Lineage specific transcriptional profiles of *Symbiodinium spp.* unaltered by heat stress in a coral host

Barshis DJ\textsuperscript{1*}, Ladner JT\textsuperscript{1}, Oliver TA\textsuperscript{1}, and Stephen R. Palumbi\textsuperscript{1}

\textsuperscript{1}Department of Biology, Stanford University, Hopkins Marine Station, Pacific Grove, CA 93950, USA
Abstract

Dinoflagellates of the genus *Symbiodinium* form an endosymbiosis with reef building corals, in which photosynthetically-derived nutrients comprise the majority of the coral energy budget. An extraordinary amount of functional and genetic diversity is contained within the coral-associated *Symbiodinium*, with some phylotypes (i.e. genotypic groupings), conferring enhanced stress tolerance to host corals. Recent advances in DNA sequencing technologies have enabled transcriptome-wide profiling of the stress response of the cnidarian coral host, however, a comprehensive understanding of the molecular response to stress of coral-associated *Symbiodinium*, as well as differences amongst physiologically susceptible and tolerant types remains largely unexplored. Here, we examine the transcriptome-wide response to heat stress via RNA-Seq of two types of *Symbiodinium*, the putatively thermotolerant type D2 and the more susceptible type C3K, resident within the same coral host species, *Acropora hyacinthus*. Contrary to previous findings with coral hosts, we find no detectable change in gene expression across the dinoflagellate transcriptome after 3 days of elevated thermal exposure, despite physical evidence of symbiosis breakdown. However, hundreds of genes identified as orthologs between the C and D types (up to 35%) exhibited significant expression differences within treatments (i.e. attributable solely to type, not heat exposure). These include many genes related to known thermotolerance mechanisms including heat shock proteins and chloroplastic membrane components. Additionally, both the between-treatment similarities and between-type differences remained pervasive after 12-18 months of common garden acclimation as well as in mixed *Symbiodinium* assemblages within the same coral host colony.
Introduction

The endosymbiotic association between photosynthetic dinoflagellates of the genus *Symbiodinium* and reef-building corals is highly productive, allowing corals to build extensive reef frameworks in largely nutrient poor tropical oceans (Muscatine and Porter 1977; Hallock 1981). In recent decades, this symbiosis has been increasingly threatened by anthropogenic disturbances, leading some to suggest that reefs will be irrevocably damaged within this century (Hughes et al. 2003; Hoegh-Guldberg et al. 2007; Pandolfi et al. 2011). Increasing frequency and severity of widespread coral bleaching events (*i.e.* the disassociation of *Symbiodinium* and their coral hosts) associated with global climate change has resulted in reductions in live coral cover across the globe, exacerbating current declines in coral health caused by local-scale stressors (Carpenter et al. 2008). However, substantial variance in susceptibility to bleaching exists among populations of corals due to regional differences in thermal environments, contrasting histories of disturbance, and diverse genotypic combinations of coral and *Symbiodinium* (Marshall and Baird 2000; Baker et al. 2004; Sotka and Thacker 2005; Oliver and Palumbi 2011).

Within *Symbiodinium*, there are currently nine distinct phylogenetic clades (or subgenera; Pochon et al. 2006; Pochon and Gates 2010), with a number of functional physiological differences amongst different phylotypes. In particular, association with some phylotypes of *Symbiodinium* clade D can provide enhanced thermal tolerance to host corals when dominated by these types (reviewed in Stat and Gates 2011). Increased prevalence of clade D *Symbiodinium* has been found in habitats characterized by unusually high temperatures and in areas recently impacted by natural bleaching events.
(Fabricius et al. 2004; Jones et al. 2008; Oliver and Palumbi 2009, 2011). Clade D hosting corals have also been shown to exhibit reduced levels of bleaching and greater maintenance of photosynthetic efficiency during thermal exposure (Berkelmans and van Oppen 2006; Oliver and Palumbi 2011). While this phenomenon is not universal (Fabricius et al. 2004; Abrego et al. 2008), and clade D is also associated with cold water *Oulastrea* corals (Lien et al. 2007), one potential mechanism for corals to persist in a warming ocean is through association with a more thermally tolerant *Symbiodinium* type (Buddemeier and Fautin 1993; Berkelmans and van Oppen 2006; Baskett et al. 2009). Additionally as substantial physiological differences exist amongst specific phylotypes within clades as well as between clades (e.g. Cantin et al. 2009; Howells et al. 2012), increased tolerance could be gained via association with a different sub-clade type or member of a different clade altogether.

Yet coral hosts are also capable of reacting to increased temperatures through acclimation or adaptation (Middlebrook et al. 2008; Bellantuono et al. 2012a; Bellantuono et al. 2012b; Barshis et al. 2013). At the molecular level, coral hosts mount a widespread gene expression response to heat exposure across a multitude of biochemical pathways (DeSalvo et al. 2008; Meyer et al. 2009; Polato et al. 2010; Bellantuono et al. 2012a; Barshis et al. 2013). In particular, Barshis *et al.* (2013) found transcription of heat responsive genes to be constitutively higher in American Samoan corals that showed greater heat tolerance. Among nearly 260 genes that changed expression during heating experiments, 63 showed elevated expression in heat resistant corals under control conditions (Barshis et al. 2013), suggesting that in the cnidarian host, differential gene
expression, generally, and expression front loading, specifically, could be central to coral heat resistance.

