Transcriptomic profiling of differential responses to drought in two freshwater mussel species, the giant floater *Pyganodon grandis* and the pondhorn *Uniomerus tetralasmus*

by

Yupeng Luo

A Thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
December 14, 2013

Keywords: Unionidae; freshwater mussel; heat stress; drought; RNA-seq

Approved by

Eric Peatman, Chair, Assistant Professor, School of Fisheries, Aquaculture, and Aquatic Sciences
James Stoeckel, Associate Professor, School of Fisheries, Aquaculture, and Aquatic Sciences
Yolanda Brady, Associate Professor, School of Fisheries, Aquaculture, and Aquatic Sciences

Abstract

The southeastern US has experienced recurrent drought during recent decades. Increasing demand for water, as precipitation decreases, exacerbates stress on the aquatic biota of the Southeast: a global hotspot for freshwater mussel, crayfish, and fish diversity. Freshwater unionid mussels are ideal candidates to study linkages between ecophysiological and behavioral responses to drought. Previous work on co-occurring mussel species suggests a coupling of physiology and behavior along a gradient ranging from intolerant species such as *Pyganodon grandis* (giant floater) that track receding waters and rarely burrow in the substrates to tolerant species such as *Uniomerus tetralasmus* (pondhorn) that rarely track receding waters, but readily burrow into the drying sediments. We utilized a next-generation sequencing-based RNA-seq approach to examine heat/desiccation-induced transcriptomic profiles of these two species in order to identify linkages between patterns of gene expression, physiology and behavior. Sequencing produced over 425 million 100 bp reads. Using the *de novo* assembly package Trinity, we assembled the short reads into 321,250 contigs from giant floater (average length 835 bp) and 385,735 contigs from pondhorn (average length 929 bp). BLAST-based annotation and gene expression analysis revealed 2,832 differentially expressed genes in giant floater and 2,758 differentially expressed genes in pondhorn. Trancriptomic responses included changes in molecular chaperones, oxidative stress profiles, cell cycling, energy metabolism, immunity, and cytoskeletal rearrangements. Comparative analyses between species indicated significantly higher induction of molecular chaperones and cytoskeletal elements in the intolerant *P. grandis* as well as important differences in genes regulating apoptosis and immunity.

Acknowledgments

I would like to express my very great appreciation to Dr. Eric Peatman for his constant encouragement, patient guidance and valuable suggestions. I would like to thank my committee members Dr. Jams Stoeckel and Dr. Yolanda Brady for their advice and assistance in all stages of this work. I offer special thanks to Dr. Chao Li for his direction and help without reservation in my research work. I would also like to thank Dr. Zhanjiang Liu, Milla Kaltenboeck and Dr. Huseyin Kucuktas for their help in providing me the resources to complete this project. Special thanks should be given to the faculty and staff of the School of Fisheries, Aquaculture and Aquatic Sciences, Auburn University and Shanghai Ocean University who provided the best exchange education program. Sincere thanks will also go to Ruijia Wang, Wilawan Thongda, Dr. Xingqiang Wang, Spencer Gowan and Honggang Zhao for their friendship and support. Last but not least, I can hardly finish my graduate study without my parents' (Dr. Zhigang Luo and Huimin Peng) continuous encouragement and selfless love, so nothing, but thank you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BAG4	BAG family molecular chaperone regulator 4
BNIP1	BCL2/adenovirus E1B interacting protein 1-like
CAPN5	Calpain 5
CRYAB	Alpha-crystallin B chain
DEF	Defensin
HSP70B2-1	Heat shock protein 70 B2,type1
HSP70B2-2	Heat shock protein 70 B2, type2
HSP90AA1	Heat shock protein HSP 90-alpha 1
HSPA12B	Heat shock 70 kDa protein 12B-like
HSPB1-1	Heat shock protein beta-1, type 1
HSPB1-3	Heat shock protein beta-1, type 3
IAP	Inhibitor of apoptosis protein
KLF5	Kruppel-like factor5
LSA-3	Liver stage antigen 3 precursor
NLRP1	NLR family, pyrin domain containing 1
TLR13	Toll-like receptor 13

I. Introduction

Background

Water resources are increasingly at the center of environmental and geopolitical conflicts across the globe. Burgeoning populations have led to large-scale urban development, land-use conversion, increased agricultural water usage, and greater industrial outputs. Deleterious alterations of aquatic ecosystems have closely followed these changes, including dam construction, channelization, and disrupted flow regimens. Coupled with climate change, these alterations have had severe, negative impacts on freshwater stream biota. These impacts are acutely seen in the southeastern United States. Severe, recurrent drought, rapid population growth, and increased agricultural usage have led to widespread hydrological disturbances [1,2]. Freshwater mussels (*Unioniformes*), endemically diverse in the region, have borne the full brunt of these changes, with 73% of mussel species currently considered either presumed extinct (13%), endangered (28%), threatened (14%) or of special concern(18%) [3]. As our understanding of mussel biology grows, so has our understanding of the importance of predictable stream flows, maintenance of tolerable temperatures, stable habitats, and host fish abundance for their welfare [2,4-6]. Unfortunately, each of these aspects has been altered over the last century. Going forward, effective conservation, restoration, and recovery of mussel populations in the southeast US will require careful monitoring and planning. Needed inputs into a recovery program include knowledge of population and community dynamics in the face of drought conditions [1].

Mussel Responses to Drought

The factors governing the impact of drought on mussels range from those relatively easy to measure (stream size) to those often imperceptible to the human eye (evolutionary adaptation). Recently Gough et al. [7] catalogued many of the abiotic aspects, which include drought severity, stream size, stream habitat composition, and water quality [8-11]. Less is known regarding species and population differences in physiological tolerance and behavioral strategies. However, recent work identified distinct guilds of mussels that are either thermally sensitive or thermally tolerant [5]. These differing tolerances may be resulting in shifts in mussel assemblage composition due to drought conditions [11]. The tolerance of a given mussel species to drought can be reasonably assumed to be governed by several factors including movement/migration patterns, response to desiccation and/or high temperatures, burrowing tendencies, and metabolic activity levels.

In a recent elegant study, Gough et al. [7] used a combination of field and laboratory experiments to examine several of these factors in co-occurring members of three different tribes of the family Unionidae. *Uniomerus tetralasmus* and *P. grandis* represented two known extremes of tolerance to emersion [12], with *U. tetralasmus* capable of living out of water at cool temperatures for up to 2 years, while *P. grandis* may die within a few weeks of desiccation. A third co-dominant species in the examined stream was *Lampsilis straminea*. The authors assessed physiological tolerance to survival at three temperatures, examined horizontal and vertical movement responses to receding water levels, and compared survival of the three species. The study results revealed that the species with the least physiological tolerance and highest mortality under drought conditions, *P. grandis*, could avoid desiccation by tracking the receding waters but rarely burrowed as water disappeared. The species with the greatest physiological tolerance and lowest mortality, *U. tetralasmus*, showed little evidence of water tracking but rapidly burrowed into the stream bottom as waters receded. The species with intermediate physiological tolerance, *L. straminea*, exhibited a more plastic

behavior pattern, initially avoiding desiccation by tracking the receding water and then tolerating desiccation via burrowing when the water disappeared. Differences in survival between the species were most pronounced in laboratory desiccation trials at 35 ℃, where 0% survival was observed at the end of week 1 in *P. grandis*, and 100% survival recorded for *U. tetralasmus* [7]. The extreme variation in physiological tolerance to desiccation of *P. grandis* and *U. tetralasmus*, coupled with their differing behavioral strategies, represent an ideal model for examining the genetic underpinnings of drought tolerance in unionid mussels.

Expression Studies in Marine Shellfish

To-date, studies examining linkages between genetics, physiology, behavior and environmental stresses, survivorship, and reproductive fitness have been limited for freshwater mussels [13-15]. Assuming that different behavioral responses are tightly linked to differential biochemical pathways elicited by drought, the identification of underlying genetic mechanisms regulating these behavioral and physiological differences would provide key insights into adaptive responses of freshwater mussels to heat stress and drought. Previous studies in marine shellfish species have explored connections between variations in gene and protein expression and differences in latitudinal adaption, differential success of native and invasive species, and resistance to summer mortality in the context of heat stress. Meistertzheim et al. [16] identified specific up- and down-regulated genes in *Crassostrea gigas* after prolonged thermal stress using suppression subtractive hybridization. Candidate genes for families of *C. gigas* selectively bred for heat tolerance have been identified by a cDNA microarray [17]. Many genes such as heat shock protein 27, collagen, peroxinectin, S-crystallin, had higher expression levels in low-surviving families than high-surviving families. Meanwhile, expression level of a putative cystatin B gene was higher in high-surviving families. Lockwood et al. developed an oligonucleotide microarray with 8,874 probes representing 4,488 different genes from invasive and native blue mussels of the genus

Mytilus [18]. Under drought stress, 1,513 genes showed temperature-dependent changes, and 96 genes were related to a putative species-specific response. In the Mediterranean blue mussel *Mytilus galloprovincialis* and native blue mussel *Mytilus trossulus* following acute heat stress (24 ℃, 28 ℃ and 32 ℃) for 1 h and then a return to 13 ℃ to recover for 24 h, 47 and 61 proteins were found to be dramatically changed respectively [19]. Fields et al. observed latitudinal variation after acute heat stress (40 ℃) in the salt marsh mussel *Geukensia demissa* using two-dimensional gel electrophoresis [20]. They found that protein expression among mussels from different locations varied substantially, with 31 of 448 proteins changing in abundance in the northernmost (Maine) group, compared with 5–11 proteins in the four southern groups [20]. Another recent study in *M. galloprovincialis* recently examined gene expression responses to the dual stressors of high temperature and copper [21].

Next-Generation Sequencing Studies

While studies such as those cited above have traditionally required time-consuming and expensive generation of molecular resources and have, therefore, been limited to a handful of key species, new ribonucleic acid sequencing (RNA-seq) approaches have dramatically expanded our ability to carry out transcriptome profiling in non-model species [22]. RNA-seq depends on next-generation sequencing to carry out massively parallel sequencing of short mRNA reads representing transcripts of expressed genes. Bioinformatic assembly of these reads into longer gene transcripts (contigs) allows for rapid and affordable development of transcriptome and molecular marker resources in historically understudied species groups such as unionid mussels [23]. In recent years, RNA-seq approaches have been employed in molluscan non-model species to explore trait-related biological questions (Table 1).

Bai et al. [24] utilized 454 to capture the transcriptome of tissues secreting purple and white nacre in the pearl mussel *Hyriopsis cumingii*. They obtained 541,269 sequences (298 bp

average size) and 440,034 sequences (293 bp average size) in tissues secreting purple and white nacre, respectively. After assembly of these two libraries together, they obtained 47,812 contigs (634 bp average size) and 22,495 annotated genes, in which 33 genes were

Table 1. Transcriptomic profiling in mussel and oyster using RNA-seq in recent years.

	Bai et al., (2012)	Gavery et al., (2012)	Wang et al., (2012)	Zhao et al., (2012)
Species	*H. cumingii*	*C. gigas*	*V. lienosa*	*P. martensii*
Platform	454 Sequencing	SOLiD	Illumina	Illumina
Reads	981,303	~45,300,000	~162,000,000	39,400,004
Assembled Contigs	47,812	18,510	778,234	102,762
Annotated unigenes	22,495	7,724	23,742	37,188
DE genes	358	2,991	1,934	NA
Extension	novel transcript, SSR, SNP, etc.	novel transcripts, etc.	SSR, etc.	novel transcript, etc.

involved in pearl or shell formation. During gene expression analysis, a total of 358 genes were identified as the differentially expressed. A set of SSR and SNPs between these two different phenotypes was also captured from the RNA-seq data.

Gavery and Roberts [25] conducted RNA-seq in the Pacific oyster (*C. gigas*) from two different sites using SOLiD3 System. The *de novo* assembly generated 18,510 sequences (276 bp average size) from 45.3 million high quality reads. A total of 7,724 sequences could be annotated by BLAST searching against the UniProtKB/Swiss-Prot database. A total of 2,991 genes were identified as differentially expressed between the different locations ($p<0.05$ and absolute fold change >2).

Wang et al. [23] utilized a RNA-seq-based approach to develop molecular resources for *Villosa lienosa* using the Illumina HiSeq 2000 platform. A total of 778,234 contigs (average length=707.5 bp) were assembled from 162 million filtered reads. They identified 23,742 unigene hits against the non-redundant (nr) database and 36,582 microsatellites with sufficient flanking region for primer design. Following a heat shock exposure, a total of 1,934 contigs showed significant differential expression ($p < 0.05$, mapped reads >5 and absolute fold change >2).

Zhao et al. [26] utilized RNA-seq to identify genes that are potentially related to biomineralization in the pearl oyster *Pinctada martensii* using Illumina HiSeq 2000. In that study, a total of 39,400,004 reads were generated and assembled into 102,762 contigs. Of these contigs, 37,188 unigenes were annotated based on the Swiss-Prot databases using BlastX. They also identified 51 biomineralization-related unigenes and 268 immune-related unigenes that were not previously detected in *P. martensii*.

Heat shock Protein Responses in Shellfish

While studies of transcriptional responses to heat stress in shellfish have varied in their findings, and a broad array of pathways clearly play a role in heat stress responses, a central component in each case has been alterations in the expression of heat shock protein genes. Heat shock proteins (HSPs) are a kind of molecular chaperone with varying molecular weight (*c.* 16-100 kDa) [27] which can be classified into five major conserved families: HSP60, HSP70, HSP90, HSP100 and small heat shock proteins (sHSP). HSPs are also known as stress proteins and extrinsic chaperones [27], which maybe a key part of the heat shock response, primarily induced by heat shock factor (HSF) [28], and are recognized as key components contributing to cellular homeostasis in cells under both optimal and adverse growth conditions [29]. Important housekeeping functions of molecular chaperones discovered to date include (1) import of proteins into cellular compartments; (2) folding of proteins in the cytosol, endoplasmic reticulum and mitochondria; (3) degradation of unstable proteins; (4) dissolution of protein complexes; (5) prevention of protein aggregation; (6) control of regulatory proteins and; (7) and refolding of misfolded proteins [30]. HSP60 is expressed in the mitochondria and predominantly plays a role in refolding proteins, preventing aggregation of denatured proteins and proapoptotic [31,32]. HSP70 genes are reported to be essential to thermo-tolerance [31], anti-apoptotic activity [33], protein folding [31] and denatured protein refolding [32]. HSP90 proteins play an important role in

preventing formation of aggregates of unfolded proteins [34] and misfolded proteins [35], regulation of steroid hormone receptors [36], protein translocation [31] and regulation in cell cycle and proliferation [37-42]. HSP100 genes constitute a poorly studied family with the proposed function of increased tolerance to high temperatures, promotion of proteolysis of specific cellular substrates and regulation of transcription [43]. Small HSPs, those less than 40 kDa in weight, suppress aggregation and heat inactivation of proteins, provide thermo-tolerance to protect cell from stress and also have anti-apoptotic activity through inhibition of caspase activation. Small HSPs also played multiple roles in stabilizing cytoskeletal elements (actin microfilaments, intermediate filaments and microtubules) and enhancing the cell's ability to combat oxidative stress [44-49]. The major stress pathways in which HSPs play a role are shown in Figure 1 (adapted from [35]).

