Abstract

Title of Dissertation:	DEVELOPMENT OF GENE EXPRESSION-BASED BIOMARKERS OF EXPOSURE TO METALS AND PESTICIDES IN THE FRESHWATER AMPHIPOD *HYALELLA AZTECA*
	Ryan Christopher Gott Doctor of Philosophy, 2016
Dissertation Co-directed By:	Professor William O. Lamp Department of Entomology
	Professor David J. Hawthorne Department of Entomology

Ecological risk assessment (ERA) is a framework for monitoring risks of exposure and adverse effects of environmental stressors to populations or communities of interest. One tool of ERA is the biomarker, which is a characteristic of an organism that reliably indicates exposure to or effects of a stressor like chemical pollution. Traditional biomarkers which rely on characteristics at the tissue level and higher often detect only acute exposures to stressors. Sensitive molecular biomarkers may detect lower stressor levels than traditional biomarkers, which helps inform risk mitigation and restoration efforts before populations and communities are irreversibly affected. In this study I developed gene expression-based molecular biomarkers of exposure to metals and

insecticides in the model toxicological freshwater amphipod *Hyalella azteca*. My goals were to not only create sensitive molecular biomarkers for these chemicals, but also to show the utility and versatility of *H. azteca* in molecular studies for toxicology and risk assessment. I sequenced and assembled the *H. azteca* transcriptome to identify reference and stress-response gene transcripts suitable for expression monitoring. I exposed *H. azteca* to sub-lethal concentrations of metals (cadmium and copper) and insecticides (DDT, permethrin, and imidacloprid). Reference genes used to create normalization factors were determined for each exposure using the programs BestKeeper, GeNorm, and NormFinder. Both metals increased expression of a nuclear transcription factor (*Cnc*), an ABC transporter (*Mrp4*), and a heat shock protein (*Hsp90*), giving evidence of general metal exposure signature. Cadmium uniquely increased expression of a DNA repair protein (*Rad51*) and increased *Mrp4* expression more than copper (7-fold increase compared to 2-fold increase). Together these may be unique biomarkers distinguishing cadmium and copper exposures. DDT increased expression of *Hsp90*, *Mrp4*, and the immune response gene *Lgbp*. Permethrin increased expression of a cytochrome P450 (*Cyp2j2*) and decreased expression of the immune response gene *Lectin-1*. Imidacloprid did not affect gene expression. Unique biomarkers were seen for DDT and permethrin, but the genes studied were not sensitive enough to detect imidacloprid at the levels used here. I demonstrated that gene expression in *H. azteca* detects specific chemical exposures at sub-lethal concentrations, making expression monitoring using this amphipod a useful and sensitive biomarker for risk assessment of chemical exposure.

DEVELOPMENT OF GENE EXPRESSION-BASED BIOMARKERS OF EXPOSURE TO METALS AND PESTICIDES IN THE FRESHWATER AMPHIPOD *HYALELLA AZTECA*

By

Ryan Christopher Gott

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:

Professor William O. Lamp, Co-chair
Professor David J. Hawthorne, Co-chair
Professor Jian Wang
Professor Kelly A. Hamby
Professor William W. Bowerman, Dean's Representative

Dedication

I dedicate this work to those closest to me that helped me to keep going and see everything through.

To mom: I have never had a barrier to pursuing anything that has interested me my whole life thanks to you. I know how hard you worked for the three of us growing up to allow that. I would not be who I am today without you. Thank you for everything.

To Ashton: I feel so lucky, blessed, to have met you. I so look forward to our future, to moving somewhere new, to everything. Thank you for all your support during the past 3 years. You mean the world to me.

Acknowledgements

I first want to thank my advisors, Bill Lamp and Dave Hawthorne. Bill allowed me to meander among several potential research interests and projects (which I may have done more than I should have) rather than directing me down the pre-set path of an existing project. The ability to pursue my own scientific interests ultimately made all the difference in my enthusiasm for my project and my development as a scientist. I truly believe I could not have had a better graduate experience anywhere other than the Lamp Lab. Dave has provided me with the utmost enthusiasm for my project, conversations and ideas critical to my success and research, and of course the means to sequence the transcriptome that laid the foundation for all my research. Having Bill and Dave as co-advisors is an honor. I am grateful for all the time, support, and mentorship they have given.

I would also like to thank the members of my advisory committee, both past and present. I thank my current members, Dr. Jian Wang, Dr. Kelly Hamby, and Dr. William Bowerman for their willingness to serve on my committee, their advice, and their commitment towards my successful graduate experience. I also thank past committee members Dr. Angus Murphy, Dr. Wendy Hillwalker, and Dr. Qin Wang. They each played a role in shaping and focusing my scientific interests and research project.

No graduate degree can be completed without moral support from peers as well. I'd like to thank the members of the Lamp Lab, past and present, for their advice, comments on writings and presentations, donuts, smoothies, and an overall wonderful 5 years: Rebecca Wilson, Jessica Grant, Rebecca Eckert, Jennifer Jones, Elanor Spadafora,

Alan Leslie, Dilip Venugopal, Bridget DeLay, Robert Smith, Alison Post, Claire Hirt, Sadia Naseem, Claire Weber, and Sarah Au.

I would like to acknowledge the funding that has supported my work and me for the past 5 years: the Department of Entomology Gahan Fellowship, the Maryland Water Resources Research Center Graduate Summer Fellowship, the Graduate School Summer Research Fellowship, the Charlie Mitter Travel Award, and the Maryland Agriculture Experimental Station Grant.

Graduate school makes friends forged in fire. I want to deeply thank my friends for their support; advice; and good times spent at Town Hall, in Atlantic City, with board games, going camping, collecting insects, playing DnD, watching Bad Bug Movies, and just wasting time hanging out and recuperating from life. I know that making it through these years was only possible with their support. Grace Kunkel and Ashley Jones, you are two of my favorite people. I can't wait to see where life keeps taking us. Crystal McEwen, we have grown to be fast friends, and I can't express how much your friendship means in my life both now and looking to the future. And in no particular order, Peter Coffey, Chris Taylor, Nathalie Steinhauer, and Emily Zobel: I am so glad we are such a close group of friends and that we are all making it through graduate school together. Finally to my life-long friend Colleen Nicholls: I don't know where I would be now without you.

Table of Contents

List of Tables

List of Figures

Chapter 1: Transcriptome analysis of the model freshwater amphipod *Hyalella azteca*

Abstract

Ecological risk assessment (ERA) uses multiple lines of evidence to monitor environmental health and prioritize ecological restoration efforts. The effects of environmental contaminants on organisms are a major source of data for these lines of evidence. To examine effects of toxins on the model toxicological organism *Hyalella azteca*, I sequenced the transcriptome of this freshwater amphipod. A total of 16,537 full transcripts with an N50 length of 2,156 nt remained following assembly and trimming of the transcriptome. Many of the 7,915 (47.86%) successfully annotated transcripts showed high similarity to other crustaceans, such as *Daphnia pulex* and *Penaeus monodon*, or other arthropods, such as *Tribolium castaneum* and *Stegodyphus mimosarum*. Many stress response and reference gene transcripts with potential for use as biomarkers in ERA were identified after annotation of the transcriptome. PCR products of common housekeeping gene transcripts matched those expected in terms of length and sequence for 9 of 10 genes investigated. This is the first NextGen RNA-seq transcriptome for *H. azteca* and as such will serve as a rich source of gene transcript sequence and expression data for use in both ERA and basic and applied research.

Introduction

The continuous production and discharge of pollutants by human activity drives the need to monitor the health of the environment. Ecological risk assessment (ERA) is the main process used for monitoring ecosystem health and prioritizing protection and restoration efforts (Figure 1.1). ERA investigates an ecosystem's risks of suffering adverse effects from particular stressors, including chemical, physical, and biological pollutants (U.S. EPA, 1998). These risks are often broadly defined or calculated as the following: (Exposure to stressor) x (Effect of stressor) = Risk of adverse effect. In other words, if there is no exposure to or no effect of a stressor, there is no risk. Determination of potential exposures and effects is one purpose of ERA.

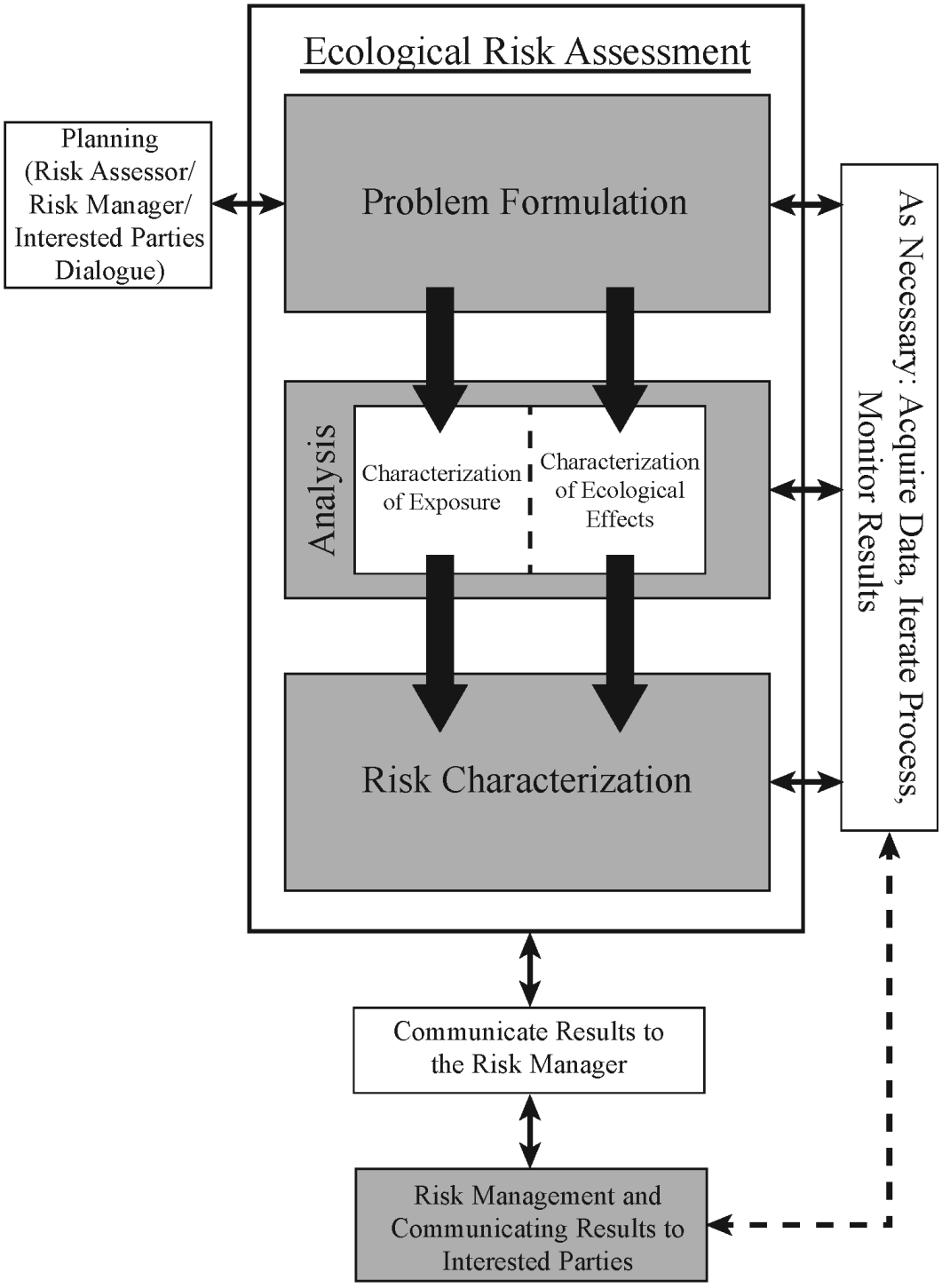

Figure 1.1: Framework for ecological risk assessment (ERA) (figure based on Figure 1-1 in U.S. EPA, 1998)

The ERA process, outlined in Figure 1.1, begins with problem formulation. During this stage the problem is clearly defined; assessment endpoints, or the components of the ecosystem being evaluated for risk, are chosen; and all available relevant information about potential stressors, the assessment endpoints, and the system under assessment are integrated to form a conceptual model (U.S. EPA, 1998). Planning for testing of confirmed or candidate toxin exposure and effects also takes place. Analysis, the second stage of ERA, consists of assays to create profiles of exposure risks and ecological effects. The final stage of ERA, risk characterization, integrates all obtained data with the uncertainties and limitations of the data to create a risk description. A risk description contains an evaluation of all lines of evidence relating stressors to assessment endpoints (U.S. EPA, 1998). The true significance of any adverse effects on assessment endpoints is determined by the risk assessor. Stakeholders and risk managers then make decisions on needed risk mitigation efforts, restoration efforts, or potentially revisiting and revising portions of the ERA for iterations of more focused testing and analysis. It is important to note is that ERA is iterative at nearly all points in the process, taking into account new findings at each stage and incorporating communication among all parties.

Characterization of exposure and ecological effects integrates expertise from many disciplines to obtain multiple lines of evidence in both field and lab settings (Figure 1.2). Field-based lines of evidence are diverse, including qualitative or quantitative descriptions of stressor sources, distribution, and transport; (e.g., Meij, 1991; Carpi, 1997; Lundstedt et al., 2007); community richness or diversity sampling, as is commonly done with macroinvertebrates (e.g., Colas et al., 2014); comparisons of an ecosystem's

community with existing indices like the Index of Biotic Integrity (Karr, 1981); tissue sub-sampling from assessment endpoint organisms for direct measurement of a stressor or for measurement of an effect of a stressor (e.g., Evers et al., 2011); and sampling of environmental media like water, soil, and sediment (e.g., Hou et al., 2013). Lab-based lines of evidence may include analysis of structure-activity relationships (SARs) (e.g., Lindel et al., 2000), quantitative structure-activity relationships (QSARs) (e.g., Altenburger et al., 2003), large-scale data base compilation and analysis (e.g., Ma et al., 2013), and toxicity testing for development of dose-response relationships or evaluation of sub-lethal effects (e.g., MacDonald et al., 2011). All of these biomarkers can contribute to a description of a given assessment endpoint's risk of being exposed to or of experiencing adverse effects from a potential toxin.

Many lines of evidence involve the measurement substances in or characteristics of cells, tissues, organs, or whole organisms that change in response to stressors (Forbes et al., 2006). These are referred to as biomarkers. Organisms or tissue samples collected in the field are processed in the lab for measurement of biomarkers. The levels or changes in certain biomarkers can reveal to what stressors the organism is exposed in its habitat. Similarly, samples from controlled stressor exposures conducted under laboratory conditions allow the association of stressors to specific biomarkers. Biomarkers commonly used in ERA include gene expression, protein titers, histopathological assessments, life history traits like survival and fecundity, behavior, and gross organismal malformations. Changes in these biomarkers in an exposed organism in relation to a reference environment or control indicate stressor exposure, effects, or both.

Biomarkers are common in freshwater toxicity testing and ERA. Specific uses include life history traits in *Daphnia magna* as a marker of polycyclic aromatic hydrocarbon (PAH) exposure, with survival and reproduction decreasing with increasing PAH exposure (Feldmannova, 2006). Glutathione-*S*-transferase, acetylcholinesterase, and catalase enzyme activities in mussels (*Mytilus edulis*) are used as indicators of general ecological health (Beliaff and Burgeot, 2002). These protein activities should be at levels similar to those in clean reference sites, with aberrations attributable to environmental stress. Histological abnormalities, single-strand DNA breaks, and stress-response gene induction increase in Senegalese sole (*Solea senegalensis*) after exposure to sediment with multiple contaminating chemical stressors (Costa et al., 2012; Goncalves et al., 2013). These biomarkers taken together provide an overall picture of environmental health rather than indicating exposure to a specific stress. Metallothionein protein titers in mussels, crabs, and annelids are reliable monitored as biomarkers of metal exposure (Amiard et a., 2006). Increasing concentrations of many different metals can induce transcription of metallothioneins. Changes in expression of sex-differentiation associated genes in fathead minnow (*Pimephales promelas*) occurs after expose to endocrine disrupting compounds (Leet et al., 2015). These biomarkers indicate exposure of non-target vertebrates to synthetic hormones, as well as the sex-altering effects of these hormones. Biomarkers can be examined at the large-scale "omic" level as well, such as proteomic analysis of *D. magna* after exposure to arsenic and cadmium as just one example (Le et al., 2013). There are many more biomarkers ranging over a huge diversity of taxa for an equally large number of stressors.

Biomarkers inform management decisions related to environmental and human health. Action is only taken if significant risk of exposure and adverse effects from a given stressor exists for assessment endpoints (U.S. EPA, 1998). Applying biomarkers to regulatory decisions is more difficult. Transitioning biomarkers from academic and management use to regulatory use is described in four components by Handy et al., 2003 as the following:

1. International agreement on how biomarkers should be incorporated into a regulatory framework,
2. Development of standard operating protocols (SOPs) for testing of biomarkers,
3. Validation of SOPs across testing laboratories;
4. Modification of legislation to include biomarkers.

The inherent variability in biomarkers, especially in relation to changing ambient environmental conditions, makes it challenging to describe thresholds for stressors that can be applied broadly (Handy et al., 2003). This highlights the importance of investigating not just the changes in a given biomarker from a stressor, but also how these changes change themselves spatially and temporally and the effects of potential confounding factors on the biomarker. Standardization of this testing would provide consistent evidence that can be more easily incorporated in regulatory law and decisions.

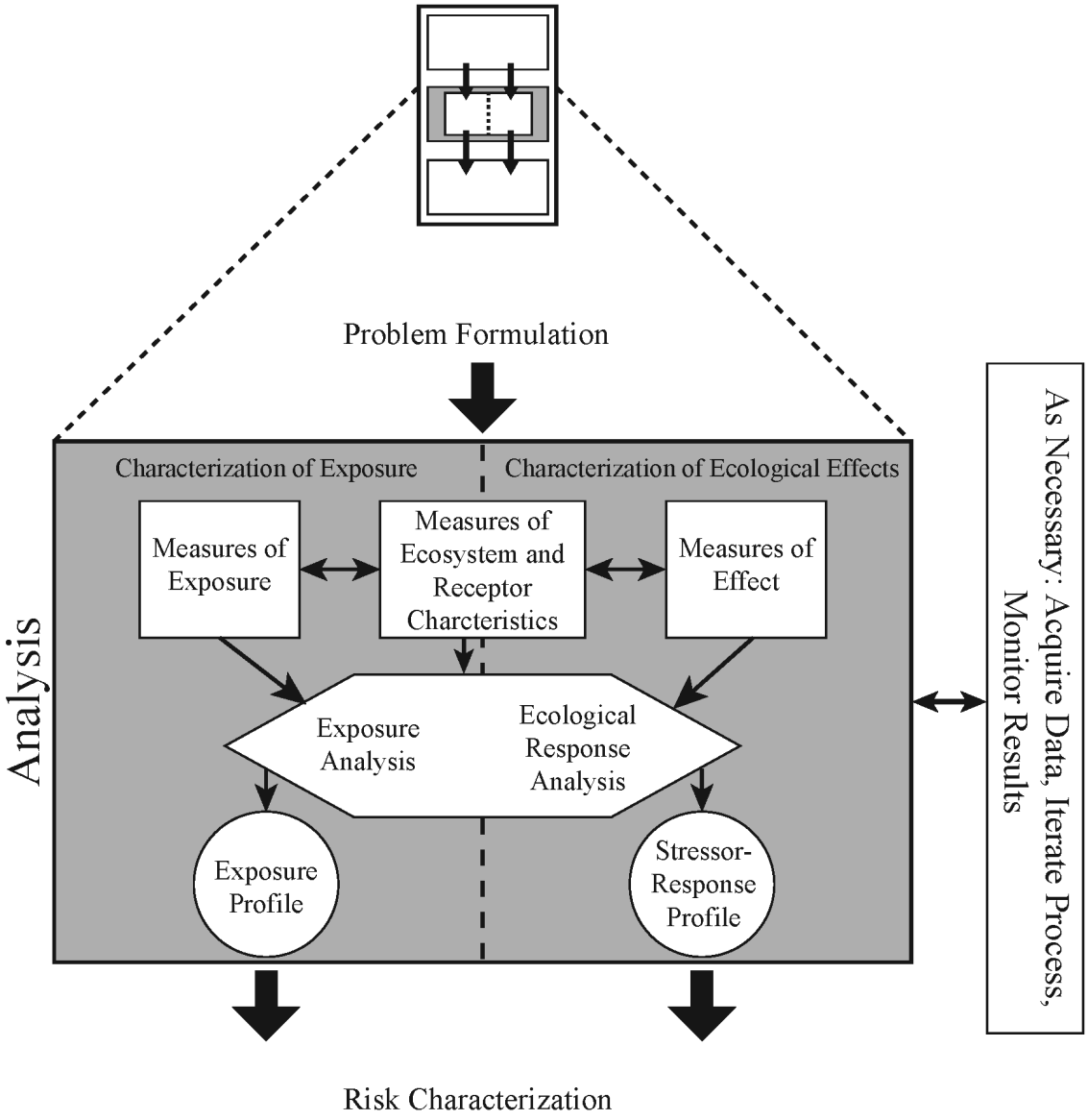

Figure 1.2: Details of the analysis phase of ERA (figure based on Figure 4-1 in U.S. EPA, 1998)

Hyalella azteca (Saussure, 1858) (Malacostraca: Amphipoda: Hyalellidae) (Lowry and Myers, 2013) is a freshwater amphipod and model organism for toxicology. *H. azteca* is common across North America into Central America (Cooper, 1965). *H. azteca* lives on the benthos and is omnivorous, feeding on any variety of decaying material and occasionally living plant tissue (Hargrave, 1970). Young pass through seven

instars before becoming reproductively mature and continue to grow indeterminately (Strong, 1972). *H. azteca* was recently revealed to likely be diverging cryptic ecomorphs across its expansive range, instigating redescription of the species (Duan et al., 1997; Duan et al., 2000; Gonzalez and Watling, 2002; Wellborn et al., 2005).

