Dual-compartmental transcriptomic + proteomic analysis of a marine endosymbiosis exposed to environmental change

ANDERSON B. MAYFIELD,*† YU-BIN WANG,‡§[1] CHII-SHIARNG CHEN,*¶**††
SHU-HWA CHEN§ and CHUNG-YEN LIN§‡‡§§

*National Museum of Marine Biology and Aquarium, 2 Houwan Road, Checheng, Pingtung 944, Taiwan, †Living Oceans Foundation, 130 Severn Avenue, Annapolis, MD 21403, USA, ‡Graduate Institute of Zoology, National Taiwan University, No. 1, Sec. 4 Roosevelt Road, Taipei 106, Taiwan, §Institute of Information Sciences, Academia Sinica, 128 Academia Road, Sec. 2, Nangang, Taipei 115, Taiwan, ¶Taiwan Coral Research Center, 2 Houwan Road, Checheng, Pingtung 944, Taiwan, **Graduate Institute of Marine Biotechnology, National Dong-Hwa University, 2 Houwan Road, Checheng, Pingtung 944, Taiwan, ††Department of Marine Biotechnology and Resources, National Sun Yat-Sen University, 70 Lianhai Road, Kaohsiung 80424, Taiwan, ‡‡Institute of Fisheries Science, National Taiwan University, No. 1, Sec. 4 Roosevelt Road, Taipei 106, Taiwan, §§Institute of Population Health Sciences, National Health Research Institutes, 5 Keyan Road, Zhunan Miaoli 350, Taiwan

Abstract

As significant anthropogenic pressures are putting undue stress on the world's oceans, there has been a concerted effort to understand how marine organisms respond to environmental change. Transcriptomic approaches, in particular, have been readily employed to document the mRNA-level response of a plethora of marine invertebrates exposed to an array of simulated stress scenarios, with the tacit and untested assumption being that the respective proteins show a corresponding trend. To better understand the degree of congruency between mRNA and protein expression in an endosymbiotic marine invertebrate, mRNAs and proteins were sequenced from the same samples of the common, Indo-Pacific coral *Seriatopora hystrix* exposed to stable or upwelling-simulating conditions for 1 week. Of the 167 proteins downregulated at variable temperature, only two were associated with mRNAs that were also differentially expressed between treatments. Of the 378 differentially expressed genes, none were associated with a differentially expressed protein. Collectively, these results highlight the inherent risk of inferring cellular behaviour based on mRNA expression data alone and challenge the current, mRNA-focused approach taken by most marine and many molecular biologists.

Keywords: acclimation, coral reef, endosymbiosis, environmental change, gene expression, proteome, transcriptome

*Received 21 April 2016; revision received 7 October 2016; accepted 10 October 2016*

Introduction

As scleractinian corals and the reefs they create are currently under threat of global climate change (Hoegh-Guldberg *et al.* 2007), as well as an array of additional anthropogenic insults (e.g. Huang *et al.* 2011), there has

been an impetus to learn more about both their basic biology (e.g. Mayfield & Gates 2007; Mayfield *et al.* 2010, 2012b; Peng *et al.* 2011; Chen *et al.* 2012, 2015) and environmental physiology (e.g. Mayfield *et al.* 2011, 2013a,c,d, 2015). For instance, recent studies in Hawaii and Taiwan have shed light on the capacity for coral acclimation to differential salinity (Mayfield *et al.* 2013d), temperature (Mayfield *et al.* 2013b, 2014a), and pCO_2 (Putnam *et al.* 2013) regimes in common, widely distributed species, notably *Pocillopora damicornis* and *Seriatopora hystrix* (Fig. 1A,B). The latter species is

Correspondence: Anderson B. Mayfield, Fax: 886 2 265 18660;
E-mail: andersonblairmayfield@gmail.com and Chung-Yen Lin, Fax: 886 2 278 24814; E-mail: cylin@iis.sinica.edu.tw
[1]Contributed equally.
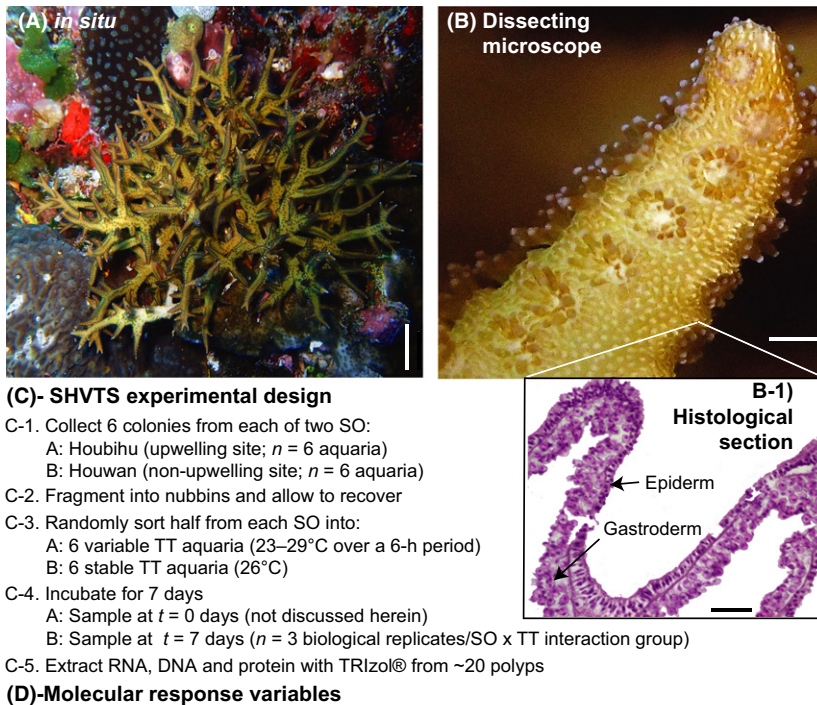
Fig. 1 Key components of the *Seriatopora hystrix* variable temperature study (SHVTS). A model coral for ecophysiological research, *S. hystrix* (A,B), an outline of the SHVTS (C) and a description of the molecular response variables measured herein and in prior works (D). Scale bars in A, B and B-1 represent 30, 3 and 0.3 mm, respectively, and the latter histological section, which was prepared as in Mayfield *et al.* (2013b), was presented to depict the dual-tissue layer structure of this coral species, as well as all other cnidarians. SO = site of origin. TT = temperature treatment. RFLP = restriction fragment length polymorphism. 2D + MS = two-dimensional gel electrophoresis followed by mass spectrometry. NA = not applicable. [Colour figure can be viewed at wileyonlinelibrary.com].

(C)- SHVTS experimental design

C-1. Collect 6 colonies from each of two SO:
   A: Houbihu (upwelling site; *n* = 6 aquaria)
   B: Houwan (non-upwelling site; *n* = 6 aquaria)

C-2. Fragment into nubbins and allow to recover

C-3. Randomly sort half from each SO into:
   A: 6 variable TT aquaria (23–29°C over a 6-h period)
   B: 6 stable TT aquaria (26°C)

C-4. Incubate for 7 days
   A: Sample at *t* = 0 days (not discussed herein)
   B: Sample at *t* = 7 days (*n* = 3 biological replicates/SO x TT interaction group)

C-5. Extract RNA, DNA and protein with TRIzol® from ~20 polyps

(D)-Molecular response variables

Biological material	Parameter assessed	Method #1	Method #2	Reference (s)
RNA	Gene expression	Real-time PCR	RNA-Seq	Mayfield *et al.* (2012a, **herein**)
DNA (host)	Genotype	PCR+sequencing	NA	Mayfield *et al.* (2014b)
DNA (*Symbiodinium*)	Genotype	RFLP	Real-time PCR	Mayfield *et al.* (2012a, 2014b)
Protein	Protein expression	Western blot	2D+MS	Mayfield *et al.* (2016)

known to be environmentally sensitive in Japan (Loya *et al.* 2001), although in Southern Taiwan it can withstand 8–10°C temperature changes brought upon by upwelling events that commonly occur in the boreal summer (Jan & Chen 2008).

Prior studies aimed to uncover the mechanisms by which these resilient seriatoporids from Southern Taiwan acclimate to variable temperatures by conducting a laboratory-based reciprocal transplant, herein referred to as the '*S. hystrix* variable temperature study' (SHVTS; Fig. 1C), whereby corals from an upwelling reef in Nanwan Bay, Houbihu, were exposed to a stable temperature profile (26 °C) while conspecifics from a nearby (~20 km), non-upwelling site, Houwan, were exposed to variable temperatures (23–29 °C over a 6-h period) for 1 week (Mayfield *et al.* 2012a). Simultaneously, controls from the two sites of origin were exposed to variable and stable temperature regimes, respectively. As expected, corals from the upwelling site grew faster at variable temperature, whereas those from the non-upwelling site grew faster at stable temperature (Mayfield *et al.* 2012a). However, even 'transplanted' corals continued to grow at normal rates, evidently unstressed at temperature regimes to which they had never before been exposed in the case of the Houwan samples incubated at the variable temperature regime.