Several studies have begun to examine the gene and protein expression patterns of HSPs in shellfish. HSP60 proteins were significantly up-regulated in cadmium-exposed gill cells in eastern oysters *Crassostrea virginica* [50] and in *Mytilus edulis* gills after heat stress, and in marine mussels gill and mantle cells following exposure to copper [51-54], indicating the induction of HSP60 may also be related to metal stress. But during the acute heat stress in gill and hepatopancreas tissues in eastern oysters, *C. virginica* [50], HSP60 and HSP90 were not consistently up-regulated by either acute heat or cadmium exposure. HSP60 may not play a key role in the direct response to heat stress, but may coordinate with other HSP (e.g. HSP70 and sHSP) to contribute to shellfish survival under heat/drought conditions.

HSP70 have showed highly correlative responses to heat stress in many shellfish species from the studies listed below. The recent completion of the *C. gigas* [55] genome revealed a startling 88 HSP70 genes. Five HSP70 genes were observed ~2,000-fold up-regulated after heat stress(25 ℃ for 7 d, 30 ℃ and 35 ℃ for 12 h). Moreover, all HSP70 genes were ~13.9-fold up-regulated in average [55]. In *M. edulis*, HSP70 expression was also significantly

induced when exposed to higher seawater temperatures [56]. In *M. galloprovincialis* and *M. trossulus*, 7 HSP70 genes were significantly differentially induced at all heat-shock temperatures (24 ℃, 28 ℃ and 32 ℃) [18]. In addition, the expression of HSP70 and HSP90

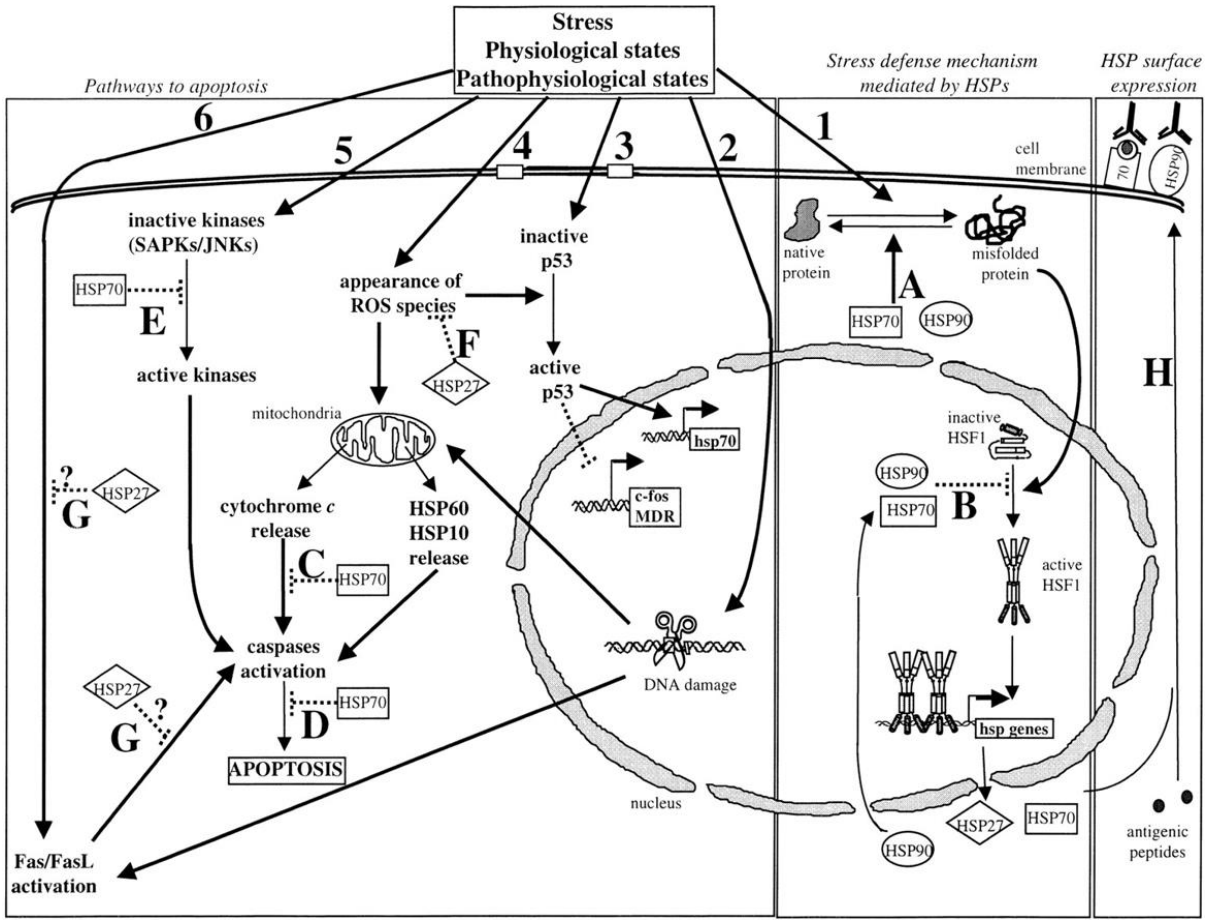

Figure 1. Schematic of the major cellular pathways activated by stress [adapted from 35].

in the digestive glands of horse bearded mussels, *M. barbatus*, went up consistent with our previous understanding of the heat shock response in mussel [57-60]. In the recent heat stress study of *V. lienosa* [23], the first large-scale transcriptome project in freshwater mussels, HSP70, heat shock cognate 70 protein-interacting protein (HSC70IP) and HSP90 were significantly up-regulated after heat stress (29 ℃ for 6 d) in the gene expression analysis based on RNA-seq.

In several studies on two different species with differing temperature tolerances, *M. galloprovincialis* and *M. trossulus,* HSP24 showed higher expression in *M. galloprovincialis,* the more thermo-tolerant species, than in *M. trossulus*, the more heat sensitive species [18,19]. This suggested that sHSP may serve as an important component in protecting cells from acute heat stress and in indicating interspecific differences in thermal tolerance [53,59,61-63]. In *V. lienosa* [23], the expression level of HSP21.4, HSP25 and HSP40 from th sHSP family increased significantly after acute heat stress (29 ℃ for 6 d). To date, no study has reported expression patterns of the HSP100 family in mussel and oyster. Additionally, all the aforementioned studies have focused on heat stress responses in shellfish that were submerged in water rather than those exposed and subject to the additional stressors of desiccation, as often seen in freshwater mussels stranded by drought. The study covered in the next section (II), therefore, covers new ground, in examining broad transcriptional changes of two unionid species with known divergent and behavioral and physiological responses to heat and desiccation.

II. Transcriptomic profiling of differential responses to drought in two freshwater mussel species, the giant floater *Pyganodon grandis* and the pondhorn *Uniomerus tetralasmus*

Introduction

As drought conditions recur with increasing frequency and severity in the Southeastern U.S., sessile aquatic organisms in freshwater ecosystems, particularly in streams, are bearing the brunt of these environmental perturbances [7,10,11]. Freshwater unionid mussel populations are already among the most endangered groups of organisms in the world [64]. Growing efforts to document the stunning diversity and varied life history strategies of unionids are also recording the measurable impacts of altered stream flows and thermal profiles on survival, recruitment, reproductive strategies, and community structure of these species [5,7,11,65,66].

The ability of individual species to tolerate drought conditions depends on several factors including severity and duration of the disturbance, stream habitat (debris, pools, etc.), and differences in behavioral and physiological responses to emersion/desiccation and heat stress [7]. Indeed, a recent study utilizing field and laboratory experiments revealed links between behavioral responses, physiological tolerances, and survival in three co-existing mussel species (*U. tetralasmus* (pondhorn); *P. grandis* (giant floater); and *L. straminea* (fatmucket)). There, the authors observed that a burrowing response in pondhorn was correlated with a higher tolerance to desiccation and higher survival (77%) compared to the water-tracking behavior and low tolerance and survival (0%) of giant floater [7].

The identification of underlying genetic mechanisms regulating these behavioral and physiological differences would provide key insights into adaptive responses of freshwater

mussels to heat stress and drought. Previous studies in marine shellfish species have explored connections between variations in gene and protein expression and differences in latitudinal adaptation, differential success of native and invasive species, and resistance to summer mortality in the context of heat stress [16-20,67]. While such studies have traditionally required time-consuming and expensive generation of molecular resources and have, therefore, been limited to a handful of key species, new RNA-seq approaches have dramatically expanded our ability to carry out transcriptome profiling in non-model species [22]. Indeed, using RNA-seq we recently identified the components of a classical heat stress response in the unionid *V. lienosa* [23]. Of even greater interest than highly conserved stress factors and responses [68], however, may be the identification of divergent responses and gene components accounting, at least in part, for heightened drought tolerance in some species or populations. Recent sequencing of the Pacific oyster genome revealed a startling diversity of genes available for adaptation to environmental stress, including 88 heat shock protein 70 (HSP70) genes (compared to 17 in humans) and 48 inhibitor of apoptosis (IAP) genes (compared to 8 in humans). The finding not only highlights the importance of these gene families in molluscs but also emphasizes the need for global expression profiling of particular species of interest rather than examination of a conserved, narrow subset of heat shock genes based on mammalian paradigms. Here, we conducted RNA-seq-based transcriptome profiling on *P. grandis* and *U. tetralasmus* exposed to experimental heat stress/desiccation. Our main objectives were: (i) to compare the ability of Trinity and Trans-ABySS to assemble high quality transcriptomes for both species; (ii) identify key shared and divergent responses to drought in unionid mussels; and (iii) determine whether there were consistent differences between the tolerant and sensitive unionid species in both individual gene and pathway level responses to drought stress. The information gained here may serve as a foundation to guide natural resource managers in assessing the adaptive potential of

freshwater mussel species in the face of climate change and should aid in the development of molecular tools to monitor stress levels of mussels in drought-stricken streams.

Materials and Methods

Experimental animal and tissue sampling

Two species of freshwater mussel, *P. grandis* and *U. tetralasmus*, were utilized in experiments. *Pyganodon grandis* have been designated a species of lowest conservation concern, and *U. tetralasmus* a species of moderate conservation concern in Alabama. *Uniomerus tetralasmus* were collected from Opintlocco Creek, located in the Tallapoosa catchment of east central Alabama, with permission from the J. W. Huskey family and under scientific collection permits from the Alabama Department of Conservation and Natural Resources. *Pyganodon grandis* were produced in experimental ponds at the South Auburn Fisheries Research Station, Auburn University as part of previous experiments by the Crustacean and Molluscan Ecology Lab. Prior to experiments, all mussels were submerged in a large water bath and maintained at 21˚C. Pond water was used throughout the study so mussels could feed on natural food sources. All mussels were placed individually in 500 mL plastic cups within the water bath. Each cup was filled with sand and mussels were allowed to position themselves naturally in the substrate. After three days of acclimation at 21˚C, temperature was increased 3˚C/d over 4 d to 33˚C for the experimental mussels, while temperature (21˚C) and water level remained constant for the duration of the experiment for the control mussels. Upon reaching 33˚C, mussels were maintained at that temperature for 2 d. On day 3 at 33˚C, the water level in the water bath was lowered below the top of the cups and the water in each cup was drained.

At 24, 48, and 72 h after initiation of desiccation/emersion, non-lethal foot tissue samples were collected from 15 individuals per species using biopsy forceps. Mussels were randomly

assigned to sampling time groups. After the 72 h samples were collected, the recovery phase of the experiment began. The remaining experimental mussels were submerged and temperatures were dropped to 21˚C by 3˚C/d for 4 d. After one day at 21˚C, 15 mussels per species were sampled (5 d recovery timepoint). Tissues were flash frozen with liquid nitrogen immediately upon collection and stored in a -80°C freezer until RNA extraction.

RNA extraction, library construction and sequencing

From each timepoint/treatment group, three replicate pools of 5 individual samples/pool were constructed. Pooled tissue samples were ground to a fine powder in the presence of liquid nitrogen. RNA was extracted from ground tissue using an RNeasy Kit (Qiagen, Valencia, California) following the manufacturer's instruction. RNA concentration and integrity of each sample were measured on an RNA Nano Bioanalysis chip. Control groups used here were initial control group. The control group (submerged 21˚C) and 72h (desiccation 33˚C) experimental groups from each species were selected for RNA-seq library construction, Illumina sequencing and expression analysis. The expression level between initial control group and 72h control group was compared by real-time PCR and no significant difference was found between them.

RNA-seq library preparation and sequencing were carried out by HudsonAlpha Genomic Services Lab (Huntsville, AL, USA). cDNA libraries were constructed with 2.14-3.25 ug of starting total RNA and utilized the IlluminaTruSeq RNA Sample Preparation Kit (Illumina), as detailed in the TruSeq protocol. The libraries were amplified with 15 cycles of PCR and contained TruSeq indexes within the adaptors, specifically indexes 1-12. Finally, amplified library yields were 30 ul of 19.8-21.4 ng/ul with an average length of ~270 bp, indicating a concentration of 110-140 nM. After KAPA quantitation and dilution, the libraries were clustered 12 per lane and sequenced on an Illumina HiSeq 2000 instrument with 100 bp PE read chemistry.

De novo assembly of sequencing reads

The *de novo* assembly of short reads was performed on the mussel short reads using both ABySS and Trinity [69,70], versions 1.3.2 and the 2012-10-05 editions, respectively. Before assembly, raw reads were trimmed by removing adaptor sequences and ambiguous nucleotides. Reads with quality scores less than 20 and length below 30 bp were all trimmed. The resulting high-quality sequences were used in the subsequent assembly.

In ABySS, briefly, the clean reads were first hashed according to a predefined k-mer length, the 'k-mers'. After capturing overlaps of length k-1 between these k-mers, the short reads were assembled into contigs. The k-mer size was set from 50 to 96, assemblies from all k-mers were merged into one assembly by Trans-ABySS (version 1.4.4).

In Trinity, briefly, the raw reads were assembled into the unique sequences of transcripts in Inchworm via greedy k-mer extension (k-mer 25). After mapping of reads to Inchworm contigs, Chrysalis incorporated reads into de Bruijn graphs and the Butterfly module processed the individual graphs to generate full-length transcripts.