Gene expression work on coral symbionts, however, has lagged behind work on their cnidarian hosts (Leggat et al. 2011b; but see Baumgarten et al. 2013). Leggat et al. (2011a), for example, found coral hosts to exhibit up to five times larger changes in gene expression in response to heat stress when compared to the same genes in their resident *Symbiodinium* across six putative stress response genes. These results suggest that coral hosts and their symbionts may have distinct manners of responding to environmental stress or different roles in a coordinated response. Additionally, most studies of gene expression in *Symbiodinium* have focused on small sets of specific genes (e.g. Rosic et al. 2010; Leggat et al. 2011a; Rosic et al. 2011; Putnam et al. 2013), although one study employing transcriptome-wide approaches has recently been conducted (Baumgarten et al. 2013).

Here, we report findings that *in hospite* (within a coral host) populations of both a putatively thermotolerant and more thermally susceptible *Symbiodinium* phylotype show little to no transcriptional response to a simulated bleaching exposure across tens of thousands of genes. However, our data do show pervasive differences in transcriptional profiles between phylotypes at orthologous gene loci that span a multitude of molecular processes. In addition, we were able to test for expression differences in corals transplanted among different back reef microclimates, and compare expression profiles for different phylotypes hosted within the same corals. In both cases, expression differences among types remained strong but differences within types among treatments were non-existent. Combined, these results suggest that physiological properties of
*Symbiodinium in hospite* may be fairly static across environments but that differences between types and clades are substantial. These results provide broad insight into the response of *Symbiodinium* to thermal stress and the multiple molecular mechanisms responsible for the enhanced thermal tolerance conferred by a clade D *Symbiodinium* type to the host coral *Acropora hyacinthus*.

**Results**

Comparative transcriptomics (RNA-Seq) represents a powerful tool for the elucidation of the molecular underpinnings of phenotypic variability as it allows for in-depth characterization of gene expression profiles across the vast majority of the transcriptome within a single experiment (Nagalakshmi et al. 2008). Here, we subjected *Acropora hyacinthus* (cryptic species E from Ladner and Palumbi 2012) corals from two locations, a highly variable (HV) back reef dominated by clade D (identified by sequence as Internal Transcribed Spacer region 2 [ITS2]-type D2) and a moderately variable (MV) back reef dominated by clade C (ITS2-type C3k; Craig et al. 2001; Barshis et al. 2010; Oliver and Palumbi 2011) to a three-day simulated bleaching exposure (ambient mean = 29.2 °C and elevated mean = 31.9 °C). *Symbiodinium* phylotypes were identified via comparing Illumina reads to previously published ITS1, ITS2, and chloroplast 23s sequences (for full details see Methods and "*Symbiodinium* characterization" in Ladner et al. 2012). As only a single phylotype per clade was identified in these hosts, we will refer to C3k as C and D2 as D throughout the text. Samples were taken after 72hrs from corals in both experimental and control conditions, and messenger RNA was extracted and sequenced using the Illumina Genome Analyzer II (Illumina Corp.) following previous
protocols (De Wit et al. 2012; Ladner et al. 2012; Barshis et al. 2013). Resulting sequences were assembled, *de novo*, and gene expression values were calculated for individual *Symbiodinium* C and D genes (for full details see Methods; De Wit et al. 2012; Ladner et al. 2012; Barshis et al. 2013).

*Symbiodinium community characterization and heat stress analysis*

A total of 20–391 non-duplicate reads (mean=212) mapped to our three clade-specific markers per library, and the estimated *Symbiodinium* type proportions were highly consistent across the three markers (avg. standard deviation across markers = 3.1%; Table S1A). Seven samples had an average estimated proportion of C ≥95% (mean=98.2%). An additional 10 samples had average estimated proportions of D ≥95% (mean=98.9%). The remaining 3 samples had estimated proportions of D of 94%, 23%, and 19% (Table S1A). An average of 1.87 million total reads and 834 thousand de-duplicated reads per sample mapped to our *Symbiodinium* C and D reference assemblies (Table S1B, C) using a mapping quality threshold of 20 and minimum read length threshold of 20bps.

Average fold change (heated to control ratio) across all transcripts (after variance and low expression filtering; see Methods) was 1.27/-1.33 (range 1.00 to 6.25/-1.00 to -5.40) and 1.18/-1.19 (range 1.00 to 3.30/-1.00 to -3.39) for C and D-types and up- and down-regulated contigs, respectively. There were 541 C and 401 D contigs (approximately 2%) that showed significant differential expression (p<0.05) prior to multiple test correction (composed of the tails of the range of fold changes described above), suggesting a possible alteration in transcription in response to the heat stress exposure. However, because of the increased likelihood of type 1 error inherent in these
large-scale genomic data sets, the confidence that this is a sign of actual differential expression is extremely low. There were no significant differences in gene expression between control and heated treatment groupings for either C nor D-type *Symbiodinium* after False Discovery Rate correction across 50,763 assembled contigs (26,986 and 23,777 for C and D respectively, FDR = 5%, minimum q-value of 0.279 for C and 1.00 for D), demonstrating no consistent transcriptional response to the elevated thermal exposure in either the heat-tolerant or susceptible phylotypes.