To uncover the molecular mechanisms by which such acclimation to variable temperatures occurred, mRNAs and proteins from these coral samples were sequenced with 'RNA-Seq' (Illumina) technology and mass spectrometry [MS; following two-dimensional (2D) gel electrophoresis], respectively (Fig. 1D). The 75 host coral and 42 endosymbiotic dinoflagellate (genus *Symbiodinium*) proteins underexpressed at variable temperature, respectively, are described in a sister publication (Mayfield *et al.* 2016) and feature representatives from a variety of cellular pathways, notably mRNA processing and lipid body (LB) formation.

It was predicted that the differentially expressed proteins (DEPs) uncovered by this proteome-scale analysis could be validated at the mRNA level with RNA-Seq followed by, in certain cases, real-time PCR (qPCR). If such congruency could be documented (sensu Greensbaum *et al.* 2003), it not only would represent a step towards understanding protein regulation in corals, but it might also justify the current state of research in the field, whereby assumptions about protein behaviour are made based on mRNA expression data alone (e.g. Barshis *et al.* 2013; Palumbi *et al.* 2014). Briefly, gene expression analyses are cheaper, faster and more sensitive than the proteome-scale methods currently available; if mRNA expression of a target gene is a valid proxy for expression of the respective protein, then only

the former need be measured, thereby saving time and money. In yeast, there is a modest degree of correlation ($r^2 = 0.73$) between mRNA and protein expression (Lu *et al.* 2007), although such is not the case in humans ($r^2 = 0.40$; Fagerberg *et al.* 2014) and *Escherichia coli* ($r^2 = 0.47$; Lu *et al.* 2007). Herein only 2 proteins out of the pool of 167 DEPs (117 described in Mayfield *et al.* 2016 + 50 identified herein; <2%) found to be downregulated in samples of the variable temperature treatment by 2D + MS were associated with mRNAs that revealed a similar temperature effect. This suggests that inferring protein expression levels from mRNA concentrations alone may be unjustified for this coral holobiont, and possibly others.

Materials and methods

The experiment

The SHVTS (Fig. 1C) was conducted in 2010 at Taiwan's National Museum of Marine Biology and Aquarium (NMMBA) and has been described in detail in prior works (e.g. Mayfield *et al.* 2013c). Briefly, six *S. hystrix* colonies from each of the two study sites were transported to NMMBA, allowed to acclimate for 3 days and fragmented into nubbins (Fig. 2A; four nubbins/colony). The 48 nubbins were allowed to acclimate from fragmentation for 3 weeks at 26 °C, randomly mixed within a seawater table for each site of origin to eliminate the effect of colony as a factor and randomly assigned to one of six aquaria for each site of origin (4 nubbins/aquarium). Half of the aquaria from each site were assigned to the variable temperature treatment (23–29 °C over a 6-h period × 4 periods/24-h cycle), while the other three were assigned to the stable temperature regime (26 °C only), and four nubbins were suspended in each of the 12 aquaria ($n = 3$ aquaria for each of the four site of origin × temperature treatment interaction groups). Aquarium temperatures were tightly controlled by a series of commercially available heaters and chillers as described in Mayfield *et al.* (2012a).

Half of the 48 nubbins were collected just prior to the initiation of the variable temperature regime (time = 0) and are not discussed herein. After 7 days of stable or variable temperature exposure, two nubbins were collected from each aquarium, and one of the two pseudo-replicates was randomly chosen for RNA-Seq analysis (12 sequenced samples). For the 2D + MS analysis (Mayfield *et al.* 2016), proteins extracted from these same 12 nubbins were pooled across sites of origin prior to electrophoresis to test only the effect of temperature on protein expression; as discussed by Mayfield *et al.* (2014b), temperature tended to have a greater effect on coral physiology than collection site (Houbihu

vs. Houwan). Furthermore, proteins from corals of the two sites were differentially soluble in the dual-protein buffering system described in Mayfield *et al.* (2016); therefore, separating them by site of origin could have inadvertently contributed to the variation in the perceived protein expression profiles.

RNA extraction, cDNA library preparation and mRNA sequencing

RNAs were extracted as in Putnam *et al.* (2013) from the same 12 samples from which the proteins analysed by 2D + MS were isolated (Mayfield *et al.* 2016). Details of the RNA extraction, cDNA library preparation and mRNA sequencing protocols (Fig. 2B) can be found in Appendix S1 (Supporting information).

Read acquisition and assembly

In addition to the 12 Taiwanese *S. hystrix* cDNA libraries sequenced from the SHVTS, raw read files from another *S. hystrix* transcriptome project were graciously provided by the Meyer Lab (Oregon State University) to aid in the initial assembly. By combining sequence data from transcriptomes of different samples of the same species, it was hoped that a more complete reference transcriptome could be generated against which the mRNAs from the 12 SHVTS samples could be aligned and quantified. In total, 42 464 176 (67%) and 21 467 731 (33%) high-quality reads were from the SHVTS and Meyer Lab, respectively, and SolexaQA (Cox *et al.* 2010) was used to check the global quality of the transcript reads and trim each to its longest contiguous read segment. Sequences <25 bp or with quality scores <20 were not considered in the assembly. After quality control (QC), the remaining 63 931 907 read pairs were assembled into 209 481 contigs >200 bp (Fig. 2C) with Trinity (Grabherr *et al.* 2011) on a super-computer running Linux. 'Transdecoder' (Trinity) was used to identify and translate open reading frames (ORFs), and protein sequences <30 amino acids (AA) were discarded. Only the longest continuous protein sequence generated from a contig was used in the alignment-based analyses described in Appendix S1 (Supporting information). Although 126 168 transcripts encoded putative protein products, a subset of 41 175 was the focus of this study (Fig. 2C) for reasons outlined below.

Annotation and taxonomy breakdown

A variety of alignment-based bioinformatics analyses were performed to attempt to assign a gene identity, as well as a compartment of origin, to each contig, and the
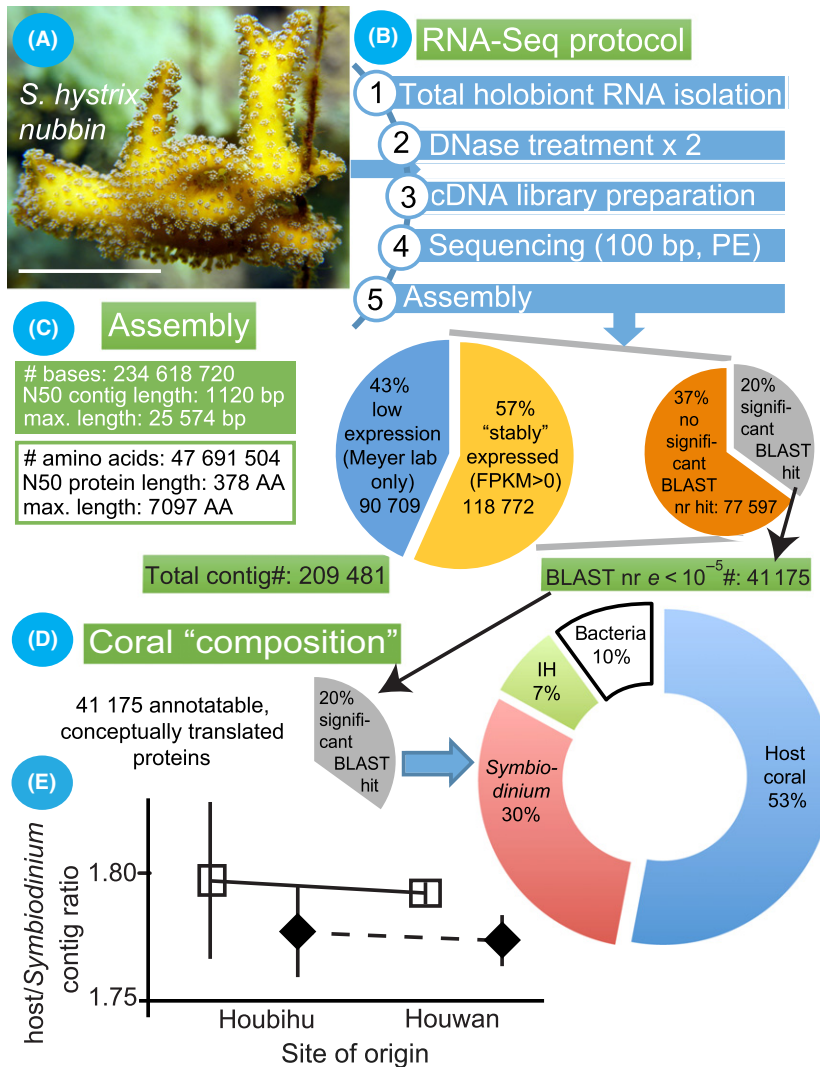
Fig. 2 Breakdown of the *Seriatopora hystrix–Symbiodinium* transcriptome. Reads derived from RNA-Seq cDNA libraries created from each of 12 samples of the *S. hystrix* (A; scale bar = 1 cm) variable temperature study (SHVTS) collected at the 7-d sampling time (~two-thirds of the assembly; B) were combined with short-read mRNA sequences from the laboratory of Dr. Eli Meyer at Oregon State University (~one-third of the assembly) and assembled (C). Of the 118 772 contigs expressed by samples of the SHVTS, 41 175 (20% of the total, 209 481-contig assembly) could be conceptually translated into proteins that aligned significantly ($e < 10^{-5}$) to those of the NCBI nr database (C). The taxonomic breakdown of these 41 175 conceptually translated open reading frames (ORFs) is also shown (D). A host coral/*Symbiodinium* contig ratio was calculated for each SHVTS sample (E) and analysed across site of origin and temperature treatment (stable and variable temperature treatments are denoted by hollow squares and filled diamonds, respectively, and are connected by solid and dotted lines, respectively.) with two-way ANOVA. In (E), error bars represent standard deviation ($n = 3$ biological replicates for each interaction group). bp = base pair. PE = paired end. AA = amino acid. FPKM = fragments per kilobases mapped. IH = intermediate homology (of putative host coral, *Symbiodinium* or bacterial origin). [Colour figure can be viewed at wileyonlinelibrary.com].

details can be found in Appendix S1 (Supporting information). For each of the 12 samples of the SHVTS, a host coral/*Symbiodinium* contig ratio was calculated, and these data were analysed by two-way ANOVA with JMP® (ver. 12) to determine whether site of origin, temperature regime or their interaction affected the mRNA breakdown of the holobiont after having confirmed both normality of data (Shapiro–Wilk W test $P > 0.05$) and homogeneity of variance (Levene's test $P > 0.05$). When these conditions were not met for any parameter, nonparametric tests based on rank transformations were instead utilized except in special cases involving an abundance of null expression levels; such expression data were analysed in a different manner (described below).