In order to reduce redundancy, the assembly results from different assemblers were passed to CD-Hit version 4.5.4 [71] and CAP3 [72] for multiple alignments and consensus building. The threshold was set as identity equal to 1 in CD-Hit, the minimal overlap length and identity equal to 100 bp and 99% in CAP3.

Gene annotation and ontology

All contigs in the resulting Trinity assembly were used as queries for BLASTX searches against the UniProtKB/SwissProt database and the National Center for Biotechnology Information (NCBI) non-redundant (nr) protein database. The cutoff Expect value (E-value) was set at 1e-5 and only the top hit result was assigned as the annotation for each contig. Gene ontology (GO) annotation analysis was processed using the UniProtKB/SwissProt results in Blast2GO (version 2.5.0), which is an automated tool for the assignment of gene

ontology terms. The final annotation file was generated after gene-ID mapping, GO term assignment, annotation augmentation and generic GO-Slim processes and categorized with regard to Biological Process, Cellular Component and Molecular Function.

Identification of differentially expressed contigs

Differentially expressed contig analysis was performed using the RNA-seq module and the expression analysis module in CLC Genomics Workbench. The high quality reads from each sample were mapped onto the Trinity reference assembly using CLC Genomics Workbench software. At least 95% of the bases were required to align to the reference and no more than two mismatches were allowed. The total mapped reads number for each transcript was determined and then normalized to detect RPKM (Reads Per Kilobase of exon model per Million mapped reads). After scaling normalization of the RPKM values, fold changes were calculated [73]. The proportions-based test was used to identify differentially expressed contigs between the control group and heat/desiccation group with three replications in each group with the threshold for significant differences set at a corrected p-value of $p<0.05$ [74]. Transcripts with absolute fold change values greater than 1.5 were defined as differentially expressed genes, which were divided into up- and down- regulated groups for further analysis.

Differentially expressed contigs were also used as queries for BLASTX searches against the *C. gigas* (Pacific oyster) protein database from NCBI with the cutoff Expect value (E-value) of 1e-5. Contigs from *P. grandis* and *U. tetralasmus* with shared annotation (nr and/or *C. gigas*) based on BLAST analysis were subjected to additional analyses to establish orthology including reciprocal BLAST, sequence alignment, and phylogenetic analysis. Differentially expressed contigs from a given species with no clear orthologous gene(s) in the other species were regarded here as unique.

Gene ontology and enrichment analysis

GO analysis and enrichment analysis between significantly expressed GO terms and overall reference assembly GO terms was processed to select overrepresented GO annotations in the differentially expressed genes using Ontologizer 2.0 using the Parent-Child-Intersection method with a Benjamini-Hochberg multiple testing correction[75,76]. GO terms for each gene were obtained by utilizing UniProtKB/SwissProt database annotations for the unigene set. The difference in the frequency of GO terms annotation in the differentially expressed genes sets were compared to the overall mussel reference assembly for each species. The threshold was set as FDR value < 0.1.

Functional groups and pathways encompassing the differentially expressed genes were identified based on GO analysis, pathway analysis based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, and manual literature review. Based on these analyses, key genes were organized into broad functional categories. Fold change differences between species for a given gene in each category were assessed for significance using a simple t-test on log-transformed RPKM values. Category-level differences in fold changes between species were assessed using a paired t-test.

Experimental validation—QPCR

Sixteen significantly expressed genes, including 7 shared genes, 4 unique genes of P. grandis and 5 unique genes of U. tetralasmus, with different expression patterns were selected for validation of the RNA-seq results using real time QPCR. Beta-actin was set as the reference gene in both species. Primers were designed based on contig sequences using Primer Premier 6 (Table 2). Tissue samples from the 72 h experimental group and the control

Table 2. Primers used for QPCR validation. Primers are listed in the 5' to 3' orientation.

Gene	Contig ID	Forward	Reverse
P. grandis			
Alpha-crystallin B chain	fl_ctg_1125	cgcagaatggacagaatg	ttacctccagttcttccg
BAG4	fl_ctg_2336	aacagcagtcagcgtctca	gttgtggtggtgtcattggt
BNIP1	fl_ctg_43	tgttagatgctgccttagtc	gaaccatagagaacgcctta
Calpain 5	fl_ctg_1914	acggaacacagagttatgc	gaggatgaagatgagccaat
Heat shock protein 70 B2,type1	fl_ctg_2493	gtggatgttgttgtgtctg	gctggaaggtcttgcttat
Heat shock protein 70 B2,type2	fl_ctg_1763	cctgtctctgtgaatcgtta	gaagaagtctcctcaatggt
Heat shock protein beta-1, type 1	fl_ctg_1863	gattgataggcacgctgat	gctggaagtaacggctaa
Heat shock protein beta-1, type 3	fl_ctg_1192	atcgtcactgaggctgat	tgagaggtatcggattgttc
Heat shock protein HSP 90-alpha 1	fl_ctg_2540	catcatcctcttctccttca	atcttgtcctcctcctctt
Kruppel-like factor 5	fl_ctg_1639	cgagaaagccaaacaagg	tgtcctcccacaacgaat
Liver stage antigen 3 precursor	fl_ctg_642	ggaactatcagcaggtagaa	cgaactccacagtcttaca
Beta actin		actctggtgatggtgtga	agcagtggttgtgaagga
U. tetralasmus			
Alpha-crystallin B chain	ph_ctg_1374	cgcagaatggacagaatg	ttacctccagttcttccg
BNIP1	ph_ctg_164	tctgaacctgactctccat	caagcaataccacctcctt
Calpain 5	ph_ctg_688	gaacagtgaacacatcttgg	ggctcatcatcatcctctg
Defensin	ph_ctg_2420	gtggtgcttgacgattatg	cgaggtggctattgtgata
Heat shock 70 kDa protein 12B-like	ph_ctg_1045	gtgttagtgcgttgtaggt	ctgttggagtgaggtattgt
Heat shock protein 70 B2, type2	ph_ctg_1662	gcagacacattgagaatacc	gacacagaccttcactacc
Heat shock protein beta-1, type 3	ph_ctg_1086	tccttccttgacagtgac	gccatgaacagattgactc
Heat shock protein HSP 90-alpha 1	ph_ctg_1044	acggagaagtgcttaacag	gatgacgaggacaagaaca
Inhibitor of apoptosis protein	ph_ctg_868	gactggtcacatggatagag	actgttctccgccttcat
kruppel-like factor 5	ph_ctg_1332	cgagaaagccaaacaagg	tgtcctcccacaacgaat
NLRP1	ph_ctg_1301	caggaaggtcaggttaatca	ggtatgtaacggagatgtca
Toll-like receptor 13	ph_ctg_1671	ccgcataatcctccgtatag	cctgacacatattccgacaa
Beta actin		tgtgctatgtggctcttg	gctgttgtaggttgtctca

Gene abbreviations are available in List of Abbreviations.

group were used as the template for QPCR. cDNA was synthesized using the iScript™ cDNA Synthesis Kit (Bio-Rad) according to manufacturer's protocol. The iScript chemistry uses a blend of oligo-dT and random hexamer primers. All the cDNA products were diluted to 250 ng/µl and utilized for the quantitative real-time PCR reactions using the SsoFast™ EvaGreen®Supermix on a CFX96 real-time PCR Detection System (Bio-Rad Laboratories, Hercules, CA). The thermal cycling profile consisted of an initial denaturation at 95 ℃ (for 30 s), followed by 40 cycles of denaturation at 94 ℃ (5 s), and an appropriate annealing/extension temperature (58 ℃, 5 s). An additional temperature ramping step was utilized to produce melting curves of the reaction from 65 ℃ to 95 ℃. Results were expressed relative to the expression levels of beta actin in each sample using the Relative Expression

Software Tool (REST) version 2009[77]. The biological replicate fluorescence intensities of the control and treatment products for each gene, as measured by crossing-point (Ct) values, were compared and converted to fold differences by the relative quantification method. Expression differences between groups were assessed for statistical significance using a randomization test in the REST software. The mRNA expression levels of all samples were normalized to the levels of beta actin gene in the same samples. Test amplifications were conducted with ensure that beta actin and target genes were within an acceptable range. A no-template control was run on all plates. QPCR analysis was repeated in triplicate runs (technical replicates) to confirm expression patterns.

Temporal expression analysis of key genes

A subset of key classical heat shock proteins and species-specific genes differentially expressed at 72 h were selected to examine their broader transcriptional profiles. Primers, reference gene, and reaction conditions were the same as described above. Relative expression profiles were captured for the experimental samples of each species at the 24h, 48h, and 5 d recovery timepoints. Data were analyzed as described above.

Results

Sequencing of short expressed reads from mussel foot tissue

Illumina-based RNA-sequencing (RNA-seq) was carried out on RNA extracted from replicated, pooled foot tissue samples from *P. grandis* and *U. tetralasmus* exposed to control and heat stress/desiccation experimental conditions. Reads from different samples were distinguished through the use of multiple identifier (MID) tags. A total of 427.5 million reads were generated on an Illumina HiSeq2000 instrument, including 201.1 million reads from *P. grandis*, and 226.4 million reads from *U. tetralasmus* (Table 3). At least 29 million reads

Table 3. Summary of Illumina expressed short reads production and filtering from *P. grandis* (A) and *U. tetralasmus* (B). Paired-end reads were generated on a HiSeq 2000 instrument.

A)

Sample	Reads (x 10^6)	Avg.length (bp)	Reads after trimming(x 10^6)	Percentage kept	Avg. length after trimming(bp)
Control 1	29.2	100	28.2	96.55%	94.2
Control 2	44.7	100	43.0	96.08%	94.0
Control 3	34.0	100	32.1	94.44%	93.5
Heat 1	31.2	100	30.0	95.91%	94.0
Heat 2	29.5	100	26.6	90.24%	90.6
Heat 3	32.5	100	31.0	95.49%	93.8
Total	201.1		190.9	94.79%	93.4

B)

Sample	Reads (x 10^6)	Avg.length (bp)	Reads after trimming(x 10^6)	Percentage kept	Avg. length after trimming(bp)
Control 1	38.8	100	37.1	95.40%	94.0
Control 2	38.7	100	37.2	96.09%	94.0
Control 3	45.1	100	43.2	95.87%	94.0
Heat 1	39.9	100	38.3	96.12%	94.2
Heat 2	34.6	100	31.2	90.43%	90.9
Heat 3	29.3	100	26.5	90.57%	91.1
Total	226.4		213.5	94.08%	93.0

were generated from each barcoded sample with an average of 35.6 million reads/library. Raw data was archived at the NCBI Sequence Read Archive (SRA) under Accession SRP026193.

De novo assembly of mussel transcriptomes

De novo assembly of RNA-seq reads can be achieved using several assembly algorithms and software programs. We had previously developed an in-house bioinformatics pipeline around Trans-ABySS and demonstrated superior performance relative to other assembly options in several species including freshwater mussel [23,78-80]. However, the recently developed Trinity software package provides several potential advantages for transcriptome

profiling in non-model species [70]. Therefore, here we sought to compare critical metrics of performance in transcriptome assembly between Tran-ABySS and Trinity.

Trans-ABySS

From *P. grandis* cleaned reads, Trans-ABySS generated 804,906 contigs including 178,956 contigs longer than 1000 bp. Average contig length was 772.6 bp while N50 was 801 bp. After filtration of redundant contigs by CD-Hit and CAP3, 56.6% of total contigs remained (455,659) with average length of 616.0 bp. The percentages of reads mapped in pairs and mapped to the final reference were 62.8% and 83.7%, respectively (Table 4).

From *U. tetralasmus* cleaned reads, Trans-ABySS generated 499,413 contigs including 147,046 contigs longer than 1000 bp. Average contig length was 963.0 bp while N50 was 1,503 bp. After filtration of redundant contigs by CD-Hit and CAP3, 52.4% of total contigs remained (261,489) with average length of 885.5 bp. The percentages of reads mapped in pairs and mapped to the final reference were 65.9% and 82.0%, respectively (Table 4).

Trinity

From *P. grandis* cleaned reads, Trinity generated 336,799 contigs including 66,959 contigs longer than 1,000 bp. Average contig length was 970.1 bp while N50 was 2,100 bp. After filtration of redundant contigs by CD-Hit and CAP3, 95.4% of total contigs remained (321,250) with average length of 835.0 bp. The percentages of reads mapped in pairs and mapped to the final reference were 79.2% and 84.5%, respectively (Table 4).

From *U. tetralasmus* cleaned reads, Trinity generated 405,996 contigs including 81,095 contigs longer than 1,000 bp. Average contig length was 968.1 bp while N50 was 2,233 bp. After filtration of redundant contigs by CD-Hit and CAP3, 95.0% of total contigs remained (385,735) with average length of 929.3 bp. The percentages of reads mapped in pairs and mapped to the final reference were 79.5% and 83.8%, respectively (Table 4).

Table 4. Summary of *de novo* assembly results of Illumina cleaned reads from *P. grandis* and *U. tetralasmus* foot tissue using Trans-ABySS and Trinity

	P. grandis		***U. tetralasmus***	
	Trans-ABySS	Trinity	Trans-ABySS	Trinity
Contigs	804,906	336,799	499,413	405,996
Large contigs (≥1000bp)	178,956	66,959	147,046	81,095
N50 (bp)	801	2,100	1,503	2,233
Average contig length	772.6	970.1	963.0	968.1
Contigs (After CD-HIT-EST+ CAP3)	455,659	321,250	261,489	385,735
Percentage contigs kept after redundant removing	56.6%	95.4%	52.4%	95.0%
Average length (bp) (After CD-HIT-EST+ CAP3)	616.0	835.0	885.5	929.3
Reads mapped in pairs (%)	62.8	79.2	65.9	79.5
Reads mapped to final reference (%)	83.7	84.5	82.0	83.8

Best assembly selection

Comparison of the assemblies generated by Tran-ABySS and Trinity (Table 4) made clear that, although Tran-ABySS generated a larger initial count of contigs, redundancy was much higher than observed with Trinity. CD-HIT/CAP3 trimmed over 40% of Trans-ABySS contigs compared with less than 5% of Trinity contigs. Additionally, Trinity produced contigs with greater N50, average length and mapped reads metrics. The greatest differences in assembly results between the two methods were observed in *P. grandis*. There, partial degradation of RNA samples in the heat stress/desiccation samples likely resulted in a Trans-ABySS assembly with numerous smaller contigs. Trinity was able to assemble transcript pieces into longer contigs with average lengths over 200 bp greater than Trans-ABySS and a high rate of paired read mapping. Considering these results, we utilized the Trinity assembly for subsequent analysis. Trinity contig assemblies for *P. grandis* and *U. tetralasmus* are available upon request.

