experimental

biomarkers of heat stress in scleractinian corals: Promises and limitations.
Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology.


Data Accessibility

All raw data, R code, and additional supplements associated with this manuscript can be accessed in the Pennsylvania State University’s ScholarSphere database [https://doi.org/10.18113/S1RP4R]. Additional supplements include RNAseq library composition, Illumina run statistics, RNAseq differential expression results, updated A. cervicoris gene transcript annotations, HTqPCR assay design, complete correlation outputs, graphs of individual genotype variation for each gene, and graphs of ‘diverse’ vs. ‘monotypic’ variance for each gene.

Author Contributions

JEP, DL, TCL, and IBB conceived of this project. JEP, EB, CL, KN, and SS performed field experiments. JEP and MKD performed laboratory experiments. JEP analyzed the data and created the figures. JEP wrote the paper. All authors contributed editorially to the final manuscript.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Text S1 Complete methodology with more genomic details.

Table S1 Host and symbiont multilocus genotypes and donor colony GPS coordinates.

Table S2 Gene Ontology (GO) term enrichment for each temperature treatment.

Table S3 Statistical output for (a) the Bayesian model and (b) the time series model.

Figure S1 Graphs of relative expression time series for all HTqPCR genes.
Figures

(c) RNaseq Table

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Host DEGs</th>
<th>Symbiont DEGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot vs. Ambient</td>
<td>949 (327)</td>
<td>028 (004)</td>
</tr>
<tr>
<td>Cold vs. Ambient</td>
<td>237 (113)</td>
<td>001 (001)</td>
</tr>
<tr>
<td>Summer 1 vs. Winter 1 (Hot)</td>
<td>020 (007)</td>
<td>034 (008)</td>
</tr>
<tr>
<td>Summer 1 vs. Winter 1 (Cold)</td>
<td>015 (005)</td>
<td>035 (003)</td>
</tr>
<tr>
<td>Summer 1 vs. Winter 1 (Ambient)</td>
<td>007 (001)</td>
<td>003 (000)</td>
</tr>
</tbody>
</table>

(d) RNaseq Principal Component Analysis (Host Only)

(e) RNaseq Venn Diagram (Host Only)

(f) RNaseq PCA Vector Statistics (Host Only)