Read quantification

Expression profiling with RNA-Seq requires mapping of reads to contigs, and a single read may map to multiple contigs. RNA-Seq by expectation maximization (RSEM; Li & Dewey 2011) is based on the expectation maximization (EM) algorithm and uses paired-end sequence quality information as part of its model to generate accurate alignment data. Herein 42 464 176 high-quality reads were mapped onto 209 481 unique transcripts using RSEM, resulting in 33 867 406 (79.8% of the high-quality reads) paired alignments that condensed into 193 499 unique transcripts (92.4% of the total assembly). An extra step was added to filter out low-expression contigs [<1 fragment per kilobases mapped (FPKM)], resulting in 118 772 (56.7% of the total assembly) 'stably expressed' transcripts (Fig. 2C). Remapping of the high-quality reads of each sample to these stably expressed transcripts allowed for a focusing on genes that were expressed above background levels. After normalization, 33 510 486 quantified fragments (78.9% of the total read count) were found to map to the assembled transcriptome. The mRNA sequences

(fasta files), expression levels (csv files) and conceptually translated ORF sequences (fasta files) of the 118 772 contigs can be downloaded from the SHVTS transcriptome website: http://symbiont.iis.sinica.edu.tw/s_hystrix. Details of the features provided by this interactive website can be found in Appendix S1 (Supporting information). Raw read files (fastq) and expression data can also be accessed from Dryad (doi:10.5061/dryad.7cv4g).

Differentially expressed genes

Because of the large number of contigs, there was a focus on differentially expressed genes (DEGs) in four groups: (i) those mRNAs expressed *only* by samples of the stable temperature treatment, (ii) those expressed *only* by samples of the variable temperature regime, (iii) those expressed *only* by samples from Houbihu, and (iv) those expressed *only* by samples from Houwan. Although significant interaction effect genes [i.e. those expressed only by samples of one site of origin at one temperature treatment (e.g. Houbihu + variable temperature only)] were tested for in the ANOVA model described below and can be searched for on the interactive SHVTS transcriptome website detailed in Appendix S1 (Supporting information), they are not discussed herein. For all four comparison categories, there was the additional condition that all six samples of the treatment of interest expressed the gene. For both RNA-Seq and qPCR data, two-way ANOVAs were performed with JMP to uncover the effects of site of origin, temperature treatment and their interaction on both host coral and *Symbiodinium* gene expression. A zero-inflated model based on the Poisson distribution was then performed with JMP to confirm significant differences ($P < 0.05$) due to half of the samples being characterized by null expression levels for the DEGs in the four aforementioned categories. qPCR primers were designed from 1, 0, 7 and 4 genes, respectively, in the aforementioned four groups [Table S1, Supporting information; an *e*-value of 10^{-2} (in contrast to the 10^{-5} value set for taxonomy assignment) was used for gene identity assignment.]. The expression of eight and six *S. hystrix* and *Symbiodinium* genes, respectively, was already measured in the 12 sequenced samples (Mayfield *et al.* 2014b), and the degree of correlation between qPCR and RNA-Seq was assessed for these 14 genes in addition to the 12 newly targeted ones. The proteins conceptually translated from these 26 mRNAs were BLASTed (BLASTp) against the SHVTS differentially expressed proteome to determine whether the respective proteins were differentially expressed.

Mascot-derived DEPs vs. the SHVTS transcriptome

As described in Mayfield *et al.* (2016), 117 DEPs were uncovered within 10 protein spots (Table S2, Supporting information) removed from 2D gels aimed *only* at uncovering temperature, and not site of origin, effects on *S. hystrix–Symbiodinium* protein expression. Several analytical approaches were taken in order to understand the relationship between this partial proteome (pooled across sites of origin) and the fully sequenced transcriptomes (unpooled), which were derived from the same samples. First, the DEPs, all of which represented proteins that were downregulated at variable temperature, were used to query the SHVTS transcriptome (tBLASTn) to acquire the respective mRNA expression data (FPKM). As the peptide sequences were generally too short (Tables S2, Supporting information) to align significantly to conceptually translated *S. hystrix–Symbiodinium* proteins, the top hit from the Mascot (Matrix Science) analysis, which was typically a full-length protein from a model organism (Mayfield *et al.* 2016), was instead used in the tBLASTn analysis, and the mRNA expression data from conceptually translated proteins that demonstrated >40% identity to this 'reference' protein were acquired and analysed by two-way ANOVA as described above. These same reference proteins were aligned (tBLASTn) against the *P. damicornis* transcriptome sequenced by Mayfield *et al.* (2014d): http://symbiont.iis.sinica.edu.tw/coral_pdltte/static/html/index.html#home), and the respective mRNA expression data (FPKM) associated with conceptually translated proteins demonstrating >40% identity were analysed by repeated-measures ANOVA to determine the effects of temperature (27 vs. 30 ℃), time (2 vs. 36 weeks) and their interaction. qPCR primers (Table S1, Supporting information) were designed for 10 *S. hystrix* genes whose (i) respective proteins were downregulated at variable temperatures and (ii) mRNA expression levels were affected significantly by site of origin, temperature regime and/or their interaction ($P < 0.05$).

MS-SCAN-derived DEPs vs. the SHVTS transcriptome

Freeware based on MS-GF+ (Kim & Pevzner 2014), MS-SCAN, was implemented on the SHVTS transcriptome server (http://symbiont.iis.sinica.edu.tw/s_hystrix/static/html/#mscan) so that raw MS .mgf files could be used directly as queries of the SHVTS transcriptome. Briefly, although Kim & Pevzner (2014) found that Mascot and MS-GF+ yielded similar results, it was hypothesized that the latter might unveil additional proteins not identified with Mascot by Mayfield *et al.* (2016). Preliminary analysis of the 10 spots with MS-SCAN yielded thousands of

predominantly small (<10 AA) reads (Table S2, Supporting information). Therefore, the longest five reads from each spot (typically >30 AA) *only* were used as queries (tBLASTn) of the SHVTS transcriptome, as the a priori-determined minimal AA length for inclusion as a 'true' DEP set forth by Mayfield *et al.* (2016) was 15. The expression data of the respective 50 mRNAs were acquired and analysed by two-way ANOVA as described above to further determine the degree of congruency between the transcriptome- and proteome-scale approaches. It was hypothesized that the majority of these 50 mRNAs would be overexpressed by samples of the stable temperature treatment, as were the respective proteins.

qPCR validation of DEGs, DEPs and a priori-selected genes

There were three sets of qPCR assays developed herein or in previous works (Table S1, Supporting information): (i) those genes found to be differentially expressed by analysis of the transcriptome alone (described above; $n = 12$), (ii) those genes whose respective proteins were found to be differentially expressed by 2D + MS and whose mRNA expression levels showed either a site of origin, temperature or interaction effect ($n = 10$ and 1 genes identified by Mascot and MS-SCAN, respectively) and (iii) genes hypothesized a priori to be differentially expressed between temperature treatments and whose expression was measured in prior works in these same samples ($n = 14$; Mayfield *et al.* 2012a, 2014b). The latter were used to (i) verify the correlation between RNA-Seq and qPCR and (ii) the congruency between 2D + MS and qPCR. The details of the qPCR analysis can be found in Appendix S1 (Supporting information).