Gene identification and annotation

Trinity contigs were used to query the UniProtKB/SwissProt and NCBI non-redundant (nr) protein databases by BLASTX (Table 5).

Table 5. Summary of gene identification and annotation of assembled mussel contigs based on BLAST homology searches against various protein databases (UniProt, nr).

	P. grandis		***U. tetralasmus***	
	UniProt	**NR**	**UniProt**	**NR**
Contigs with putative gene matches	48,448	57,469	56,170	66,381
Annotated contigs ≥500bp	28,205	32,290	33,595	38,375
Annotated contigs≥1000bp	11,773	13,243	14,391	16,017
Unigene matches	15,045	22,616	15,170	20,610
Hypothetical gene matches	0	6,120	0	6,101
Quality Unigenematches	11,278	11,877	11,324	10,699

Putative gene matches were at E-value ≤ 1e-5. Hypothetical gene matches denote those BLAST hits with uninformative annotation. Quality unigene hits denote more stringent parameters, including score≥100, E-value ≤ 1e-20.

In *P. grandis*, 57,469 contigs had significant BLASTX hits against 22,616 unique nr proteins. Of these, 6,120 contigs had top BLAST hits against hypothetical gene matches (predicted proteins and/or those with no annotation). Additionally, 11,877 unigenes could be identified based on more stringent criteria of a BLAST score ≥100 and E-value ≤ 1e-20 (quality matches). The same BLAST criteria were used to query the UniProt database.

In *U. tetralasmus*, 66,381 contigs had significant BLASTX hits against 20,610 unique nr proteins. Of these, 6,101 contigs had top BLAST hits against hypothetical gene matches (predicted proteins and/or those with no annotation). Additionally, 10,699 unigenes could be identified based on more stringent criteria of a BLAST score ≥100 and E-value ≤ 1e-20 (quality matches). The same BLAST criteria were used to query the UniProt database.

Identification and analysis of differentially expressed genes

Differential expression analyses in comparison to control samples were carried out for both the *P. grandis* and *U. tetralasmus* experimental samples (Table 6). In *P. grandis*, a total of 2,559 genes (unique annotated contigs with significant nr database BLASTX identities) including 1,499 up-regulated genes and 1,060 down-regulated genes were differentially expressed (⩾1.5-fold, corrected p-value <0.05) following 72 h of exposure to heat and

desiccation stress. Similarly, in U. tetralasmus, 2,532 genes including 1,594 up-regulated genes and 938 down-regulated genes were differentially expressed.

Table 6. Statistics of differentially expressed genes of *P. grandis* and *U. tetralasmus* annotated by nr and *C. gigas* protein database following drought challenge.

	P. grandis		***U. tetralasmus***	
	NR	***C. gigas***	**NR**	***C. gigas***
Up-regulated	1,499	1,093	1,594	1,126
Down-regulated	1,060	770	938	648
Total unigenes	2,559	1,863	2,532	1,774
Reads per contig	797	718	569	586

Values indicate contigs/genes passing cutoff values of fold change ≥1.5 (corrected $p<0.05$) and read number per contig ≥5. Reads/contig refers to average contig size.

Read coverage, critical in accurate determination of fold change, averaged 797 and 567 reads/contig for *P. grandis* and *U. tetralasmus*, respectively. We also used the differentially expressed contigs of both species as queries against annotated proteins from the recently sequenced Pacific oyster, *C. gigas* genome (the only publicly available genome among bivalve species [55]). Annotating based on hits to the *C. gigas* protein database identified 1,863 differentially expressed genes in *P. grandis* and 1,744 differentially expressed genes in *U. tetralasmus*.

While fewer contigs were annotated by *C. gigas*, its larger, comprehensive protein database aided in initial identification of "shared" genes present in sets of differentially expressed genes from both mussel species (Table 7). For example, we identified 318 and 380 significantly up-regulated contigs with the same top hit in both species using nr and *C. gigas* annotation, respectively. Similarly, 190 contigs shared the same top hit and were down-regulated in both species using *C. gigas* annotation. Finally, a subset of differentially expressed genes in both species shared the same BLAST identity but showed fold changes in opposite directions (Table 7).

Enrichment and pathway analysis

Gene ontology (GO) annotations were assigned using Blast2GO. In *P. grandis*, 7,571 GO

Table 7. Shared gene numbers of differentially expressed genes from *P. grandis* and *U. tetralasmus* annotated based on BLASTX searches against nr and *C. gigas* databases.

	After heat stress	
	NR	***C. gigas***
Shared up-regulated	318	380
Shared down-regulated	171	190
Shared identity/different direction	130	162

terms were assigned to 2,559 unique genes including 2,660 (35.1%) cellular component terms, 1,629 (21.5%) molecular functions terms and 3,282 (43.4%) biological process terms. In *U. tetralasmus*, a total of 6,764 GO terms were assigned to 2,532 unique gene matches including 2,478 (36.6%) cellular component terms, 1530 (22.6%) molecular functions terms and 2,756 (40.8%) biological process terms. The percentages of annotated sequences of the two species assigned to GO terms are shown in Figure 2.

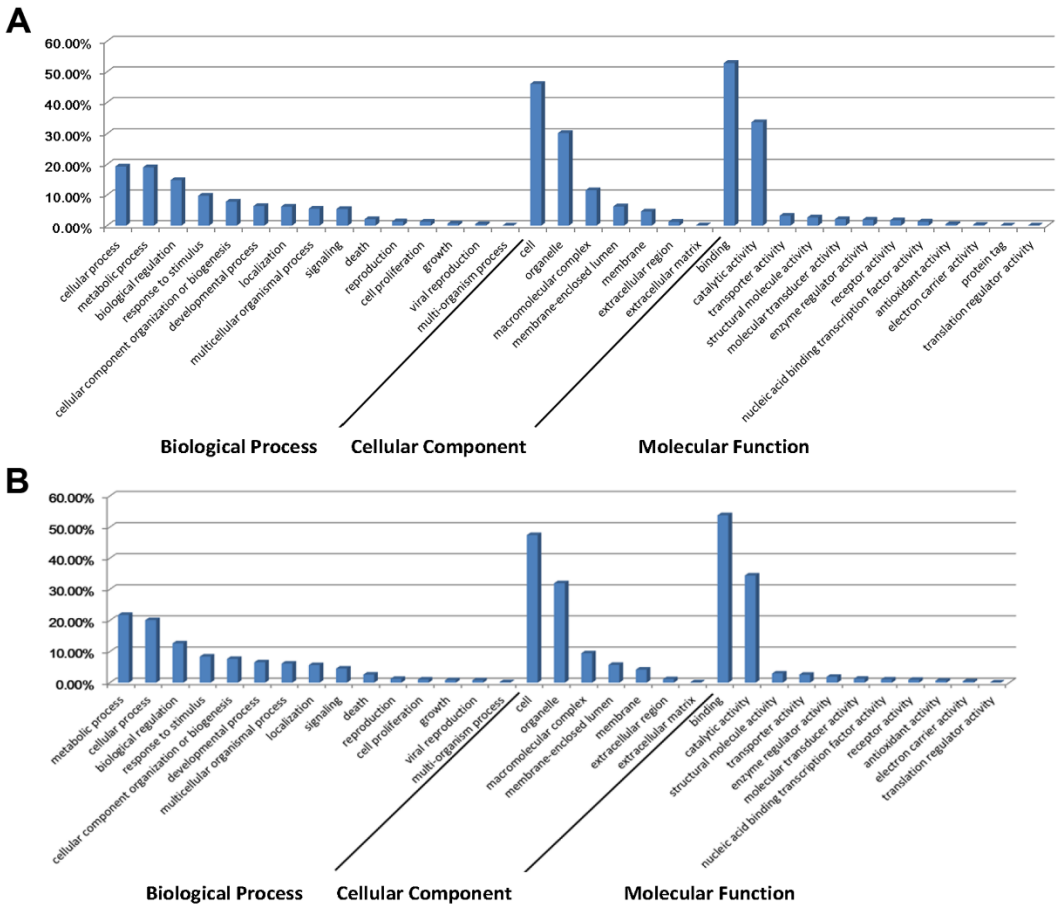

Figure 2. Gene ontology (GO) term categorization and distribution of assembled Trinity contigs encoding genes in *P. grandis* (A) and *U. tetralasmus* (B).

The differently expressed unique genes were then used as inputs to perform enrichment analysis using Ontologizer. Parent-child GO term enrichment analysis was performed to detect significantly overrepresented GO terms. A total of 23 and 50 overrepresented terms with *p* (FDR-corrected) < 0.1 were detected in *P. grandis* and *U. tetralasmus*, respectively. Potentially informative higher level GO terms were carried for further pathway analysis.

Table 8. Summary of GO term enrichment result of significantly expressed genes in *P. grandis* and *U. tetralasmus* following drought challenge.

GO ID	GO Name	Population count	Study count	p-Value
P. grandis				
GO:0005930	Axoneme	68	33	8.140E-05
GO:0044441	Cilium part	103	41	6.600E-04
GO:0048770	Pigment granule	66	30	4.143E-03
GO:0030286	Dynein complex	44	20	4.143E-03
GO:0006457	Protein folding	157	49	2.814E-02
GO:0051082	Unfolded protein binding	91	32	4.171E-02
GO:0031033	Myosin filament organization	12	8	8.194E-02
U. tetralasmus				
GO:0050792	Regulation of viral reproduction	42	20	8.470E-04
GO:2000243	Positive regulation of reproductive	39	19	2.076E-03
GO:0006457	Protein folding	153	51	2.594E-03
GO:0048524	Positive regulation of viral reproduction	31	16	5.965E-03
GO:0060548	Negative regulation of cell death	296	73	1.065E-02
GO:0097300	Programmed necrotic cell death	8	7	1.441E-02
GO:0048770	Pigment granule	56	22	2.375E-02
GO:0002764	Immune response-regulating signaling	102	29	2.783E-02
GO:0070265	Necrotic cell death	14	9	4.015E-02
GO:0051346	Negative regulation of hydrolase	158	39	4.345E-02
GO:0010941	Regulation of cell death	610	122	4.725E-02
GO:0010466	Negative regulation of peptidase	103	30	4.725E-02
GO:0071542	Dopaminergic neuron differentiation	4	4	5.958E-02
GO:0071779	G1/S transition checkpoint	31	11	6.974E-02
GO:0061135	Endopeptidase regulator activity	80	25	6.996E-02
GO:0005761	Mitochondrial ribosome	41	19	7.255E-02
GO:0004857	Enzyme inhibitor activity	138	36	7.444E-02
GO:0009068	Aspartate family amino acid catabolic	16	9	7.589E-02

p-value≤0.1 was considered significant. Population count is the number of genes associated with the term in the population set. Study count is the number of genes associated with the term in the study set.

These included indications of changes in protein folding in both species, and species-specific enrichment of cellular transport, cell structure and energy metabolism in *P. grandis* and genes negatively regulating cell death and the immune response in *U. tetralasmus* (Table 8). However, the general lack of functionally annotated genes in molluscan species limited the extent of GO analysis that could be carried out.

Based on enrichment analysis, manual annotation, and literature searches, representative key genes were organized into 7 functional categories encompassing important shared and potentially species-specific response signatures to heat stress/desiccation (Table 9).

Table 9. Key differentially expressed genes following drought challenge in *P. grandis* and *U. tetralasmus*.

Gene description	***P. grandis***		***U. tetralasmus***		**t-test *p*<0.05**
	FC	p-value	FC	p-value	
Chaperone/HSP (*p*=0.0277)					
60 kDa heat shock protein	**2.7**	4.7E-05	**2.0**	7.1E-03	
78 kDa glucose-regulated protein	**6.0**	3.6E-43	**3.8**	2.6E-31	
Alpha-crystallin B chain	**35.8**	2.8E-11	**7.2**	1.5E-152	*
DnaJ-like protein subfamily A member 1	**3.4**	1.3E-02	**2.2**	1.9E-04	
DnaJ-like protein subfamily A member 2	**2.0**	2.5E-15	**1.6**	2.2E-05	*
Endoplasmin	**3.1**	7.1E-35	**2.1**	1.8E-08	*
Heat shock 10kDa protein	**5.3**	4.9E-03	**2.1**	4.7E-06	*
Heat shock 70 kDa protein 12A, type 1	**-2.7**	1.5E-02	**-5.8**	4.2E-04	
Heat shock 70 kDa protein 14	**4.6**	4.5E-05	***1.3***	*4.5E-01*	*
Heat shock protein 68	**24.6**	1.7E-02	**15.9**	1.2E-06	
Heat shock protein 70 B2, type2	**4.61**	0	**1.66**	4.3E-21	*
Heat shock protein beta-1, type 2	**21.5**	1.6E-10	**10.6**	1.8E-243	*
Heat shock protein beta-1, type 3	**58.7**	5.0E-07	**21.3**	9.3E-78	*
Heat shock protein HSP 90-alpha 1	**9.5**	5.5E-18	**9.5**	8.9E-65	
Stress-70 protein, mitochondrial	**3.4**	7.0E-03	**1.6**	3.8E-08	
BAG molecular chaperone regulator 4	**27.9**	6.0E-16			
Calnexin	**6.1**	2.4E-08			
Calreticulin	**2.7**	1.1E-16			
DnaJ-like protein subfamily B member 4	**6.3**	1.1E-41			
Heat shock 70 kDa protein 12A, type 2	**3.6**	2.7E-02			
Heat shock factor protein 1	**6.1**	2.3E-02			
Heat shock protein 70 B2, type1	**8.9**	2.6E-03			
Heat shock protein beta-1, type 1	**173.2**	2.9E-10			
HSPB1-associated protein 1	**4.0**	1.3E-05			
Heat shock 70 kDa protein 12B			**7.4**	3.2E-06	
Antioxidant/Oxidative Stress Response (*p*=0.1534)					
Calpain-5	**1.7**	1.1E-02	**4.4**	1.1E-05	