H. azteca is used in many freshwater and sediment toxicity tests, making it an important species for ERA, especially for characterization of exposure and effects during the analysis phase (U.S. EPA, 2000; Environment Canada, 2013). While we know much about changes in *H. azteca* life history parameters like growth, reproduction, behavior, and survival in response to many stressors, we know nearly nothing about potential molecular biomarkers. The small amount of previous molecular work on *H. azteca* focuses on monitoring protein titers or activities (e.g., Gomez-Olivan et al., 2012).

Life history biomarkers are relatively insensitive compared to molecular level responses, reliably indicating only large exposure levels that already could have irreversible ecological impacts. In zebrafish (*Danio rerio*) exposed to the insecticide permethrin, significant changes in gene expression occurred after a 7-day exposure to a concentration an order of magnitude lower than the 4-day LC50 (Jin et al., 2009; Zhang et al., 2010). *Daphnia pulex* exposed to only LC01 cadmium concentrations show significant upregulation of metal stress-response genes (Shaw et al., 2007). *D. magna* exposed to cadmium at the LC0.5 concentration also displayed significant changes in expression of multiple genes (Soetaert et al., 2007). Nearly 250 genes changed in expression in the amphipod *Melita plumulosa* at sublethal copper exposures below the EC10 value for reproduction (Hook et al., 2014). Clearly at low exposure levels

monitoring only life history can lead to incorrect conclusions about the true level of exposure and effects occurring in an assessment endpoint. Expression level changes can indicate stressors at magnitudes much lower than those detected by survival, histology, or other life history or tissue-level biomarkers.

Since our current knowledge is limited largely to life history parameter changes in *H. azteca*, furthering our molecular insight in this amphipod is a critical goal in advancing sensitive characterization of exposure and effects during the ERA process. Transcriptome analysis previously proved useful for toxicological studies in the crustaceans *D. magna* (Connon et al., 2008; Watanabe et al., 2008) and *Tigriopus japonicas* (Ki et al., 2009) as well as the mussel *Mytilus galloprovincialis* (Dondero et al., 2011). The work presented here describes the sequencing and assembling of the *H. azteca* transcriptome to identify genes that can be used for gene expression analysis following exposure to environmentally relevant pollutants.

Developing gene expression biomarkers in *H. azteca* poses a challenge because until very recently little genetic information outside ribosomal RNA and cytochrome oxidase genes (e.g., Major et al., 2013) was available for *H. azteca* despite its importance as a model toxicological organism. These genes, though useful for phylogenetic studies (Duan et al., 1997; Duan et al., 2000; Witt and Hebert, 2000), are ineffective as tools for toxicity assessment. Fortunately *H. azteca* is currently experiencing a genetic renaissance. The i5K Genome Sequencing Initiative for Insects and Other Arthropods sequenced the *H. azteca* genome. Additionally some microarray expression analysis has assessed insecticide resistance (Weston et al., 2013) and toxicity of zinc oxide

nanoparticles (Poynton et al., 2013), revealing a first glimpse into the genetic workings of this amphipod. However known problems with microarray work, including reliability and sensitivity, make it sub-optimal compared to the new technology of RNA-seq (Draghici et al., 2006; Wang et al., 2009; Marioni et al., 2008; Mantione et al., 2014).

To identify the genes actively transcribed in un-stressed adult *H. azteca*, I sequenced and assembled a large percentage of the transcriptome of the most commonly used U.S. Laboratory strain of *H. azteca* using Illumina RNA-seq technology. I validated the assembly using PCR and primers developed for 10 common "housekeeping" genes identified from the transcriptome. My goal in assembling the *H. azteca* transcriptome is to identify "housekeeping" and stress-response gene transcripts to use in gene expression monitoring to develop biomarkers of exposure. These gene expression-based biomarkers can be used as sensitive indicators of exposure to stressors. Molecular biomarkers will be useful during ERA, whether the assessment endpoint is *H. azteca* or *H. azteca* is used as a surrogate species for testing contamination of environmental samples. Detecting sublethal levels of contamination will inform risk management decisions to prevent higher ecological-scale effects.

The overarching goals of this dissertation work are to develop *H. azteca* as a molecular toxicology model species and to demonstrate the utility and versatility of molecular biomarkers in *H. azteca*. The description and analysis of the transcriptome of *H. azteca* constitute Chapter 1 of this dissertation. Chapter 2 consists of developing housekeeping genes identified here in Chapter 1 into quality control references for monitoring gene expression changes in *H. azteca* after exposure to sub-lethal

concentrations of the metals cadmium and copper. Chapter 3 describes measuring changes in the expression of stress-response genes as a biomarker of exposure to the same metals as Chapter 2. Finally Chapter 4 demonstrates the versatility of gene expression as a biomarker in *H. azteca* by applying the same concepts from Chapters 2 and 3 to exposure studies involving the pesticides DDT, permethrin, and imidacloprid.

Materials and Methods

RNA extracted from two samples of 20 lab culture *H. azteca* were sequenced using RNA-seq technology on an Illumina HiSeq1500 machine at the University of Maryland Sequencing Core. Raw reads were assessed for quality and assembled into transcripts. Transcripts were annotated using BLAST2GO software and searched for common "housekeeping" reference genes and stress-response genes. Potential reference genes were amplified using PCR and sequenced to validate the transcriptome assembly.

Research animals

H. azteca (Figure 1.3) of the U.S. Laboratory strain were purchased from Aquatic BioSystems (Fort Collins, CO) in 2012. This commercially available strain is commonly used in toxicity testing. The original source of this strain of *H. azteca* is not known, but it is most related to populations in Oklahoma (Major et al., 2013). The amphipod colonies (15 gallon glass aquarium, 24" x 12" x 12") contained approximately 8 liters of tap water treated with chemical conditioner to remove ammonia, chloramines, metals, and other contaminants; air stones; and 4" x4" squares of synthetic mesh fabric substrate. Treated tap water, rather than reconstituted or synthetic freshwater, was used to maintain colonies as it supported adequate survival and reproduction of the amphipods. Reconstituted water

is shown to support poor survival of *H. azteca* in long-term chronic exposure tests of up to 28 days and therefore would be inappropriate for maintaining colonies for years at a time (Kemble et al., 1998; McNulty et al., 1999). Aquaria were in an environmental chamber maintained at 26°C with lights set to a 16:8 light:dark cycle. Colonies were fed crushed Tetramin brand tropical fish flake food (guaranteed analysis 47% crude protein from fish meal and shrimp meal, 10% crude fat, 3% max crude fiber, 6% max moisture, 1% phosphorus) three times weekly. The *H. azteca* colony has been maintained without addition of individuals since establishment of the colony in 2012, creating more genetically similar individuals.

Figure 1.3: *H. azteca* lateral (left) and dorsal (right) views

H. azteca transcriptome sequencing

I collected four samples of 20 *H. azteca* of mixed age and sex from the laboratory colony using a number 60 standard sieve, stored samples in 200 µL RNAlater (Life Technologies), and froze them in liquid nitrogen. Before processing samples were briefly

thawed, and surface RNAlater was removed. I then quickly placed samples in a cooled mortar with a small volume of liquid nitrogen and crushed them into a fine powder using a pestle until all the liquid nitrogen was evaporated. I then added Trizol reagent (500 μL) (Life Technologies) warmed to 35°C to the mortar and mixed with the powder to halt RNA degradation as the sample thawed. The mixture was then moved into a 1.5 mL microcentrifuge tube and incubated at room temperature for five minutes with frequent vortexing. I then added chloroform (100 μL), followed by 15 seconds of vortexing, 1 minute of room-temperature incubation, and another 15 seconds of vortexing. The mixture was then centrifuged at 12,000 rpm for 10 minutes to separate phases.

The top aqueous phase containing the RNA (about 200 μL) was carefully removed using a micropipette and placed in a new 1.5 mL microcentrifuge tube. The remaining organic phase containing DNA and protein was frozen at -80°C. I then added Buffer RLT from an RNeasy Kit (700 μL) (Qiagen) and 100% molecular-grade ethanol (500 μL) (Sigma-Aldrich) to the 200 μL of aqueous phase, followed by vortexing. I then used the sample in the manufacturer-provided RNeasy Kit RNA extraction protocol starting at step 4. Final elution of RNA was performed with 30 μL of sterile water run through the RNA extraction column twice. I submitted samples to the University of Maryland Institute for Bioscience & Biotechnology Research (IBBR) Sequencing Core for bioanalysis, and the two best samples were used for RNA sequencing (RNA-seq). Illumina library creation was performed by the Sequencing Core using a 200 bp library insert size. Paired-end RNA-seq was performed on an Illumina HiSeq1500 instrument using 100 bp read lengths. The two samples were loaded in different lanes of the flow cell to account for lane effects.

Transcriptome assembly and analysis

I performed transcriptome quality assessment and assembly on sequences using procedures similar to those described by Lenz et al. (2014) on a Macbook Pro personal lap top computer (OS X v.10.9.5, 2.5 GHz Intel Core i5). FastQC v.0.11.2 software (Andrews, 2014a) was used first to check the quality of the sequences and identify probable contamination sequences and over-represented sequences, such as primers and sequencing adapters. FastQ Screen v.0.4.4 (Andrews, 2014b) and Bowtie 2 v.2.2.3 (Langmead and Salzberg, 2012) were used to check for contamination. Paired forward and reverse reads from each sample were checked against genomes downloaded from GenBank and indexed using Bowtie 2 (Table 1.1).

Table 1.1: Indexed genomes used to scan the *H. azteca* transcriptome for possible contamination

Organism	NCBI Accession/RefSeq Assembly ID	BioProject Number
Hyalella azteca	JQDR00000000	PRJNA243935
Daphnia pulex	ACJG00000000	PRJNA12756
Homo sapiens GRCh38	GCF_000001405.26	PRJNA31257
Escherichia coli	GCF_000005845.2	PRJNA57779
Saccharomyces cerevisiae	GCF_000146045.2	PRJNA128

RNA-seq adapters and possible contamination sequences were trimmed out using Trimmomatic v.0.32 (Bolger et al., 2014). I aligned remaining sequences using TopHat2 v.2.0.9 (Kim et al., 2013) and Bowtie 2 v.2.1.0 within the iPlant Collaborative Discovery Environment (http://www.iplantcollaborative.org/ci/discovery-environment, Goff et al., 2011) using the partially assembled *H. azteca* genome as reference. All default options were used except Phred+33 quality scoring was used in place of Phred+64. I performed sequence assembly using CuffLinks2 v.2.1.1 (Trapnell et al., 2012) within the iPlant

Collaborative Discovery Environment using the unannotated *H. azteca* genome as reference. All default parameters were used.

The .gtf output file from CuffLinks2 was reformatted to a .fasta file using gffread within CuffLinks2 and then fed into the Blast2GO v.3.0.7 desktop software (Conesa et al., 2005) for comparing transcripts against NCBI databases using blastx, mapping, annotation, assignment of gene ontology (GO) terms, and analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa and Goto, 2000). Blasting was performed using a cut off E-value of 10^{-3}. All procedures in Blast2GO were performed with default settings. GO term distribution was determined only for GO terms covered by at least 10% of the total transcripts being analyzed. I performed Benchmarking Universal Single-Copy Orthologs (BUSCO) analysis using assembled transcripts, which provides a picture of the completeness of the transcriptome by searching the transcripts against a database of known arthropod orthologs (http://busco.ezlab.org/, Simao et al., 2015). BUSCO analysis was performed using BUSCO_v1.1b1.py, EMBOSS-6.5.7 (http://emboss.sourceforge.net/download/, Rice et al., 2000), hmmer-3.1b2 (http://hmmer.janelia.org/, Mistry et al., 2013), and ncbi-blast-2.2.31+ (http://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/, Camacho et al., 2015). Codes used in transcriptome assembly and BUSCO analysis may be found in Appendix A.

Transcriptome validation

To validate the transcriptome assembly, I designed primers for 10 common housekeeping gene transcripts based off of the assembled transcriptome using online design tools from IDT DNA and GenScript (Table 1.2; Appendix B). Total RNA was

extracted from two samples of *H. azteca* as described above and quantified using a NanoDrop Lite spectrophotometer (Thermo Scientific). A total of 1 µg of RNA from both samples was reverse transcribed into cDNA using a Transcriptor First Strand cDNA Synthesis Kit (Roche) following the manufacturer's protocol. I then used cDNA in a 20 µL PCR reaction consisting of the following: 10 µL of 2x PCR Master Mix (SydLabs), 1 µL cDNA, 1 µL of 100 µM forward primer, 1 µL of 100 µM reverse primer, 7 µL PCR grade water. PCR was run with the following cycles: 1) 1 cycle of 95ºC for 2 minutes; 2) 25 cycles of 95ºC for 15 seconds, 57ºC for 15 seconds, and 72ºC for 30 seconds; 3) 1 cycle of 72ºC for 7 minutes; 4) 4ºC indefinitely until PCR products were stored at -20ºC. All reactions were paired with a negative control with 1 µL of water added instead of cDNA. I ran PCR products on a 4% agarose gel at 110 V for 40 minutes along with a 10 bp DNA ladder (Life Technologies) and used ethidium bromide and a UV camera to visualize the gel. Products were sent to GENEWIZ, Inc. for Sanger multi-fluorescent dideoxynucleotide chain termination sequencing using the forward primer of each primer set as the sequencing primer. Experimental PCR product sequences were then aligned to their corresponding expected amplicon sequence (Appendix C) and analyzed for percent similarity using EMBOSS Water (Rice et al., 2000).

Table 1.2: Housekeeping gene transcripts identified in the *H. azteca* transcriptome and used for transcriptome assembly validation

Gene	Abbreviation	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Amplicon length (nt)
40s ribosomal protein s23	*RpS23*	CAATGCTGAGT CACGAATCG	TGTTTCGCCTCA ATACCTCTG	127
60s ribosomal protein l10	*RpL10*	GATTTTCGTTTC TGGGAAGTGG	TTAATTAGTGG ACCGTGCTCAG	132
Beta actin	*Act*	GCCATCGCACA GGATATCTAAG	TGTGCTGATTG AGGATTCTGG	140
Glyceraldehyde-3-phosphate dehydrogenase	*Gapdh*	GGATTACCTTC GCAGTACAGTG	TTGCAGTTTAC AGAGGACAGTC	150
Tubulin alpha-1a chain	*αTub*	CCTACCACGAG CAGATTTCAG	AGCGTTGACAT CCTTGGG	149
Matrix metalloproteinase-15-like	*Mmp*	TTGGTTAGGTA CAGTTGTGGG	ATCCTTATCGG CCTCAGTTTAC	146
Syntaxin-5-like isoform x2	*Syx*	CAGCAAACCTA ATGGCGATG	GTAAGCCCCTC CACTGTTG	76
TATA-box-binding protein	*Tbp*	CCATGTTTGTG CTTTCTCCAG	GGCTCTGAAGA CAACCGTATG	149
Ubiquitin	*Ubi*	CATGGTATACG CCCTAATTCCG	GTCCATGTCCA TAGGAACGAC	148
Ubiquitin-conjugating enzyme e2 d2	*Ubc*	TGGGTGTGGGT GTAAAACTC	CAAATGTTGCG GTCAGAAGC	119

Results

H. azteca transcriptome assembly and analysis

Bioanalysis showed extracted RNA to be of good quality (Table 1.3). Samples HA1 and HA3 were chosen for sequencing based on their higher RNA integrity numbers and/or higher RNA concentration. Sequencing of the two samples yielded a combined 63,443,870 raw reads. FastQC showed all raw sequences to be of high quality based on Phred+33 scores (Figure 1.4) and identified 8 and 10 overrepresented sequences of adapters or possible contamination in HA1 and HA3, respectively (Table 1.4). FastQ Screen detected minimal contamination in the samples, with many sequences aligning to either of the two parts of the indexed genome of *H. azteca* (Figure 1.5). The *H. azteca* genome was indexed in two parts due to its state of ongoing assembly and annotation

during assembly of this transcriptome. Trimmomatic removed sequences detected by FastQC. Few sequences fell below the minimum quality requirement during trimming (Table 1.5). 36,546,804 reads remained after alignment, 8,141,792 (22.3%) of which were singletons. After assembly of aligned reads, a total of 16,537 full transcripts (25,968,498 bp) were produced with an N50 length of 2,156 nt. The transcriptome RNA-seq reads can be found in GenBank under BioProject PRJNA277380 (Accession SRP066301).