Congruency vs. correlation

There was an interest in knowing whether the three techniques utilized herein – RNA-Seq, qPCR and 2D + MS – yielded similar results, and, more generally, whether mRNA-level changes manifested at the protein level (and vice versa). As the 2D + MS data were qualitative and pooled across collection sites, whereas the mRNA data were quantitative and unpooled, only 'congruency', rather than correlation, could be assessed between the mRNA- and protein-targeting techniques. 'Congruency' was defined to occur when two techniques gave the same result, regardless of whether or not the data correlated significantly or differed significantly between treatments. The congruency of only temperature regime-induced expression changes, and not site of origin-induced ones, could be determined between 2D + MS and the two mRNA-targeting approaches. In contrast, qPCR data could be regressed directly against RNA-Seq values, and it was hypothesized that a significant, positive, linear correlation would be obtained between RNA-Seq- and qPCR-derived data. Congruency was assessed between (i) RNA-Seq and 2D + MS and (ii) qPCR and 2D + MS.
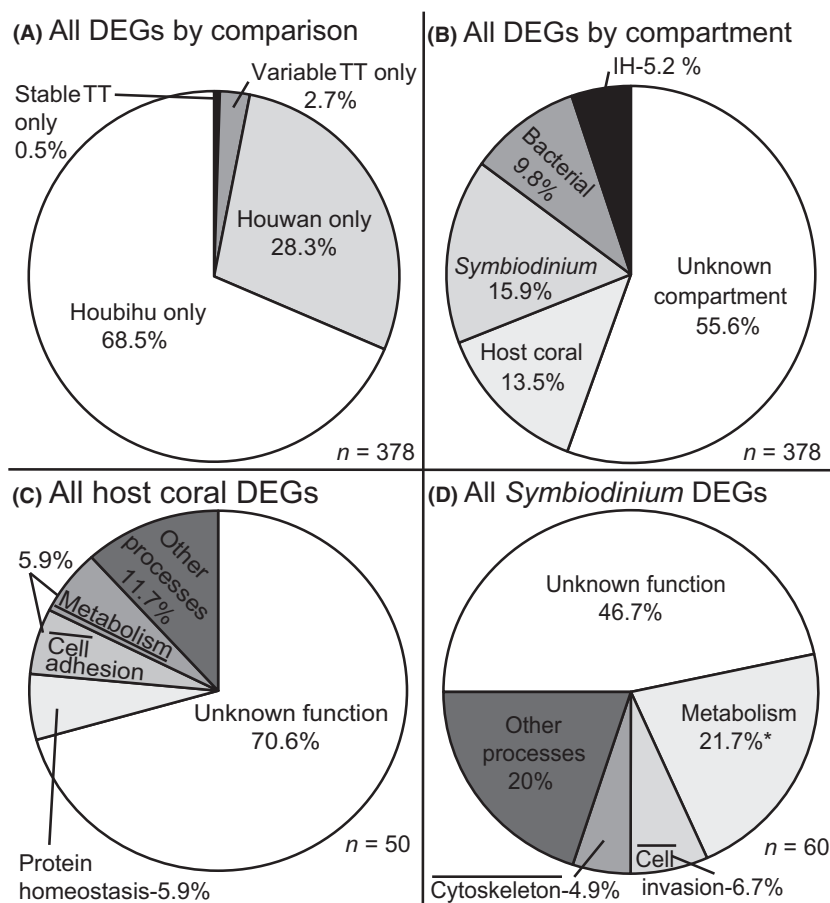
Results

Transcriptome breakdown

Of the 118 772 stably expressed contigs (NCBI BioProject PRJNA310803), 70 983 ORFs were identified, and the corresponding protein sequences aligned significantly to 11 813 proteins in the PFAM database. Furthermore, 22 195 protein sequences aligned significantly to 6136 proteins in the KEGG database, and 51 241 protein sequences aligned significantly to 35 510 proteins in the NCBI nr database. Regarding the latter group, the compartment of origin (Fig. 2D) was determined for the 41 175 proteins translated from the stably expressed contigs, and sequences aligned to those of 3234 different organisms: 21 949, 12 194, 3928 and 3104 contigs were found to be of host coral, *Symbiodinium*, bacterial and intermediate homology (IH) origin, respectively (Fig. 2D). There was no significant effect of site, temperature or their interaction on the host coral/*Symbiodinium* contig ratio (Fig. 2E), which averaged 1.79 [±0.03 (SD for this and all values henceforth)]; therefore, RNA-Seq FPKM data were *not* normalized to the *Symbiodinium* biomass 'proxy' advocated in prior works (e.g. Mayfield *et al.* 2013a).

DEGs identified by RNA-Seq

Across the 118 772 stably expressed contigs, 2, 10, 259 and 107 mRNAs were expressed by all six samples of the stable temperature profile only, the variable temperature regime only, Houbihu only and Houwan only, respectively, and not by the six samples of the opposing treatment groups (Fig. 3A). Of these 378 DEGs (Fig. 3B), a similar percentage was of host coral (Fig. 3C) and *Symbiodinium* (Fig. 3D) origin, although there was a relatively higher proportion of DEGs involved in metabolism in the *Symbiodinium* DEG pool compared with the host coral one. None of the respective 378 proteins were identified in the 167-protein differentially expressed proteome (117 DEPs from Mascot and 50 from MS-SCAN).

Of the two genes expressed only by corals of the stable temperature treatment (Fig. 4A), only one, a fibronectin of likely host origin, could be confidently identified. The respective protein was not uncovered by

(A) All DEGs by comparison

Stable TT
only
0.5%

Variable TT only
2.7%

Houwan only
28.3%

Houbihu only
68.5%

n = 378

(B) All DEGs by compartment

IH-5.2 %

Bacterial
9.8%

Symbiodinium
15.9%

Unknown
compartment
55.6%

Host coral
13.5%

n = 378

(C) All host coral DEGs

5.9%

Other
processes
11.7%

Metabolism

Cell
adhesion

Unknown function
70.6%

Protein
homeostasis-5.9%

n = 50

(D) All *Symbiodinium* DEGs

Unknown function
46.7%

Other
processes
20%

Metabolism
21.7%*

Cell
invasion-6.7%

Cytoskeleton-4.9%

n = 60

Fig. 3 Pie graphs of differentially expressed genes (DEGs)-1: DEGs sorted by comparison (A), compartment (B), host coral only (all four comparisons; C) and *Symbiodinium* only (all four comparisons; D). In (B), there was a similar proportion of DEGs of host coral (50/378 = 13.5%) and *Symbiodinium* (60/378 = 15.9%) origin (two-sample proportion test, $P > 0.05$). When an asterisk has been placed next to a KEGG-assigned functional group in (D), the relative proportion of the respective category was significantly higher in the *Symbiodinium* compartment (two-sample proportion test, $P < 0.05$). When a line has been placed over or under a functional group name in (C) and (D), the respective functional group was over- or under-represented, respectively, relative to the entire assembly ($P < 0.05$). For a breakdown of annotatable mRNAs for each compartment within each of the four comparisons of interest, please see Fig. 4. TT = temperature treatment. IH = intermediate homology.
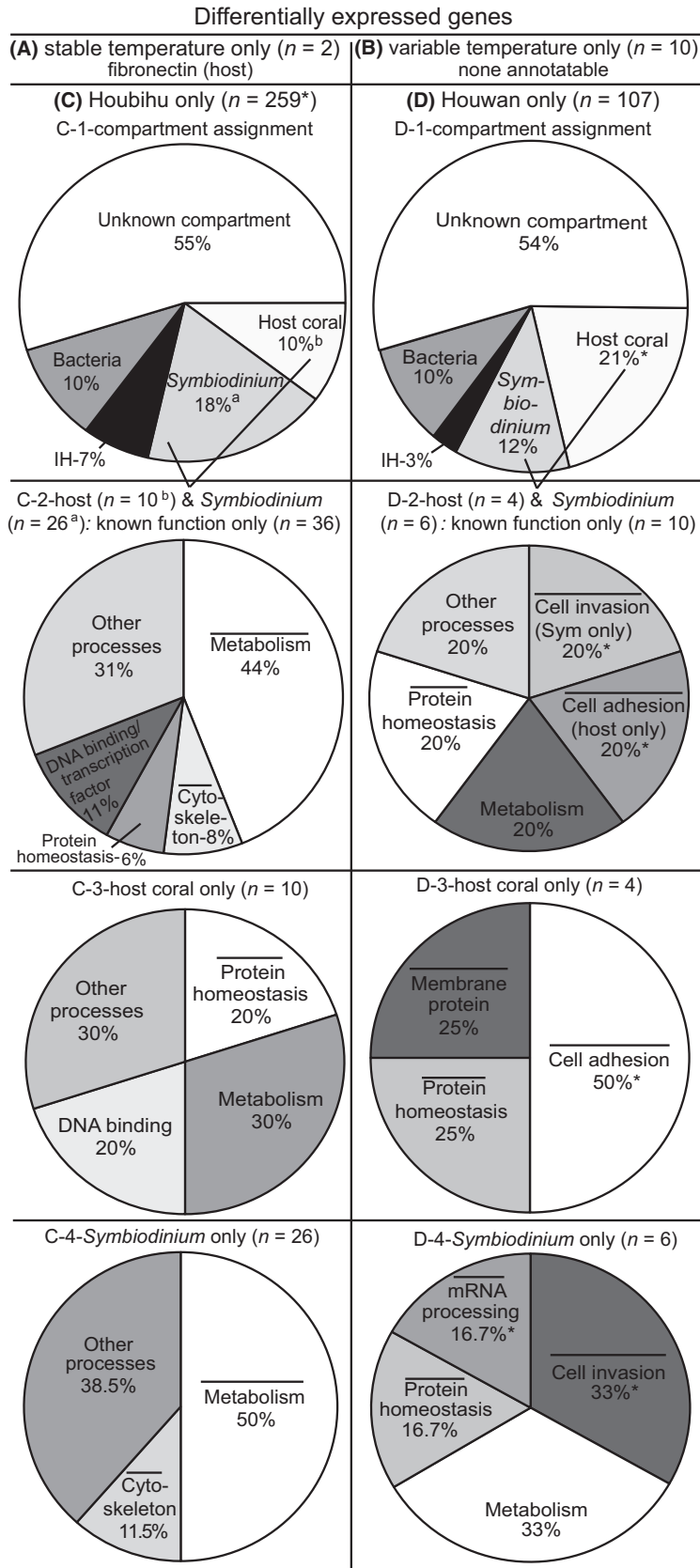
2D + MS, nor did qPCR (Fig. S1A–C, Supporting information) verify this RNA-Seq-based temperature treatment effect (Table 1). Of the 10 DEGs expressed only by the samples of the variable temperature regime (Fig. 4B), none could be confidently assigned an identity by querying the NCBI, PFAM or KEGG databases. BLASTx analysis of these 10 cDNA sequences did not yield significant hits against the 167-protein differentially expressed proteome, suggesting that the respective proteins were not sequenced.