Ceruloplasmin	**4.8**	6.8E-31	**-17.3**	1.6E-02	*
Dual oxidase	**-7.0**	2.1E-03	**-4.4**	2.6E-02	
Glutathione peroxidase 1	**2.9**	2.6E-14	**2.2**	9.9E-12	
Pantetheinase	**2.6**	1.0E-04	***1.2***	*6.5E-01*	
Protein toll	**10.9**	2.7E-04	**3.1**	4.9E-02	*
Selenide, water dikinase	***1.3***	*2.4E-02*	**2.3**	7.7E-20	*
Selenoprotein M	**-3.7**	3.3E-02	**-2.4**	4.2E-04	
Selenoprotein P, plasma, 1b	**-2.1**	1.3E-04	**-2.1**	0	
Selenoprotein T	**3.0**	6.9E-10	**2.0**	3.6E-04	
Superoxide dismutase	**-3.2**	1.4E-04	**-3.4**	2.2E-02	
Thioredoxin reductase 3	**-1.8**	2.1E-02	**-2.0**	7.8E-06	
X-box-binding protein 1	**33.0**	2.0E-67	**2.5**	1.1E-06	*
Elongation factor 1-alpha	**2.5**	1.4E-02			
Ubiquitin-protein ligase E3C	**9.8**	1.5E-03			
Calcium homeostasis endoplasmic reticulum			**6.1**	5.9E-03	
Selenoprotein W			**-1.7**	3.6E-07	
Cell proliferation/Apoptosis (*p*=0.2148)					
AP-2 complex subunit mu-1	**-2.0**	2.3E-02	**21.5**	1.6E-05	*
Apoptosis regulator R1	**1.7**	6.0E-03	**2.1**	1.7E-02	
BNIP1	**2.8**	4.1E-02	**11.0**	1.3E-13	*
BNIP3	**3.8**	1.6E-07	**3.2**	2.3E-13	
Bcl-2-like protein 1	**2.8**	1.7E-27	**3.4**	5.4E-06	
Cathepsin L-like cysteine proteinase	**2.3**	3.6E-08	**-2.1**	5.0E-02	*
Cold shock domain-containing protein E1	**1.8**	2.3E-10	**2.0**	2.6E-04	
Corticosteroid 11-beta-dehydrogenase 2	**32.2**	1.2E-06	**7.4**	5.0E-02	*
Cyclin-I	**2.0**	6.2E-09	**2.0**	1.7E-05	
E3 ubiquitin-protein ligase MIB2	**-6.2**	2.1E-02	**3.6**	4.6E-05	*
GADD45 alpha	**28.2**	1.2E-05	***-1.2***	*6.3E-01*	*
Inhibitor of apoptosis protein	**14.7**	7.8E-24	**11.9**	6.7E-03	
Krueppel-like factor 5	**25.8**	8.1E-60	**3.7**	3.0E-04	*
Polycomb complex protein BMI-1	**2.7**	7.7E-06	**2.9**	2.0E-08	
Protein BTG1	**9.1**	8.4E-175	**1.5**	1.0E-02	*
Serine/threonine-protein kinase Pim-3	**3.7**	6.6E-13	**-2.4**	1.3E-08	*
Stress-induced-phosphoprotein 1	**3.5**	3.5E-08	**2.5**	9.8E-08	*
TMBIM4	**2.8**	1.3E-02	**2.0**	3.1E-12	*
Ubiquitin	**14.0**	5.0E-13	**2.0**	1.0E-31	*
Antigen KI-67	**-4.9**	1.3E-02			
Caspase-10	**8.2**	2.8E-05			
CDH1-D	**2.4**	2.8E-06			
MAP kinase signal-integrating kinase 1	**3.9**	1.8E-10			
Protein SET			**-2.6**	2.7E-04	
Putative inhibitor of apoptosis			**86.8**	6.5E-26	
Ubiquitin-conjugating E2 variant 1 (UEV1A)			**4.5**	5.1E-14	
Energy Metabolism (*p*=0.1482)					
5'-AMP-activated protein kinase beta-2	**-4.3**	1.2E-05	**-2.1**	2.9E-02	*
ATP synthase F0 subunit 6	**-7.1**	5.8E-08	***-3.3***	*2.7E-01*	
Cholecystokinin receptor	**3.8**	1.1E-04	**-3.3**	4.9E-02	*
Cytochrome b	**-5.6**	0	**-2.7**	2.1E-02	

Cytochrome c oxidase subunit I	**-3.4**	0	**-1.6**	1.3E-13	*
Cytochrome c oxidase subunit II	**-3.3**	0	**-1.9**	3.1E-04	
Cytochrome c oxidase subunit III	**-4.5**	0	***-2.4***	*3.1E-01*	
Glutamine synthetase 2 cytoplasmic	**5.3**	1.5E-24	**2.0**	1.6E-24	*
Group XIIA secretory phospholipase A2	**2.8**	1.6E-03	**10.9**	5.0E-05	*
NADH dehydrogenase subunit 1	**-6.0**	0	***-1.3***	*8.8E-01*	*
NADH dehydrogenase subunit 4	**-2.0**	1.5E-03	***-1.1***	*9.5E-01*	
NADH dehydrogenase subunit 5	**-3.8**	2.5E-12	**1.7**	1.5E-03	*
Succinyl-CoA ligase	**-2.2**	2.7E-03	**-3.5**	6.2E-06	
ATF5 activating transcription factor 5	**2.4**	9.5E-07			
Immune (*p*=0.7569)					
B-cell lymphoma 3-encoded protein	**6.2**	1.2E-02	**-2.4**	4.0E-06	*
Histone H4 transcription factor	**6.0**	5.8E-08	***1.9***	*5.5E-01*	*
Interferon regulatory factor 2	**-2.5**	1.4E-06	**-4.2**	3.3E-10	
Interleukin-1 receptor-associated kinase 1BP1	***1.1***	*6.0E-01*	**-1.7**	1.3E-02	*
Lipopolysaccharide-binding protein	**-2.8**	9.7E-03	**-2.0**	1.2E-03	
LPS-induced tumor necrosis factor-alpha	**-4.3**	2.0E-03	**-2.3**	4.9E-02	
Lysozyme	**2.1**	4.1E-02	**2.3**	3.0E-10	
NLR family, pyrin domain containing 3	**4.5**	9.8E-04	**2.7**	1.5E-03	
NF-kappa-B inhibitor cactus	**3.9**	1.9E-18	**2.1**	2.5E-04	*
Peptidoglycan-recognition protein SC2	**3.4**	3.0E-03	**3.4**	1.6E-13	
Phospholipid scramblase 2, partial	**-4.0**	2.0E-02	**14.3**	5.4E-06	*
Serine/threonine-protein kinase TBK1	**7.0**	3.2E-31	**11.3**	3.4E-03	
Slit-like protein 1 protein	**6.5**	7.0E-06	***1.2***	*5.2E-01*	*
TNFAIP3-interacting protein 2	**4.7**	4.3E-04	**2.5**	2.2E-02	
TNF receptor-associated factor 2	***1.9***	*5.4E-01*	**4.5**	1.7E-03	*
Toll-like receptor 2 type-2	**3.8**	9.9E-06	**4.0**	1.3E-10	
Toll-like receptor 3	**-3.0**	3.4E-03	**4.6**	7.1E-03	*
14-3-3 protein epsilon	**-4.2**	9.8E-13			
Myeloid differentiation primary response 88	**3.6**	3.0E-03			
Tax1-binding protein 1-like protein B	**-1.5**	8.1E-03			
14-3-3 protein zeta			**-7.8**	1.3E-04	
Complement C3			**-1.6**	1.7E-06	
Defensin			**100.1**	1.9E-06	
EIF4EBP1			**2.5**	1.1E-07	
NLR family, pyrin domain containing 1			**13.2**	2.2E-03	
TNF receptor-associated factor 3			**3.3**	4.5E-03	
Toll-like receptor 13			**-10.7**	1.8E-02	
Cytoskeletal (*p*=0.0149)					
Collagen alpha-2(I) chain	**-2.3**	1.0E-02	**-1.9**	1.5E-05	
Fibrillin-1-like	**-2.3**	2.0E-02	**-6.5**	1.2E-03	*
Microtubule-associated protein 2	**40.1**	4.4E-08	**5.9**	6.1E-04	*
Myosin catalytic light chain LC-1	***1.2***	*8.4E-02*	**-13.0**	2.1E-04	*
Protocadherin Fat 4	**1.9**	8.5E-03	**-16.5**	2.8E-07	*
Ras-like GTP-binding protein Rho1	**15.5**	3.8E-03	**-2.7**	4.9E-02	*
Tubulin alpha-1C chain	**-2.4**	1.3E-10	**-2.2**	5.4E-09	
Tubulin beta chain	**33.1**	7.9E-31	**3.6**	1.2E-03	*
Dynein heavy chain 6, axonemal	**-2.8**	7.6E-04			

Tubulin alpha chain	**-3.4**	3.7E-08			
Other (*p*=0.8545)					
Chitin synthase 1	**-3.5**	2.1E-04	**-4.3**	3.5E-06	
Chloride intracellular channel exc-4	**5.2**	6.5E-22	***1.4***	*1.8E-02*	*
Far upstream element-binding protein 3	**-31.2**	0	**-1.9**	1.8E-02	*
Fibulin-2	**-3.9**	6.4E-03	**-3.8**	5.0E-08	
Hemicentin-1	**15.8**	4.9E-21	**-2.4**	0	*
Krueppel-like factor 11	**3.9**	4.1E-07	**3.6**	5.0E-02	
Outer dense fiber protein 3	**-12.2**	8.1E-04	**4.0**	4.9E-02	*
Putative accessory gland protein	**21.0**	5.2E-20	**2.2**	4.5E-05	*
Tribbles-like protein 2	**10.4**	2.3E-164	**1.9**	3.0E-09	*
Water and ammonia transporting aquaporin	**2.3**	1.5E-04	**-1.6**	2.4E-05	*
Epithelial membrane protein 2	**17.2**	3.2E-07			
Liver stage antigen 3 precursor	**-19.1**	1.3E-15			
Semenogelin II precursor	**-16.5**	0			
Dexras1			**24.8**	1.7E-02	

Genes were organized in broad functional categories. Absence of contig, fold change, and p-value information for a given gene indicates that no orthologous sequence was detected in that particular species. Italicized contig identifiers indicate that the gene did not fit criteria for significance (≥1.5 fold change, $p<0.05$) in the particular species. * indicates significantly different fold changes (FC) between species for a given gene based on a t-test of replicated fold change values ($p<0.05$). Differences in fold changes between species in each category were assessed for significance using a paired t-test ($p<0.05$) with values given next to the category titles.

These included chaperone/heat shock proteins, antioxidant/oxidative stress response, cell proliferation/apoptosis, energy metabolism, immune, cytoskeletal and uncategorized (other). While imputed putative functional roles of these genes are covered in-depth below (Discussion), several general trends were apparent. First, while a robust heat-shock response was mounted in both species, the drought-susceptible *P. grandis* manifested significantly higher and broader up-regulation of molecular chaperones at 72 h (p=0.0277, paired t-test). Second, significantly different category-level fold changes were observed in genes regulating cytoskeletal arrangements (p=0.0149), with higher up-regulation of cell fiber elements observed for *P. grandis*. Third, individual key genes within each category had significantly different responses to drought stress, including, for example, those regulating inhibition of apoptosis and immunity (Table 9).

Validation of RNA-seq profiles by QPCR

In order to validate the differentially expressed genes identified by RNA-seq, we selected 16 genes for QPCR confirmation from *P. grandis* and *U. tetralasmus* (Table 10). Selection of genes was based on putative functions in response to heat stress/desiccation and pathway analyses. Melting curve analysis revealed a single product for all tested genes. Fold changes from QPCR were compared with the RNA-seq expression analysis results. QPCR results for *P. grandis* were significantly correlated with RNA-seq results (avg. correlation coefficient, R=0.87; p-value <0.001; Figure 3A). Similarly, with the exception of calpain 5, QPCR results for *U. tetralasmus* closely matched those observed by RNA-seq results (avg. correlation coefficient, R=0.89; p-value <0.001; Figure 3B). Overall, the QPCR results indicated the reliability and accuracy of the Trinity reference assembly and the RNA-seq-based transcriptome expression analysis.

Table 10. Fold change values of QPCR validation as presented in Figure 3.

Gene	**72h fold change after drought**			
	P. grandis		***U. tetralasmus***	
	Realtime	RNA-seq	Realtime	RNA-seq
Alpha-crystallin B chain	38.18	35.81	28.61	7.23
BCL2/adenovirus E1B interacting protein 1	1.12	2.79	3.04	11.03
Calpain 5	1.50	1.70	-1.77	4.37
Heat shock protein 70 B2,type2	3.52	4.61	2.15	1.66
Heat shock protein beta-1, type 3	122.45	58.71	74.60	21.32
Heat shock protein HSP 90-alpha 1	10.76	9.54	4.76	9.51
Kruppel-like factor5	35.45	25.83	1.48	3.74
BAG family molecular chaperone regulator 4	10.54	27.88		
Heat shock protein 70 B2,type1	6.06	8.93		
Heat shock protein beta-1, type 1	127.18	173.23		
Liver stage antigen 3 precursor	-13.58	-19.10		
Defensin			712.55	100.10
Heat shock 70 kDa protein 12B-like			4.15	7.39
Inhibitor of apoptosis protein			217.02	86.77
NLR family, pyrin domain containing 1			5.08	13.21
Toll-like receptor 13			-3.70	-10.65
Correlation between Methods	0.87		0.89	

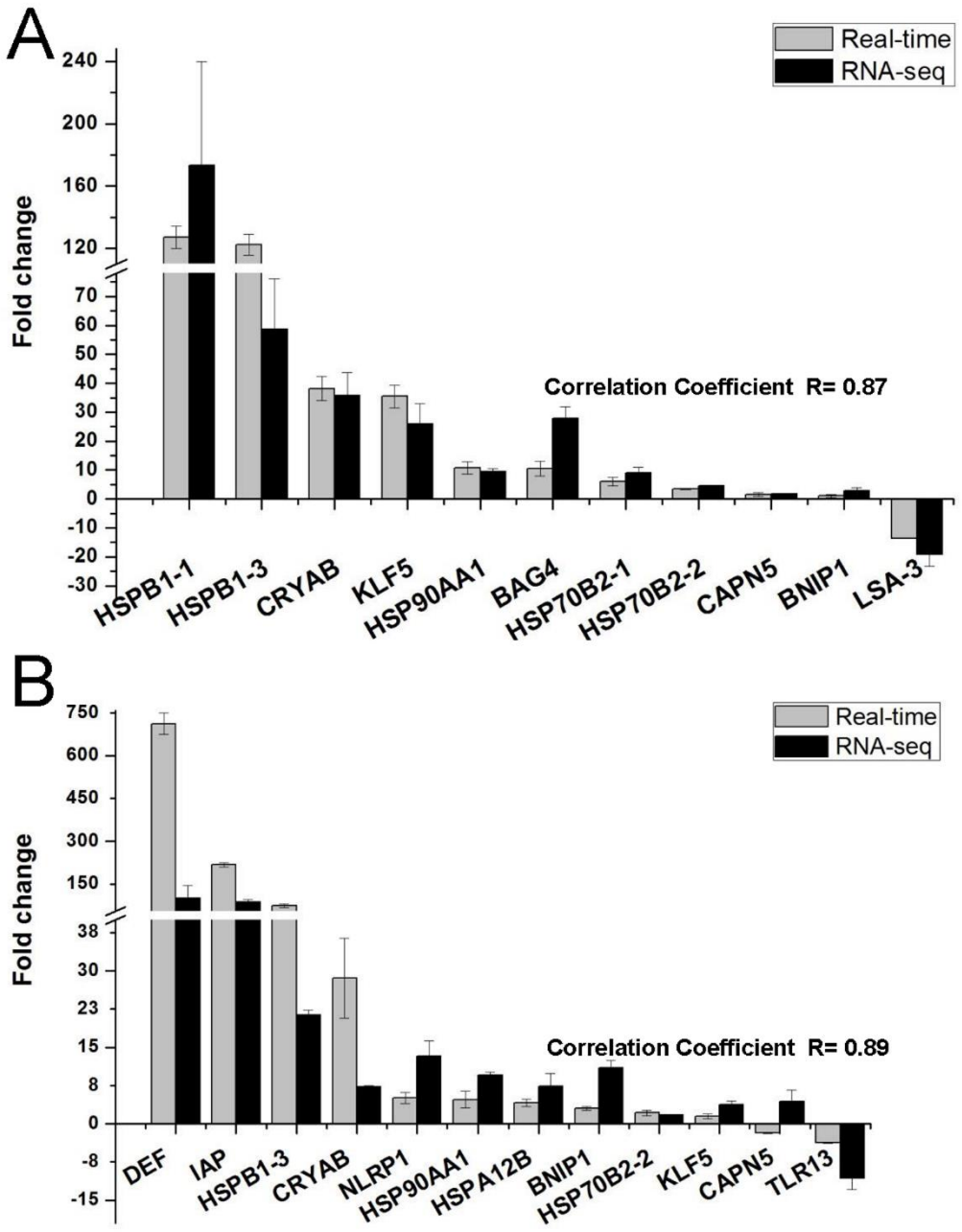

Figure 3. Comparison of fold changes between RNA-seq and QPCR results in selected genes of *P. grandis* (A) and *U. tetralasmus* (B) at 72 h of heat/desiccation exposure. QPCR fold changes are relative to control samples and normalized by changes in beta-actin values. Gene abbreviations are available in List of Abbreviations. Contig IDs are available in Table 2.