In contrast to the response to heat stress, a comparison within treatments between expression profiles of C and D types at 4,840 orthologous loci revealed substantial differences in gene expression (Figs 1, 2). There were 463 (35.8%) differentially expressed genes between types out of 1,347 orthologs after variance and low expression filtering (FDR = 5 %). Many of these genes and gene pathways correspond to previously hypothesized physiological differences among the clade types such as increased carbohydrate production of certain C lineages (Table S2; Little et al. 2004; Stat et al. 2008) and potentially greater thermostability of components of the photosystem in heat tolerant types (Table S2; Tchernov et al. 2004). In fact, one of the genes identified as a candidate for functional adaptive sequence divergence in our previous study (Photosystem I reaction center subunit II; 4-taxon ortholog 707 from Ladner et al. 2012) also shows significant increased expression (2.05 fold) in C when compared to D (ortho_4225 from Table S2). There were no significant orthologs with differential expression due to heat treatment, or with a significant interaction between type and heat exposure (FDR=5%).
Functional enrichment analyses showed the set of genes with differential expression due to type to be significantly enriched for gene ontology (GO) categories that highlight the role of the plastid envelope under cellular component (CC), and carbon fixation and the dark reactions of photosynthesis under biological processes (BP) (Table 1). This further supports a potential role of chloroplast and thylakoid membrane thermostability and maintenance of photosynthesis in the increased tolerance of this particular clade D member. In total, there were 4 cellular components, 8 biological processes, and 4 molecular functions that were significantly enriched in the set of genes with differential expression due to type (Table 1). There were many similar categories represented when enrichment was tested in those genes that had significantly higher expression in D (3 BP, 3 CC, and 9 MF categories; Table S3) as well those that had significantly higher expression in C (5 BP, 6 CC, and 4 MF categories; Table S3). As a whole, these gene expression differences suggest that the C and D Symbiodinium types here may be operating in fundamentally different ways, i.e. transcriptionally hardwired for physiological performance differences in a multitude of cellular processes, including previously hypothesized differences in thermostability of chloroplast components and thylakoid membranes.

Differences within the same local environmental conditions

An important consideration is that the observed differences among types may be attributable to the samples originating from two different locations with different microclimates (D from the HV pool and C from the MV pool; see Methods), thus potentially arising from influence of the local environment and not the genetic identity of the Symbiodinium themselves. To investigate this possibility, we conducted a reciprocal
transplantation experiment between the HV and MV pool using three additional C and D colonies (n=6 total in each environment), and examined *Symbiodinium* gene expression profiles after 12 months of common garden acclimation in both pools. Following acclimation, we still found a large number of differentially expressed genes across the 4,840 orthologous loci between C and D. There were 100 (13.8%) differentially expressed genes between types out of 726 orthologs after variance and low expression filtering (FDR = 5 %), 37 of these 100 also showed significant between-type expression differences in the heat stress experiment (Table S4). No genes showed a significant response to transplantation, nor any with a significant interaction between type and transplantation similar to the heat stress results above.

* Differences within the same host genetic background

Another potential factor that could explain the gene expression differences observed here could be the influence of host genetic background on the expression patterns of resident symbionts. In fact, within the genus *Acropora*, certain clade D *Symbiodinium* types have been shown to be associated with both increases and decreases in thermal tolerances of the coral holobiont (coral host and associated symbiont community) compared to certain clade C types depending on the host species (Abrego et al 2009). In our heat stress experiment, two coral individuals hosted a mixture of C vs. D *Symbiodinium* under control conditions (18.9% and 23.4% D and 81.1% and 76.6% C for corals 68c and 45c respectively). In these two individual samples, 113 differentially expressed genes out of 3,067 after variance and low expression filtering were evident between types within the same identical host background (Table S5). While the number of samples is small, the fact that a large number of orthologs still showed highly
divergent expression profiles within the same coral individuals further supports the evidence from the heat stress and transplant experiments that there are fundamental, widespread transcriptional differences between these C and D *Symbiodinium* types.

**Discussion**

The data from the present study show little to no consistent response in *Symbiodinium* gene expression to elevated thermal conditions across thousands of genes for two divergent phylotypes from two different clades. This is in stark contrast to our previous analysis of coral host gene expression from the same experiment, with the same statistical power, where heat stress significantly altered coral gene expression (up to 27-fold) across hundreds of cnidarian genes (Barshis et al. 2013). It is worth noting that by the end of the heat exposure (24hrs after sampling, 96hrs total exposure time), the majority of corals showed clear signs of bleaching and some partial mortality.

Conventional wisdom would thus suggest that the heat treatment should have caused substantial physiological stress in both the coral and *Symbiodinium*, and while a stress response was apparent in the gene expression of the coral host (Barshis et al. 2013), there was no detectable gene expression response of the symbionts. However, there were consistent and widespread differences between the transcriptional profiles of the two different *Symbiodinium* types, many of which persisted after 12 months of common garden acclimation and were evident within the same two host genetic backgrounds. These data thus present a type of “transcriptional conundrum” for *Symbiodinium*: little to no short-term response, yet fundamental differences in steady state transcriptional profiles of divergent genotypes.
Little to no transcriptional change in *Symbiodinium* in response to stress

The lack of a transcriptional response to experimental treatment suggests that *Symbiodinium* cells might respond to changes in the external environment differently than the majority of studied organisms. Changes in external environmental conditions have been shown to elicit strong transcriptional responses in organisms across the tree of life, from yeasts (Gasch et al. 2000), to fish (Podrabsky and Somero 2004), a variety of plants (Mustroph et al. 2010), insects (Teets et al. 2012), molluscs (Lockwood et al. 2010), crustaceans (Teranishi and Stillman 2007), birds (Cheviron et al. 2008), mammals (Mancia et al. 2008), and corals (DeSalvo et al. 2008; Meyer et al. 2011).