Table 1.3: Bioanalysis results for all four submitted samples of *H. azteca* RNA

Sample	RNA concentration (ng/μL)	RNA area	RNA integrity number
HA1	221	479	6.1
HA2	1922	4160.1	4.7
HA3	2286	4946.8	5.3
HA4	1941	4200.3	5.3

Table 1.4: Sequences of adapters and contamination identified by FastQC and removed by Trimmomatic in each sample

Sample	Sequence name	Sequence
HA1	PE1	TACACTCTTTCCCTACACGACGCTCTTCCGATCT
	PE1_rc	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTA
	PE2_TruSeq7_1	GATCGGAAGAGCACACGTCTGAACTCCAGTCACGAGTG GATATCTCGTAT
	PE2_TruSeq7_1_rc	ATACGAGATATCCACTCGTGACTGGAGTTCAGACGTGT GCTCTTCCGATC
	PE3_TruSeq7_2	AGATCGGAAGAGCACACGTCTGAACTCCAGTCACGAGT GGATATCTCGTA
	PE3_TruSeq7_2_rc	TACGAGATATCCACTCGTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT
	HA1_1	GAAAAATTTAGTGTAGTCTTTAATATACCTTTTGTCTCA GGGTTTATTTA
	HA1_1_rc	TAAATAAACCCTGAGACAAAAGGTATATTAAAGACTAC ACTAAATTTTTC
HA3	PE1_TruSeq25_1	GATCGGAAGAGCACACGTCTGAACTCCAGTCACACTGA TATATCTCGTAT
	PE1_TruSeq25_1_rc	ATACGAGATATATCAGTGTGACTGGAGTTCAGACGTGT GCTCTTCCGATC
	PE2_TruSeq25_2	AGATCGGAAGAGCACACGTCTGAACTCCAGTCACACTG ATATATCTCGTA
	PE2_TruSeq25_2_rc	TACGAGATATATCAGTGTGACTGGAGTTCAGACGTGTG CTCTTCCGATCT
	SE_primer1	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGAT CTCGGTGGTCGCC
	SE_primer1_rc	GGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCT
	HA3_1	GAAAAATTTAGTGTAGTCTTTAATATACCTTTTGTCTCA GGGTTTATTTA
	HA3_1_rc	TAAATAAACCCTGAGACAAAAGGTATATTAAAGACTAC ACTAAATTTTTC
	HA3_2	CTCGGCTCTACTTTATTATAAAACCAAGCCTATATATTA TTTGACTGGGG
	HA3_2_rc	CCCCAGTCAAATAATATATAGGCTTGGTTTTATAATAA AGTAGAGCCGAG

Table 1.5: Transcript numbers for input raw reads and transcripts remaining after trimming

Output sequence file name	No. of input read pairs	Both surviving (number, %)	Forward only surviving (number, %)	Reverse only surviving (number, %)	Dropped (number, %)
HA1_001	4000000	3651878, 91.3%	178160, 4.45%	99963, 2.5%	69999, 1.75%
HA1_002	4000000	3606936, 90.17%	192705, 4.82%	122117, 3.05%	78242, 1.96%
HA1_003	4000000	3552195, 88.8%	209079, 5.23%	148006, 3.7%	90720, 2.27%
HA1_004	4000000	3568383, 89.21%	201237, 5.03%	146372, 3.66%	84008, 2.1%
HA1_005	4000000	3512791, 87.82%	237217, 5.93%	163896, 4.1%	86096, 2.15%
HA1_006	4000000	3609005, 90.23%	211367, 5.28%	106543, 2.66%	73085, 1.83%
HA1_007	4000000	3561600, 89.04%	193834, 4.85%	160538, 4.01%	84028, 2.1%
HA1_008	2410340	2087212, 86.59%	135589, 5.63%	125777, 5.22%	61762, 2.56%
HA3_001	4000000	3194383, 79.86%	402250, 10.06%	301205, 7.53%	102162, 2.55%
HA3_002	4000000	3122593, 78.06%	416362, 10.41%	337862, 8.45%	123183, 3.08%
HA3_003	4000000	3121193, 78.03%	424071, 10.6%	331496, 8.29%	123240, 3.08%
HA3_004	4000000	3126752, 78.17%	425239, 10.63%	329209, 8.23%	118800, 2.97%
HA3_005	4000000	3108627, 77.72%	453107, 11.33%	318907, 7.97%	119359, 2.98%
HA3_006	4000000	3066906, 76.67%	449466, 11.24%	349372, 8.73%	134256, 3.36%
HA3_007	4000000	3094658, 77.37%	441411, 11.04%	337119, 8.43%	126812, 3.17%
HA3_008	4000000	3150278, 78.76%	411129, 10.28%	329031, 8.23%	109562, 2.74%
HA3_009	1033530	725592, 70.21%	121268, 11.73%	130716, 12.65%	55954, 5.41%

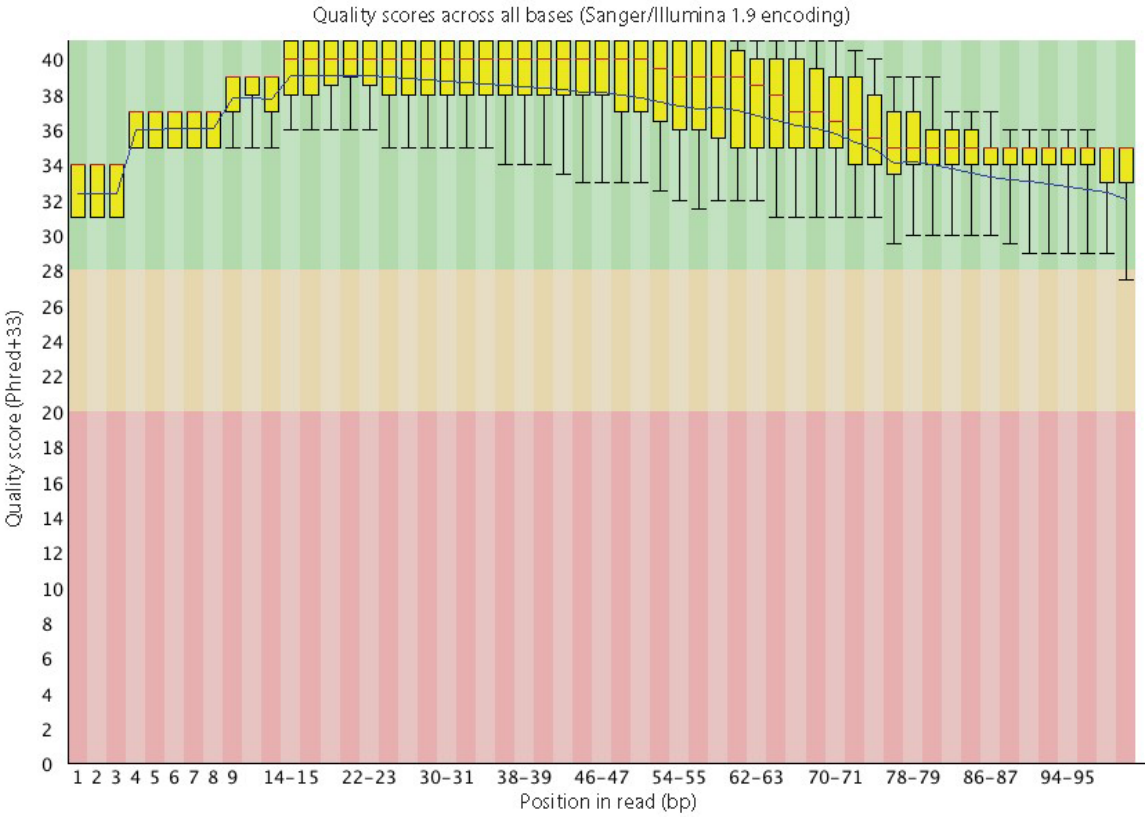

Figure 1.4: An output graph from FastQC representative of those for all reads

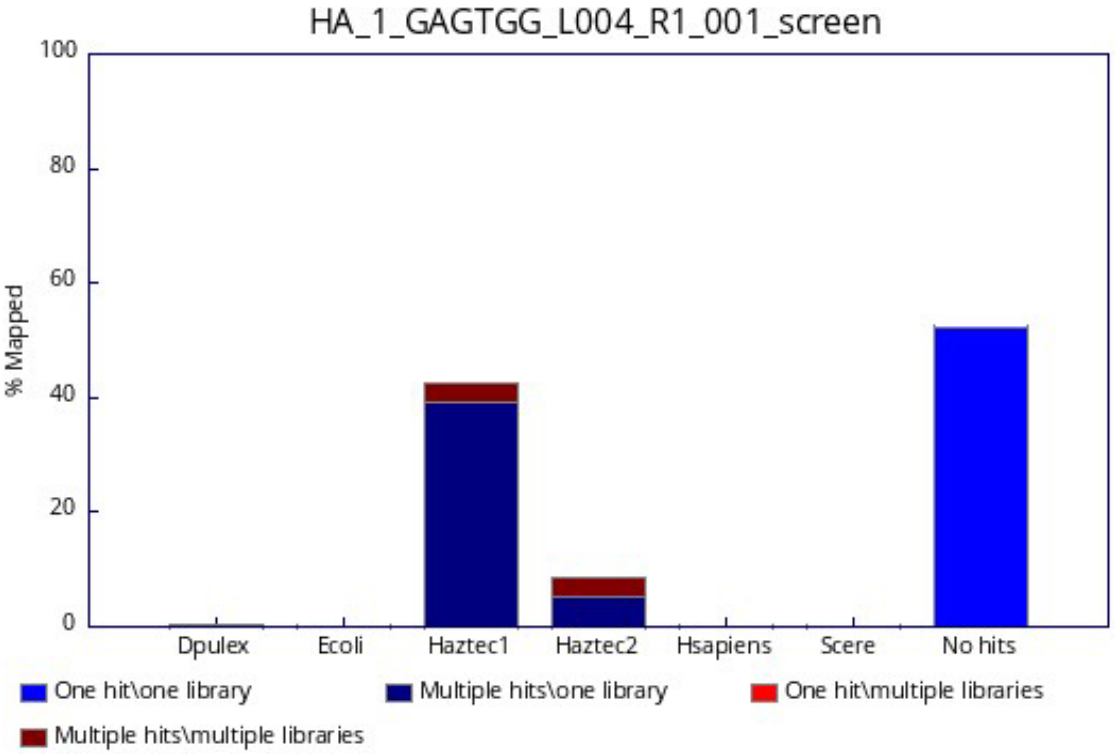

Figure 1.5: An output graph from FastQ Screen representative of those for all reads

Blast2GO returned a total of 7,915 (47.86%) transcripts with hits from blast of the original 16,537 transcripts. A total of 8,576 (51.86%) transcripts had no hits in blast, and 46 (0.28%) were not blasted due to the length of the transcripts (> 8,000 nt). Blast2GO successfully mapped 6,874 (86.85%) of the 7,915 transcripts with blast hits and annotated 5,541 (80.6%) of the mapped transcripts. The distribution of the top hits from Blast2GO, only those with the highest degree of similarity between the transcript and the hit, contains many arthropods, including the model crustacean *D. pulex* and the model beetle *Tribolium castaneum* (Figure 1.6). GO term analysis determined the majority of transcripts to be related to cellular components of organelles (2,053 transcripts), the membrane (1,159), and macromolecular complexes (1,463); molecular functions of binding (3,315) and catalytic activity (2,731); and biological processes of metabolic

activity (3,367), cellular processes (3,207), and single-organism processes (2,728) (Figures 1.7, 1.8, 1.9).

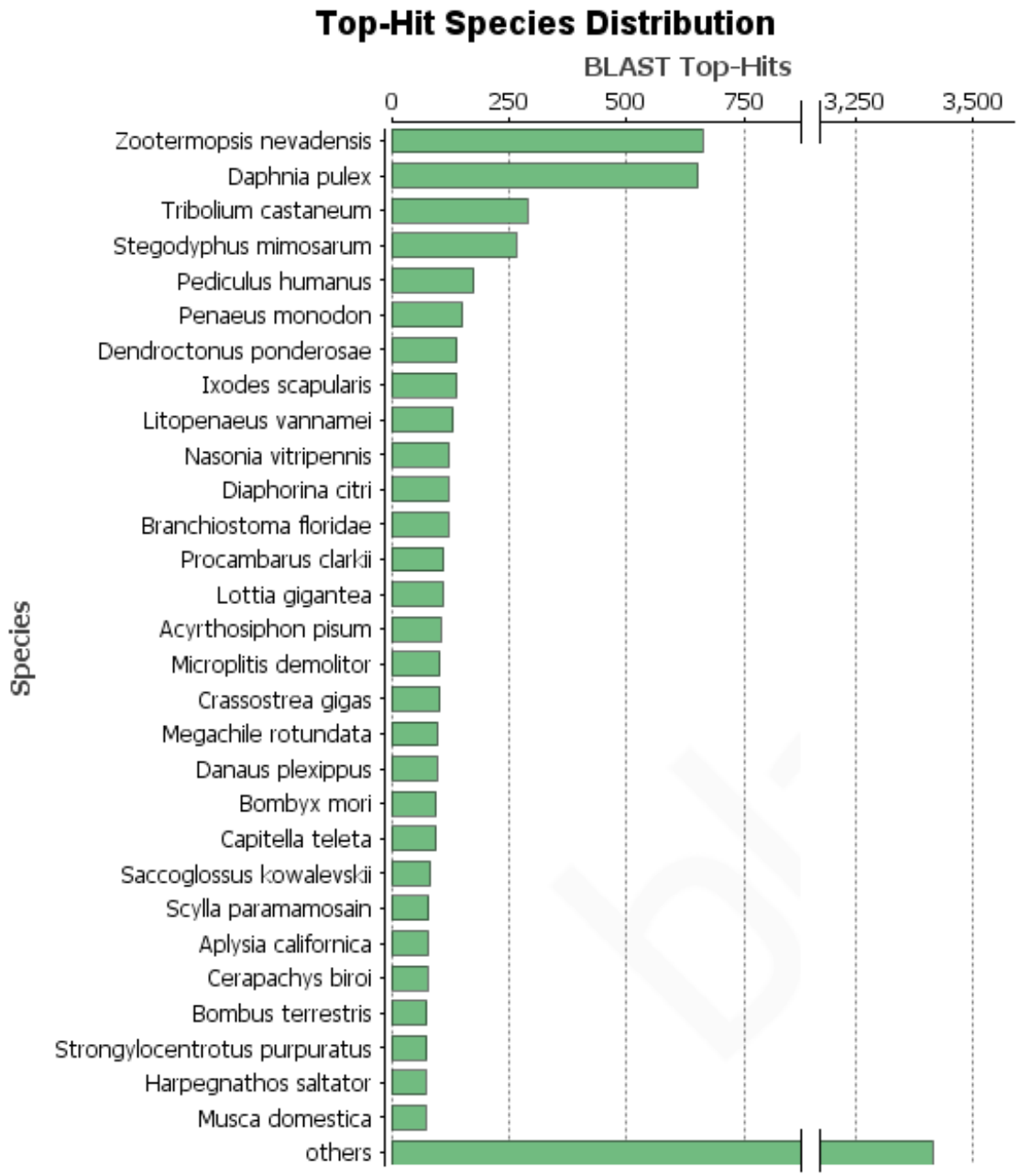

Figure 1.6: Distribution of top species hits from Blast2GO annotation

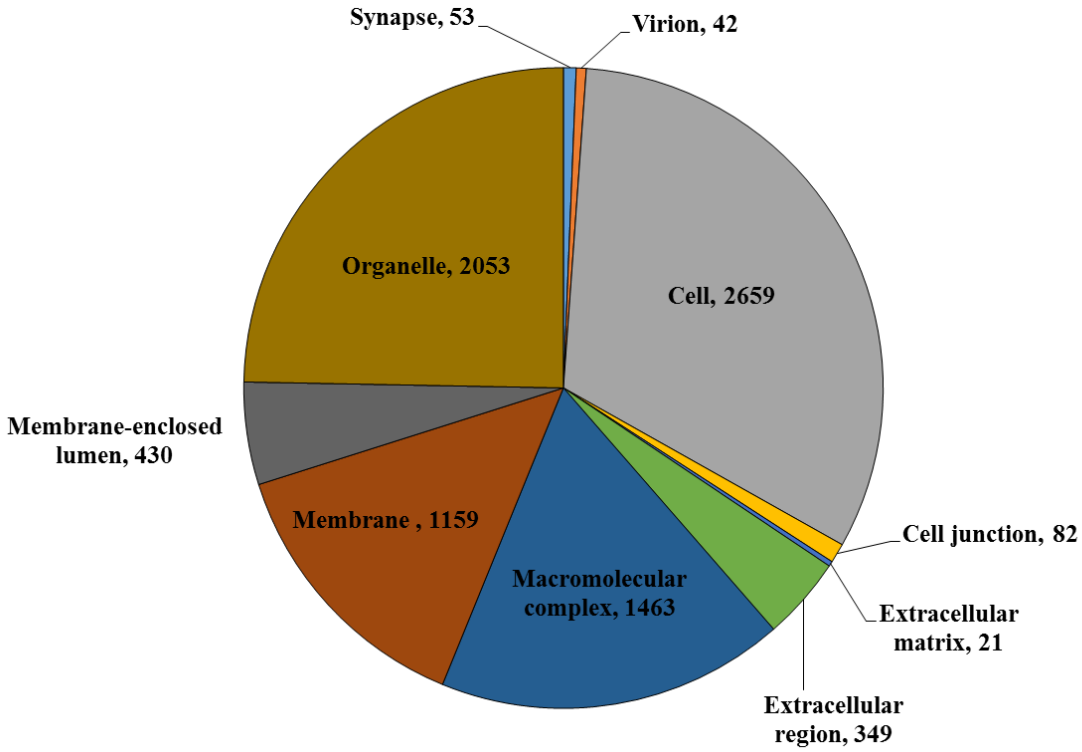

Figure 1.7: GO terms associated with cellular components in the *H. azteca* transcriptome

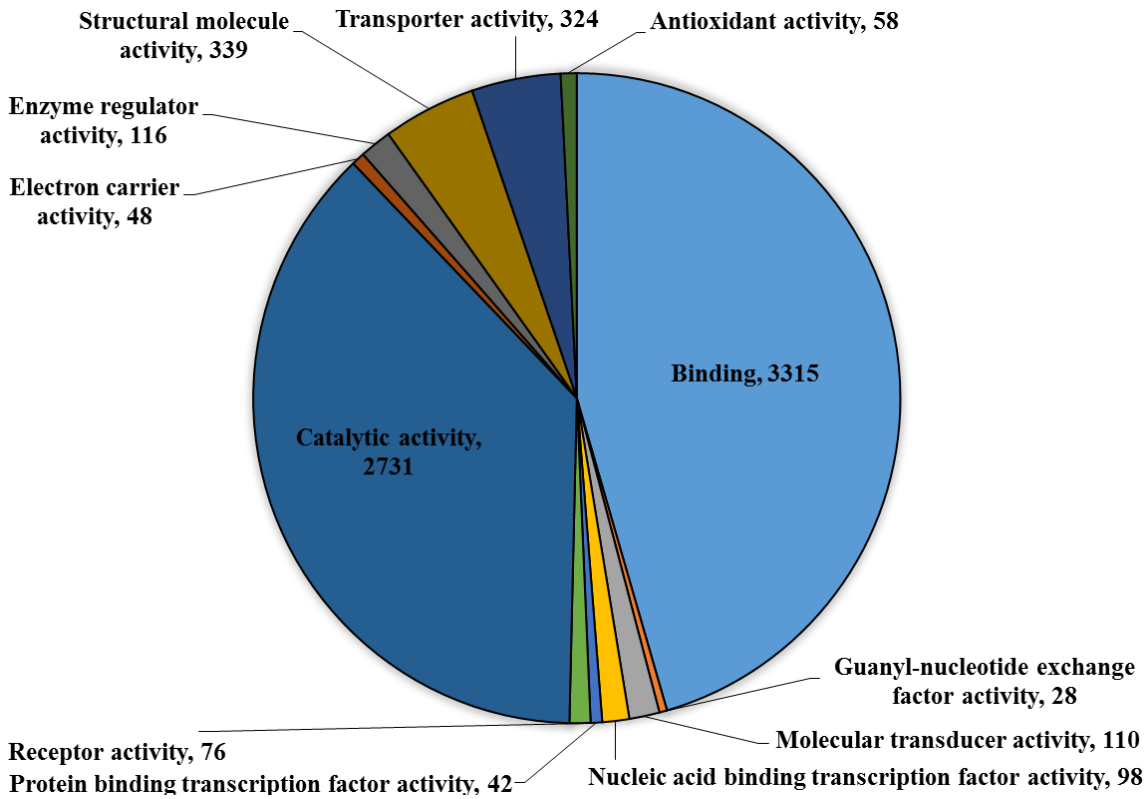

Figure 1.8: GO terms associated with molecular functions in the *H. azteca* transcriptome

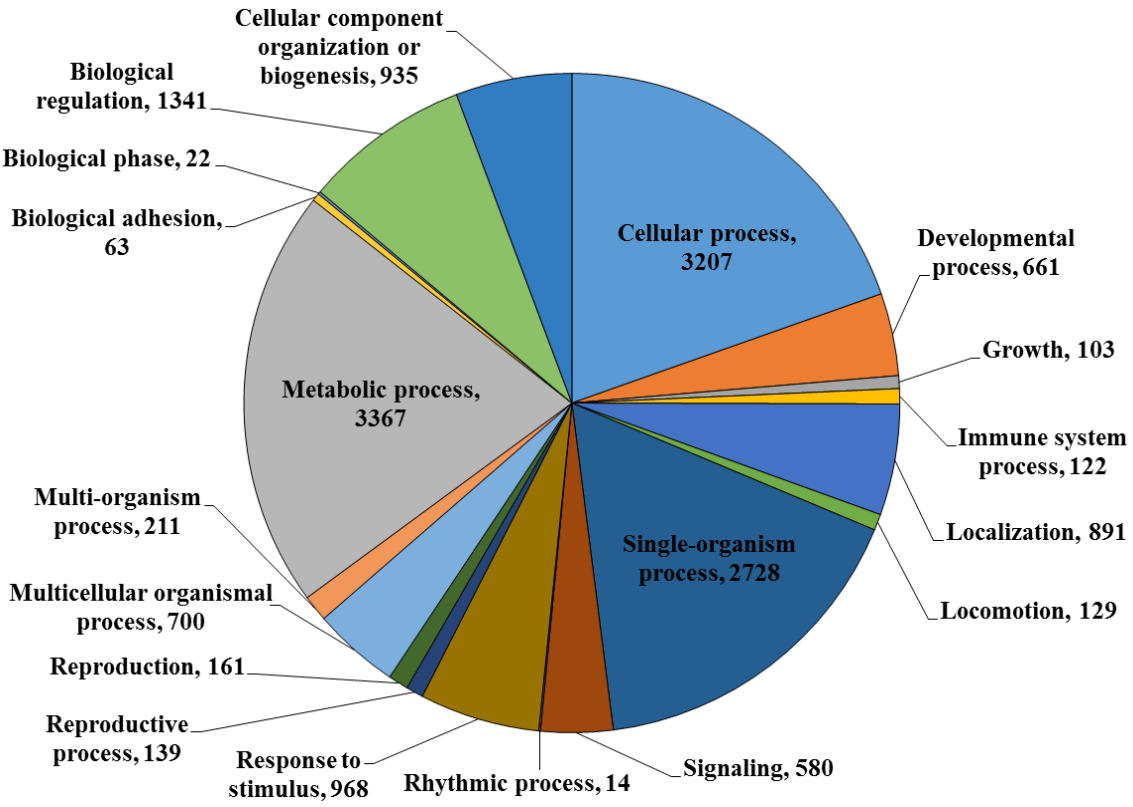

Figure 1.9: GO terms associated with biological processes in the *H. azteca* transcriptome

KEGG pathway analysis returned the highest number of transcripts as involved in purine metabolism (471 transcripts) and thiamine metabolism (399) pathways (Table 1.6). Among many other nucleic acid, amino acid, and fatty acid biosynthesis and degradation pathways were several potential stress response pathways, including "Drug metabolism - other enzymes" (31), "Drug metabolism - cytochrome P450" (16), and "Metabolism of xenobiotics by cytochrome P450" (16).