Temporal expression of key heat/desiccation signatures

Given our focus on the 72 h timepoint for RNA-seq expression analysis, we sought to extend the profiles of several key genes to earlier (0 h, 24 h, 48 h) timepoints as well as the

recovery phase (5 d). While financial constraints limited our ability to examine these timepoints comprehensively by RNA-seq, we wished to examine whether key genes identified at 72 h could serve as sensitive markers for stress at other timepoints. We selected four heat shock genes that were up-regulated to differing extents in both species and a gene found to be up-regulated uniquely in each species (Figure 4). We found that heat shock

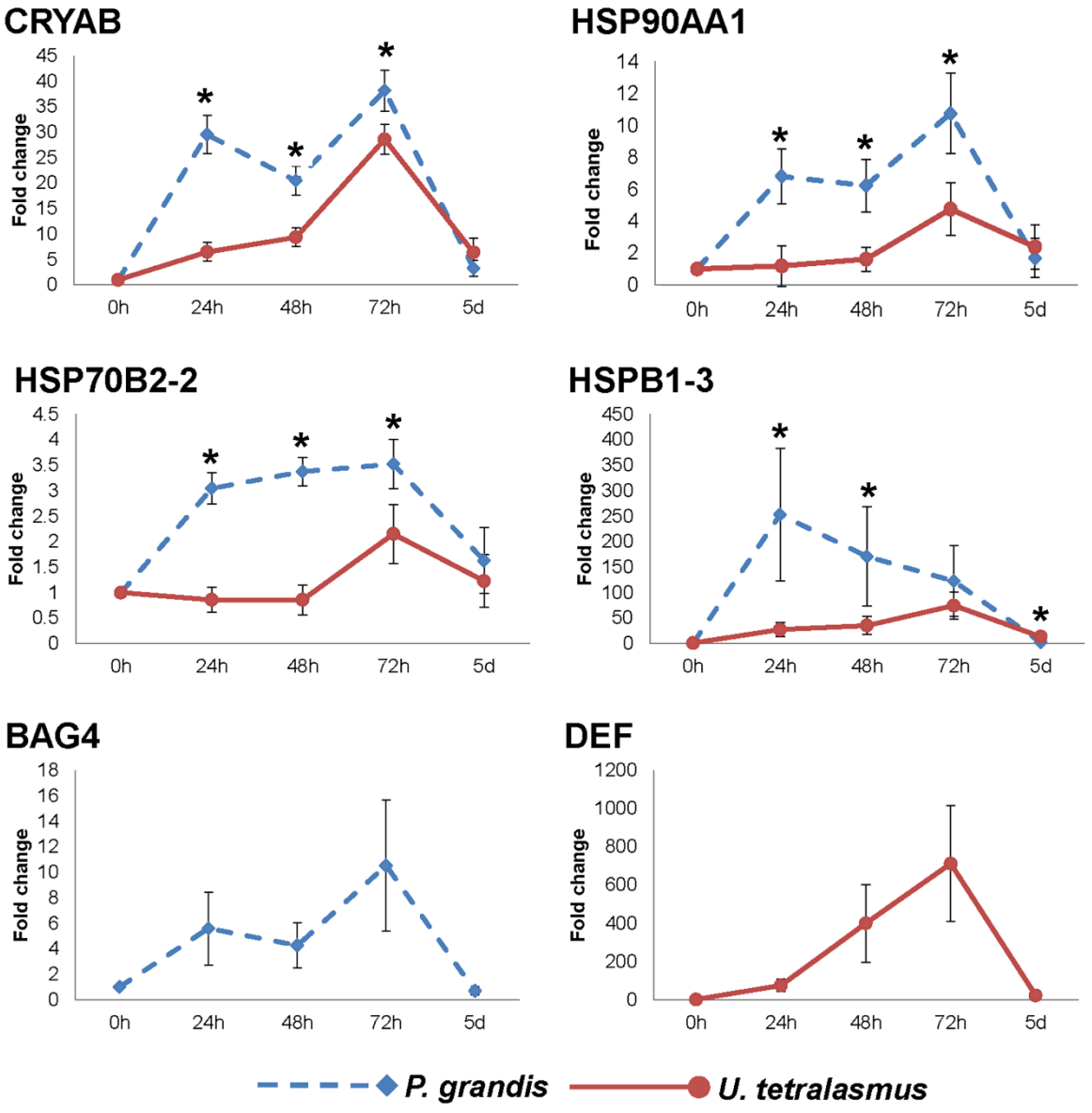

Figure 4. Temporal profiling of key genes by QPCR in *P. grandis* and *U. tetralasmus*. Fold changes are relative to control values (set as 1) and normalized by changes in beta-actin values. * indicates significantly different fold changes between species ($p<0.05$). Error bars reflect standard error of the mean. Gene abbreviations are available in List of Abbreviations. Contig IDs are available in Table 2.

gene responses in *P. grandis* showed faster and higher levels of induction than observed in *U. tetralasmus*. All four shared heat shock proteins (CRYAB, HSP90AA1, HSP70B2-2, HSPB1-3) showed significant induction of expression in the more sensitive *P. grandis* by 24 h, while up-regulation of these genes in *U. tetralasmus* at the same timepoint was either relatively modest or not observed. Expression of the four shared genes trended higher throughout the heat and desiccation timepoints, with the exception of HSPB1-3 in *P. grandis* which peaked at 24 h and then declined. Similar patterns of up-regulation were also observed for BAG molecular chaperone regulator 4 (BAG4) in *P. grandis* and defensin (DEF) in *U. tetralasmus*. All examined genes had declined from their 72 h expression levels in the recovery timepoint, with most genes approaching homeostatic (0 h) levels. Fold changes (relative to 0 h) were significantly higher in *P. grandis* than *U. tetralasmus* at most tested timepoints (Figure 4). Based on this data subset, it seems likely that that other key genes identified at 72 h also responded sensitively to heat/desiccation at earlier timepoints and returned to normal levels as experimental stressors were relaxed (recovery period).

Discussion

Recent studies of mussel behavioral and physiological responses to drought [2,3] have revealed an array of strategies to maximize survival, growth, and reproduction. Mussel species inhabiting small streams during drought periods can be stranded by receding waters, forcing them to burrow into cooler, wetter sediments or to track waters to deeper pools, when available. These differing behavioral approaches (tracking, burrowing, or tracking then burrowing) likely result in differential fluctuations in population sizes and stream bed distributions among species based on annual variations in the hydrologic cycle. Physiologically, stranded mussels face the simultaneous stressors of increased temperature and tissue desiccation which can, ultimately, impact survival but, more immediately, regulate

the evolved avoidance strategies of each species. We sought here to examine the transcriptional regulation of differing tolerances to drought in two mussel species representing physiological and behavioral extremes. *Uniomerus tetralasmus* (pondhorn) can live up to 2 years out of water at cool temperatures, while *P. grandis* (floater) can tolerate only a few weeks of desiccation [12]. Recent field and laboratory studies by our group revealed burrowing behavior in *U. tetralasmus* and tracking behavior in *P. grandis* [7]. In this study, we carried out RNA-seq transcriptome profiling on *P. grandis* and *U. tetralasmus* exposed to experimental conditions simulating those of a drought-induced stranding event, i.e. heat and desiccation.

Prior to initiation of this research, few genetic resources were available for either *P. grandis* or *U. tetralasmus*, with only a handful of mitochondrial sequences currently in public databases. Therefore, prior to examining transcriptional responses in either species, we needed to sequence and assemble high-quality transcriptomes. Use of Trinity [18] provided superior assembly when compared with Trans-ABySS and generated contigs with average lengths greater than 800 bp. Although it is difficult to predict gene numbers based on comparison with other invertebrate species known for high rates of mutation and diversification [81], estimates of gene numbers from the draft genomes of the Pacific oyster (*C. gigas*) and the pearl oyster (*Pinctada fucata*) are available. These bivalve species encode 28,027 and 23,257 genes, respectively [55,82], compared with 22,616 and 20,610 predicted unigene annotations in *P. grandis* and *U. tetralasmus*, respectively (Table 5). Based on predicted low rates of sequence conservation across species, we estimate that the *P. grandis* and *U. tetralasmus de novo* assemblies captured at least 70% of the coding genes of each species.

Sampled tissues and pooling strategies inevitably impact the size, scope, and nature of captured transcriptomes and associated downstream expression profiles. Here, we sampled

tissues from the foot of individual mussels using a biopsy punch and pooled individual samples to construct replicate pools for RNA-seq analysis as in previous studies [23,78-80]. While pooling may mask some of the individual variation in homeostatic and induced gene expression profiles, it also serves to normalize outliers and direct focus to key shared responses. Putative biomarkers developed here can be field-validated on individual samples in future studies.

The mussel foot is a tough, extendible muscular organ used primarily for locomotion and, in juveniles, for anchoring the animal to substrate via byssal threads. No reports of gene expression in the foot of unionid mussels are available to-date, although flow regimen has recently been reported to impact byssal thread production in juvenile unionids [83]. In the zebra mussel, *Dreissena polymorpha*, gene and protein expression of the foot organ have been studied in the context of adhesion [84-86]. Our primary consideration in choosing to examine the foot transcriptome was to obtain samples rapidly and in a non-lethal manner [87]. Given that 70% of freshwater mussels are currently imperiled [64], future wide-scale assessments of individual or population health in the field will require sampling techniques that are rapid and minimally invasive. We, therefore, wanted to obtain relevant signatures of heat/desiccation stress from a tissue available without sacrifice or extensive manipulation of the organism. Additionally, expression signatures in the foot may reflect the resulting differential behavioral impulses for locomotion, adhesion and/or burrowing. While beyond the scope of the present study, a comprehensive tissue expression atlas of certain representative, non-endangered mussel species should be considered in the future.

Several transcriptomic and proteomic studies in bivalves over the last decade have expanded our understanding of the components of thermal stress responses in these species. The most comprehensive studies have been carried out in marine mussels [18-20,67] and in the Pacific oyster [16,17,55]. With the exception of Zhang et al. 2012 [30], other studies have

utilized microarray technology, potentially missing rare, previously un-sequenced transcripts. Recently, we conducted a smaller-scale heat stress study in the unionid mussel *V. lienosa* utilizing RNA-seq [23]. Examined together, several key components of bivalve heat stress responses can be seen in these studies. These include changes in molecular chaperones, oxidative stress responses, changes in cell turnover and death, changes in energy metabolism, immune and inflammatory responses, and cytoskeletal reorganization. We observed similar dysregulation of genes involved in these pathways in our datasets from *P. grandis* and *U. tetralasmus* (Table 9). Below we highlight putative functions of selected key genes from these categories which may underlie shared and species-specific responses to drought.

Chaperones/Heat Shock Proteins

Up-regulation of molecular chaperones, particularly HSP70 family members, has long been understood as a classic indicator of protein damage due to heat [20,88]. Heat shock proteins (HSPs) interact with stress-denatured proteins, preventing their aggregation. An array of molecular chaperones was induced in both species by exposure to 72 h of elevated temperatures and desiccation (Table 9). Only one chaperone (HSP70-12A1) was down-regulated in both species. A general pattern of higher expression of shared chaperones in the less-tolerant mussel species was present, with 8 of 15 shared chaperones showing a significantly higher fold increase by *P. grandis* compared to *U. tetralasmus,* but none showing a significantly lower fold increase. For example, a small heat shock protein, HSPB1-type 3 was induced 58.7-fold in *P. grandis* compared with 21.3-fold in *U. tetralasmus*. An additional nine HSP family members were unique to *P. grandis*, including HSPB1-type 1 (173.2-fold). These differences were confirmed in selected HSP genes analyzed by QPCR (Figure 2). Further temporal profiling of four shared molecular chaperones showed that species differences in chaperone expression were not confined to the 72 h timepoint but were also present at 24 h and 48 h. *Uniomerus tetralasmus* HSPs showed

limited induction prior to 72 h, while *P. grandis* HSPs were rapidly induced by 24 h. Induction of the HSP response was temporal in nature, however, as expression levels of the tested genes all had returned to baseline levels by 5 d after cessation of the emersion/heat exposure (Figure 4). In previous comparisons of marine mussel species and populations with differing heat tolerances, it was observed that differences in small HSP (HSPB-type) up-regulation were among the most significant signatures of heat tolerance/sensitivity [18-20,67]. There, small HSPs were induced at lower temperatures and/or to higher levels in heat-sensitive mussel groups. Small HSPs, often denoted as HSPBs, are a family of heat shock proteins involved in regulation of apoptosis, oxidative stress, and cytoskeletal regulation [89,90]. Here also, HSPB family members were highly induced in *P. grandis* relative to *U. tetralasmus* indicating their potential importance in determining drought tolerance in unionid mussels.