In previous studies of *Symbiodinium* and other dinoflagellates, muted changes in gene expression appear to be common, though strong transcriptional responses have been found on occasion in response to certain treatments. For example, *Symbiodinium* gene expression changes across 8 genes thought to be important in the stress response were less than 2-fold different compared to up to 10-fold differences in similar coral host genes over an 8-day heat exposure (Leggat et al. 2011a). *Symbiodinium* in larvae of the host coral *Pocillopora damicornis* exhibited no change in expression of 5 putative stress genes despite 9 days of elevated temperature and pCO2 exposure and significant changes in protein levels of Rubisco (Putnam et al. 2013). Similarly, less than 2.4-fold changes were observed in heat shock protein 90 (Hsp90) transcripts in the dinoflagellate *Prorocentrum minimum* following heat exposure, though significant upregulation (up to 8.9-fold) was observed after exposure to copper (Guo and Ki 2012). In contrast, Rosic et al. (2010) found up to 4 fold differences in *Symbiodinium* cytochrome P450 gene expression in response to both a rapid and gradual heat exposure and, in a later study, up to a 57%
(1.57 fold) increase and 89% (1.89 fold) decrease in Hsp70 and Hsp90 in *Symbiodinium* under heat stress, both in culture and when *Symbiodinium* were *in hospite* within a coral host (Rosic et al. 2011). Additionally, a recent study by Baumgarten et al. (2013) found that the majority of stresses (salinity, cold, and dark stresses) elicited few changes in gene expression, although many genes responded significantly to heat stress (34°C for 12hrs; 351 differentially expressed genes [DEGs]) and heat shock (36°C for 12hrs; 2,465 DEGs; up to 288 fold upregulation and 345 fold downregulation).

In terms of the present data set in comparison to three common stress genes examined in previous studies, across all 128 putative matches to Hsp70, Hsp90, or CytP450 family genes, only 2 were significant for heat stress at an uncorrected p<0.05, while none were significant for heat stress in any comparison at an FDR level of 5% (Table 2). However, 11 of the 15 orthologs that were putative matches to Hsp70, Hsp90, or CytP450 family genes were significantly different in the C vs. D comparison (Table 2). With the exception of one putative Cytp450 in C (foldchange -3.38 ns; c_sym_77768), one putative Hsp90 in D (foldchange -2.38 ns after FDR; d_sym_7208), and one putative Hsp90 ortholog (foldchange 2.35 ns for heat; ortho_4785), maximum absolute fold change for any of the 125 remaining potential Hsp70/Hsp90/CytP450 genes was 1.88 (or 88%). While these shifts in *Symbiodinium* expression are in the range observed by Rosic et al. (2011), the lack of statistical significance of *Symbiodinium* expression changes, concurrent with strongly significant changes in the coral host in the same experiment (Barshis et al. 2013), suggest very weak or non-existent *Symbiodinium* expression responses in comparison to their coral hosts. The findings herein also corroborate those of Putnam et al. (2013) where no change in expression was observed for 5 putative stress
response genes (including Hsp70) in *Symbiodinium* within coral larvae of *Pocillopora damicornis* exposed to 9 days of elevated temperature and pCO2. It is also interesting to note that out of the 2,658 unique loci that responded significantly to increased heat in Baumgarten et al. (2013), only one (Locus_28763) had a putative annotation as a heat shock protein and no CytP450 annotations were evident. Baumgarten et al. (2013) however, did identify substantial transcriptional regulation in response to heat of many minicircle-encoded photosynthesis proteins, which was not observed in the present study.

One potential explanation for the lack of short-term gene expression response that we observed is that the primary mechanisms to regulate the size and composition of the available protein pool in these *Symbiodinium* may act after transcription and/or after translation. Bayer et al. (2012) showed a paucity of transcription factors in the transcriptomes of a clade A and clade B *Symbiodinium* which could signify low capacity for transcriptional regulation in these taxa. For instance, in the bioluminescent dinoflagellate *Gonyaulax polyedra*, Fagan et al. (1999) found that circadian differences in glyceraldehyde-3-phosphate dehydrogenase (GADPH) protein concentration were not related to mRNA levels, which remained constant throughout the daily cycle. Additionally, Baumgarten et al. (2013), found the presence of numerous small RNAs (smRNA) that correlated in abundance with ~3,500 transcript targets, suggesting that smRNA post-transcriptional regulation (*i.e.* gene silencing) could act on a variety of cellular processes in *Symbiodinium*. Mismatches between protein and mRNA levels are a common feature of many different organisms (Lu et al. 2007), however, studies of both prokaryotic and eukaryotic taxa have estimated that a large proportion of the variance in protein abundance (*e.g.* 34 - 83% for yeast) can be explained by mRNA abundance.
(Abreu et al. 2009). Thus, given what we know from other taxa, the widespread stability in mRNA levels seen herein across thousands of Symbiodinium genes during putatively stressful conditions remains perplexing.

Alternatively, the transcriptional stability of Symbiodinium observed during our experimental treatments could signify a corresponding stability in Symbiodinium protein composition. This could result from some type of in hospite buffering or protection of Symbiodinium cells residing within host tissues. For instance, in the symbiotic anemone Anemonia viridis and coral Stylophora pistillata, cultured Symbiodinium cells showed novel superoxide dismutase (SOD) isoforms when they were in culture compared to those found when in hospite within a coral host, the difference putatively attributed to host cnidarian buffering (Richier et al. 2005). The lack of transcriptional response to stress of the Symbiodinium observed in the present study could thus signify some sort of host buffering of the intracellular environment of in hospite symbionts, essentially protecting them from physiological stress. It is worth noting though, that Rosic et al. (2011) found comparable patterns of Hsp70 and Hsp90 expression in Symbiodinium both in symbiosis with a coral and in culture with significant changes during heat stress in both situations, suggesting little effect of residing within host tissues on the expression levels of these genes.