Of the 259 gene mRNAs expressed only by samples of Houbihu (Fig. 4C), 27, 47, 26 and 17, were of host coral, *Symbiodinium*, bacterial and IH origin (Fig. 4C-1), respectively, and none of the respective proteins were identified by 2D + MS. Of the Houbihu-only DEG pool (Fig. 4C-1 and 4C-2), a significantly higher proportion of the genes were of *Symbiodinium* relative to host coral origin (two-sample proportion test, $Z = 2.73$, $P < 0.001$). Of the 27 genes that could be confidently assigned to being of host coral origin, 10 (37%) aligned significantly to functionally characterized proteins (Fig. 4C-3). Of this host subset, multiple genes were shown to belong to three functional categories: metabolism ($n = 3$; two galactosyltransferases and a tankyrase also known to

function as a galactosyltransferase), protein homeostasis/QC ($n = 2$; protein-glutamine gamma-glutamyltransferase 4-like and a protein containing a kelch-like motif) and DNA-binding/transcription factor [$n = 2$; GDNF-inducible zinc finger protein 1 (*gizfp1*-like) and tigger transposable element-derived protein 6-like (*tted6*-like)]. qPCR verified the RNA-Seq-derived findings for an egg protein (Fig. S1D–F, Supporting information) in additional to these two transcription factors (Fig. S1G–I and S1J–L [Supporting information] for *gizfp1*-like and *tted6*-like, respectively; Table 1).

Of the 47 genes expressed only by *Symbiodinium* populations residing within corals originally collected from Houbihu, 26 encoded characterized proteins (Fig. 4C-4), and qPCR assays were designed for four genes (Table S1, Supporting information) – axonemal dynein heavy chain 7 (*cildyn7*; cytoskeleton; Fig. S2A–C, Supporting information), a hypothetical protein (contig 8516; Fig. S2D–F, Supporting information), meprin and TRAF-C homology (MATH) domain (Fig. S2G–I, Supporting information; protein QC) and a starch-binding domain-containing protein (Fig. S2J–L, Supporting information; metabolism). These two mRNA-targeting techniques yielded congruent results for all four genes

Differentially expressed genes

(A) stable temperature only (*n* = 2)
fibronectin (host)

(B) variable temperature only (*n* = 10)
none annotatable

(C) Houbihu only (*n* = 259*)
C-1-compartment assignment

Unknown compartment 55%
Host coral 10%[b]
Symbiodinium 18%[a]
Bacteria 10%
IH-7%

(D) Houwan only (*n* = 107)
D-1-compartment assignment

Unknown compartment 54%
Host coral 21%*
Symbiodinium 12%
Bacteria 10%
IH-3%

C-2-host (*n* = 10[b]) & *Symbiodinium* (*n* = 26[a]): known function only (*n* = 36)

Other processes 31%
Metabolism 44%
DNA binding/transcription factor 11%
Cytoskeleton-8%
Protein homeostasis-6%

D-2-host (*n* = 4) & *Symbiodinium* (*n* = 6): known function only (*n* = 10)

Other processes 20%
Cell invasion (Sym only) 20%*
Protein homeostasis 20%
Cell adhesion (host only) 20%*
Metabolism 20%

C-3-host coral only (*n* = 10)

Protein homeostasis 20%
Other processes 30%
Metabolism 30%
DNA binding 20%

D-3-host coral only (*n* = 4)

Membrane protein 25%
Cell adhesion 50%*
Protein homeostasis 25%

C-4-*Symbiodinium* only (*n* = 26)

Other processes 38.5%
Metabolism 50%
Cytoskeleton 11.5%

D-4-*Symbiodinium* only (*n* = 6)

mRNA processing 16.7%*
Protein homeostasis 16.7%
Cell invasion 33%*
Metabolism 33%

Fig. 4 Pie graphs of differentially expressed genes (DEGs)-2: DEGs were sorted into those expressed only by samples of the stable temperature treatment (A), those expressed only by samples of the variable temperature regime (B), those expressed only by samples of the upwelling site Houbihu (C) and those expressed only by samples of the non-upwelling site Houwan (D). Alignment-based homology searches were used to assign contigs of the latter two groups (C-1 and D-1, respectively) to one of five compartments: unknown (no significant BLAST hit), host coral, *Symbiodinium*, bacteria or intermediate homology (IH; of host coral, *Symbiodinium* or bacterial origin). Pie graphs were then made to show the breakdown of the functional categories (derived from comparison with the KEGG database) for genes whose respective proteins have been functionally characterized for both Houbihu-only (C-2) and Houwan-only (D-2) contigs. The breakdown of annotatable proteins was then shown for each of the two predominant compartments of *Seriatopora hystrix*: the host coral itself (C-3 and D-3 for Houbihu- and Houwan-only expressed contigs, respectively) and *Symbiodinium* (C-4 and D-4, respectively). When the value of a parameter differed between sites of origin, an asterisk (*) has been placed next to the higher of the two values. When one eukaryotic compartment of the endosymbiosis was over-represented relative to the other (two-sample proportion test, *P* < 0.05), lower-case letters ('a' and 'b') have been used to represent the over- and under-represented compartment, respectively. When a line has been placed over or under a functional group in C-2 to C-4 and D-2 to D-4, the respective functional group was over- or under-represented, respectively, relative to the entire assembly (*P* < 0.05). IH = intermediate homology.

Table 1 Correlation/congruency between one protein and two mRNA quantification techniques. Of the 12, 11 and 14 RNA-Seq-, 2D + MS- and a priori-selected genes, respectively, targeted by real-time PCR (qPCR), a statistically significant correlation between RNA-Seq- and qPCR-derived data was verified for 8 (75%), 1 (9%) and 2 (14%) genes, respectively [11 of the 37 target genes (30%)]; the host and *Symbiodinium* compartments contributed 6 (24% of the 25 host genes) and 5 (42% of the 12 *Symbiodinium* genes) genes, respectively. The average r^2 value between RNA-Seq and qPCR was statistically similar across the 25 host and 12 *Symbiodinium* genes [0.22 ± 0.24 (SD) and 0.30 ± 0.33, respectively]. The 8 (75%), 0 and 7 (50%) genes, respectively, for which such techniques yielded congruent results have been highlighted in blue except when all three techniques yielded congruent results (highlighted in red). qPCR and 2D + MS yielded a congruent result for 4 (33%), 0 and 7 (50%) genes/proteins, respectively, in the aforementioned three categories, and these are highlighted in yellow except for when all three techniques yielded congruent results, these are highlighted in green. 2D + MS and RNA-Seq yielded congruent results for 0, 2 (18%) and 12 (86%) genes/proteins, respectively, and, excluding the six molecules for which all three techniques yielded congruent results (no site of origin, temperature treatment or interaction effects) for one and five *Symbiodinium* and host genes/proteins, respectively (16% of the 37 genes targeted by qPCR). See Fig. 5 for a graphical display of these results [Colour table can be viewed at wileyonlinelibrary.com]