BUSCO analysis found 375 complete single-copy orthologs in the transcriptome, 14% of the 2,675 arthropod orthologs searched against. Additionally there were 108 (4%) complete duplicated orthologs and 237 (8.9%) fragmented orthologs.

Table 1.6: KEGG pathways identified for *H. azteca* transcripts

Pathway	Number of transcripts	Number of enzymes
Purine metabolism	471	26
Thiamine metabolism	399	2
Aminobenzoate degradation	81	4
Pyrimidine metabolism	67	18
Phenylalanine metabolism	59	9
Citrate cycle (TCA cycle)	58	14
Glycolysis/gluconeogenesis	56	20
Oxidative phosphorylation	53	7
Amino sugar and nucleotide sugar metabolism	52	19
Pyruvate metabolism	50	16
T cell receptor signaling pathway	50	2
Carbon fixation pathways in prokaryotes	47	15
Phenylpropanoid biosynthesis	46	1
Aminoacyl-tRNA biosynthesis	45	15
Lysine degradation	40	9
Glutathione metabolism	39	10
Pentose phosphate pathway	33	14
Carbon fixation pathways in photosynthetic organisms	32	11
Starch and sucrose metabolism	32	15
Fructose and mannose metabolism	32	15
Drug metabolism - other enzymes	31	10
Arginine and proline metabolism	31	14
Methane metabolism	30	9
Tryptophan metabolism	30	11
Alanine, aspartate, and glutamate metabolism	26	13
Glyoxylate and dicarboxylate metabolism	25	9
Fatty acid degradation	24	11
Tyrosine metabolism	23	11
Alpha-linolenic acid metabolism	22	5
Valine, leucine, and isoleucine degradation	22	11
Cysteine and methionine metabolism	22	10
Porphyrin and chlorophyll metabolism	21	10
Glycerophospholipid metabolism	21	9
Glycine, serine, and threonine metabolism	21	12
Galactose metabolism	21	9
Pentose and glucuronate interconversions	21	11
Butanoate metabolism	17	7

Pathway	Number of transcripts	Number of enzymes
Propanoate metabolism	17	10
Drug metabolism - cytochrome P450	16	5
Metabolism of xenobiotics by cytochrome P450	16	5
Caprolactam degradation	16	4
Arachidonic acid metabolism	14	7
Phosphatidylinositol signaling system	14	8
Fatty acid elongation	14	6
Inositol phosphate metabolism	13	8
Biosynthesis of unsaturated fatty acids	12	4
Geraniol degradation	12	3
Beta-alanine metabolism	12	7
Sphingolipid metabolism	12	6
Selenocompound metabolism	12	6
Nitrogen metabolism	11	5
One carbon pool by folate	11	8
Glycerolipid metabolism	11	5
Riboflavin metabolism	11	2
Ascorbate and aldarate metabolism	11	4
Toluene degradation	10	1
Primary bile acid biosynthesis	10	1
Ether lipid metabolism	10	3
Histidine metabolism	10	4
N-glycan biosynthesis	9	5
Nicotinate and nicotinamide metabolism	9	6
Sulfur metabolism	8	5
Terpenoid backbone biosynthesis	8	5
Retinol metabolism	8	4
Linoleic acid metabolism	7	1
Chloroalkane and chloroalkene degradation	7	3
Ubiquinone and other terpenoid-quinone biosynthesis	7	2
Other glycan degradation	7	3
Limonene and pinene degradation	7	2
Benzoate degradation	7	3
Pantothenate and CoA biosynthesis	6	4
Various type of N-glycan biosynthesis	6	3
Phenylalanine, tyrosine, and tryptophan biosynthesis	6	5
Lysine biosynthesis	6	2
Fatty acid biosynthesis	6	3

Pathway	Number of transcripts	Number of enzymes
Styrene degradation	6	2
Tropane, piperidine, and pyridine alkaloid biosynthesis	5	4
Novobiocin biosynthesis	5	4
Isoquinoline alkaloid biosynthesis	5	3
Steroid hormone biosynthesis	5	3
Taurine and hypotaurine metabolism	4	2
Carbapenem biosynthesis	4	2
Streptomycin biosynthesis	4	4
Other types of O-glycan biosynthesis	4	3
Mucin type O-glycan biosynthesis	4	2
Ethylbenzene degradation	4	1
Folate biosynthesis	4	3
Naphthalene degradation	3	1
Valine, leucine, and isoleucine biosynthesis	3	2
C5-branched dibasic acid metabolism	3	2
mTOR signaling pathway	3	1
D-glutamine and D-glutamate metabolism	2	1
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	2	1
Synthesis and degradation of ketone bodies	2	2
Vitamin B6 metabolism	2	1
Aflatoxin biosynthesis	2	1
Tetracyline biosynthesis	2	1
Glycosaminoglycan biosynthesis - heparan sulfate/heparin	2	1
Glycosaminoglycan biosynthesis - chondroitan sulfate/dermatan sulfate	2	1
Glycosaminoglycan degradation	2	2
Steroid degradation	1	1
Caffeine metabolism	1	1
Lipoic acid metabolism	1	1
Butirosin and neomycin biosynthesis	1	2
Indole alkaloid biosynthesis	1	1
Penicillin and cephalosporin biosynthesis	1	1
Glycosphingolipid biosynthesis - globoseries	1	1
Steroid biosynthesis	1	1
Peptidoglycan biosynthesis	1	1
Flavone and flavonol biosynthesis	1	1
Biosynthesis of ansamycins	1	1

Transcriptome validation

PCR products for primers *RpS23*, *RpL10*, *Gapdh*, *αTub*, *Mmp*, *Tbp*, *Ubi*, and *Ubc* appeared at the expected positions on the agarose gel (Figure 1.10 - bright band in lanes 1 and 6 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *RpS23* negative control, 3) *RpS23* sample 1, 4) *αTub* negative control, 5) *αTub* sample 1, 6) 10 bp DNA ladder, 7) *Act* negative control, 8) *Act* sample 1, 9) *Gapdh* negative control, 10) *Gapdh* sample 1) (Figure 1.11 - bright band in lanes 1 and 14 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Syx* negative control, 3) *Syx* sample 1, 4) *Syx* sample 2, 5) *Ubi* negative control, 6) *Ubi* sample 1, 7) *Ubi* sample 2, 8) *Tbp* negative control, 9) *Tbp* sample 1, 10) *Tbp* sample 2, 11) *Mmp* negative control, 12) *Mmp* sample 1, 13) *Mmp* sample 2, 14) 10 bp DNA ladder, 15) *RpL10* negative control, 16) *RpL10* sample 1, 17) *RpL10* sample 2, 18) *Ubc* negative control, 19) *Ubc* sample 1, 20) *Ubc* sample 2, 21) *RpS23* negative control, 22) *RpS23* sample 2, 23) *αTub* negative control, 24) *αTub* sample 2, 25) *Gapdh* negative control, 26) *Gapdh* sample 2).

Initial primers for both *Act* and *Syx* did not produce bands on gels, so additional primers were designed and tested (Table 1.7, Figure 1.12 - bright band in lanes 1 and 8 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Act2* negative control, 3) *Act2* sample 1, 4) *Act2* sample 2, 5) *Act3* negative control, 6) *Act3* sample 1, 7) *Act3* sample 2, 8) 10 bp DNA ladder, 9) *Syx2* negative control, 10) *Syx2* sample 1, 11) *Syx2* sample 2, 12) *Syx3* negative control, 13) *Syx3* sample 1, 14) *Syx3* sample 2, 15) *Syx4* negative control, 16) *Syx4* sample 1, 17) *Syx4* sample 2). Of these additional primers, only *Act2*, *Syx2*, and *Syx3* produced sequencable bands.

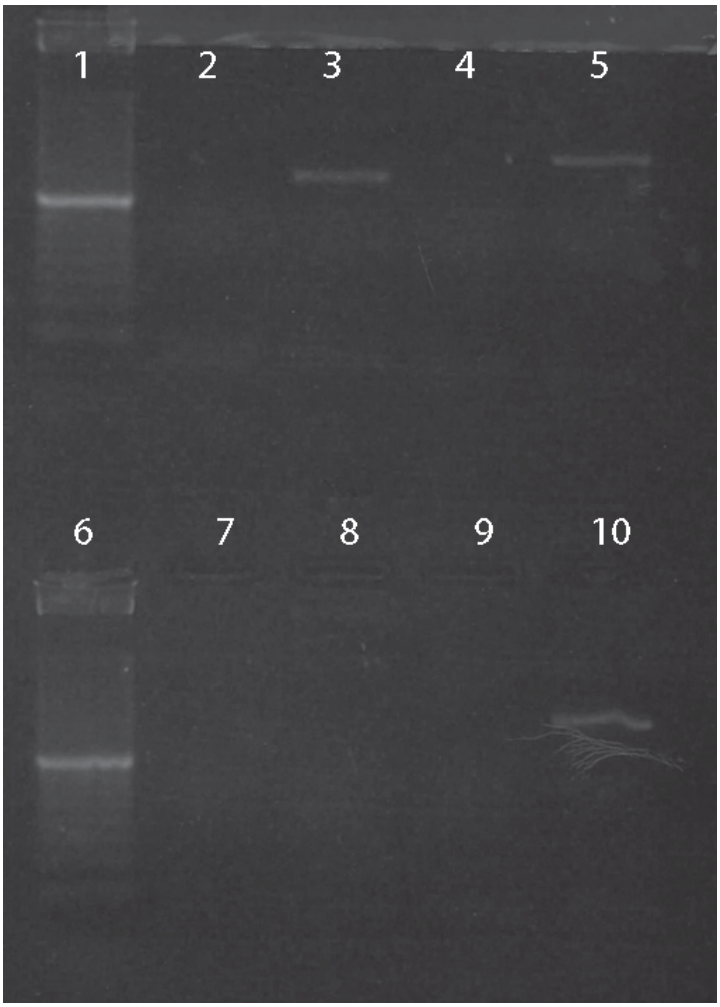

Figure 1.10: PCR products obtained from sample 1 of *H. azteca* cDNA using primers designed for *RpS23*, *Tub*, and *Gapdh*

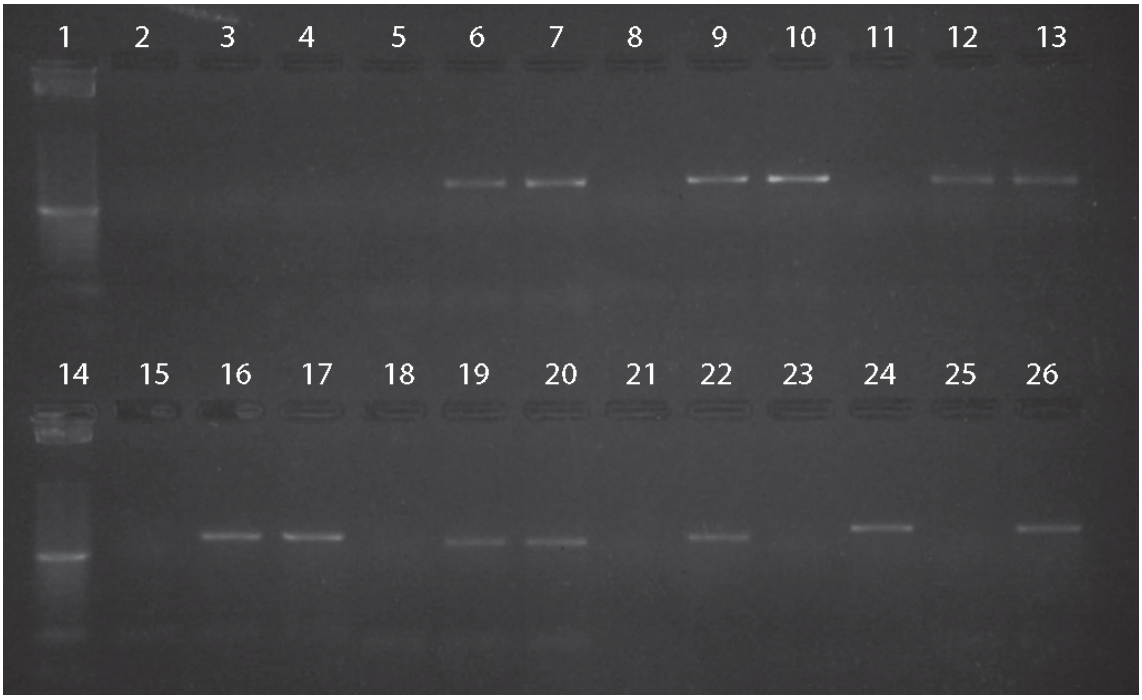

Figure 1.11: PCR products obtained from both samples of *H. azteca* cDNA using primers for *Syx*, *Ubi*, *Tbp*, *Mmp*, *RpL10*, and *Ubc*. PCR products also shown for *RpS23*, *Tub*, and *Gapdh* using the second sample of cDNA

Table 1.7: Names and sequences of the additional primers designed for beta actin and syntaxin

Transcript name	Primer abbreviation	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Amplicon length (nt)
Beta actin	*Act2*	AAAGTTGACGCC CGGAACTA	ACTGGTGGGTG ATGTTCACG	130
	Act3	TTCATCAAGCCTG TCGTTCC	TAGTTCCGGGC GTCAACTTT	126
Syntaxin-5-like isoform x2	*Syx2*	ACTTCGATGGCGT AAGCGTT	CTGCAGCAGCG AGTCTTGTT	92
	Syx3	CGCTGAGACCAT GAAGACCA	TTCGTCTTGCTC CTTGACCA	91
	Syx4	AGGAGAACTGCC AGCAAACC	CTGCAGCAGCG AGTCTTGTT	132

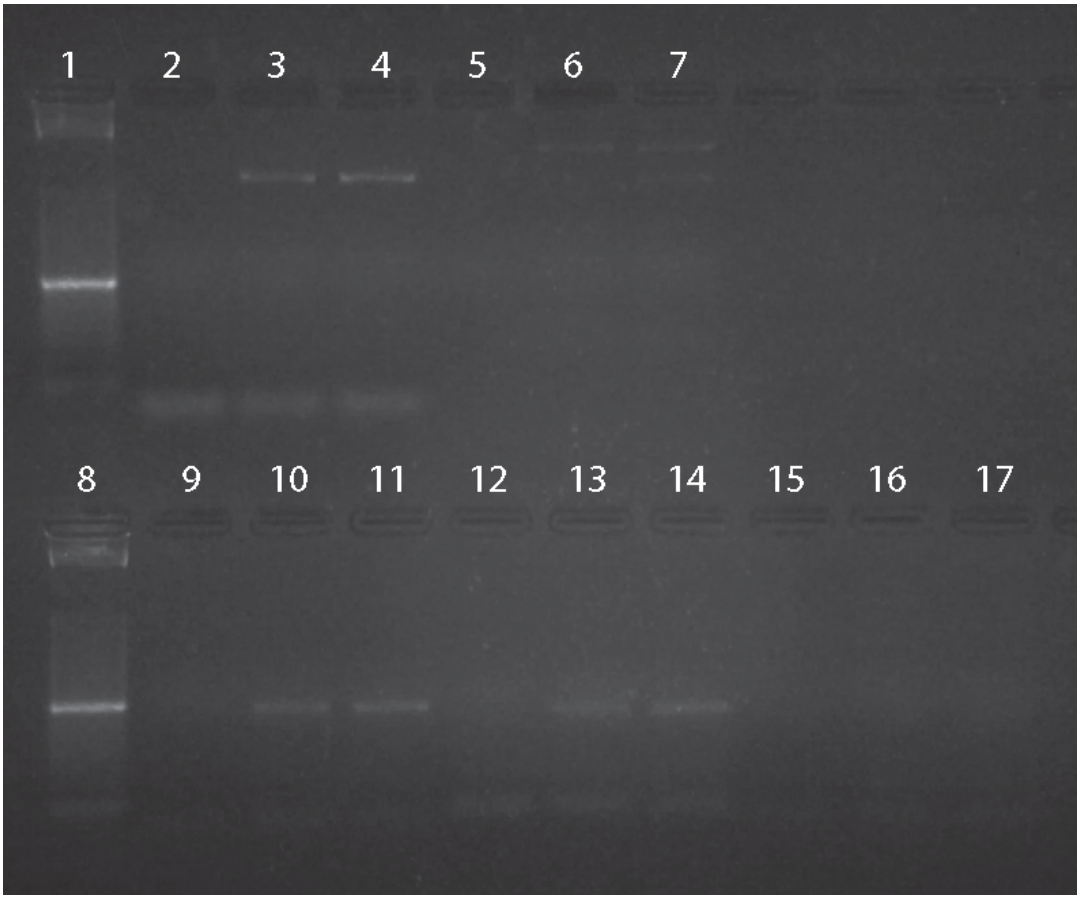

Figure 1.12: PCR products obtained from both samples of *H. azteca* cDNA using additional primers for *Act* and *Syx*

Sequencing was performed on all products except for those of *Act*, *Act3*, *Syx* and *Syx4* primer sets. Sequencing revealed high similarity between products and the expected amplicon sequences for most primers (Table 1.8, Table 1.9). Products of *Act2* were of low similarity to the expected amplicon, but the product of *Act2* from sample 1 of cDNA was of a length similar to that of the product seen during gel electrophoresis. Products of *Syx3* were of inconsistent length and quality.