In the absence of significant in hospite buffering, transcriptional and protein stability could instead be an outcome of reaction by the coral host. The coral bleaching process results, in part, in digestion and/or expulsion of a large proportion of symbiont cells (Gates et al. 1992). As we sampled our corals after 72hrs of heat stress, by which time some individuals exhibited visible paling of the tissues, it could be that the reactive
Symbiodinium cells had been removed by then and the one’s still in hospite at that stage were more tolerant of the heat stress and had not yet developed signs of a cellular stress response that would be evident at the gene expression level. Similarly, if Symbiodinium expression in response to stress occurs more rapidly than host expression changes, our later sampling point may have missed the majority of expression changes in the symbionts. However, our previous study found substantial expression differences in the coral host at the same timepoint, and this sampling timepoint was chosen to examine the onset of the bleaching response, as well as differential tolerance between the corals hosting these D versus C type Symbiodinium (Oliver and Palumbi 2011; Barshis et al. 2013).

Deep-rooted expression differences between phylotypes

Regardless of the specific mechanisms behind the static gene expression levels following heat stress and transplantation, the widespread differences between types seen across the ortholog pairs suggests deep-rooted physiological effects of the different evolutionary histories of these two clades. Our previous work identified multiple candidates for adaptive evolutionary divergence between these clades (Ladner et al. 2012), one of which also showed significant increased expression in C when compared to D types herein. Additionally, 12 out of 15 (80%) of the putative Hsp70/Hsp90/CytP450 orthologs also showed significant differences among types while remaining unresponsive to heat treatment (Table 2). Previously hypothesized differences in Symbiodinium physiologies such as differences in thylakoid membrane stabilities, carbohydrate production, and contribution to coral growth also appear to be reflected in the specific genes that show expression differences between these C and D types and the categories
represented in the functional enrichment analyses (Tchernov et al. 2004; Stat et al. 2008; Stat and Gates 2011). The expression results herein, together with our previous sequence-based analysis (Ladner et al. 2012) suggest that the evolutionary divergence of *Symbiodinium* has resulted in a multitude of differences in both the sequences of specific genes and the machinery that controls their expression.

The lack of short-term transcriptional response to heat stress coupled with the pervasive between type differences suggests that *Symbiodinium* types may be constrained in their ability to respond to novel environmental conditions via a change in gene expression. It is worth noting, however, that in the reciprocal transplantation study, a smaller number of genes (13.8% or 100) showed significant expression differences following the common garden acclimation as compared to the heat stress exposure. Additionally, less than half (37 out of 100) of these were also significant in the between-type heat stress comparison. The smaller number of differentially expressed genes as well as the specific significant genes novel to the transplant experiment could signify some long-term acclimatization of gene expression after the 12 months in a common garden. It could also represent a reduced power to detect significant differences, as the transplant sample size (n=3 per treatment/type) was smaller than that of the original heat stress experiment (n=4-6 per treatment/type). While it is tempting to speculate on the ability for *Symbiodinium* transcriptional acclimation, additional research into the temporal dynamics of long term acclimatization potential in *Symbiodinium* gene expression will be needed to determine to what degree *Symbiodinium* may be able to adjust gene expression over longer time-scales in response to shifting environmental conditions.

Conclusions
As a whole, these findings present a number of fundamental hypotheses about the functional genomics of thermal stress in reef-building corals. It is well established that a multitude of cellular processes respond to thermal stress in the host cnidarian (e.g. DeSalvo et al. 2008; Meyer et al. 2011; Barshis et al. 2013), however, the transcriptional response of *Symbiodinium* to thermal stress appears either non-existent (e.g. Leggat et al. 2011a; Putnam et al. 2013; this study) or highly treatment specific (Baumgarten et al. 2013) in most studies of gene expression to date. Whether *Symbiodinium* transcriptional architecture is fundamentally different from the majority of eukaryotic organisms, their gene expression response is tuned to particular stressors, buffered or determined by their host partner, and/or evolutionarily constrained based on phylogenetic history all remain to be determined. Additionally, as the types examined here represent only one phylotype per clade, future research is needed to elucidate the degree to which closely related *Symbiodinium* types (phylotypes of the same clade) differ in steady state transcription compared to more divergent (phylotypes of different clades) pairings. Advances in technology are bringing the study of *Symbiodinium* genomics and transcriptomics out of its nascency (Leggat et al. 2011b; Bayer et al. 2012; Baumgarten et al. 2013) and opening the doors for future research and potential answers to these fundamental questions.