Gene	Fig.	Compartment	RNA-Seq result	2D + MS result	qPCR result	RNA-Seq vs. qPCR r^2	Linear regression test t-value
Differentially expressed genes emerging from transcriptome analysis (n = 12; Figs S1 and S2, Supporting information)							
fibronectin	S1A-C	Host coral	Stab only	NS	NS	0.00	−0.07
egg protein	S1D-F	Host coral	HBH only	NS	HBH only	0.54	3.40**
GDNF-inducible zinc finger protein 1-like	S1G-I	Host coral	HBH only	NS	HBH only	0.78	6.11***
tigger transposable element-derived protein 6-like	S1J-L	Host coral	HBH only	NS	HBH only	0.73	5.18***
iGLON5	S1M-O	Host coral	HWN only	NS	NS	0.03	−0.59
sacsin	S1P-R	Host coral	HWN only	NS	NS	0.00	−0.18
axonemal dynein heavy chain 7	S2A-C	*Symbiodinium*	HBH only	NS	HBH only	0.73	5.17***
hypothetical protein (contig 8516)	S2D-F	*Symbiodinium*	HBH only	NS	HBH only	0.61	3.96**
meprin and TRAF-C homology	S2G-I	*Symbiodinium*	HBH only	NS	HBH only	0.73	5.22***
starch-binding domain-containing protein	S2J-L	*Symbiodinium*	HBH only	NS	HBH only	0.65	4.35**
OTU domain-containing protein 3	S2M-O	*Symbiodinium*	HWN only	NS	HWN only	0.65	4.30**
vacuolar ATP synthase subunit b	S2P-R	*Symbiodinium*	HWN only	NS	NS	0.00	−0.10
Differentially expressed proteins emerging from proteome analysis (n = 11; Fig. S3, Supporting information)							
chromodomain helicase DNA-binding protein 1	S3A-C	Host coral	stab > var	stab > var	NS	0.05	0.74
ciliary dynein heavy chain	S3D-F	Host coral	HWN > HBH	stab > var	NS	0.01	0.30
dumpy CG33196-PB	S3G-I	Host coral	HWN > HBH	stab > var	NS	0.16	−1.39
fanconi anaemia group J protein	S3J-L	Host coral	HBH > HWN	stab > var	NS	0.07	0.90
fibroblast growth factor receptor a	S3M-O	Host coral	HBH > HWN	stab > var	NS	0.06	0.81
pre-mRNA-splicing factor SYF1	S3P-R	Host coral	HBH > HWN	stab > var	NS	0.02	0.39
DM domain protein	S3S-U	Host coral	HBH > HWN	stab > var	NS	0.13	1.20

Table 1 Continued

Gene	Fig.	Compartment	RNA-Seq result	2D + MS result	qPCR result	RNA-Seq vs. qPCR r^2	Linear regression test t-value
porphobilinogen deaminase	S3V–X	Host coral	HBH > HWN	stab > var	HBH-stab > HWN-stab	0.05	0.72
serine/arginine repetitivematrix protein 2	S3Y–AA	Host coral	HBH > HWN	stab > var	NS	0.49	3.09**
serine/arginine-richsplicing factor 7	S3BB–DD	Host coral	HBH > HWN	stab > var	NS	0.20	1.56
transformer-2 protein homologue beta-like	S3EE–GG	Host coral	stab > var	stab > var	NS	0.21	1.65
a priori-selected genes [n = 14; Fig. S4 (host coral) and Fig. S5 (*Symbiodinium*), Supporting information]							
heat-shock protein 70†	S4A–B	Host coral	stab > var	NS	stab > var	0.66	4.40**
β-actin†	S4C–D	Host coral	NS	NS	var > stab	0.32	2.15
α-tubulin†	S4E–F	Host coral	NS	NS	NS	0.10	1.07
tropomyosin†	S4G–H	Host coral	NS	NS	NS	0.07	0.84
ezrin†	S4I–J	Host coral	NS	NS	NS	0.25	1.82
phospholipase-α2†	S4K–L	Host coral	NS	NS	HWN-var > HBH var	0.36	2.38*
transient receptorcation channel†	S4M–N	Host coral	NS^d	NS	NS	0.00	−0.09
organic anion transportprotein†	S4O–P	Host coral	NS^f	NS	NS	0.02	0.63
ribulose-1,5-bisphosphate carboxylase/oxygenase‡	S5A–B	*Symbiodinium*	NS	NS	var > stab	0.01	−0.37
photosystem I (subunit III)‡	S5C–D	*Symbiodinium*	NS	NS	var > stab	0.05	0.74
phosphoglycolate phosphatase‡	S5E–F	*Symbiodinium*	NS	NS	var > stab	0.01	−0.38
ascorbate peroxidase‡	S5G–H	*Symbiodinium*	NS	NS	NS	0.02	0.44
nitrate transporter 2†	S5I–J	*Symbiodinium*	HWN > HBH	NS	NS	0.04	0.61
heat-shock protein 70†	S5K–L	*Symbiodinium*	NS	NS	var > stab	0.08	−0.91

Stab, stable temperature treatment; var, variable temperature treatment; HWN, Houwan; HBH, Houbihu; NS, not significant.

†Data from Mayfield *et al.* (2014b).

‡Data from Mayfield *et al.* (2012a).

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

(Table 1). Furthermore, there were significant, positive, linear correlations between RNA-Seq- and qPCR-derived data for all four genes (Fig. S2, Supporting information; avg. $r^2 = 0.68 \pm 0.04$).

Of the 107 mRNAs expressed only by samples of Houwan, 23, 13, 11 and 3 could be confidently assigned to be host coral, *Symbiodinium*, bacterial and IH origin, respectively, based on querying of online bioinformatics resources (Fig. 4D-1). Four of these 23 coral genes and six of these 13 *Symbiodinium* genes could be confidently ascribed a protein identity based on querying the PFAM and KEGG databases (Fig. 4D-2). Of the four host coral proteins (Fig. 4D-3), two, iGLON5 and immunoglobulin (Ig) domain, are involved in cell adhesion, while sacsin and lamina-associated polypeptide 1C are involved in protein homeostasis/QC and nuclear envelope structure/formation, respectively. qPCR assays were designed for iGLON5 (Fig. S1M–O, Supporting information) and sacsin (Fig. S1P–R, Supporting information), and neither gene was characterized by a significant correlation between RNA-Seq and qPCR, nor did qPCR and RNA-Seq yield congruent results.

Of the six *Symbiodinium* genes expressed only by samples from Houwan (Fig. 4D-4), two (33%) possessed domains known to be present in proteins that allow for the invasion of host cells. One contig possessed a bacterial-like Toll/interleukin receptor (TIR) domain, and the second possessed a bacterial-like Ig domain. Of the remaining identifiable proteins, one, a CAF1 family ribonuclease, is involved in mRNA processing, as were a large number of the *Symbiodinium* and host coral DEPs (Mayfield *et al.* 2016). A significant, positive correlation was documented between RNA-Seq- and qPCR-derived data for the lone gene involved in protein homeostasis/QC: an OTU domain-containing cysteine protease (Fig. S2M–O, Supporting information). The final two annotatable *Symbiodinium* genes expressed only by samples of Houwan were involved in metabolism: alanine ligase and a vacuolar ATP synthase (subunit B); qPCR primers were designed for the latter (Fig. S2P–R, Supporting information), and no correlation was found between the RNA-Seq- and qPCR-derived data. Of the 12 DEGs identified by RNA-Seq for which qPCR assays were created, qPCR verified RNA-Seq-derived findings for five of six *Symbiodinium* and three of five host coral genes; this difference in percentages between compartments was not statistically significant (two-sample proportion test, $P > 0.05$).

DEPs identified by 2D+MS

The top hit sequence for each of the 117 DEPs identified by Mascot was BLASTed against the SHVTS and 'P. damicornis' long-term temperature experiment (PDLTTE; Mayfield *et al.* 2014d) transcriptomes to search for homologues, and 70 and 58, respectively, were identified (Table S3, Supporting information); 11 and 34 mRNAs, respectively, whose respective proteins were found to be differentially expressed were characterized by site of origin, temperature treatment or interaction effects in the SHVTS and by temperature, time or temperature × time interaction effects in the PDLTTE. The expression of only one of these 11 SHVTS mRNAs showed a congruent response to that of the respective protein: a host coral transformer-2 protein homologue beta-like protein known to function in mRNA processing and splicing was expressed at higher mRNA and protein levels in samples of the stable temperature treatment (Table S3, Supporting information).

The MS data associated with the 10 sequenced protein spots were compared directly with the SHVTS transcriptome with MS-SCAN, and the five longest MS-derived peptide reads for each of the 10 protein spots were chosen for detailed analysis of the expression of the respective mRNAs (Table S4, Supporting information). Of these 50 proteins, there were significantly more host coral genes queried ($n = 22$; 44%) than *Symbiodinium* ($n = 11$; 22%; two-sample proportion test, $Z = 2.34$, $P < 0.05$); this equates to 97 and 53 host coral and *Symbiodinium* DEPs across both Mascot (75 and 42, respectively) and MS-SCAN-based approaches (1.8:1 host coral/*Symbiodinium* protein ratio).

Only one of the 50 MS-SCAN-identified proteins (2%) demonstrated congruency between 2D + MS and RNA-Seq (Table S4, Supporting information); this host coral chromodomain helicase DNA-binding protein (spot 1) was expressed at higher levels by corals of the stable temperature treatment at both the mRNA and protein levels, although qPCR did not validate this finding (Fig. S3AA–CC, Supporting information). qPCR also failed to validate protein-level expression differences for the remaining 10, Mascot-derived mRNAs encoding DEPs (Fig. S3D–GG, Supporting information and Table 1), including the aforementioned transformer-2 protein homologue beta-like protein found to be expressed at higher mRNA and protein levels by RNA-Seq and 2D + MS, respectively, in samples of the stable temperature treatment (Fig. S3EE–GG, Supporting information). Furthermore, there was a statistically significant correlation between RNA-Seq- and qPCR-derived data for only one of these 11 genes: serine/arginine (s/a) repetitive matrix protein 2 (Fig. S3AA, Supporting information).