Table 1.8: Sequences of amplicons obtained through PCR

Primer set	Sample	Product sequence
RpS23	1	AGTAGTTACGGTGCCCCGTATGTTTGCGGGTGAAGCCTGTGTCAGATTGA TCTTTTTTTTCAGAGGTATTGAGGCAAAACA
	2	GNCTAG-AGTAGTAGTTACGGTGCCCCGTATGTTTGCGGGTGAAGCCTGT GTCAGATTGATCTATTTTTTCAGAGGTTTTGAGGCAAAACA
RpL10	1	ACG-CCCGACTACGAGAAGANGCTAGCGTCAGGCCATATCAAGTCTGATG GTGTCAATGTCAAGTACATCCCTGAGCACGGTCCACTAATTAA
	2	CCNGACTACGAGAAGANGCTAGCGTCAGGCCATATCAAGTCTGATGGTGTCAATG TCAAGTACATCCCTGAGCACGGTCCACTAATTAA
Act2	1	CCTCANTCAGCACACACAGTCTGAGCTCATCAAGATCCTCTTCGAAAAGTACGGCA TCCATACGTGCTTTGTTTCTGATATATANATTGTTCCTCAAGAGAATTGCTGGACTG GTATTAACATGCGCAAGTTTCATAATCGTTACGTAGAAAACATTCTATATAATTTTC CCTTTGAAAACTTGTCATTATAAATTGACATTTACTATTCAATTTATGTGTTTGTTA CTATAAACTTTATGTGGTTATTGCAATAAACTTAGGTTGTAGCTAAAAACTTACTA ATTTACGGCTTTAAGGTTCGGGGTGAAGGGCGTGAACATCACCCACCAG
	2	ATCNGCACACACAGTCTGAGCTCATCAAGATCCTCTTCGAAAAGTACGG----------CATC
Gapdh	1	AGCTAGT-NNGGANATATCTGNT-ANCATATATTTTTCCGTTTTCTTGGGACCCCAA GAATTAATTGATTTCATCTTTCGGTCTAACTGACTGTCCTCTGTAAACTGCAA
	2	CNAGTGNNGG-NNTANCTGTTANACATATATTTTTCCGTTTTCTTGGGACCCCAAGA ATTAATTGATTTCATCTTTCGGTCTAACNGACTGTCCTCTGTAAACTGCAA
αTub	1	CTTGTTTNGA-NCNNCCNACCAGATGGTGAAGTGCGACCCGCGTCATGGAAAGTA CATGGCGTGTTGCCTGCTTTTCCGAGGTGACGTCGTGCCCAAGGATGTCAACGCT
	2	TTGNTTTGAG-CCGCCNACCAGATGGTGAAGTGCGACCCGCGCTCATGGAAAGTACATGGCGTGTT GCCTGCTTTTCCGAGGTGACGTCGTGCCCAAGGATGTCAACGCT
Mmp	1	GGGAGGTTGANTTGGCTCTCTTGTTGTTGTTGTTGTTGTTGGAGTCGTGGAGGTGCT GGTTGTGGTTGTCGTATCTGGAGTAGTAAACTGAGGCCGATAAGG
	2	TGATTTGGCTCTCTTGTTGTTGTTGTTGTTGNAGTCGTCGAGGAGCTGNTTGTG GTTGTCGTATCTGGAGTAGTAAACTGATGCCGATGATGAT
Syx2	1	GNTTACCNGCAGCAGCTGGCGCTGATGGAGGAACAAGACTCGC
	2	CAGCAGCTGGCGCTGANGGAGGAACAAGACTCGCTGCTGCAG
Syx3	1	CTCGGCNCGATGTTCACGCAGCTCGCTGCTATGGTCAAGGAGCAAGACGAA
	2	GTGCAA
Tbp	1	CAGAGCTCGGCTGTTTTATTGATGAGTAGCATGAAGCTTCAGAGATCTCTCCTCGG CATCCCACATAGCGGCGTATCCAGCCATACGGTTGTCTTCAGAGCC
	2	GNNGCAGAGCTCGGCTGTTTTATTGATGAGTAGCATGAAGCTTCAGAGATCTCTCC TCGGCATCCCACATAGCGGCGTATCCAGCCATACGGTTGTCTTCAGAGCC
Ubi	1	ANGTTCTTCCGGCAGGACTGGGAGGGATGTAATTCCCTTGGAGTGCTCTTCATCGG TGCCNCC--NCCCCCCAGGNGAAAAGAAGTCCCTCCGATGGACAT
	2	TCTGG---NNGTTCTTCCGGCAGGACTGGGAGGGATGTAATTCCCTTGGAGTGCTCT TCATC---GGCGCCAACNCCCCCCAGGAGAAAAGAAGTCCCTCCGATG---ATGGTC
Ubc	1	CNCCGAGTCTTGAGATTTCNCGAACTCCCACGAACCTNCGCGTTTGAATGGTGTTC AGTGCTTCTGACCGCAACATTTG
	2	CTCC--NCGAGTCTTNNGATTTCNCGAACTCCCACGAACCTGCGCGTTTGAATGGTG TTCAGTGCTTCTGACCGCAACATTTG

Table 1.9: Alignment of transcriptome transcript sequences and those obtained through PCR

Primer set	Sample	Expected product length (nt)	Sequenced product length (nt)	Percent similarity (%)	Gaps (%)
RpS23	1	127	81	91.4	1.2
	2	127	91	90.1	2.2
RpL10	1	132	93	96.8	1.1
	2	132	89	96.6	0
Act2	1	130	332	24.4	75.3
	2	130	68	75.0	22.1
Gapdh	1	150	110	92.7	1.8
	2	150	108	91.7	0.9
αTub	1	149	110	94.5	0.9
	2	149	110	96.4	1.8
Mmp	1	146	102	99.0	0
	2	146	97	92.8	0
Syx2	1	92	43	95.3	0
	2	92	42	97.6	0
Syx3	1	91	51	98.0	0
	2	91	6	83.3	0
Tbp	1	149	102	100.0	0
	2	149	106	98.1	0
Ubi	1	148	101	82.2	4.0
	2	148	114	77.2	10.5
Ubc	1	119	79	94.9	1.3
	2	119	83	91.6	3.6

Discussion

The transcriptome of *H. azteca* was sequenced and assembled to not only locate potential reference and stress-response genes for future study, but also to create the first publicly available "omic" data set for this amphipod. The transcriptome assembly was successful, providing many of the target genes. Reference gene transcript assemblies were validated through PCR, returning amplicon sequences matching those from the transcriptome.

The assembled transcriptome of *H. azteca* is shown to be of good quality. The distribution of gene identities is comparable to that of other crustaceans, including the marine amphipod *Parhyale hawaiensis* (Zeng et al., 2011) and the white-leg shrimp

Litopenaeus vannamei (Li et al., 2012a). Multiple iterations of *H. azteca* transcriptome assembly using stricter parameters may have detected and removed more potential contamination and lower quality sequences and produced additional or different transcripts. A re-assembly after the completion of the annotation of the *H. azteca* genome would provide a much more complete transcriptome. For example BUSCO analysis found only 14% of expected complete single-copy orthologs, a number that would be greatly improved by a transcriptome assembly guided by a complete genome containing the missing expected orthologous genes (Simao et al., 2015). This low BUSCO output was also seen in a recent genome analysis of *H. azteca* of the same U.S. Laboratory strain used here, indicating loss of these genes from the genome and potentially an overall reduced genome size (H. Poynton, University of Massachusetts Boston, Boston, MS, U.S.A., personal communication). Additionally the combination of multiple assembly approaches for the transcriptome could provide a higher quality final assembly resulting in more complete transcripts. This method is especially useful in non-model organisms with little genetic information (Feldmesser et al., 2014). However, given the exploratory nature of this transcriptome these additional assemblies were not necessary.

Probable contaminants were observed only in KEGG pathway analysis and included transcripts for streptomycin and tetracycline biosynthesis likely from actinobacteria (Nett et al., 2009); transcripts for carbon fixation from contamination by prokaryotes; and transcripts for penicillin, cephalosporin, and aflatoxin biosynthesis likely from fungal contamination (Samson et al., 1987; Geiser et al., 1998). It is possible that these transcripts are not from contamination of samples *per se*, but rather are from symbiotic organisms within the amphipods. Additional evidence for this possibility

comes from the KEGG pathways for chlorophyll breakdown and carbon fixation in photosynthetic organisms identified in the transcriptome. These transcripts likely are present in the amphipods due to the green algae they consume in their culturing tank. Bacteria and fungi may grow with these algae as a biofilm (Sabater et al., 2007). These many co-ingested microorganisms could become gut symbionts of *H. azteca*, as is seen in the freshwater amphipod *Gammarus pulex* (Harris, 1993), and consequently would be detected during RNA-seq. Detection of symbionts like bacteria and fungi as well as infectious agents like viruses within arthropods is common during transcriptome analyses (Pauchet et al., 2009; Mittapalli et al., 2010).

PCR products were of the expected sizes had high sequence similarity to the predicted product based off transcript sequences. This gives confidence that the transcriptome was assembled properly, yielding reliable transcripts, and ensuring that the intended target transcripts were being amplified specifically by the developed primer sets. The exceptions to this were *Act2* and *Syx3* in which large gaps in the expected amplicon sequence were required in order to align with the product sequence. Also sequencing results were inconsistent between cDNA samples for these primers as well. These difficulties could be due to alternative splicing of these transcripts or to gene duplication, resulting in paralogs with similar but not identical sequences. Lacking a fully assembled and annotated genome, transcript splice sites cannot be identified and avoided during primer development, making these transcripts uncertain amplification targets. Sequencing of PCR products was also performed with only forward primers. If a second round of sequencing were performed with reverse primers and the resulting sequences were aligned with those from the forward primers, a more complete amplicon may be obtained.

The use of different primers nested within the original primers may have also allowed more accurate sequencing by allowing primers to anneal and sequencing to start downstream of the start of the sequence.

The first goal in sequencing the *H. azteca* transcriptome was to locate housekeeping genes to use for quality control during gene expression measurement. Nine of the ten housekeeping genes used to verify the transcriptome successfully amplified during PCR. These genes are also commonly used as reference genes for quality control of gene expression (Dheda et al., 2004; Kozera and Rapacz, 2013). The ubiquitin gene is used in soft-shell clams (*Mya arenaria*) (Araya et al., 2008); TATA-box binding protein in *Daphnia pulex* (Spanier et al., 2010); GAPDH, ubiquitin conjugating enzyme, and beta-actin in *D. magna* (Heckmann et al., 2006); and beta-actin and GAPDH in various amphipods (Protopopova et al., 2014). Additionally, though ultimately not used as reference genes, the housekeeping genes alpha-tubulin, syntaxin, and matrix metalloproteinase were also explored in *D. pulex* as they were here in *H. azteca* (Spanier et al., 2010). Other candidate reference genes found in the *H. azteca* transcriptome which were not examined here include X-box binding protein 1, cyclophilin, succinate dehydrogenase, tryptophanyl-tRNA synthetase, ADP-ribosylation factors 1 and 4, translation elongation factor 1α, heterogeneous nuclear ribonucleoprotein A2/B1, and ribosomal proteins L7, L27, and S18. These genes were explored as reference genes for gene expression monitoring in *D. pulex* (Spanier et al., 2010), *D. magna* (Heckmann et al., 2006), the Pacific oyster (Du et al., 2013), and the Colorado potato beetle *Leptinotarsa decemlineata* (Shi et al., 2013), and could be examined for *H. azteca* as needed.

Many stress response-related gene transcripts with potential for use in toxicity testing and environmental risk assessment were found after annotation of the transcriptome. These included genes whose products are involved in detoxification and general stress response genes. Phase I detoxification transcripts with products associated with modification included cytochrome P450 monoxygenases, alcohol dehydrogenases, aldehyde dehydrogenases, peroxidases, and esterases. Phase II conjugation-associated transcripts included those coding for methytransferases, sulfotransferases, N-acetyltransferases, UDP-glucuronosyltransferases, and glutathione S-transferases. Phase III excretion-associated transcripts included ATP-binding cassette transporters of the multidrug resistance protein family (MRP, ABCB) and the multidrug resistance-associated protein family (MDR, ABCC) as well as many members of the solute carrier (SLC) superfamily of transport proteins. General stress response gene transcripts such as those for a variety of heat shock proteins as well as immune response genes were in abundance.

More stress-response gene transcripts may have been detected through the use of comparative transcriptomics in which the transcriptome of organisms exposed to a stressor is compared to that of unexposed organisms (e.g., Bonizzoni et al., 2012) or similarly through the development and use of microarrays (e.g., Watanabe et al., 2008). These methods can provide a large number of differentially expressed genes for more detailed experimentation or analysis and can be especially useful for non-model organisms (Miao et al., 2015; Oppenheim et al., 2015). The cost associated with doing such a comparison and analysis, however, can be prohibitive. A small scale sequencing such as that performed here provides a lower cost alternative but at the expense of not

determining transcript sequences and differentially expressed genes simultaneously. However, research already performed on differential expression of genes and proteins in the crustaceans *D. pulex* and *D. magna* (Shaw et al., 2007; Lyu et al., 2014) and work done on proteins in *H. azteca* and other amphipods provides a base line for future work on differential expression in *H. azteca* (Werner and Nagel, 1997; Correia et al., 2002).

The transcripts identified here and many others can be explored as biomarkers for physical, biological, and chemical stressors investigated during ERA. Differences in their expression levels can indicate stressors at magnitudes much lower than those detected by survival, histology, or other life history or tissue-level biomarkers. However, gene expression must be linked to higher-level impacts through adverse outcome pathways (AOP). AOPs are needed to connect initiating events like the interaction of a toxin and its receptor at the molecular level to ecologically relevant effects, which are the ultimate goal in ERA (Ankley et al., 2010; Kramer et al., 2011). AOPs expand a molecular biomarker's utility from that of solely an indicator of exposure to one of both exposure and of effects with real consequences. Once these connections are known, molecular responses can be used in lieu of apical organismal responses. Molecular responses can provide more specific information about the stressor's identity and magnitude while using fewer test organisms.

The housekeeping gene transcripts as well as stress response gene transcripts identified in the *H. azteca* transcriptome will be further explored for use in gene expression monitoring assays for detection of environmental stressors. The presence of many housekeeping genes also present in transcriptomes of *D. pulex* and *D. magna* allow

not only use of these genes in gene expression experiments in *H. azteca*, but also comparison to these other crustaceans. Which housekeeping genes serve as the best reference genes is determined in Chapters 2 and 4 and can be compared to those that proved best in similar exposure experiments with *D. pulex* and *D. magna*. This allows better comparison of changes in stress-response genes of interest across these species as well, as is shown in Chapters 3 and 4.

Chapter 2: Normalization factor development for gene expression studies in *Hyalella azteca* after exposure to cadmium and copper

Abstract

Biomarkers are an important tool in ecological risk assessment (ERA) to demonstrate exposure to or effects of stressors in the environment. Sensitive indicators of exposure may be changes in protein titers or activities and gene expression changes. Gene expression is a highly sensitive response. Because of this sensitivity, internal controls must be developed to prevent false conclusions from biomarkers based on gene expression. Normalization of the expression of genes of interest to that of reference genes, which do not change in expression even after stressor exposure, is an important quality control step. Here I developed normalization factors (NFs) based on multiple reference genes for sub-lethal exposures to cadmium and copper in *H. azteca*. I measured the expression of five housekeeping genes with and without exposure to cadmium or copper using RT-qPCR and used the resulting cycle threshold (Ct) values in three programs (BestKeeper, GeNorm, and NormFinder) designed for analysis of candidate reference genes. The housekeeping genes tubulin alpha-1a chain (*αTub*) and matrix metalloproteinase (*Mmp*) formed the NF for cadmium exposure, while *αTub*, *Mmp*, and ubiquitin (*Ubi*) formed the NF for copper exposure. Knowing the most stable reference genes for all different exposure scenarios ensures accurate estimation of expression responses in genes of interest, such as stress-response genes. This allows the development of reliable and accurate biomarkers for exposure and effect assessments during ERA.

Introduction

Ecological risk assessment (ERA) is a formal, guided process used for monitoring ecosystem health and prioritizing protection and restoration efforts. The main goals of ERA are to determine an ecosystem's risks of being exposed to and suffering adverse effects of stressors (U.S. EPA, 1998). The characterizations of exposure and of effects integrate expertise from many disciplines to obtain data on assessment endpoints, called lines of evidence, in both field and lab settings. Many lines of evidence come from the use of biomarkers, measurable substances in or characteristics of cells, tissues, organs, or whole organisms that change in response to stressor exposure (Forbes et al., 2006). The transcriptome of the freshwater amphipod *Hyalella azteca* was sequenced and assembled as described in Chapter 1 to develop sensitive molecular biomarkers for metal exposure. Here in Chapter 2, I use that transcriptome to investigate gene expression changes in response to metal exposure, specifically to develop rigorous internal controls to provide more accurate measurement of responses.

The biomarkers developed here for *H. azteca* use transcript abundance to assess changes in gene expression. Reverse transcription quantitative PCR (RT-qPCR) is a widely used and sensitive method for measuring transcript abundance (Heid et al., 1996; Bustin, 2002). An important component of this process is the selection of reference genes whose expression is used to normalize that of the genes of interest. Normalization is the attempt to account for variability in expression of genes due to technical differences of interest across samples, such as slight differences in sample extraction and automated data acquisition (Calza and Pawitan, 2010). After normalization the true biological variation due to the applied treatment remains. Commonly used reference genes include

many of the constitutive "housekeeping" genes like actin, tubulin, ribosomal proteins, ubiquitin, TATA-box-binding proteins, and others (Kozera and Rapacz, 2013). It is critical that these reference genes express stably and do not themselves respond to the stressor being investigated. Otherwise the effect of the stressor may be overestimated or underestimated. Recently it has been found that many of the historically used reference genes do in fact differ their expression levels in different tissues or under certain stressors, making them invalid as reference genes and sometimes even distorting expression results regarding genes of interest (e.g., Hauton et al., 2009; Dheda et al., 2004; Spanier et al., 2010; Radonic et al., 2004).

One solution to the issue of varying expression of reference genes is an initial screening of multiple candidate reference genes to determine which do not respond to the stressors investigated or which respond the least. Two or more stable reference genes are then measured for expression along with the genes of interest. The geometric mean of the expression of all the reference genes is calculated and used as a normalization factor (NF) to normalize the expression of the genes of interest (Vandesompele et al., 2002). This method has already been employed for more accurate gene expression analysis for a variety of stressor exposures in *Daphnia magna* (Heckmann et al., 2006), *Daphnia pulex* (Spanier et al., 2010), soft-shell clams *Mya arenaria* (Araya et al., 2008), Pacific oyster *Crassostrea gigas* (Du et al., 2013), disk abalone *Haliotis discus discus* (Wan et al., 2011), red flour beetle *Tribolium castaneum* (Lord et al., 2010), Colorado potato beetle (Shi et al., 2013), and the European honey bee *Apis mellifera* (Scharlaken et al., 2008). The use of NFs is also recommended in the Minimum Information for Publication of Quantitative Real-time PCR Experiments (MIQE) guidelines (Bustin et al., 2009).

The strategy of NFs is applied here to gene expression in *H. azteca* for exposures to cadmium and copper. Cadmium and copper are both on the U.S. EPA list of priority pollutants (U.S. EPA, 2012). Cadmium is released into the environment by mining and refinement of cadmium and zinc ores, in phosphate fertilizers, fossil fuel combustion, waste incineration, improper disposal of cadmium-containing items such as batteries, and natural sources (ATSDR, 2012). The maximum cadmium concentration allowed in drinking water is 5 μg/L (U.S. EPA, 1991; U.S. EPA, 2002). Higher concentrations begin the risk of adverse effects including neurotoxicity, reproductive toxicity, and genotoxicity (ATSDR, 2012). While cadmium in surface water is usually <1 μg/L, the concentration can skyrocket in the vicinity of contamination sources like waste sites or mines (Elinder, 1992; Pinot et al., 2000; ATSDR, 2012). Cadmium in water exists as both a freely dissolved hydrated form and adsorbed to sediment and organic material. The adsorption of cadmium lends to its relatively long residence time, which can be up to 4 to 10 years in freshwater sediment (Wester et al., 1992). Adsorbed cadmium is then easily accumulated and concentrated by organisms like *H. azteca* that are associated with sediment and organic material (Borgmann et al, 1993; Guan and Wang, 2006; Shuhaimi-Othman and Pascoe, 2007). There is also the potential for cadmium to be accumulated by aquatic plants or introduced into the aquatic system in cadmium-laden terrestrial plant material (Peralta-Videa et al., 2009). The detritus formed by these sources are actively consumed by omnivorous scavengers like *H. azteca*, adding another exposure source for these organisms. *H. azteca* is a common prey item, so there is potential for biomagnification of cadmium through higher trophic levels (Barwick and Maher, 2003; Croteau et al., 2005).

Copper is released into the environment primarily by mining and refining of copper ore and copper alloys, fossil fuel combustion, waste incineration, wood production, urban storm water runoff, and natural sources (ATSDR, 2004). Copper is typically found in surface water at a concentration of 10 μg/L, but this value can range widely from only 0.5 μg/L to over 10,000 μg/L and up to 69,000 μg/L in heavily contaminated areas near copper mining operations (ATSDR, 2004; EPA, 1991; Rosner, 1998). In water the majority of copper speciates to Cu(II) and copper metal and quickly adsorbs to organic matter, clay, hydrous iron, and manganese oxides (ATSDR, 2004). Residence times of copper in freshwater lakes have been estimated at 15 to 101 years in some lakes (Georgopoulos et al., 2001). As with cadmium this long residence makes copper easily encountered by benthic organisms like *H. azteca* (Borgmann et al., 1993; Borgmann and Norwood, 1997; Shuhaimi-Othman and Pascoe, 2007).

Here I describe experiments comparing expression of candidate reference genes between exposures to cadmium or copper and controls using three programs (GeNorm, NormFinder, and BestKeeper). These programs identify the genes that change in their relative expression the least after exposure to these metals. GeNorm calculates stability values, called M values, using the arithmetic mean of all pairwise variation values between a single candidate reference gene and all other genes (Vandesompele et al., 2002). Stability here represents the amount of variation in a gene's expression between the control and the treatment. M values are calculated for each candidate reference gene individually. The gene with the largest M value is the least stable gene and is dropped from the analysis. New M values are then calculated again for all remaining candidate

genes. The process is repeated until only two genes remain. No further analysis of these individual genes is possible since the M values rely on pairwise comparisons.