**Methods**

*Heat stress experiment*

Small branchlets (~2 cm³) of *Acropora hyacinthus* (cryptic species E from Ladner and Palumbi 2012) were collected from 16 different coral colonies from two backreef pools on the south side of Ofu Island, American Samoa (14°11’S, 169°36’W). Ten
colonies of *A. hyacinthus* were sampled from a larger, moderately variable pool (MV pool) and six colonies were sampled from a smaller, highly variable pool (HV pool; for temperature profiles see Oliver and Palumbi 2011, pool 400 and 300 respectively). Replicate samples (n=2) of the same colony were randomly placed in one of two experimental tanks per condition. The ambient/control condition ranged from 26.8 – 34.5°C (mean = 29.2°C, n = 2 tanks) while the heat stress condition was elevated by ~2.7°C over ambient conditions (27 – 37.6°C, mean = 31.9°C, n = 2 tanks). Coral health was monitored every 6-12 hours and samples were taken at 12:00 after 72hrs of exposure to the experimental conditions. The 72hr time point was chosen based on our previous study where a 72hr exposure to elevated temperatures induced initial mortality in the MV samples while HV samples remained alive and appeared resilient (figure 5 from Oliver and Palumbi 2011).

*Reciprocal transplantation experiment*

Corals were transplanted from their native locations into three sites each in the HV Pool and the MV Pool on August 30, 2010. Six colonies were transplanted from the HV Pool (AH02, AH04, AH06, AH07, AH09, AH75) and twelve from the MV Pool (AH27, AH55, AH69, AH68, AH65, AH61, AH64, AH40, AH28, AH70, AH11, AH31). At noon on August 22, 2011, we selected and sampled six colonies that had survived transplantation into both the HV and MV pools for analysis of gene expression patterns; three had originally come from the HV pool (AH02, AH07, AH75) and three had originated in the MV pool (AH40, AH68, AH70). One branch from each colony was sampled and placed, within 15 minutes, into RNALater (Life Technologies). All samples were stored at -80°C until subsequent analysis.
**Symbiodinium phylotype identification**

We counted the abundance of clade specific reads of *Symbiodinium* for each sample (i.e., mRNA library) at three loci that are divergent between clades: the internal transcribed spacer regions 1 and 2 (ITS1, ITS2) and chloroplast 23S rRNA (cp23S; Table S2 from Ladner et al. 2012). Initial analyses, using sequences representing the all clades of *Symbiodinium*, indicated that our sampled reads matched only two phylotypes, ITS-types C3k (100% match) and D2 (1 bp different; data not shown). This is consistent with previous work on the same coral species and locations using direct sequencing of ITS1 and cp23S (Oliver and Palumbi 2011).

Each read file was mapped to these six clade-specific sequences (two clades x three markers) using *BWA* (Li and Durbin 2009); for paired-end lanes, only the forward end was used to avoid bias among single-end and paired-end samples. Default settings were used except that we allowed for ~10% mismatches between individual reads and the reference (−n 0.005). These settings allow for sequence variability/sequencing error within populations while still preventing reads from aligning to the incorrect clade. Clade proportions at each locus were calculated based on the number of well-mapped, non-duplicate reads (≥25 bp, mapping quality ≥30) to each of the clade-specific sequences (for full description see Ladner et al. 2012).

Each of these rRNA sequences is known to be multi-copy thus copy number differences between our *Symbiodinium* types, could impact the number of rRNA sequences in our samples. However, we independently estimated proportions of the two clades from loci on two different genomes (nuclear and chloroplast), and in order to further validate these proportions, we also re-estimated the proportions of each clade by
mapping all sequencing reads to the full transcriptomes that resulted from the de novo assembly and found no major differences in estimated proportion (Table S1).

**mRNA sequencing and analysis**

For the heat stress samples, total RNA was extracted as described previously (Ladner et al. 2012; Barshis et al. 2013). Briefly, approximately 150-200mg of coral tissue and skeleton was placed in 1 ml of TRIzol (GibcoBRL/Invitrogen, Carlsbad, CA, USA) and homogenized for 2 min by vortexing with ~100 µl of 0.5 mm Zirconia/Silica Beads (BioSpec Products, Inc., Bartlesville, OK, USA). Resulting tissue/TRIzol slurry was removed by centrifugation and the standard TRIzol extraction was performed according to manufacturer’s specifications. Between 40 ng and 1 µg of mRNA was used in Illumina library construction as in Beck et al. (2010) but random hexamer primers were used to increase transcriptome coverage. Sample processing for the reciprocal transplantation samples differed from the heat stress experiment samples in the following ways: total RNA was extracted using the Ambion RNAqueous 4-PCR kit, cDNA libraries were constructed using the Illumina TruSeq mRNA Sample preparation kit, and all 12 libraries (two from each colony) were pooled for one lane of single end 36 bp sequencing on an Illumina HiSeq platform at the Center for Genomics and Personalized Medicine at Stanford University.

**Reference assembly and ortholog identification**

Assemblies for both *Symbiodinium* clades C and D and the orthologous gene pairs are those from Ladner et al. (2012). Only the C and D orthologs (n = 4840) identified by Ladner et al. (2012) were used in the current analyses. Briefly, separate de novo assemblies were conducted for the two clades of *Symbiodinium*, using CLC Genomics
Workbench (v. 4, CLC Bio) and including only the sequence libraries with \( \geq 95\% \) of a single \textit{Symbiodinium} clade (averaged across the three clade-specific loci, see above). A small amount of non-focal clade contamination may have occurred, however contigs were constructed using the consensus sequence at each base position to minimize this effect. We identified putative \textit{Symbiodinium} contigs for each assembly based on sequence similarity to a variety of dinoflagellate references and non-\textit{Symbiodinium} contamination was removed via a similar procedure using cnidarian and microbial references (for full description see Ladner et al. 2012).