<table>
<thead>
<tr>
<th>Environmental Variable</th>
<th>$r^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursery</td>
<td>0.75</td>
<td>0.002</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.63</td>
<td>0.014</td>
</tr>
<tr>
<td>Season (Not Displayed)</td>
<td>0.14</td>
<td>0.498</td>
</tr>
</tbody>
</table>
Fig. 1 RNAseq experimental overview. (a) Map of the Florida Keys depicting the location of the three nurseries used in the study. (b) Monthly sea surface temperature observations (SST) near the Upper Keys nursery from 2011-2012 (data from NOAA National Data Buoy Center; http://www.ndbc.noaa.gov/). Bars indicate sampling points, with red corresponding to Summer and blue corresponding to Winter. (c) The number of differentially expressed genes (DEGs) among coral hosts and algal symbionts for each main contrast in the RNAseq experiment (left value: total number of DEGs; right value in parentheses: number of well-annotated DEGs). (d) Principal component analysis (PCA) of all host DEGs with significant environmental vectors superimposed. The vectors point in the direction of the greatest change in the variable of interest, while the vector length is proportional to the correlation between the ordination and the variable. (e) Venn diagram depicting overlap in the total number of host DEGs (top value) or well-annotated DEGs (bottom value in parentheses) that were up- or down-regulated in the hot or cold treatment relative to ambient conditions. Overlapping regions correspond to allied or antagonistic expression patterns. (f) Statistics for the PCA environmental vectors.
Fig. 2 HTqPCR experimental overview. (a) Table of gene targets for the high-throughput qPCR (HTqPCR) experiment, including annotation information and expected expression patterns based on the RNAseq experiment. (b) Thermal shock responses for
coral host genes included on the HTqPCR array averaged across all genotypes and seasons. Points represent host expression responses (log₂ fold changes; LFC) relative to ambient controls for cold (blue) and hot (red) temperature treatments. Error bars represent 95% credible intervals from the Bayesian linear mixed model. Significant divergence from 0 is evident when the bars do not cross the horizontal reference at y = 0; significant divergence between cold and hot treatments for a given gene is evident when the bars from different treatments do not cross each other.
Fig. 3 Host relative expression levels by season based on the HTqPCR experiment. Temperature treatment is indicated by color: ambient (black), cold (blue), and hot (red). Representative examples are provided for (a) allied downregulation, (b) antagonism, (c)
uncoupled heat stress response, (d) seasonal heat acclimatization, (e) highly variable expression, and (f) low variance control genes. Expression patterns were explored with a time series analysis. Significance for the main effects of time (Season), cold stress (Cold vs. Ambient), and heat stress (Hot vs. Ambient) are indicated (NS = Not Significant; Sig = Significant; $p < 0.05$). Gray shading corresponds to 95% confidence intervals.
Fig. 4 Relative expression variation (heat shock protein 16.2) among all colonies, nurseries, and seasons. Bars represent host expression responses (log₂ fold changes; LFC) relative to ambient controls for cold (blue) and hot (red) temperature treatments. Colonies are grouped by nursery (separated by black vertical lines) and host genotypes along the x-axis. The first letter represents nursery (L = ‘Lower Keys’, U = ‘Upper Keys’, M = ‘Miami’), while numbers identify unique genotypes within each nursery. Dominant symbiont (S. ‘fitti’) strain identities are also provided in brackets using lower case letters to indicate nursery and numbers for unique genotypes within a nursery (independent from host labels). Where expression values were determined for 3 replicate
ramets of the same host genet, the values for each colony are plotted separately (but grouped together between solid gray lines) and identified by the same host genotype name appended with ‘a,’ ‘b,’ and ‘c’ suffixes. Clonal symbiont strains are also identified with shared names labeled with suffixes, as well as shared text color (unique strains are labeled in black). Colonies U03, U10a, M05, M08, and M10c are missing some data due to failed amplification.
Fig. 5 Relationships between host expression and colony performance. Heatmaps depict Pearson correlation coefficients ($\rho$) between coral holobiont phenotypes and expression levels in ambient, cold, and hot temperature treatments at all seasons. Correlations are presented for host relative expression level and (a) annual growth rate, (b) annual mortality, and (c) a bleaching frequency index ranging from 1 (mild) to 3 (severe). Plots of linear fits for the gene target A19_TIR1 in the hot treatment at Winter 2 (highlighted in green) illustrate a range of $\rho$ values and the data that produce them. To aid visualization, relatively weak correlations ($|\rho| < 0.3$) appear white in the heatmaps.
Fig. 6 Influence of symbionts on host expression and colony performance. (a)
Representative boxplots of variance in host expression responses (log$_2$ fold changes; LFC) for target A15_wnt4 when comparing two groups of three colonies with different levels of intraspecific diversity in the dominant symbiont, *Symbiodinium* ‘fitti’ (=ITS2 type A3Caribbean). ‘Diverse’ indicates three *S. ‘fitti’* strains (one per colony); ‘monotypic’ indicates one *S. ‘fitti’* strain (shared in all three colonies). Symbiont strains were unique to each group. Also presented are heatmaps of Pearson correlation coefficients ($\rho$) between abundances of background symbionts (Clades B, C, and D) and (b) host relative expression levels or (c) holobiont phenotypes at all seasons. Highlighted in green are four heat stress genes featuring high correlations with the abundance of Clade D (most likely *S. trenchii* = ITS2 type D1a) during Summer 1. To aid visualization, relatively weak correlations ($|\rho| < 0.3$) appear white in the heatmaps.