NormFinder uses a model-based approach, allowing the user to assign samples to groups prior to analysis (Andersen et al., 2004). NormFinder estimates intragroup and intergroup variation and uses these to calculate a stability value for each individual gene. Stability values represent the variation in gene expression, so lower values indicate more stable expression. BestKeeper calculates multiple pairwise correlations for all candidate reference genes (Pfaffl et al., 2004). The most highly correlated candidate genes are combined into an index, called the BestKeeper index. Each individual candidate gene is then analyzed for correlation to the index. The candidate genes showing the highest correlations to the index are the most stably expressed reference genes.

Advantages and disadvantages exist between these three programs. BestKeeper provides basic descriptive statistics for all genes and correlations to the BestKeeper index (Pfaffl et al., 2004). BestKeeper does not provide rankings of genes or suggestions of which genes should be used in the NF. The user is left to decide the ranking of genes based on the calculated statistics. Moreover, the BestKeeper index and the statistics associated with it change as the user removes highly unstable genes. This means the user must perform multiple iterations of BestKeeper manually. The process is left open to user errors like removal of stable genes that should actually be part of the NF.

GeNorm uses a pairwise method of comparing gene stabilities, providing a ranking of the genes to the user (Vandesompele et al., 2002). However, this means the final two most stable genes are not differentiated from one another. The final decision on

using only one or both of these genes is left to the user's interpretation of descriptive statistics like variation of gene expression. GeNorm also suggests a minimum number of genes to include in the NF. NormFinder provides rankings and stability values for all genes, without the drawback of leaving two undifferentiated genes like GeNorm. NormFinder follows a "net zero" approach to suggesting genes for NFs, balancing genes with equal but opposite expression changes (Andersen et al., 2004). The use of all three programs together and the simultaneous interpretation of all results allow the user to balance the advantages and disadvantages of each program. The different approaches of the programs often provide different results for NFs. Having three sets of results allows the user to determine the best NF using the consensus seen between the three programs (e.g., Spanier et al., 2010).

In this chapter I measure several candidate reference genes for expression stability after exposure to the metals cadmium and copper. The goal is to identify the most stably expressed reference genes. These genes will be used in Chapter 3 as normalization factors in exposure studies for cadmium and copper to measure changes in genes of interest involved in stress response.

Materials and Methods

H. azteca were exposed to sub-lethal concentrations of cadmium and copper for 24 hours. RNA was then extracted and reverse transcribed. Expression of candidate reference genes was measured in all samples using RT-qPCR. Expression of each gene was assessed for stability using the programs BestKeeper, GeNorm, and NormFinder. The most stable genes in each treatment were used to form normalization factors (NFs).

NFs are used to normalize expression of genes of interest in treatments to that in controls in gene expression studies.

Candidate reference genes

Candidate reference genes include those tested in Chapter 1 and verified by PCR and sequencing: 40s ribosomal protein s23 (*RpS23*), 60s ribosomal protein l10 (*RpL10*), glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*), tubulin alpha-1a chain (*αTub*), matrix metalloproteinase-15-like (*Mmp*), syntaxin-5-like isoform x2 (*Syx*), TATA-box-binding protein (*Tbp*), ubiquitin (*Ubi*), and ubiquitin-conjugating enzyme e2 d2 (*Ubc*). Primers confirmed to produce single, repeatable amplification products for these gene transcripts in Chapter 1 were further tested for amplification efficiency and expression stability to assess their suitability as reference genes.

I extracted RNA from four pooled samples of 20 adult *H. azteca* of mixed sex and two samples of 10 adult *H. azteca* using the method described in Chapter 1, with an additional on-column DNA digestion using an RNase-free DNase (Qiagen), following manufacturer's instructions. I then performed cDNA synthesis as described in Chapter 1 for all samples and their –RT controls. Amplification of candidate reference genes for each sample were added to wells of a LightCycler 480 Multiwell 96-well plates (Roche) in triplicate with the following reaction mixture: 10 µL of 2x SYBR Green I, 0.5 µL of 10 µM of each primer, 5 µL of 1/10-diluted cDNA/-RT, and 4 µL of PCR grade water. Proper water controls were performed. Amplification was performed on a LightCycler 480 qPCR machine (Roche) with the following: 1) a single pre-incubation period of 95 °C for 10 min, 2) an amplification period consisting of 30 cycles of 95 °C for 15 s, 61 °C

for 30 s, and 72 °C for 30 s, with a single fluorescence measurement performed at the end of each cycle, 3) a single melting curve analysis of 95°C for 0.05 s followed by an increase of 0.5 °C/cycle from 65 °C to 97 °C, with 5 fluorescence readings/°C, and 4) a cooling period of 40 °C for 1 min. I calculated efficiencies for each primer set from the individual RT-qPCR reactions using the real-time PCR Miner algorithm (Zhao and Fernald, 2005; http://ewindup.info/miner/).

Cadmium and copper exposures

Cadmium and copper were chosen from the metals listed on the U.S. EPA's list of priority pollutants based on pre-existing toxicological data for *H. azteca* (U.S. EPA, 2012). Given a 24-hour LC50 of 55 µg/L for cadmium exposure as cadmium chloride for *H. azteca* obtained from the U.S. EPA's ECOTOX database, I chose an exposure concentration of 1.2 µg/L, 4% of the LC50 (Werner and Nagel, 1997). While 24-hour exposure data were not available, the average 48-hour LC50 for copper exposure as copper sulfate for *H. azteca* was calculated as 123.25 µg/L based on values obtained from the ECOTOX database (Suedel et al., 1996; Huggett et al., 2001; Mastin and Rodgers, 2000; Cherry et al., 2002). To obtain sub-lethal effects for only a 24-hour exposure, I chose a concentration of 3.6 µg/L, 6% of the LC50, for copper. Cadmium chloride (Sigma Aldrich) and cupric sulfate (Sigma Aldrich) were made as 1 mg/L stock solutions in deionized water. These low concentrations of both metals based on published LC50 values were used to obtain sublethal effects. The goal was to test reference genes at sublethal exposure concentrations at which measurement of mortality or other life history parameters would not differentiate between a control and an exposure. Changes detected

in stress response genes and normalized with reference genes at sublethal concentrations would thus serve as more sensitive biomarkers of exposure than life history parameters.

Experimental units consisted of round glass specimen dishes (10 cm diameter, 5.5 cm height, approximately 350 mL maximum volume) with 200 mL of reconstituted fresh water (pH 7.84, conductivity 307 µS/cM) made in accordance to U.S. EPA (2000) as follows: 10 L deionized water, 0.5 g $CaSO_4$ (Colorado Scientific), 0.5 g $CaCl_2$ (J. T. Baker), 0.3 g $MgSO_4$ (Sigma-Aldrich), 0.96 g $NaHCO_3$ (Sigma-Aldrich), and 0.04 g KCl (Sigma-Aldrich). Each dish had an air bubble and a 4" x 4" piece of plastic mesh substrate. Reconstituted water was aerated for 24 hours. Cadmium, copper, and control exposures were replicated in triplicate, giving a total of nine dishes. Twenty adult *H. azteca* caught on a number 60 standard sieve were added to each replicate dish (180 total *H. azteca* used) with approximately 2 mg of crushed Tetramin tropical fish food and acclimated for 24 hours.

After the 24-hour acclimation period, I removed 232 or 720 µL of water in the cadmium and copper exposures, respectively, and replaced it with the same volume of the appropriate 1 mg/L stock solution of either cadmium chloride or copper sulfate. This yielded the desired concentrations of 1.2 and 3.6 µg/L for the cadmium chloride and copper sulfate exposures, respectively. All exposures started at 08:00 and ended at 08:00 the following day, a typical 24 hour exposure time for aquatic toxicity testing looking at gene expression (e.g., Heckmann et al., 2006; Poynton et al., 2011). I also assembled equal numbers of replicates of both chemical exposures and controls and allowed them to run for 48 hours to test for effects of the metals on survival. During these exposures a

16:8 light:dark cycle was used, with light beginning at 07:00 each day. The total number of surviving *H. azteca* in each 24 and 48-hour replicate was recorded. The proportion of amphipods surviving from each replicate was arcsine square root transformed and analyzed using the non-parametric Kruskal-Wallis analysis of variance in SAS statistical software (v.9.4). This non-parametric equivalent to the ANOVA was chosen because the small sample size would not allow accurate examination of assumptions of ANOVA like conforming to a normal distribution or of other assumptions of parametric tests in general. Non-parametric tests are not constrained by these assumptions, though they lose power compared to a parametric test if the data do conform to parametric assumptions.

Reference gene expression analysis

I collected ten *H. azteca* from each 24-hour replicate dish by straining through a number 60 standard sieve. The ten *H. azteca* were placed into microcentrifuge tubes containing 200 µL of RNAlater (Sigma Aldrich), flash frozen in liquid nitrogen, and stored at -80°C until further processing. RNA extraction was performed using the protocol described in Chapter 1 with an additional on-column DNA digestion of the extracted total RNA using an RNase-free DNase (Qiagen), following the manufacturer's instructions. I sent frozen aliquots of each RNA sample (5 µL) for bioanalysis at the University of Maryland Genomics Core. Bioanalysis was performed on a Agilent 2100 Bioanalyzer to determine RNA quantity and quality via electrophoresis. RNA samples were stored at -80°C until cDNA synthesis. cDNA synthesis was performed as described in Chapter 1. Simultaneous reactions were performed for each sample without reverse transcriptase (-RT) as a control.

I performed RT-qPCR on a Roche LightCycler 480 Real-Time PCR machine in the Genomics Core at the University of Maryland. Each well of the LightCycler 480 Multiwell 96-well plates (Roche) consisted of 10 µL of 2x LightCycler 480 SYBR Green I Master mix (Roche), 0.5 µL of 10 µM of each primer, 5 µL of 1/10-diluted cDNA, and 4 µL of PCR grade water. Three technical replicates of each biological replicate were performed. Appropriate cDNA-free –RT controls were used for all treatments, each with three technical replicates. Template-free water controls were run, also with three replicates. The optimized RT-qPCR protocol used was as follows: 1) a pre-incubation period of 95 °C for 10 min, 2) an amplification period consisting of 30 cycles of 95 °C for 15 s, 61 °C for 30 s, and 72 °C for 30 s, with a single fluorescence measurement performed at the end of each cycle, 3) a single melting curve analysis of 95 °C for 0.05 s followed by an increase of 0.5 °C/cycle from 65 °C to 97 °C, with 5 fluorescence readings/°C, and 4) a cooling period of 40 °C for 1 min. I determined cycle threshold (Ct) values using fit point analysis on manufacturer provided LightCycler 480 Software release 1.5.0 SP3 (version 1.5.0.39).

The mean of the Ct values of all technical replicates served as the Ct value of each biological replicate. I calculated delta Ct values (ΔCt) using equation (1) for each replicate for each gene, where ε is the efficiency of the primer for the gene being analyzed, $Ct_{gene,min}$ is the minimum Ct value of the gene being analyzed across all replicates, and Ct_x is the Ct value of the biological replicate being analyzed. This served to calibrate all values against the lowest expression value. I analyzed results for stability of reference gene transcripts using BestKeeper (Pfaffl et al., 2004), GeNorm (Vandesompele et al., 2002), and NormFinder (Andersen et al., 2004). While Ct values

were used in BestKeeper, ΔCt values were used in GeNorm and NormFinder. I then used the most stable genes to calculate a normalization factor (NF) for each treatment-control pair, i.e. one NF each for the cadmium and copper treatments, one NF for the control as calculated using the most stable genes for cadmium, and one NF for the control as calculated using the most stable genes for cadmium. I calculated NFs as the geometric mean of the ΔCt values of all biological replicates. The geometric mean has better control for outlying values and abundance differences between genes (Vandesompele et al., 2002).

$$(1)\ \Delta Ct_x = \varepsilon^{Ct_{gene,min} - Ct_x}$$

I calculated the standard error of the mean of the NF (NFsem) incorporating proper error propagation using equation (2), where $\Delta Ctsem_x$ is the standard error of the mean of the ΔCt value of the first gene in the NF, ΔCt_x is the ΔCt value of the first gene in the NF, $\Delta Ctsem_i$ is the standard error of the mean of the last gene in the NF, and ΔCt_i is the ΔCt value of the last gene in the NF.

$$(2)\ NFsem = NF * \sqrt{\left(\frac{\Delta Ctsem_x}{3 * \Delta Ct_x}\right)^2 + \cdots + \left(\frac{\Delta Ctsem_i}{3 * \Delta Ct_i}\right)_i^2}$$

Results

Primer efficiencies

Ct values did not differ between samples derived from ten versus twenty *H. azteca*. Evaluation of primer efficiencies produced single strong melting peaks in cDNA samples for all primers, indicating lack of contamination or primer dimers (Figure 2.1).

Some melting peaks appeared from –RT controls and water reactions. This is likely due to genomic DNA contamination. Primer efficiencies ranged from 79% to 87%, below the optimum efficiency range of 90-105% (Table 2.1). Primer efficiencies were significantly different from one another (two-way ANOVA, α=0.05, F=10.69 (8, 18), p<0.0001). The highest efficiency primer, *RpS23*, and the lowest efficiency primers, *RpL10*, *Gapdh*, and *Syx2*, were removed from the analysis one at a time as analysis of the differences of least squares means indicated these four efficiencies were driving the significant difference. Analysis was repeated after each removal. After all four primer sets were removed from analysis, no significant differences existed between the efficiencies of the remaining primers (two-way ANOVA, α=0.05, F=0.6 (4, 10), p=0.67). *RpS23*, *RpL10*, *Gapdh*, and *Syx2* were therefore excluded as reference genes from all subsequent experiments. Performing future experiments with only primers with equivalent efficiencies allows simpler analysis of gene expression downstream. The mean efficiency (± SD) of the remaining primers was 84.8 ± 0.84%.

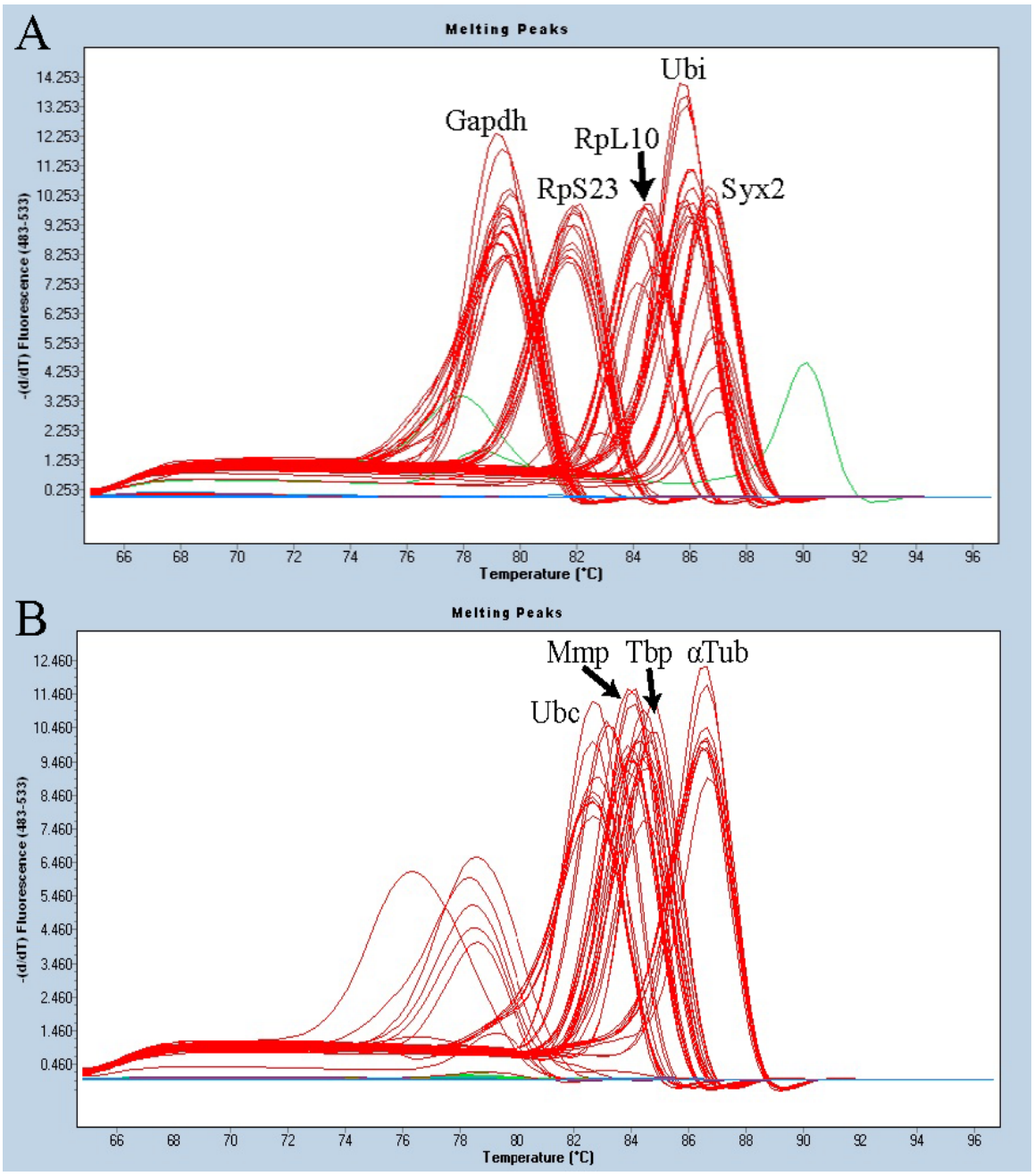

Figure 2.1: Melting curve peaks for all primer sets

Table 2.1: Amplification efficiencies of reference primers

Primer set	Efficiency % (mean ± standard deviation)
RpS23	87 ± 0.6
RpL10	79 ± 0.3
Gapdh	81 ± 0.6
αTub	85 ± 0.4
Mmp	84 ± 1.6
Syx2	82 ± 0.3
Tbp	86 ± 0.9
Ubi	85 ± 0.7
Ubc	84 ± 0.1

H. azteca survival

Mean survival (± SD) in the 24 hour control, cadmium, and copper exposures was 96.7 ± 2.9%, 100 ± 0%, and 93.3 ± 2.9%, respectively, and in the 48 hour exposures was 93.3 ± 7.63%, 98.3 ± 2.9%, and 93.3 ± 2.9%, respectively. No significant effect of the metals on survival was found at either 24 hours (Kruskal-Wallis analysis of variance, X^2=5.63, df=2, *P*=0.086) or 48 hours (Kruskal-Wallis analysis of variance, X^2=2.44, df=2, *P*=0.46).

Normalization factor analyses

Extracted samples of RNA were of high concentration and good quality according to bioanalysis (Table 2.2). Descriptive statistics of RT-qPCR results are of the threshold (Ct) values (Table 2.3). The Ct value is the PCR cycle at which the signal from the sample became higher than the background noise present during data acquisition from source such as 96-well plate autofluorescence. The standard deviation of the Ct value represents the stability of a gene's expression across samples from the same treatment. Lower standard deviations show a gene that is similarly expressed across all samples.

Table 2.2: Bioanalysis results of total RNA concentration and quality

Treatment	Replicate	RNA concentration (ng/μL)	RNA integrity number (RIN)
Control	1	1,513	7.3
	2	720	7.4
	3	1,244	7.8
Cadmium	1	1,134	6.2
	2	914	6.5
	3	1,167	6.5
Copper	1	989	8.3
	2	1,476	6.7
	3	1,367	6.0

Table 2.3: Mean Ct values of all genes in control, cadmiu, and copper exposures. SD: standard deviation, SE: standard error of the mean

	Gene														
	αTub			*Mmp*			*Tbp*			*Ubi*			*Ubc*		
Treatment	Control	Cd	Cu	Control	Cd	Cu	Control	Cd	Cu	Control	Cd	Cu	Control	Cd	Cu
n	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Geometric mean	18.65	20.60	19.95	22.77	24.82	24.07	26.49	28.75	27.66	21.36	23.53	22.78	17.71	18.71	18.77
Arithmetic mean	18.65	20.60	20.0	22.77	24.82	24.10	26.49	28.77	27.71	21.36	23.55	22.83	17.71	18.73	18.79
Min	18.22	19.95	18.63	22.5	24.42	22.85	26.21	27.77	26.06	20.84	22.42	21.72	17.51	18.03	17.84
Max	19.11	21.45	21.86	23.20	25.22	25.93	26.87	30	30	21.80	24.75	24.91	17.84	19.66	20.03
SD	0.45	0.76	1.67	0.38	0.40	1.62	0.34	1.13	2.05	0.49	1.17	1.81	0.18	0.84	1.12
SE	0.26	0.44	0.96	0.22	0.23	0.93	0.20	0.65	1.18	0.28	0.68	1.04	0.10	0.48	0.65

For the cadmium exposure, the BestKeeper repeated pair-wise correlation analyses revealed *αTub*, *Mmp*, and *Ubi* to have the highest r^2 values when correlated to the BestKeeper index (Table 2.4). GeNorm and NormFinder both found *αTub* and *Mmp* to be the most stably expressed genes for the cadmium exposure (Table 2.6). For the copper exposure BestKeeper found *αTub*, *Mmp*, and *Ubi* to have the highest r^2 values when correlated to the BestKeeper index (Table 2.5). *Mmp* and *αTub* were most stable under GeNorm analysis for the copper exposure, while *Mmp* and *Ubi* were most stable under NormFinder analysis (Table 2.6). All genes ranked by their stability and suitability as reference genes are found in Table 2.7.