Open reading frames (ORFs) and protein sequences were determined for each contig using \textit{OrfPredictor} (Min et al. 2005). To minimize erroneous protein predictions, we utilized sequences only if they showed significant similarity to a known protein sequence (BLASTX to NCBI’s nr, e-value \( \leq 1 \times 10^{-5} \)) or the predicted protein sequence \( \geq 200 \) amino acids long. \textit{InParanoid} v4.1 (Remm et al. 2001) was then used to identify orthologous sequences between clades based on protein-level similarity (for full description see Ladner et al. 2012).

\textit{Sequence mapping and differential expression analysis}

Data was mapped using \textit{BWA} (aln -n 0.005 -k 5 -l; Li and Durbin 2009) to a combined reference assembly consisting of: (a) 33,496 \textit{Acropora hyacinthus} contigs from Barshis et al. (2013), (b) 26,986 \textit{Symbiodinium} clade C contigs from Ladner et al. (2012), and (c) 23,777 \textit{Symbiodinium} clade D contigs from Ladner et al. (2012). This combined assembly was used as a mapping reference to avoid taxonomical mis-assignment of individual reads \( \text{(i.e. since multiply mapping reads were excluded from expression analyses, any reads that could not be definitively assigned to either coral or a specific} \)
Symbiodinium clade were not counted). Duplicate reads were identified using Picard v1.43 (MarkDuplicates.jar; http://picard.sourceforge.net/), and read counts for each contig were compiled from .SAM files for uniquely, well-mapped, non-duplicate reads (≥ 20 bp, mapping quality ≥ 20) using a custom python script. This study is only focused on those reads that mapped exclusively to the two Symbiodinium assemblies. The Acropora hyacinthus-specific data can be found in Barshis et al. (2012).

Data were normalized for variation in sequencing depth using the RPKM method (2013) and imported into DESeq (Mortazavi et al. 2008) for differential expression analysis. Multiple normalization strategies were implemented (including the strategy built into the DESeq program) but no major differences in results were observed, hence only the RPKM data are presented herein. We implemented the RPKM normalization as it is the one of the only strategies that could account for length bias in mapping efficiency within the ortholog analyses, and non-ortholog analyses were treated similarly for purposes of clarity. Differentially expressed genes were identified based on the negative binomial distribution and tests were conducted using the nbinomTest function for the within-type heat stress comparisons, and using the 2-way GLM nbinomGLMTest function to examine the influence of type, treatment, and the type~treatment interaction (all as fixed factors) for the ortholog analyses within the DESeq package (Anders and Huber 2010). For the between type ortholog comparisons, only the reads matching the dominant type were used in the analyses (i.e. if an experimental sample was estimated to have ≥ 90% C reads, only those reads matching the C-type assembly were counted and analyzed in the cross-type comparison). The background type counts (i.e. D counts from a C dominant coral) were only used for the two “mixed” type samples (corals 68c and
45c) examined in the with-host analysis. Low expression (within-group average normalized expression < 5) contigs were excluded from analyses to avoid potential artifact caused by assembly and/or sequencing errors, and high inter-individual variability contigs (within-group mean expression < 1 SD) were also excluded so that statistical comparisons would not be overly influenced by outlier individuals. The false discovery rate (FDR) was controlled at 5 % according to the method of Benjamini and Hochberg (Anders and Huber 2010). The expression heat map (Figure 1) was generated using the heatmap.2 function of the R package gplots. Hierarchical clustering was performed using the hclust function in R (1995; p.adjust in R) and plotted using the as.phlyo option in the R package ape (Suzuki and Shimodaira 2006). Statistically over-represented gene ontology (GO) categories were determined using default statistical tests and multiple-testing adjustments in GOEAST (v3.0-8; Paradis et al. 2004).

**Acknowledgments**

We thank P. Craig, T. Clark, C. Caruso, and the rest of the staff at the National Park of American Samoa for access to field sites and logistical help; C.S. McKeon, J. Pringle, J. DeNofrio for valuable field assistance; G. Ruiz-Jones and R. Bay for field work and preliminary data analysis; and two reviewers and the editor for providing comments that improved the manuscript. This work was supported by Conservation International, BioX Stanford, the Schmidt Ocean Institute, and the Gordon and Betty Moore Foundation. Transcriptome sequences and short read files can be found under NCBI BioProject PRJNA177515 for the coral assembly and short read files, PRJNA175601
for the C *Symbiodinium* assembly and PRJNA175600 for the D *Symbiodinium* assembly. Updated annotations for the *Symbiodinium* assemblies can be found at palumbi.stanford.edu/data and/or ww2.odu.edu/~dbarshis/lab/Links-Data-Scripts.html.

**Figures**

**Figure 1.** Heat map of the 483 differentially expressed orthologs between type D and type C. Heat map generated using the heatmap.2 function of the R package gplots. Color values represent the expression of that contig in each sample divided by the average expression for that contig across all samples. Sample naming scheme is as follows. Colony#_treatment_dominanttype (i.e. 01_cont_D is colony #1, the control treatment, and D dominant). Mix denotes the C reads from colonies that had < 90% of reads assigning to a single type.