Antioxidant/Oxidative Stress Response

Heat shock causes increased production of reactive oxygen radicals leading to toxic byproducts which damage cells [91]. In response, the organism produces a variety of antioxidant enzymes and protective agents which can detoxify free radicals. Among these are glutathione peroxidase (up-regulated in both species here) and superoxide dismutase (down-regulated in both species). Selenoproteins, known for their roles in antioxidant protection were also differentially expressed in both species [92]. Overall, no significant differences were observed in genes regulating the oxidative stress responses between *P. grandis* and *U. tetralasmus*.

Cell Proliferation/Apoptosis

Accumulation of oxidative damage can induce a number of pro-apoptotic signaling pathways. Conversely, protective/adaptive mechanisms can seek to block excessive apoptotic

signals and stimulate cell proliferation and a return to homeostasis [93,94]. Many of the components of these processes were observed in the sets of differentially expressed genes from both mussel species (Table 9). However, differential regulation of several genes may indicate crucial differences impacting drought tolerance. Genes needed for protein degradation such as ubiquitin were higher in the sensitive *P. grandis* (14.0-fold up-regulated) than in *U. tetralasmus* (2.0-fold). GADD45, a key indicator of DNA damage, was up-regulated 28.2-fold in *P. grandis* but was not induced in *U. tetralasmus*. GADD45 was similarly up-regulated in *M. galloprovincialis* following exposure to metal salts [95]. Similarly, caspase-10, a component of the execution phase of cell apoptosis was up-regulated 8.2-fold in *P. grandis* but unchanged in *U. tetralasmus*. Interestingly, an inhibitor of apoptosis (IAP) gene showed the second highest up-regulation (86.8-fold) of all differentially expressed genes in the drought-tolerant *U. tetralasmus*. No clear IAP orthologue was identified from *P. grandis*. Lastly, ubiquitin-conjugating enzyme E2 variant 1 (UEV1A) was differentially expressed only in *U. tetralasmus* (Table 9). In mammals, IAP genes interact with UEV1A and other ubiquitin ligases such as MIB2 (also differentially expressed here) to inhibit stress-induced apoptosis and facilitate cell proliferation/survival [96-98]. These pathways are mediated through NF-κB activation and, therefore, often dovetail with modulation of components of the innate immune system [99,100]. Taken together, the sensitive *P. grandis* showed induced pro-apoptotic signaling pathways and indicators of DNA damage, whereas the tolerant species *U. tetralasmus* showed greater inhibition of apoptosis and little indication of DNA damage. These molecular differences are likely an important component of the observed physiological differences in drought tolerance between the two species. Further studies are needed in molluscan species to confirm the conserved function of these components in regulation of cell survival.

Immune

One area of concern for drought-exposed mussels is that stress responses induced by emersion and/or rising air/water temperatures will impact immune health increasing susceptibilities to pathogens encountered in their environment [17,83]. Indeed, we observed the dysregulation of a number of important mediators of immunity in both mussel species (Table 9). These included the down-regulation of innate cytokines such as TNF-α and the down-regulation of pathogen-recognition receptors such as Toll-like receptor 3. Several components of the NF-κB signaling pathway including TRAF2, TRAF3, and TBK1 were either present only in *U. tetralasmus* or were induced to a greater extent, potentially regulating cell survival mechanisms as discussed above.

Of particular interest in the immune category was the strong up-regulation of a *U. tetralasmus* defensin (up 100.1-fold; highest FC among pondhorn genes). An orthologous defensin was not identified among differentially expressed genes in *P. grandis*. Defensins are a family of peptides with antimicrobial activities against a range of bacterial and fungal pathogens. While their traditional immune roles have been characterized in marine mussels and oysters [101,102], little is known about their functions in freshwater mussel. Recently, however, Xu and Faisal [103] described a defensin from zebra mussel, *D. polymorpha*, with antimicrobial activity. Notably, zebra mussel defensin expression was the highest in the foot of all tested tissues and cell types and expression levels corresponded with status of byssogenesis. The authors there speculated that defensin may be providing biological protection against microbial degradation of byssal threads or may be playing a more fundamental role in byssogenesis. Here, in adult unionids, observed differences in defensin up-regulation may reflect differential use of the foot in burrowing. The burrowing species *U. tetralasmus* may scratch or injure its foot in burrowing, resulting in protective secretion of the antimicrobial defensin. By QPCR analysis, up-regulation of *U. tetralasmus* defensin was

even higher than that indicated by RNA-seq (Figure 4) and expression levels rose with the duration of the challenge (Figure 4). Future studies should examine potential functions of unionid defensins in the context of immune status, burrowing, byssal thread production, and drought tolerance.

Cytoskeletal

The buildup of free radicals associated with cellular stress leads to damage of the actin cytoskeleton. Changes to actin filament dynamics impact junctional integrity, cell cycling, and cellular movement. Gene and protein changes in cytoskeletal elements have been reported previously in response to heat stress in marine mussels [19,20]. In response to heat stress, cells modulate expression of these elements and enhance production of small HSPs to combat these effects [104]. In addition to the higher expression of small HSPs in *P. grandis*, larger fold changes were seen in genes associated with assembly of microtubules, one of the fibers composing the cytoskeleton. For example, tubulin beta chain was up-regulated 33.1-fold in *P. grandis*, but only 3.6-fold in *U. tetralasmus* (Table 9). These differences are another indication of higher levels of cellular stress in *P. grandis*.

Other

We noted a handful of additional genes with putative functions outside one of the larger response categories. These included genes with functions in ion transport, reproduction, shell structure, and circadian rhythms (Table 9). Interestingly, in both species drought stress led to reduced expression of chitin synthases, key enzymes involved in synthesis of bivalve shells, indicating potential impacts on growth and health from prolonged drought exposure even in "tolerant" species [105]. Finally, dexras1, a gene which, in mammals, is responsible for transducing signals maintaining circadian rhythms of behavior and physiology [106], was sharply induced in *U. tetralasmus* alone (24.8-fold up-regulation). Given the ability of

stranded mussels of this species to endure long periods of emersion, it would be of great interest to determine whether dexras1 is responsible for the biological reprogramming needed for tolerating harsh environmental conditions.

Conclusions

Freshwater mussels in the southeastern US are imperiled by increasingly frequent drought conditions. Here we examined the molecular underpinnings of differing behavioral and physiological responses to drought in two co-occurring mussel species. RNA-seq transcriptome profiling successfully captured and assembled high quality transcriptomes for *P. grandis* and *U. tetralasmus* and facilitated comparison of gene expression profiles following heat/desiccation exposure. Shared molecular responses included changes in molecular chaperones, oxidative stress profiles, cell cycling, energy metabolism, immunity, and cytoskeletal rearrangements. Significantly higher induction of molecular chaperones and cytoskeletal elements indicated higher levels of cellular stress in *P. grandis* while our results also highlighted individual genes potentially contributing to drought tolerance in *U. tetralasmus*. From this foundation, future studies should seek to validate putative biomarkers revealed here in individual mussels from *P. grandis* and *U. tetralasmus* collected in field settings, as well as examining their utility in assessing stress levels in a broader array of unionid species.

III. Future Directions

While the work described here significantly advances our understanding of the molecular processes underlying behavioral and physiological responses of unionid mussels to drought conditions, considerable work remains to translate this information for widespread applicability in the field. Studies following this one should address gaps in our knowledge in several areas as detailed below.

Timepoint Analysis

RNA-seq analysis was only carried out at the 72 h timepoint in this study. While temporal expression analysis was conducted by QPCR on a small handful of genes at earlier and later timepoints, a broader picture could be gained by either RNA-seq on additional timepoints or by running QPCR on a larger number of genes at these sampling points. Differences between mussel species at earlier timepoints may serve as powerful biomarkers. Examining the entire transcriptome over multiple timepoints would allow identification of genes whose expression patterns are likely co-regulated and would point to critical functional pathways.

Individual and Species Variation

In the present study, replicates were composed of pooled samples from 5 individuals, a common approach in RNA-seq studies. However, pooling may mask individual variation in expression of some key genes. Therefore, candidate genes proposed for follow-up studies needed to be tested on individual mussel samples to identify those with consistent, statistically significant expression differences on the individual level. Additionally, future studies must address whether key genes identified here are also involved in the likely diverse

of behavioral and physiological responses found across Unionidae. These studies would necessitate additional field and laboratory trials with additional species as well as generation of transcriptomes for these species by RNA-seq.

Tissue Variability

Future studies should address variability between expression signatures obtained by different non-lethal sampling methods from the foot, adductor muscle, or hemolymph. Standardization of protocols for tissue amounts and precise sampling locations are also needed to ensure repeatability in biomarker assays.

Assay Consistency

Ultimately, transcriptome expression profiles, such as those here, need to be narrowed, validated, refined, and translated into robust, user-friendly assays which could be utilized by field biologists with results obtainable within a week's time. One approach may be to develop conserved protein colorimetric assays for key genes of interest. In some cases, existing antibody reagents for highly conserved HSPs may have utility. However, in other cases, novel protein assays may need to be developed for sensitive, genera-specific markers.

References

1. Peterson JT, Wisniewski JM, Shea CP, Jackson CR (2011) Estimation of mussel population response to hydrologic alteration in a southeastern US stream. Environmental management 48: 109-122.
2. Allen DC, Galbraith HS, Vaughn CC, Spooner DE (2013) A Tale of Two Rivers: Implications of Water Management Practices for Mussel Biodiversity Outcomes During Droughts. Ambio: 1-11.
3. Neves RJ, Bogan AE, Williams JD, Ahlstedt SA, Hartfield PW (1997) Status of aquatic mollusks in the southeastern United States: a downward spiral of diversity. Aquatic fauna in peril: the southeastern perspective Special Publication 1: 43-85.
4. Galbraith H, Vaughn C (2011) Effects of reservoir management on abundance, condition, parasitism and reproductive traits of downstream mussels. River Research and Applications 27: 193-201.
5. Spooner DE, Vaughn CC (2008) A trait-based approach to species' roles in stream ecosystems: climate change, community structure, and material cycling. Oecologia 158: 307-317.
6. Vaughn CC (2012) Life history traits and abundance can predict local colonisation and extinction rates of freshwater mussels. Freshwater Biology 57: 982-992.
7. Gough HM, Gascho Landis AM, Stoeckel JA (2012) Behaviour and physiology are linked in the responses of freshwater mussels to drought. Freshwater Biology 57: 2356-2366.

8. Gagnon PM, Golladay SW, Michener WK, Freeman MC (2004) Drought responses of freshwater mussels (Unionidae) in coastal plain tributaries of the Flint River Basin, Georgia. Journal of Freshwater Ecology 19: 667-679.

9. Golladay SW, Gagnon P, Kearns M, Battle JM, Hicks DW (2004) Response of freshwater mussel assemblages (Bivalvia: Unionidae) to a record drought in the Gulf Coastal Plain of southwestern Georgia. Journal of the North American Benthological Society 23: 494-506.

10. Haag WR, Warren Jr ML (2008) Effects of severe drought on freshwater mussel assemblages. Transactions of the American Fisheries Society 137: 1165-1178.

11. Galbraith HS, Spooner DE, Vaughn CC (2010) Synergistic effects of regional climate patterns and local water management on freshwater mussel communities. Biological Conservation 143: 1175-1183.

12. Holland DF (1991) Prolonged Emersion Tolerance in Freshwater Mussels (Bivalvia-Unionidae): Interspecific Comparison of Behavioral Strategies and Water Loss Rates: University of Texas at Arlington.

13. Woodward G, Perkins DM, Brown LE (2010) Climate change and freshwater ecosystems: impacts across multiple levels of organization. Philosophical Transactions of the Royal Society B: Biological Sciences 365: 2093-2106.

14. Spooner DE, Vaughn CC, Galbraith HS (2012) Species traits and environmental conditions govern the relationship between biodiversity effects across trophic levels. Oecologia 168: 533-548.

15. Tomanek L (2012) Introduction to the symposium "Comparative proteomics of environmental and pollution stress". Integr Comp Biol 52: 622-625.

16. Meistertzheim AL, Tanguy A, Moraga D, Thebault MT (2007) Identification of differentially expressed genes of the Pacific oyster *Crassostrea gigas* exposed to prolonged thermal stress. FEBS J 274: 6392-6402.

17. Lang RP, Bayne CJ, Camara MD, Cunningham C, Jenny MJ, et al. (2009) Transcriptome profiling of selectively bred Pacific oyster *Crassostrea gigas* families that differ in tolerance of heat shock. Mar Biotechnol (NY) 11: 650-668.

18. Lockwood BL, Sanders JG, Somero GN (2010) Transcriptomic responses to heat stress in invasive and native blue mussels (genus *Mytilus*): molecular correlates of invasive success. J Exp Biol 213: 3548-3558.

19. Tomanek L, Zuzow MJ (2010) The proteomic response of the mussel congeners *Mytilus galloprovincialis* and *M. trossulus* to acute heat stress: implications for thermal tolerance limits and metabolic costs of thermal stress. J Exp Biol 213: 3559-3574.

20. Fields PA, Cox KM, Karch KR (2012) Latitudinal variation in protein expression after heat stress in the salt marsh mussel *Geukensia demissa*. Integr Comp Biol 52: 636-647.

21. Negri A, Oliveri C, Sforzini S, Mignione F, Viarengo A, et al. (2013) Transcriptional Response of the Mussel *Mytilus galloprovincialis* (Lam.) following Exposure to Heat Stress and Copper. PLoS One 8: e66802.

22. Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity (Edinb) 107: 1-15.

23. Wang R, Li C, Stoeckel J, Moyer G, Liu Z, et al. (2012) Rapid development of molecular resources for a freshwater mussel, *Villosa lienosa* (Bivalvia: Unionidae), using an RNA-seq-based approach. Freshwater Science 31: 695-708.

24. Bai Z, Zheng H, Lin J, Wang G, Li J (2013) Comparative Analysis of the Transcriptome in Tissues Secreting Purple and White Nacre in the Pearl Mussel *Hyriopsis cumingii.* PLoS One 8: e53617.

25. Gavery MR, Roberts SB (2012) Characterizing short read sequencing for gene discovery and RNA-Seq analysis in *Crassostrea gigas*. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 7: 94-99.

26. Zhao X, Wang Q, Jiao Y, Huang R, Deng Y, et al. (2012) Identification of genes potentially related to biomineralization and immunity by transcriptome analysis of pearl sac in pearl oyster *Pinctada martensii*. Marine Biotechnology 14: 730-739.

27. Roberts R, Agius C, Saliba C, Bossier P, Sung Y (2010) Heat shock proteins (chaperones) in fish and shellfish and their potential role in relation to fish health: a review. J Fish Dis 33: 789-801.