The genes chosen to form the NF for each exposure were selected based on the consensus on their stability between all three programs. *αTub* and *Mmp* formed the NF for the cadmium exposure, and *αTub*, *Mmp*, and *Ubi* formed the NF for the copper exposure. NFs were calculated for the exposure and the control separately for each metal because the genes comprising the NF differed between the two treatments. NFs were 0.29 and 0.83 for the cadmium treatment and its control, respectively, and 0.39 and 0.81 for the copper treatment and its control, respectively.

Table 2.4: Coefficients of correlation (r values) and *P*-values from the BestKeeper repeated pair-wise correlation analysis for the cadmium exposure

Vs.	*αTub*	*Mmp*	*Tbp*	*Ubi*	*Ubc*
Mmp	0.975	-	-	-	-
P-value	0.001	-	-	-	-
Tbp	0.818	0.886	-	-	-
P-value	0.047	0.019	-	-	-
Ubi	0.960	0.934	0.802	-	-
P-value	0.002	0.006	0.055	-	-
Ubc	0.916	0.842	0.656	0.970	-
P-value	0.010	0.036	0.157	0.001	-
BestKeeper index	0.984	0.978	0.879	0.982	0.922
Coefficient of determination (r^2) vs. BestKeeper	0.97	0.96	0.77	0.96	0.85
P-value	0.001	0.001	0.021	0.001	0.009

Table 2.5: Coefficients of correlation (r values) and *P*-values from the BestKeeper repeated pair-wise correlation analysis for the copper exposure

Vs.	*αTub*	*Mmp*	*Tbp*	*Ubi*	*Ubc*
Mmp	0.996	-	-	-	-
P-value	0.001	-	-	-	-
Tbp	0.897	0.914	-	-	-
P-value	0.015	0.011	-	-	-
Ubi	0.956	0.977	0.948	-	-
P-value	0.003	0.001	0.004	-	-
Ubc	0.878	0.907	0.966	0.959	-
P-value	0.021	0.013	0.002	0.002	-
BestKeeper index	0.972	0.985	0.966	0.992	0.962
Coefficient of determination (r^2) vs. BestKeeper	0.95	0.97	0.93	0.98	0.93
P-value	0.001	0.001	0.002	0.001	0.002

Table 2.6: Expression stability values from GeNorm and NormFinder analyses

	GeNorm		NormFinder	
Gene	Cadmium	Copper	Cadmium	Copper
αTub	0.414	0.363	0.097	0.114
Mmp	0.419	0.318	0.079	0.052
Tbp	0.695	0.467	0.275	0.130
Ubi	0.505	0.355	0.204	0.083
Ubc	0.597	0.485	0.279	0.159

Table 2.7: Ranking of all references genes by expression stability

	Cadmium			Copper		
	BestKeeper	GeNorm	NormFinder	BestKeeper	GeNorm	NormFinder
Most stable gene	*Ubc*	*αTub/Mmp*	*Mmp*	*Ubc*	*αTub/Mmp*	*Mmp*
	αTub	*Ubi*	*αTub*	*Mmp*	*Ubi*	*Ubi*
	Mmp	*Ubc*	*Ubi*	*αTub*	*Tbp*	*αTub*
	Ubi	*Tbp*	*Tbp*	*Ubi*	*Ubc*	*Tbp*
Least stable gene	*Tbp*		*Ubc*	*Tbp*		*Ubc*

Discussion

Candidate reference genes for gene expression normalization were measured in *H. azteca* exposed to cadmium and copper. The software programs BestKeeper, GeNorm, and NormFinder determined the most stably expressed candidate genes. These genes were used together to form normalization factors (NFs) to be used to normalize expression of stress-response genes after cadmium and copper exposure. Normalization helps eliminate the variability introduced from non-biological sources, an important tool in the sensitive process of gene expression measurement. *αTub* and *Mmp* formed the NF for the cadmium exposure, and *αTub*, *Mmp*, and *Ubi* formed the NF for the copper exposure.

The NFs developed here are based on consensus findings between three different reference gene analysis programs: BestKeeper, GeNorm, and NormFinder. All three look at the variation in the expression of the candidate reference genes as a measure of stability but in different ways. The results from GeNorm and NormFinder were much more similar to one another than to those from BestKeeper. This is caused by a couple key differences in the programs. BestKeeper looks only at the variability in expression of a single gene without distinguishing between or comparing treatments and controls. GeNorm compares genes to one another also without distinguishing. NormFinder

compares genes across groups, taking treatments and controls into account. *Ubi* has very low variation in its expression, so it is highly ranked by BestKeeper. However, the expression of *Ubi* is very different when compared to the other genes, giving it a low rank by GeNorm and NormFinder. The difference in gene stability can be attributable to the difference in input values as well. While GeNorm and NormFinder both use relative ΔCt values, BestKeeper uses Ct values. Because this similarity in input values allowed direct comparison between the two programs, the results from GeNorm and NormFinder were given greater weight when deciding which reference genes to use in the NFs.

NFs developed here are useful for expression studies but only under the conditions investigated. The stability of gene expression can change rapidly when concentrations, stressors, ambient conditions, or the test organism are changed. When European honey bees (*Apis mellifera*) were exposed to the biological stressor *Escherichia coli*, an NF consisting of *Act*, *GAPDH*, and *RPS18* was used (Scharlaken et al., 2008). *RPS18*, *RPL4*, *ARF1*, and *ARF4* were most stable when comparing different life stages of Colorado potato beetle (*Leptinotarsa decelmineata*) (Shi et al., 2013). *RPS3*, *RPS18*, and *RPL13a* served as the NF for the red flour beetle (*Tribolium castaneum*) after exposure to the entomopathogenic fungus *Beauveria bassiana* (Lord et al., 2010). The NF for *Vibrio spendidus* exposure in soft-shell clams (*Mya arenaria*) consisted of *EF-1*, *RPS18*, and *Ubi* (Araya et al., 2008), while that for ostreid herpesvirus stress in Pacific oyster (*Crassostrea gigas*) consisted of *RPL7* and *RPS18*. In *Daphnia pulex* genes *Xbp1*, *Tbp*, *CAPON*, and *Syx16* formed the NF for expression studies after exposure to predator stress from *Chaoborus* (Spanier et al., 2010). In *D. magna* after exposure to ibuprofen, *GAPDH*, *Ubc*, and *Act* served as the NF (Heckmann et al., 2006).

Some potential reference genes appear useful in more varied situations that others, such as *RPS18*, as seen above. Many of the reference genes investigated for *H. azteca* were also considered for use or used in NFs in these other species under different exposure conditions. While the sets of genes investigated stays fairly constant between these different species and exposures, the genes that prove to be stable varies widely. This underlines the importance of performing expression analysis of reference genes under every new set of experimental conditions.

When using whole tissues or whole specimens as I did here with *H. azteca*, traditional methods of mRNA level normalization like normalizing to sample size, such as the number of cells or tissue volume, is not possible (Vandesompele et al., 2002). Normalization has therefore moved towards using the expression of a stably expressed housekeeping gene. Given the increasing sensitivity of gene expression measurement technology, many housekeeping genes previously used as stable reference genes now are known to change in expression under certain conditions (e.g., Heckmann et al., 2006; Hauton et al., 2009). Thus these "classic" reference genes may not always serve as effective reference genes. This led to the development of NFs using multiple housekeeping genes.

NFs provide a robust method of normalization that can be used reliably in sensitive applications. The NF strategy does have drawbacks however. Measuring multiple reference genes can add significant cost and time to sample processing. Limited amount of sample RNA can also limit the use of multiple reference genes. Finally NFs use only the most stable combination among the reference genes investigated. These

genes may less stable than other unmeasured reference genes (Huggett et al., 2005). Currently the one alternative that potentially matches NFs in terms of reliability is the use of artificial molecules synthesized and introduced into the samples at known concentrations (Bustin and Nolan, 2004). The development of these artificial molecules is costly in both money and time. For these reasons NFs remain the current gold standard for expression normalization.

These NFs for cadmium and copper exposure serve important quality control purposes in the measurement of gene expression biomarkers in *H. azteca*. Ensuring accurate quantification of expression responses is critical for using these biomarkers for detecting sub-lethal levels of contaminants in the environment. Normalization of gene expression using unstable reference genes can lead to over or underestimating the magnitude of the biomarker response, which in turn leads to incorrect risk management decisions (e.g., Spanier et al., 2010). The appropriate use of these NFs in measuring stress response gene biomarkers is demonstrated in Chapter 3 for cadmium and copper. The versatility of these biomarkers is again demonstrated in Chapter 4 for exposures to insecticides.

Chapter 3: Alterations of gene expression in *Hyalella azteca* as biomarkers of exposure to cadmium and copper

Abstract

Biomarkers based on changes in gene expression promise to detect much lower levels of chemical stressors than biomarkers based on cell or tissue damage or changes in survival. Cadmium and copper are known to induce expression changes in detoxification genes involved in eliminating reactive oxygen species created by these metals. These genes include those for superoxide dismutase, catalase, and glutathione *S*-transferase. These molecular biomarkers have been tested in aquatic organisms under aqueous exposure, but no matching work has been performed with organisms associated with sediment. Sediments bind and accumulate metals, giving benthic organisms potentially higher exposures than organisms in the water column. Here the epibenthic amphipod *Hyalella azteca* was exposed to sublethal concentrations of cadmium and copper. Expression of stress-response genes in *H. azteca* was measured after a 24-hour exposure. Both cadmium and copper exposure induced expression of a cap n' collar protein gene (*Cnc*), the ABC transporter Mrp4 (*Mrp4*), and heat shock protein 90 (*Hsp90*). Cadmium uniquely induced expression of a DNA repair protein RAD51 gene (*Rad51*) and also induced 3.5-fold higher expression of *Mrp4* than copper did. The expression of these genes can serve as biomarkers of exposure to these metals in *H. azteca*, providing a sensitive method for detecting very low environmental levels of cadmium and copper. The importance of these biomarkers for ecological risk assessment and future work to develop the biomarkers in *H. azteca* are discussed.

Introduction

Metals released from human activities are now a nearly ubiquitous environmental contaminant. Multiple trophic levels are exposed to metals through direct accumulation and trophic transfer (Watras et al., 1998; Nussey et al., 2000; Vinodhini and Narayanan, 2008). Though generalizations about the behavior of metals in aquatic food webs are difficult because of differences between organisms, habitats, and metals, there is evidence for some specific situations. Cadmium and copper are both of concern in the environment because of their increasing prevalence from human activities and their potential toxicity (ATSDR, 2004; Pan et al., 2010; ATSDR, 2012). Both metals show evidence of bioaccumulation and biomagnification in certain environments or in short food chains (Croteau et al., 2005; Mathews and Fisher, 2008), but not in others (Barwick and Maher, 2003). Given their potential for bioaccumulation and biomagnification, we must detect very low exposures of cadmium and copper to mitigate the risk of adverse effects.

Cadmium acts as a calcium antagonist, disrupting calcium uptake (Verbost et al., 1987; Verbost et al., 1988; Verbost et al., 1989). The most famous case of cadmium's calcium antagonism is itai-itai ("ouch-ouch") disease, which appeared in Japan around 1912 as mine effluent high in cadmium began contaminating local water resources. Named for the cries of those afflicted, itai-itai disease was characterized by decreased bone mass due to decalcification by cadmium, bone deformities, and fractures. Similar to the effects on human bone, this decreased calcium uptake can also alter the biomineralization of calcium carbonate in the arthropod exoskeleton, potentially leading to malformations and inhibited growth (Machado and Lopes-Lima, 2011). Cadmium catalyzes the formation of reactive oxygen species (ROS), such as hydrogen peroxide

(Pinot et al., 2000). This oxidative stress leads to lipid peroxidation, membrane and tissue damage, impaired cellular functions, and DNA breaks (Stohs and Bagchi, 1995; Dally and Hartwig, 1997; Stohs et al., 2000).

Molecular responses to cadmium exposure are intended to eliminate cadmium from the body and prevent and repair damage. ABC transporters within the subfamilies ABCB and ABCC efflux cadmium from cells to make it available for elimination (Oh et al., 2009; Della Torre et al., 2012; Bourdineaud et al., 2015). Metallothionein and ferritin bind cadmium to prevent further damage and can transport cadmium to elimination sites (Shaw et al., 2007; Fan et al., 2008; Connon et al., 2008). Superoxide dismutases (SODs) and catalases which breakdown the radicals produced by cadmium are often upregulated (Connon et al., 2008; Lyu et al., 2014; Liping et al., 2014). Glutathione *S*-transferases (GSTs), antioxidant proteins, also respond to oxidative stress to eliminate radical compounds (Poynton et al., 2007; Lee et al., 2008). A wide variety of heat shock proteins (HSPs) are also upregulated. HSPs respond to many stressors and are generally involved in chaperone activity and translocation of proteins into organelles, both activities to sequester other proteins from cadmium damage (Werner and Nagel, 1997; Liping et al., 2014; Zhang et al., 2015).

Copper also catalyzes the formation of ROS, causing much of the same damage as with cadmium: lipid membrane peroxidation, tissue damage, and formation of DNA strand breaks (Stohs and Bagchi, 1995). The response to copper-induced damage is also similar to that of cadmium. Upregulation is seen in metallothionein, ferritin, HSPs, GSTs, and SOD (Correia et al., 2002; Lee et al., 2008; Poynton et al., 2008a; Poynton et al.,

2008b; Ki et al., 2009). An interesting decrease in expression of immune response genes is seen after copper exposure including genes for lectins and beta-1,3-glucan binding protein (LGBP), a change that could leave an exposed organism more susceptible to other stressors (Poynton et al., 2007; Poynton et al., 2008a).

Changes in expression of genes involved in the response to stressors and their damage are more sensitive biomarkers of exposure than the damage itself. Expression induction or repression can occur at levels of the stressor much lower than those required for visible damage to cells or tissues or for changes in survival or behavior. Here I aim to identify stress-response genes in *H. azteca* that respond to exposure to cadmium and copper. Many of the genes described above as responding to cadmium or copper in other organisms will be analyzed for expression changes following exposure to sub-lethal concentrations of these metals. Increased expression is expected in genes for catalase, GST, SOD, and ferritin after exposure to both cadmium and copper. Decreased expression in the immune response-related lectin and LGBP genes is anticipated for exposure to copper only. Analysis of changes in the expression of these genes will reveal which genes respond in a general manner to both metals and which show responses specific to one metal versus the other. These expression changes can serve as biomarkers of exposure and important lines of evidence in ERA for these metals.

Materials and Methods

H. azteca were exposed to sub-lethal concentrations of cadmium and copper for 24 hours. RNA was then extracted and reverse transcribed. Expression of metal stress-related genes was measured in controls and treatments using RT-qPCR. Expression was

normalized using the NFs developed in Chapter 2. The normalized expression of each gene was compared between treatments and the control to develop fold changes in expression.

Gene of interest primer design and testing

I designed RT-qPCR primers for genes of interest (GOI) using online design tools from IDT DNA and GenScript (Table 3.1; Appendix D). I tested the primers using the same two samples of *H. azteca* cDNA used in Chapter 1 to test reference primers. PCR reactions consisted of the following: 10 µL of 2x PCR Master Mix (SydLabs), 1 µL cDNA, 1 µL of 100 µM forward primer, 1 µL of 100 µM reverse primer, 7 µL PCR grade water. PCR was run with the following cycles: 1) 1 cycle of 95ºC for 2 minutes; 2) 25 cycles of 95ºC for 15 seconds, 57ºC for 15 seconds, and 72ºC for 30 seconds; 3) 1 cycle of 72ºC for 7 minutes; 4) held 4ºC indefinitely until PCR products were stored at -20ºC. All reactions were paired with a negative control with 1 µL of water added instead of cDNA. I ran PCR products on a 4% agarose gel at 110 V for 40 minutes along with a 10 bp DNA ladder (Life Technologies) and visualized the gel using ethidium bromide and a UV camera system. Products were sent to GENEWIZ, Inc. for Sanger sequencing using the forward primer of each primer set as the sequencing primer. I aligned experimental PCR product sequences to their corresponding expected amplicon sequences (Appendix E) and analyzed for percent similarity using EMBOSS Water (Rice et al., 2000). I was unable to include genes for metallothioneins (MTs), important contributors to metal sequestration and excretion, because no MTs were captured in the transcriptome analysis (Oh et al., 2009). Upon completion of the *H. azteca* genome, additional analysis of samples for any annotated MTs would be of great utility to the present study, as they

have previously been implicated in cadmium resistance in another amphipod species *Gammarus pulex* (Stuhlbacher and Maltby, 1992).

Table 3.1: GOI transcripts identified in the *H. azteca* transcriptome and their corresponding primers

Transcript name	Abbreviation	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Amplicon length (nt)
Segmentation protein cap n' collar	*Cnc*	AAATGCTATCT GTGGAGGTGG	CGAAGAGACCA TCAAGAGCAG	149
Kelch-like ECH-associated protein	*Keap*	GCCTCCGCTAT CTTTTCTTGG	GGGCTCCATCT CATTATCTACG	137
Glutathione S-transferase	*Gst*	TCCTTGCCATC AATCCAGAG	TCTTTGCCGTAC TGGTTCAC	109
Heat shock protein 70	*Hsp70*	ACCATTCCCAC TAAACACTCG	TCGTAAACCTG GATGAGCAC	77
Ferritin	*Fer*	TCCTTTCTAAA ATCCTGGCCTC	GATCCATTCCT CGTCTCACTTG	81
Catalase	*Cat*	CACACCTTCAA ACTCGTCAAC	AGTTCCTTGAT GCCCTGATC	83
C-type lectin 1	*Lectin-1*	GGCTCAGTAC GTTTACCTCAA C	ACTACATTTGT GCCCTGTCTC	122
C-type lectin 4	*Lectin-4*	AAGTTCCTTGA GAGCCGTAAC	GGGATTTTCTG ATTGGTGTTCG	140
Multidrug resistance-associated protein 4	*Mrp4*	GGAAGCGATA AAGACTCAGA GG	ATGGACCCCGA TATTTGTACG	139
Multidrug resistance-associated protein 5	*Mrp5*	TTTCTCAAGGG ACATGGACG	AGGATCACTAA CATTGGCAGG	141
Copper zinc superoxide dismutase	*Sod*	GCTCTACCCAT CTGACAATGA G	TGATGTGCCCC GTGATATG	134
Lipopolysaccharide and beta-1,3-glucan binding protein	*Lgbp*	AGATGGCACC CTTTGATCG	TGGGAAGTAAC CATTGACACC	74
Cytochrome P450 2c9-like isoform x1	*Cyp2c9*	CGACATTTTCC ACGATTTCCAG	TTGTTACGCCG AGCTTCTC	104
Cytochrome P450 18a1-like	*Cyp18a1*	AAGTTCTGTCG TGGCTCTTC	CGAAATTCTTG CGGTTCATCC	72
Alpha amylase	*Amy*	GGCACCTTGTC TATGAACCTG	CGTTGGTAATT GTGATCTGCG	118
DNA repair protein RAD51	*Rad51*	CTCCGATGAG GTGAACAAGA G	ACGGCTTCCTT AAACTGAGG	115

RNA was extracted from three pooled samples of twenty mixed-sex *H. azteca*. cDNA synthesis was performed with appropriate -RT controls as described in Chapter 1.

cDNA and -RT controls were added to wells of a Roche LightCycler 480 96-well plate in duplicate. The mixture for each reaction was 10 µL of 2x SYBR Green I, 0.5 µL of 10 µM of each primer, 5 µL of diluted cDNA, and 4 µL of PCR grade water. Six replicates of a water negative control were performed for each primer. RT-qPCR cycles used on the LightCycler 480 RT-qPCR machine were as follows: 1) a single pre-incubation period of 95 °C for 10 min, 2) an amplification period consisting of 35 cycles of 95 °C for 15 s, 61 °C for 30 s, and 72 °C for 30 s, with a single fluorescence measurement performed at the end of each cycle, 3) a single melting curve analysis of 95°C for 0.05 s followed by an increase of 0.5 °C/cycle from 65 °C to 97 °C, with 5 fluorescence readings/°C, and 4) a cooling period of 40 °C for 1 min. I calculated primer efficiencies for each primer set from the individual RT-qPCR reactions using the real-time PCR Miner algorithm (Zhao and Fernald, 2005; http://ewindup.info/miner/). I then analyzed primer efficiencies of GOI primers and reference primers from Chapter 2 together in a two-tailed ANOVA to test for significant differences.