**Figure 2.** Hierarchical clustering of gene expression profiles across the 4800 type C and type D orthologs. Clustering was primarily driven by type, not treatment or colony. Sample names are the same as denoted in Figure 1.
<table>
<thead>
<tr>
<th>GOID</th>
<th>Ontology</th>
<th>Term</th>
<th>Number of Contigs</th>
<th>FDR-P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0042542</td>
<td>biological_process</td>
<td>response to hydrogen peroxide</td>
<td>5</td>
<td>0.035</td>
</tr>
<tr>
<td>GO:0030435</td>
<td>biological_process</td>
<td>sporulation resulting in formation of a cellular spore</td>
<td>5</td>
<td>0.081</td>
</tr>
<tr>
<td>GO:0015977</td>
<td>biological_process</td>
<td>carbon fixation</td>
<td>8</td>
<td>0.007</td>
</tr>
<tr>
<td>GO:0019685</td>
<td>biological_process</td>
<td>photosynthesis, dark reaction</td>
<td>8</td>
<td>0.007</td>
</tr>
<tr>
<td>GO:0051253</td>
<td>biological_process</td>
<td>negative regulation of RNA metabolic process</td>
<td>5</td>
<td>0.035</td>
</tr>
<tr>
<td>GO:2000113</td>
<td>biological_process</td>
<td>negative regulation of cellular macromolecule biosynthetic process</td>
<td>7</td>
<td>0.029</td>
</tr>
<tr>
<td>GO:0043484</td>
<td>biological_process</td>
<td>regulation of RNA splicing</td>
<td>5</td>
<td>0.059</td>
</tr>
<tr>
<td>GO:0050684</td>
<td>biological_process</td>
<td>regulation of mRNA processing</td>
<td>4</td>
<td>0.081</td>
</tr>
<tr>
<td>GO:0009526</td>
<td>cellular_component</td>
<td>plastid envelope</td>
<td>10</td>
<td>0.074</td>
</tr>
<tr>
<td>GO:0071013</td>
<td>cellular_component</td>
<td>catalytic step 2 spliceosome</td>
<td>7</td>
<td>0.070</td>
</tr>
<tr>
<td>GO:0045111</td>
<td>cellular_component</td>
<td>intermediate filament cytoskeleton</td>
<td>4</td>
<td>0.074</td>
</tr>
<tr>
<td>GO:0071011</td>
<td>cellular_component</td>
<td>precatalytic spliceosome</td>
<td>5</td>
<td>0.050</td>
</tr>
<tr>
<td>GO:0042802</td>
<td>molecular_function</td>
<td>identical protein binding</td>
<td>10</td>
<td>0.058</td>
</tr>
<tr>
<td>GO:0051082</td>
<td>molecular_function</td>
<td>unfolded protein binding</td>
<td>13</td>
<td>0.023</td>
</tr>
<tr>
<td>GO:0003729</td>
<td>molecular_function</td>
<td>mRNA binding</td>
<td>6</td>
<td>0.029</td>
</tr>
<tr>
<td>GO:0016830</td>
<td>molecular_function</td>
<td>carbon-carbon lyase activity</td>
<td>11</td>
<td>0.011</td>
</tr>
</tbody>
</table>
Table 2 Gene expression summary for all contigs with putative annotation as Hsp70, Hsp90, or CytP450. Links and accession numbers for sequences and annotations are listed in the Acknowledgements section.

<table>
<thead>
<tr>
<th>Clade</th>
<th>Gene</th>
<th>min</th>
<th>max</th>
<th>down</th>
<th>up</th>
<th>min w/o highest</th>
<th>Significance p&lt;0.05</th>
<th>Significance FDR_5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>CytP450</td>
<td>-3.38</td>
<td>1.33</td>
<td>12</td>
<td>15</td>
<td>-1.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>Hsp90</td>
<td>-1.18</td>
<td>1.49</td>
<td>10</td>
<td>2</td>
<td>NA</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>Hsp70</td>
<td>-1.88</td>
<td>1.46</td>
<td>8</td>
<td>6</td>
<td>NA</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>CytP450</td>
<td>-1.03</td>
<td>1.36</td>
<td>2</td>
<td>15</td>
<td>NA</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>Hsp90</td>
<td>-2.38</td>
<td>1.16</td>
<td>13</td>
<td>5</td>
<td>-1.34</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>Hsp70</td>
<td>-1.41</td>
<td>1.30</td>
<td>18</td>
<td>7</td>
<td>NA</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>orthos(D/C)</td>
<td>CytP450</td>
<td>-1.03/-2.09</td>
<td>1.06/1.17</td>
<td>1/3</td>
<td>3/1</td>
<td>NA</td>
<td>2 for type</td>
<td>2 for type</td>
</tr>
<tr>
<td>orthos(D/C)</td>
<td>Hsp90</td>
<td>-1.34/-1.32</td>
<td>1.08/2.35</td>
<td>5/3</td>
<td>3/5</td>
<td>NA</td>
<td>7 for type</td>
<td>7 for type</td>
</tr>
<tr>
<td>orthos(D/C)</td>
<td>Hsp70</td>
<td>-1.19/-2.15</td>
<td>1.17/1.24</td>
<td>2/1</td>
<td>1/2</td>
<td>NA</td>
<td>2 for type</td>
<td>2 for type</td>
</tr>
</tbody>
</table>
Supplementary Materials

Table S1A-C. Read count and mapping statistics for A) *Symbiodinium* phylotype identification, B) total quality mapped reads, and C) de-duplicated quality mapped reads used for differential expression analyses.

Table S2. Full results and annotation list for the 463 genes with significant expression differences due to type in the heat stress analysis.

Table S3. Individual functional enrichment analyses for the genes with significant expression difference due to type and higher expression in D or higher expression in C separately.

Table S4. Full results and annotation list for the 100 genes with significant expression differences due to type in the reciprocal transplant experiment.

Table S5. Full results and annotation list for the 113 genes with significant expression differences due to type in the within host analysis.
Literature Cited


Figure 1