28. Wu C (1995) Heat shock transcription factors: structure and regulation. Annual review of cell and developmental biology 11: 441-469.

29. Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends in plant science 9: 244-252.

30. Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 chaperone machines. Cell 92: 351-366.

31. Kregel KC (2002) Invited review: heat shock proteins: modifying factors in physiological stress responses and acquired thermotolerance. Journal of Applied Physiology 92: 2177-2186.

32. Garrido C, Gurbuxani S, Ravagnan L, Kroemer G (2001) Heat shock proteins: endogenous modulators of apoptotic cell death. Biochem Biophys Res Commun 286: 433-442.

33. Mosser DD, Caron AW, Bourget L, Denis-Larose C, Massie B (1997) Role of the human heat shock protein hsp70 in protection against stress-induced apoptosis. Molecular and cellular biology 17: 5317-5327.

34. Muralidharan S, Mandrekar P (2013) Cellular stress response and innate immune signaling: integrating pathways in host defense and inflammation. Journal of leukocyte biology. Journal of leukocyte biology 94: 1-18.

35. Jolly C, Morimoto RI (2000) Role of the heat shock response and molecular chaperones in oncogenesis and cell death. J Natl Cancer Inst 92: 1564-1572.

36. Xu Y, Lindquist S (1993) Heat-shock protein hsp90 governs the activity of pp60v-src kinase. Proceedings of the National Academy of Sciences 90: 7074-7078.

37. Prodromou C, Roe SM, O'Brien R, Ladbury JE, Piper PW, et al. (1997) Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. Cell 90: 65-75.

38. Stebbins CE, Russo AA, Schneider C, Rosen N, Hartl FU, et al. (1997) Crystal structure of an Hsp90–geldanamycin complex: targeting of a protein chaperone by an antitumor agent. Cell 89: 239-250.

39. Whitesell L, Mimnaugh EG, De Costa B, Myers CE, Neckers LM (1994) Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. Proceedings of the National Academy of Sciences 91: 8324-8328.

40. Smith DF, Whitesell L, Nair SC, Chen S, Prapapanich V, et al. (1995) Progesterone receptor structure and function altered by geldanamycin, an hsp90-binding agent. Molecular and cellular biology 15: 6804-6812.

41. Whitesell L, Cook P (1996) Stable and specific binding of heat shock protein 90 by geldanamycin disrupts glucocorticoid receptor function in intact cells. Molecular Endocrinology 10: 705-712.

42. Scheibel T, Buchner J (1998) The Hsp90 complex—a super-chaperone machine as a novel drug target. Biochemical pharmacology 56: 675-682.

43. Schirmer EC, Glover JR, Singer MA, Lindquist S (1996) HSP100/Clp proteins: a common mechanism explains diverse functions. Trends in biochemical sciences 21: 289-296.

44. Arrigo A-P (2007) The cellular "networking" of mammalian Hsp27 and its functions in the control of protein folding, redox state and apoptosis. Molecular aspects of the stress response: chaperones, membranes and networks: Springer. pp. 14-26.

45. Concannon C, Gorman A, Samali A (2003) On the role of Hsp27 in regulating apoptosis. Apoptosis 8: 61-70.

46. Richards EH, Hickey E, Weber L, Masters JR (1996) Effect of overexpression of the small heat shock protein HSP27 on the heat and drug sensitivities of human testis tumor cells. Cancer Res 56: 2446-2451.

47. Trautinger F, Kokesch C, Herbacek I, Knobler RM, Kindås-Mügge I (1997) Overexpression of the small heat shock protein, hsp27, confers resistance to hyperthermia, but not to oxidative stress and UV-induced cell death, in a stably transfected squamous cell carcinoma cell line. Journal of Photochemistry and Photobiology B: Biology 39: 90-95.

48. Pandey P, Farber R, Nakazawa A, Kumar S, Bharti A, et al. (2000) Hsp27 functions as a negative regulator of cytochrome c-dependent activation of procaspase-3. Oncogene 19: 1975-1981.

49. Charette SJ, Lavoie JN, Lambert H, Landry J (2000) Inhibition of Daxx-mediated apoptosis by heat shock protein 27. Molecular and cellular biology 20: 7602-7612.

50. Ivanina A, Taylor C, Sokolova I (2009) Effects of elevated temperature and cadmium exposure on stress protein response in eastern oysters *Crassostrea virginica* (Gmelin). Aquatic toxicology 91: 245-254.

51. Kammenga J, Arts M, Oude-Breuil W (1998) HSP60 as a potential biomarker of toxic stress in the nematode Plectus acuminatus. Archives of environmental contamination and toxicology 34: 253-258.

52. Sanders B, Martin L (1993) Stress proteins as biomarkers of contaminant exposure in archived environmental samples. Science of the total environment 139: 459-470.

53. Sanders BM, Hope C, Pascoe VM, Martin LS (1991) Characterization of the stress protein response in two species of *Collisella limpets* with different temperature tolerances. Physiological Zoology: 1471-1489.

54. Sanders BM, Martin LS, Nelson WG, Phelps DK, Welch W (1991) Relationships between accumulation of a 60 kDa stress protein and scope-for-growth in *Mytilus edulis* exposed to a range of copper concentrations. Marine Environmental Research 31: 81-97.

55. Zhang G, Fang X, Guo X, Li L, Luo R, et al. (2012) The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490: 49-54.

56. Lesser MP, Bailey MA, Merselis DG, Morrison JR (2010) Physiological response of the blue mussel *Mytilus edulis* to differences in food and temperature in the Gulf of Maine. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology 156: 541-551.

57. Hofmann GE (2005) Patterns of Hsp gene expression in ectothermic marine organisms on small to large biogeographic scales. Integr Comp Biol 45: 247-255.

58. Ioannou S, Anestis A, Pörtner HO, Michaelidis B (2009) Seasonal patterns of metabolism and the heat shock response (HSR) in farmed mussels *Mytilus galloprovincialis*. J Exp Mar Bio Ecol 381: 136-144.

59. Tomanek L (2010) Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. J Exp Biol 213: 971-979.

60. Dimitriadis VK, Gougoula C, Anestis A, Pörtner HO, Michaelidis B (2012) Monitoring the biochemical and cellular responses of marine bivalves during thermal stress by using biomarkers. Marine Environmental Research 73: 70-77.

61. Tomanek L, Somero GN (1999) Evolutionary and acclimation-induced variation in the heat-shock responses of congeneric marine snails (genus *Tegula*) from different thermal habitats: implications for limits of thermotolerance and biogeography. Journal of Experimental Biology 202: 2925-2936.

62. Norris CE, Hightower LE (2002) Discovery of two distinct small heat shock protein (HSP) families in the desert fish *Poeciliopsis*. Small Stress Proteins: Springer. pp. 19-35.

63. Tomanek L (2005) Two-dimensional gel analysis of the heat-shock response in marine snails (genus *Tegula*): interspecific variation in protein expression and acclimation ability. Journal of Experimental Biology 208: 3133-3143.

64. Lydeard C, Cowie RH, Ponder WF, Bogan AE, Bouchet P, et al. (2004) The global decline of nonmarine mollusks. BioScience 54: 321-330.

65. Strayer DL, Downing JA, Haag WR, King TL, Layzer JB, et al. (2004) Changing perspectives on pearly mussels, North America's most imperiled animals. BioScience 54: 429-439.

66. Vaughn CC, Nichols SJ, Spooner DE (2008) Community and foodweb ecology of freshwater mussels. Journal of the North American Benthological Society 27: 409-423.

67. Tomanek L (2012) Environmental proteomics of the mussel *Mytilus*: implications for tolerance to stress and change in limits of biogeographic ranges in response to climate change. Integr Comp Biol 52: 648-664.

68. Roberts RJ, Agius C, Saliba C, Bossier P, Sung YY (2010) Heat shock proteins (chaperones) in fish and shellfish and their potential role in relation to fish health: a review. J Fish Dis 33: 789-801.

69. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, et al. (2009) ABySS: a parallel assembler for short read sequence data. Genome Res 19: 1117-1123.

70. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644-652.

71. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658-1659.

72. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9: 868-877.

73. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11: R25.

74. Baggerly KA, Deng L, Morris JS, Aldaz CM (2003) Differential expression in SAGE: accounting for normal between-library variation. Bioinformatics 19: 1477-1483.

75. Bauer S, Grossmann S, Vingron M, Robinson PN (2008) Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration. Bioinformatics 24: 1650-1651.

76. Grossmann S, Bauer S, Robinson PN, Vingron M (2007) Improved detection of overrepresentation of Gene-Ontology annotations with parent–child analysis. Bioinformatics 23: 3024-3031.

77. Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30: e36-e36.

78. Peatman E, Li C, Peterson B, Straus DL, Bradley Farmer D, et al. (2013) Basal polarization of the mucosal compartment in Flavobacterium columnare susceptible and resistant channel catfish *Ictalurus punctatus*. Mol Immunol http://dx.doi.org/10.1016

/j.molimm.2013.04.014..

79. Li C, Zhang Y, Wang R, Lu J, Nandi S, et al. (2012) RNA-seq analysis of mucosal immune responses reveals signatures of intestinal barrier disruption and pathogen entry following *Edwardsiella ictaluri* infection in channel catfish, *Ictalurus punctatus*. Fish & Shellfish Immunology 32: 816-827.
80. Sun F, Peatman E, Li C, Liu S, Jiang Y, et al. (2012) Transcriptomic signatures of attachment, NF-κB suppression and IFN stimulation in the catfish gill following columnaris bacterial infection. Developmental & Comparative Immunology 38: 169-180.
81. Wit P, Palumbi SR (2012) Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol Ecol 22: 2884-2897.
82. Takeuchi T, Kawashima T, Koyanagi R, Gyoja F, Tanaka M, et al. (2012) Draft genome of the pearl oyster *Pinctada fucata*: a platform for understanding bivalve biology. DNA Res 19: 117-130.
83. Archambault JM, Gregory Cope W, Kwak TJ (2013) Burrowing, byssus, and biomarkers: behavioral and physiological indicators of sublethal thermal stress in freshwater mussels (Unionidae). Marine and Freshwater Behaviour and Physiology: 1-22.
84. Xu W, Faisal M (2009) Development of a cDNA microarray of zebra mussel (*Dreissena polymorpha*) foot and its use in understanding the early stage of underwater adhesion. Gene 436: 71-80.
85. Xu W, Faisal M (2010) Gene expression profiling during the byssogenesis of zebra mussel (*Dreissena polymorpha*). Mol Genet Genomics 283: 327-339.
86. Gantayet A, Ohana L, Sone ED (2013) Byssal proteins of the freshwater zebra mussel, *Dreissena polymorpha.* Biofouling 29: 77-85.

87. Spicer Jr JF, Sobieski RJ, Crupper SS (2007) Testing a non-lethal DNA isolation technique for freshwater mussels that is compatible with polymerase chain reaction (PCR) based studies. Transactions of the Kansas Academy of Science 110: 30-34.

88. Somero GN (1995) Proteins and temperature. Annu Rev Physiol 57: 43-68.

89. Stromer T, Ehrnsperger M, Gaestel M, Buchner J (2003) Analysis of the interaction of small heat shock proteins with unfolding proteins. J Biol Chem 278: 18015-18021.

90. Mymrikov EV, Seit-Nebi AS, Gusev NB (2011) Large potentials of small heat shock proteins. Physiol Rev 91: 1123-1159.

91. Bruskov VI, Masalimov Zh K, Chernikov AV (2002) Heat-induced generation of reactive oxygen species in water. Dokl Biochem Biophys 384: 181-184.

92. Arbogast S, Ferreiro A (2010) Selenoproteins and protection against oxidative stress: selenoprotein N as a novel player at the crossroads of redox signaling and calcium homeostasis. Antioxid Redox Signal 12: 893-904.

93. Kourtis N, Tavernarakis N (2011) Cellular stress response pathways and ageing: intricate molecular relationships. EMBO J 30: 2520-2531.

94. Zhang L, Li L, Zhang G (2011) Gene discovery, comparative analysis and expression profile reveal the complexity of the *Crassostrea gigas* apoptosis system. Dev Comp Immunol 35: 603-610.

95. Varotto L, Domeneghetti S, Rosani U, Manfrin C, Cajaraville MP, et al. (2013) DNA damage and transcriptional changes in the gills of *Mytilus galloprovincialis* exposed to nanomolar doses of combined metal salts (Cd, Cu, Hg). PLoS One 8: e54602.

96. Syed NA, Andersen PL, Warrington RC, Xiao W (2006) Uev1A, a ubiquitin conjugating enzyme variant, inhibits stress-induced apoptosis through NF-kappaB activation. Apoptosis 11: 2147-2157.

97. Bertrand MJ, Milutinovic S, Dickson KM, Ho WC, Boudreault A, et al. (2008) cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. Mol Cell 30: 689-700.

98. Stempin CC, Chi L, Giraldo-Vela JP, High AA, Hacker H, et al. (2011) The E3 ubiquitin ligase mind bomb-2 (MIB2) protein controls B-cell CLL/lymphoma 10 (BCL10)-dependent NF-kappaB activation. J Biol Chem 286: 37147-37157.

99. Oudshoorn D, Versteeg GA, Kikkert M (2012) Regulation of the innate immune system by ubiquitin and ubiquitin-like modifiers. Cytokine Growth Factor Rev 23: 273-282.

100. Schmukle AC, Walczak H (2012) No one can whistle a symphony alone–how different ubiquitin linkages cooperate to orchestrate NF-κB activity. Journal of Cell Science 125: 549-559.

101. Mitta G, Vandenbulcke F, Roch P (2000) Original involvement of antimicrobial peptides in mussel innate immunity. FEBS Lett 486: 185-190.

102. Gueguen Y, Herpin A, Aumelas A, Garnier J, Fievet J, et al. (2006) Characterization of a defensin from the oyster *Crassostrea gigas*. Recombinant production, folding, solution structure, antimicrobial activities, and gene expression. J Biol Chem 281: 313-323.

103. Xu W, Faisal M (2010) Defensin of the zebra mussel (*Dreissena polymorpha*): molecular structure, in vitro expression, antimicrobial activity, and potential functions. Mol Immunol 47: 2138-2147.

104. Doshi BM, Hightower LE, Lee J (2010) HSPB1, actin filament dynamics, and aging cells. Ann N Y Acad Sci 1197: 76-84.

105. Cummings V, Hewitt J, Van Rooyen A, Currie K, Beard S, et al. (2011) Ocean acidification at high latitudes: potential effects on functioning of the Antarctic bivalve *Laternula elliptica*. PLoS One 6: e16069.

106. Cheng HY, Obrietan K (2006) Dexras1: shaping the responsiveness of the circadian clock. Semin Cell Dev Biol 17: 345-351.