DEG vs. DEP congruency

Of the eight host coral and six *Symbiodinium* genes hypothesized a priori to be differentially expressed

between temperature regimes (Figs S4 and S5, respectively, Supporting information), only one, a host coral heat-shock protein 70 (*hsp70*; Fig. S4A,B, Supporting information), was found to be differentially expressed between temperature profiles by RNA-Seq (Table 1). In contrast, qPCR found two host coral [*hsp70* and β-actin (*actb*; Fig. S4C,D, Supporting information)] and four *Symbiodinium* [ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbcL*; Fig. S5A,B, Supporting information), photosystem I (subunit III; *psI*; Fig. S5C,D, Supporting information), phosphoglycolate phosphatase (*pgpase*; Fig. S5E,F, Supporting information) and *hsp70* (Fig. S5K–L, Supporting information)] genes to be differentially expressed. None of the respective 14 proteins were sequenced in over-expressed protein spots removed from the stable temperature treatment 2D gel.

Of the 18 genes/proteins documented to be differentially expressed between temperature treatments by any technique out of the 37-molecule subset (Fig. 5A), there was the least amount of overlap between RNA-Seq and qPCR [one gene only (host *hsp70*)]. When looking at site of origin effects (Fig. 5B), RNA-Seq and qPCR tended to yield congruent results (41% of the 22 genes). When considering 'double negative' findings (i.e. no effects documented by either technique) in addition to significant site of origin and temperature treatment effects (Fig. 5C), qPCR and RNA-Seq agreed for 15 of the 37 genes (41%). However, the average r^2 across all 37 genes was only 0.24 ± 0.27, and only 11 genes (30%) were characterized by statistically significant correlations between these two techniques.

As mentioned above, transformer-2 protein homologue beta and chromatin helicase DNA-binding protein were the only two proteins out of the 167 DEPs identified (<2%) that were associated with mRNAs that showed a significant temperature treatment effect (Fig. 5A). Congruency between 2D + MS and RNA-Seq was documented for 14 molecules out of the 37 total (38%; Fig. 5C): the two aforementioned and 12 double negatives. When looking at congruency between 2D + MS and qPCR, 11 genes/proteins yielded congruent results: all double negatives (Fig. 5C and Table 1). RNA-Seq, qPCR and 2D + MS yielded congruent results for six molecules (Fig. 5C): one *Symbiodinium* gene/protein (ascorbate peroxidase [*apx1*]) and five host coral genes/proteins [tropomyosin (*trp1*), α-tubulin (*tuba*), ezrin, organic anion transporter (*oatp*) and transient receptor cation channel (*trcc*)]; this compartmental difference was not statistically significant, and all six genes/proteins represent 'triple negatives' (i.e. no site of origin, temperature treatment or interaction effect documented by any of the three techniques). No statistically significant site
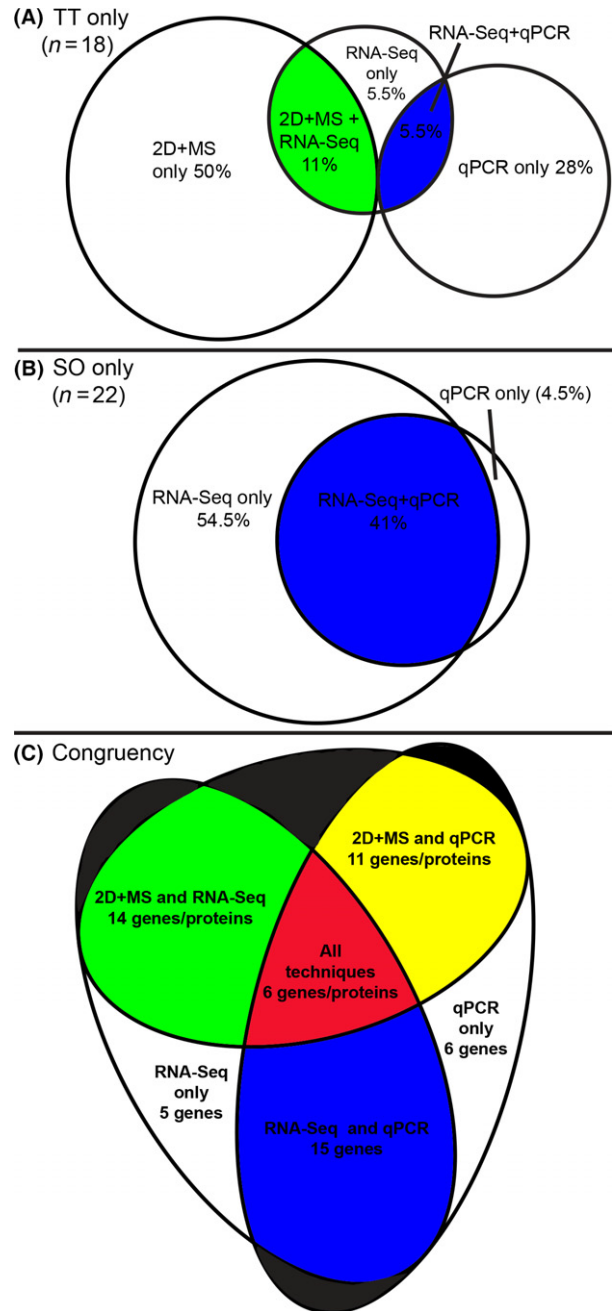
Fig. 5 Venn diagrams showing congruency across the three techniques utilized in this study. When two ellipses overlap, the respective techniques yielded a congruent result. The 18 genes/proteins that were affected by temperature treatment (TT) alone in the *Seriatopora hystrix* variable temperature study (SHVTS) were compared across all three techniques (A). The 22 genes that were affected by site of origin (SO) alone (B) were compared only between RNA-Seq and real-time PCR (qPCR), given that proteins were pooled across SO for each TT prior to electrophoresis. When the congruency between all three techniques was assessed (C), the numbers of genes/proteins, rather than the percentages, have been shown, although the areas of the ellipses nevertheless reflect the latter. [Colour figure can be viewed at wileyonlinelibrary.com].

of origin, temperature treatment or interaction effect was documented by all three techniques for the same gene/protein (Fig. 5A).

Discussion

The two previous efforts in which an mRNA vs. protein analysis within the same endosymbiotic anthozoan sample was conducted yielded contrasting results; Putnam et al. (2013) found no statistically significant correlation between Symbiodinium rbcL mRNA and RBCL protein expression within P. damicornis, whereas Mayfield et al. (2014c) did indeed observe such a positive, linear correlation between these two molecules in Symbiodinium populations residing within the common sea anemone Exaiptasia pulchella. Herein, except for the cases in which both or all techniques yielded negative findings for a gene/protein (i.e. no site of origin, temperature treatment or interaction effect), there was generally a low degree of congruency between mRNA and protein expression, and both mRNA-targeting techniques, RNA-Seq and qPCR, tended to support this finding. Specifically, not a single differentially expressed Symbiodinium mRNA was associated with a protein that also differed in expression between treatments out of the 53 identified Symbiodinium DEPs, and only two such host proteins out of the 97 coral DEPs met these criteria: a chromodomain helicase DNA-binding protein involved in transcription [the lone molecule (4.5%) demonstrating congruency between mRNA and protein expression in the 22-protein, host-only MS-SCAN DEP pool] and a transformer-2 protein homologue beta-like protein, which is involved in mRNA processing [the lone molecule (1.3%) demonstrating congruency between mRNA and protein expression in the 75-protein, host-only Mascot DEP pool]. In total, these two molecules were the only ones that were significantly affected by temperature regime and showed a congruent response between RNA-Seq and 2D + MS out of the 378 DEGs and 167 DEPs (0.4%), and neither RNA-Seq-derived finding was verified by qPCR.

In addition to the transformer-2 protein homologue beta-like protein, a large number of proteins in the DEP pool were hypothesized to function in mRNA processing and splicing (Mayfield et al. 2016), potentially implying that mRNA processing is a cellular process that is affected by variable temperature exposure in the S. hystrix–Symbiodinium holobiont; such may help in explaining why there was such low congruence between mRNA and protein expression. Indeed, many of the most differentially expressed proteins identified, caleosins and oleosins (LB-targeted proteins; Mayfield et al. 2016), were not even expressed at the mRNA level, and, across both compartments, the respective mRNAs of 47 DEPs (28%

of the 167-DEP total) were not expressed. In addition to alternative splicing and extensive mRNA processing, this may suggest that these proteins have long half-lives and therefore do not undergo high rates of translation or turnover.

A significant degree of mRNA processing may also explain the immense size of both the S. hystrix–Symbiodinium (118 000 contigs) and P. damicornis (236 000) transcriptomes. Although dinoflagellates do indeed possess large and complex genomes (Lin et al. 2015), it is doubtful whether ~60 000–100 000 genes exist in the genomes of these two corals (the other ~60 000–140 000 genes being from Symbiodinium). It is more likely, based on the Acropora digitifera (Shinzato et al. 2011) and Symbiodinium kawaguti genomes, respectively, that approximately half as many genes, at most, should be found in each compartment; this suggests a significant degree of alternative splicing that may explain, in part, the lack of congruency between gene and protein expression, as well as between RNA-Seq and qPCR in the ~60% of instances in which the latter did not correlate positively and significantly with the former. qPCR amplicons typically only span a portion of the length of the target gene, so it is possible that the primers bound to and amplified different splice variants from the ones for which the RNA-Seq data were obtained.

It appears that S. hystrix is approximately 65% animal and 35% microbial from an mRNA perspective. This 1.8:1 mRNA breakdown, which was identical at the protein-level (97 host DEPs/53 Symbiodinium DEPs = 1.8), is similar to the confamilial pocilloporid P. damicornis (60/40%; Mayfield et al. 2014d) and reemphasizes the notion that it is important to consider both eukaryotic compartments when attempting to develop a molecular model by which a coral holobiont acclimates to environmental change. When creating such schematics of acclimation or stress for pocilloporid corals, and likely other marine invertebrates and endosymbioses, as well, it will also be important to focus not only on gene mRNAs, but also on the respective proteins; only the latter partake in cellular work, and, as suggested herein and in studies of humans and bacteria (Lu et al. 2007; Fagerberg et al. 2014), there is not always a strong, significant correlation or degree of congruency between expression of these two molecules. Only by characterizing protein behaviour can the cellular biology underlying the physiological response of corals to the rapidly changing conditions to which they are being subjected ultimately be elucidated.

Acknowledgements

References

Barshis DJ, Ladner JT, Oliver TA *et al.* (2013) Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 1387–1392.

Chen WNU, Kang HJ, Weis VM, Mayfield AB, Fang LS, Chen CS (2012) Diel rhythmicity of lipid body formation in a coral-*Symbiodinium* endosymbiosis. *Coral Reefs*, **31**, 521–534.

Chen HK, Song SN, Wang LH *et al.* (2015) A compartmental comparison of major lipid species in a coral-*Symbiodinium* endosymbiosis: evidence that the coral host regulates lipogenesis of its cytosolic lipid bodies. *PLoS One*, e0132519.

Cox MP, Peterson DA, Biggs PJ (2010) SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics*, **11**, 485.

Fagerberg L, Hallstrom BM, Oksvold P *et al.* (2014) Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular and Cellular Proteomics*, **13**, 397–406.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA- Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.

Greensbaum D, Colangelo C, Williams K, Gerstein M (2003) Comparing protein abundance and mRNA expression on a genomic scale. *Genome Biology*, **4**, 117.

Hoegh-Guldberg O, Mumby PJ, Hooten AJ *et al.* (2007) Coral reefs under rapid climate change and ocean acidification. *Science*, **318**, 1737–1742.

Huang YCA, Hsieh HJ, Meng PJ *et al.* (2011) Nutrient enrichment caused by marine cage culture and its influence on subtropical coral communities in turbid waters. *Marine Ecology Progress Series*, **423**, 83–93.

Jan S, Chen CTA (2008) Potential biogeochemical effects from vigorous internal tides generated in Luzon Strait: a case study at the southernmost coast of Taiwan. *Journal of Geophysical Research*, **114**, 1–14.

Kim S, Pevzner PA (2014) MS-GF+ makes progress towards a universal database search tool for proteomics. *Nature Communications*, **5**, 5277.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.

Lin S, Cheng S, Song B *et al.* (2015) The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science*, **350**, 691–694.

Loya Y, Sakai K, Yamazato K, Nakano Y, Sambali H, van Woesik R (2001) Coral bleaching: the winners and the losers. *Ecology Letters*, **4**, 122–131.

Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnology*, **25**, 117–124.

Mayfield AB, Gates RD (2007) Osmoregulation in anthozoan-dinoflagellate symbiosis. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, **147A**, 1–10.

Mayfield AB, Hsiao YY, Fan TY, Chen CS, Gates RD (2010) Evaluating the temporal stability of stress-activated protein kinase and cytoskeleton gene expression in the Pacific corals *Pocillopora damicornis* and *Seriatopora hystrix*. *Journal of Experimental Marine Biology and Ecology*, **395**, 215–222.

Mayfield AB, Wang LH, Tang PC *et al.* (2011) Assessing the impacts of experimentally elevated temperature on the biological composition and molecular chaperone gene expression of a reef coral. *PLoS One*, e26529.

Mayfield AB, Chan PS, Putnam HM, Chen CS, Fan TY (2012a) The effects of a variable temperature regime on the physiology of the reef-building coral *Seriatopora hystrix*: results from a laboratory-based reciprocal transplant. *The Journal of Experimental Biology*, **215**, 4183–4195.

Mayfield AB, Hsiao YY, Fan TY, Chen CS (2012b) Temporal variation in RNA/DNA and protein/DNA ratios in four anthozoan-dinoflagellate endosymbioses of the Indo-Pacific: implications for molecular diagnostics. *Platax*, **9**, 1–24.

Mayfield AB, Chen M, Meng PJ *et al.* (2013a) The physiological response of the reef coral *Pocillopora damicornis* to elevated temperature: results from coral reef mesocosm experiments in Southern Taiwan. *Marine Environmental Research*, **86**, 1–11.

Mayfield AB, Fan TY, Chen CS (2013b) Physiological acclimation to elevated temperature in a reef-building coral from an upwelling environment. *Coral Reefs*, **32**, 909–921.

Mayfield AB, Fan TY, Chen CS (2013c) The physiological impact of *ex situ* transplantation on the Taiwanese reef-building coral *Seriatopora hystrix*. *Journal of Marine Biology*, Article ID 569369.

Mayfield AB, Fan TY, Chen CS (2013d) Real-time PCR-based gene expression analysis in the model reef-building coral *Pocillopora damicornis*: insight from a salinity stress study. *Platax*, **10**, 1–29.

Mayfield AB, Chen CS, Liu PJ (2014a) Decreased green fluorescent protein-like chromoprotein gene expression in specimens of the reef-building coral *Pocillopora damicornis* undergoing high temperature-induced bleaching. *Platax*, **11**, 1–23.

Mayfield AB, Chen YH, Dai CF, Chen CS (2014b) The effects of temperature on gene expression in the Indo-Pacific reef-building coral *Seriatopora hystrix*: insight from aquarium

studies in Southern Taiwan. *International Journal of Marine Science*, **4**, 1–23.

Mayfield AB, Hsiao YY, Chen HK, Chen CS (2014c) Rubisco expression in the dinoflagellate *Symbiodinium* sp. is influenced by both photoperiod and endosymbiotic lifestyle. *Marine Biotechnology*, **16**, 371–384.

Mayfield AB, Wang YB, Chen CS, Chen SH, Lin CY (2014d) Compartment-specific transcriptomics in a reef-building coral exposed to elevated temperatures. *Molecular Ecology*, **23**, 5816–5830.

Mayfield AB, Bruckner AW, Chen CH, Chen CS (2015) A survey of pocilloporids and their endosymbiotic dinoflagellate communities in the Austral and Cook Islands of the South Pacific. *Platax*, **12**, 1–17.

Mayfield AB, Chen YJ, Lu CY, Chen CS (2016) Proteins involved in acclimation to variable temperature in the reef-building coral *Seriatopora hystrix*. In: *Coral Reefs: Ecosystems, Environmental Impact and Current Threats* (ed. Ortiz S), pp. 1–72. NOVA Publishing, New York.

Palumbi SR, Barshis DJ, Traylor-Knowles N, Bay RA (2014) Mechanisms of reef coral resistance to future climate change. *Science*, **334**, 895–898.

Peng SE, Chen WNU, Chen HK *et al.* (2011) Lipid bodies in coral-dinoflagellate endosymbiosis: ultrastructural and proteomic analyses. *Proteomics*, **17**, 3540–3555.

Putnam HM, Mayfield AB, Fan TY, Chen CS, Gates RD (2013) The physiological and molecular responses of larvae from the reef-building coral *Pocillopora damicornis* exposed to near-future increases in temperature and *p*CO2. *Marine Biology*, **160**, 2157–2173.

Shinzato C, Shoguci E, Kawashima T *et al.* (2011) Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature*, **476**, 320–323.

Data accessibility

The mRNA sequences, conceptually translated protein sequences, MS-derived protein sequences, raw MS data and mRNA expression data are all openly available on the following website: http://symbiont.iis.sinica.edu.tw/s_hystrix/static/html/#stat. Additionally, the mRNA sequences for all 12 samples, as well as a tabulated spreadsheet including their respective expression levels, have been deposited at Dryad (doi:10.5061/dryad.7cv4g). The mRNA sequences have been uploaded to NCBI's BioProject database (PRJNA310803). Finally, images of the experimental setup and the sequenced nubbins can be found at http://coralreefdiagnostics.com under the "Past Projects-SHVTS" subheading.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Differentially expressed host coral genes identified by RNA-Seq.

Fig. S2 Differentially expressed *Symbiodinium* genes identified by RNA-Seq.

Fig. S3 mRNA expression of differentially expressed host coral proteins identified by 2D + MS.

Fig. S4 a priori-selected host coral genes.

Fig. S5 a priori-selected *Symbiodinium* genes.

Table S1 Real-time PCR assays.

Table S2 2D gel electrophoresis protein spot information and MS-SCAN analysis.

Table S3 Querying of two transcriptomes with 117 differentially expressed coral proteins.

Table S4 A comparison of the five longest mass spectrometry (MS)-derived peptides from each of 10 protein spots against the 'Seriatopora hystrix variable temperature study' (SHVTS) transcriptome via MS-SCAN.

Appendix S1 Supplemental methods.