Cadmium and copper exposure

The same samples of cDNA from *H. azteca* exposed to cadmium or copper in Chapter 2 were used for analysis of GOI. This decreased the variation that would be introduced by analyzing reference genes and GOI across different experimental exposures. Details of exposures, including organisms, chemicals, and treatments, are found in Chapter 2.

Expression analysis

The same optimized RT-qPCR protocol described above for measuring primer efficiencies was used for measurement of GOI expression. All biological replicates were plated with three technical replicates. Cycle threshold (Ct) values were calculated using fit point analysis on LightCycler 480 Software release 1.5.0 SP3 (version 1.5.0.39). The means of Ct values of all technical replicates were calculated as the Ct values for each biological replicate. I calculated delta (Ct) values (ΔCt) using equation (1) for each replicate for each gene, where ε is the efficiency of the primer for the gene being analyzed, $Ct_{gene,min}$ is the minimum Ct value of the gene being analyzed across all replicates, and Ct_x is the Ct value of the biological replicate being analyzed. This served to calibrate all values against the lowest expression value.

$$(1)\ \Delta Ct = \varepsilon^{Ct_{gene,min} - Ct_x}$$

I calculated normalized expression levels (NELs) by dividing each ΔCt value by the corresponding normalization factor (NF) from Chapter 2. The standard error of the mean (sem) of all NELs was calculated incorporating proper error propagation using equation (2), where $NELsem_x$ is the standard error of the mean of the NEL for the gene being analyzed, NEL_x is the NEL of the same gene, NFsem is the standard error of the mean of the NF, $\Delta Ctsem_x$ is the standard error of the mean of the ΔCt value of the same gene, and ΔCt_x is the ΔCt value of the same gene.

$$(2)\ NELsem_x = NEL_x * \sqrt{\left(\frac{NFsem}{NF}\right)^2 + \left(\frac{\Delta Ctsem_x}{\Delta Ct_x}\right)^2}$$

The relative fold change (RFC) in expression of each gene was then calculated relative to the control using equation (3). The standard error of the mean of each RFC (RFCsem) was calculated incorporating proper error propagation using equation (4).

$$(3)\ RFC_{gene} = \frac{NEL_{gene,treatment}}{NEL_{gene,control}}$$

$$(4)\ RFCsem_x = \sqrt{NELsem^2_{x,treatment} + NELsem^2_{x,control}}$$

Relative expression levels (RELs) were calculated for each gene as the NEL of that gene's expression in each individual replicate divided by the geometric mean of the expression of that gene in the control. The geometric mean of the RELs of all replicates of a treatment for a single gene is equal to the RFC of that gene for that treatment. RELs allow for statistical analysis of expression values using all replicates. RELs were analyzed with a Wilcoxon rank-sum test in SAS using the proc npar1way procedure.

Results

Gene of interest primer design

PCR products for GOI primers *Cnc*, *Gst*, *Fer*, *Lectin-1*, *Lectin-4*, *Mrp4*, *Sod*, *Lgbp*, and *Amy* appeared at the expected positions on the agarose gel (Figure 3.1 - bright band in lanes 1 and 14 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Cnc* negative control, 3) *Cnc* sample 1, 4) *Cnc* sample 2, 5) *Keap* negative control, 6) *Keap* sample 1, 7) *Keap* sample 2, 8) *Gst* negative control, 9) *Gst* sample 1, 10) *Gst* sample 2, 11) *Hsp70* negative control, 12) *Hsp70* sample 1, 13) *Hsp70* sample 2, 14) 10 bp DNA ladder, 15) *Fer* negative control, 16) *Fer* sample 1, 17) *Fer* sample 2, 18) *Cat* negative control, 19) *Cat*

sample 1, 20) *Cat* sample 2, 21) *Lectin-1* negative control, 22) *Lectin-1* sample 1, 23) *Lectin-1* sample 2, 24) *Lectin-4* negative control, 25) *Lectin-4* sample 1, 26) *Lectin-4* sample 2) (Figure 3.2 - bright band in lanes 1 and 14 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Mrp4* negative control, 3) *Mrp4* sample 1, 4) *Mrp4* sample 2, 5) *Mrp5* negative control, 6) *Mrp5* sample 1, 7) *Mrp5* sample 2, 8) *Sod* negative control, 9) *Sod* sample 1, 10) *Sod* sample 2, 11) *Lgbp* negative control, 12) *Lgbp* sample 1, 13) *Lgbp* sample 2, 14) 10 bp DNA ladder, 15) *Cyp2c9* negative control, 16) *Cyp2c9* sample 1, 17) *Cyp2c9* sample 2, 18) *Cyp18a1* negative control, 19) *Cyp18a1* sample 1, 20) *Cyp18a1* sample 2, 21) *Amy* negative control, 22) *Amy* sample 1, 23) *Amy* sample 2, 24) *Rad51* negative control, 25) *Rad51* sample 1, 26) *Rad51* sample 2).

Initial primers for *Keap*, *Hsp70*, *Cat*, *Mrp5*, *Cyp2c9*, *Cyp18a1*, and *Rad50* did not produce bands. New primers were designed for *Keap*, *Cat*, and *Mrp5*, as well as new transcripts for cytochrome P450 2j2 (*Cyp2j2*), heat shock 70 kDa protein 4 (*Hsp704*), and DNA repair protein RAD51 (*Rad51*) (Table 3.2). Of these new primers, those for *Hsp704* and *Rad51* worked (Figure 3.3 - bright band in lanes 1 and 14 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Keap2* negative control, 3) *Keap2* sample 1, 4) *Keap2* sample 2, 5) *Cyp2j2* negative control, 6) *Cyp2j2* sample 1, 7) *Cyp2j2* sample 2, 8) *Cat2* negative control, 9) *Cat2* sample 1, 10) *Cat2* sample 2, 11) *Mrp5-2* negative control, 12) *Mrp5-2* sample 1, 13) *Mrp5-2* sample 2, 14) 10 bp DNA ladder, 15) *Hsp704* negative control, 16) *Hsp704* sample 1, 17) *Hsp704* sample 2, 18) *Rad51* negative control, 19) *Rad51* sample 1, 20) *Rad51* sample 2).

A final round of primers were designed for *Keap*, *Cat*, and *Mrp5*, along with primers for new transcripts for heat shock protein 90 (*Hsp90*) and cytochrome P450 (*Cyp*) (Table 3.3). Of these final primers, only those for *Cat* and *Hsp90* worked (Figure 3.4 - bright band in lanes 1 and 14 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Mrp5-3* negative control, 3) *Mrp5-3* sample 1, 4) *Mrp5-3* sample 2, 5) *Cat3* negative control, 6) *Cat3* sample 1, 7) *Cat3* sample 2, 8) *Cat4* negative control, 9) *Cat4* sample 1, 10) *Cat4* sample 2, 11) *Hsp90* negative control, 12) *Hsp90* sample 1, 13) *Hsp90* sample 2, 14) 10 bp DNA ladder, 15) *Cyp* negative control, 16) *Cyp* sample 1, 17) *Cyp* sample 2, 18) *Keap3* negative control, 19) *Keap3* sample 1, 20) *Keap3* sample 2, 21) *Keap4* negative control, 22) *Keap4* sample 1, 23) *Keap4* sample 2). Sequenced PCR products are found in Table 3.4, and alignment results of PCR products and expected amplicons are found in Table 3.5.

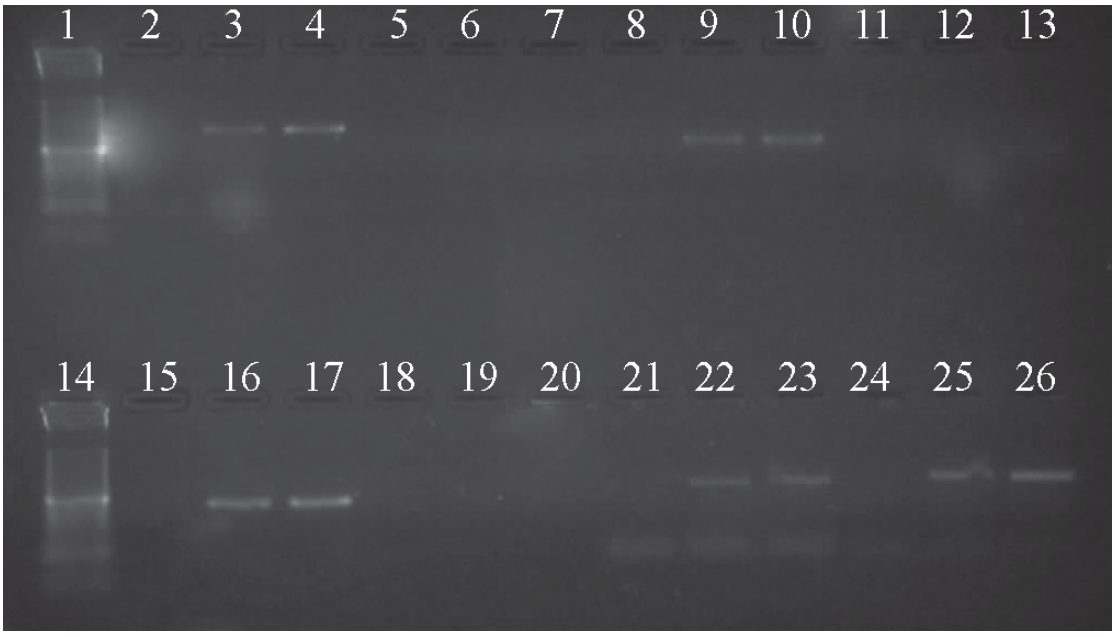

Figure 3.1: PCR products from *H. azteca* cDNA using primers for *Cnc*, *Keap*, *Gst*, *Hsp70*, *Fer*, *Cat*, *Lectin-1*, and *Lectin-4*

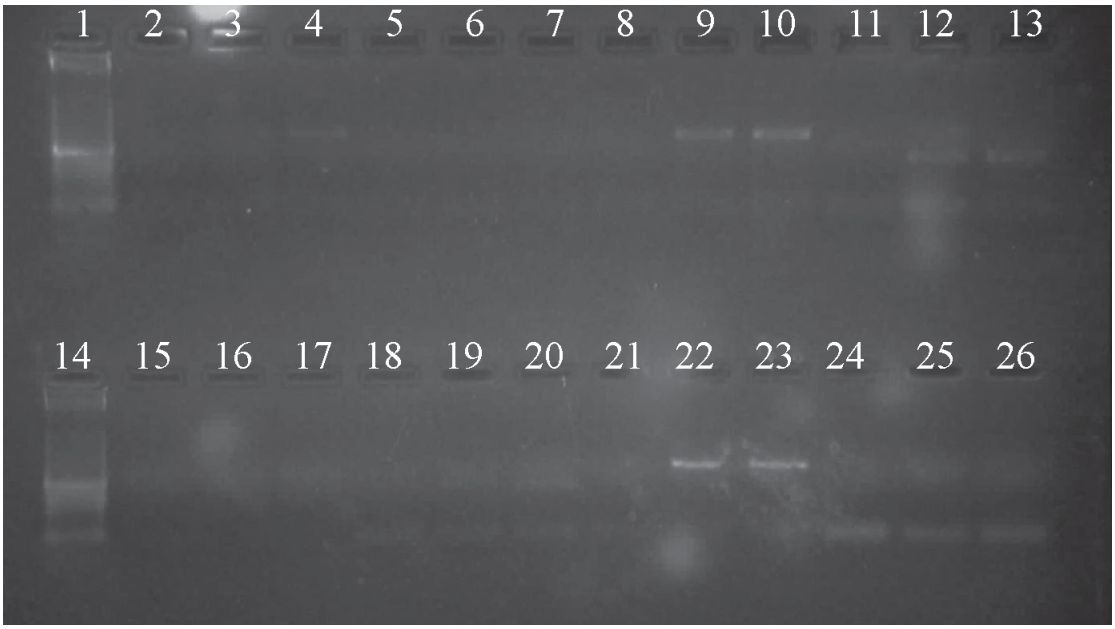

Figure 3.2: PCR products from *H. azteca* cDNA using primers for *Mrp4*, *Mrp5*, *Sod*, *Lgbp*, *Cyp2c9*, *Cyp18a1*, *Amy*, and *Rad51*

Table 3.2: Names and sequences of additional primers designed for *Keap*, *Cyp2j2*, *Cat*, *Mrp5*, *Hsp704*, and *Rad51*

Transcript name	Abbreviation	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Amplicon length (nt)
Kelch-like ECH-associated protein	*Keap2*	CCTCCGGGTAA TCGAAGCTA	GAAGTGGAAG AACCGGGAGA	122
Cytochrome P450 2j2-like	*Cyp2j2*	CATCTTCACAG CACGAAATGG	ACTCCCCTCTT TTCATACTGC	149
Catalase	*Cat2*	CGTTCGACGTC ACCAAAGTT	CTCCACCTCGG CGAAATAGT	104
Multidrug resistance-associated protein 5 (ABCC5, MRP5)	*Mrp5-2*	CCTTCACGACC AAGTCCTCA	CGTCCATGTCC CTTGAGAAA	104
Heat shock 70 kDa protein 4	*Hsp704*	ATGACGAAAAG AGGGAGAAGC	CGAGCAAACG TCCATGAATTG	143
DNA repair protein RAD51	*Rad51*	GATACTCCCAA AACTCGCAAAG	GAGATCGTAA GTTCATCCGCC	115

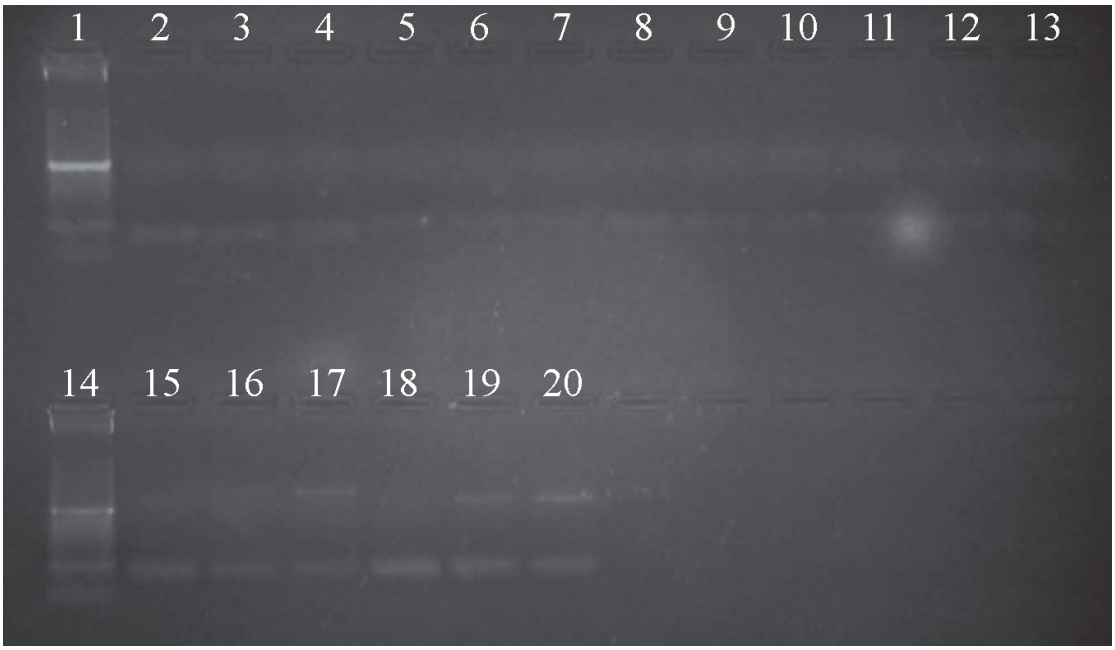

Figure 3.3: PCR products from *H. azteca* cDNA using primers for *Keap*, *Cyp2j2*, *Cat*, *Mrp5*, *Hsp704*, and *Rad51*

Table 3.3: Names and sequences of additonal primers designed for *Mrp5*, *Cat*, *Hsp90*, *Cyp*, and *Keap*

Transcript name	Abbreviation	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Amplicon length (nt)
Multidrug resistance-associated protein 5 (ABCC5, MRP5)	*Mrp5-3*	GGACATGGACG AAGTGGATG	CATACAAGCA GCAGCGAACA	93
Catalase	*Cat3*	TCTTCATTCGG GATCCTGTG	GCACACCTGG TGAGTGGTCT	134
	Cat4	CGACAACATAG CAGGCCACT	AGTTTAGCGAT GCGTTCCCT	120
Heat shock protein 90	*Hsp90*	AAACTCGTCTG ATGCTCTGG	AATGCCACTGT CCCGAATAG	139
Cytochrome P450	*Cyp*	GAAGAGATGGA GGTGAAGACG	CCTCGGGTGG AACTTAAAGG	72
Kelch-like ECH-associated protein	*Keap3*	CTGGTCGACTG CAGGACATT	TCAACGGGAA ACCTGTTCCT	87
	Keap4	AACAGGTTTCC CGTTGAAGC	CGCTCTCAAA GAACTTCGCA	90

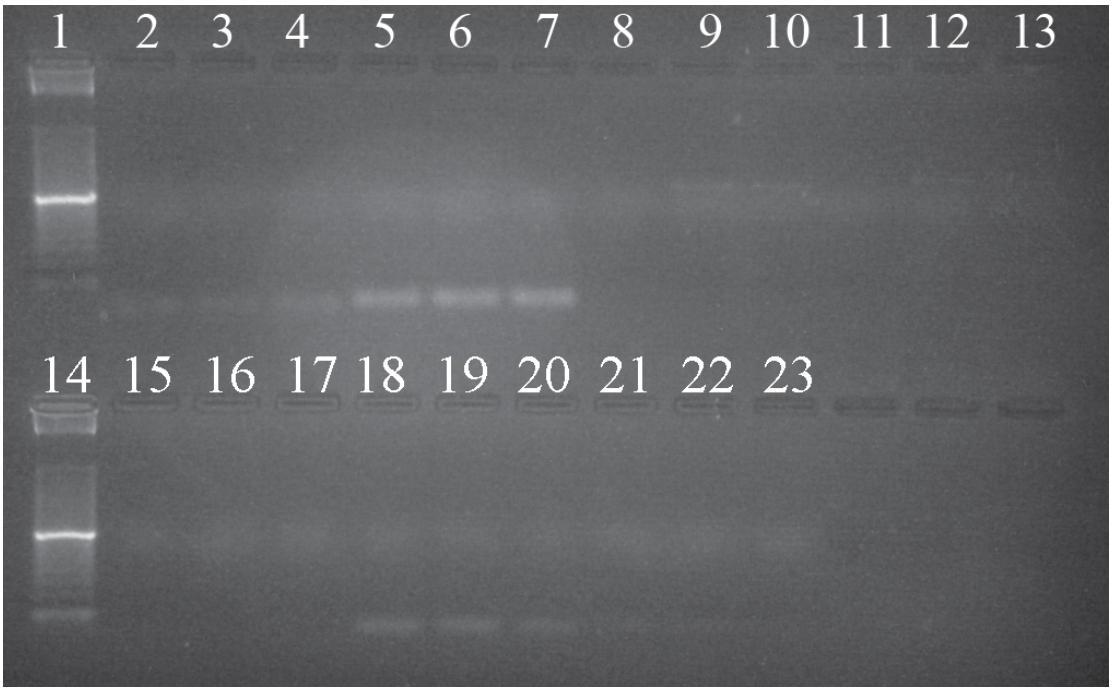

Figure 3.4: PCR products from *H. azteca* cDNA using primers for *Mrp5*, *Cat*, *Keap*, *Hsp90*, and *Cyp*

Table 3.4: Sequences of amplicons obtained through PCR

Primer set	Sample	Product sequence
Cnc	1	AGTTTCTTCCTCTAGACTGTTACTTGTATATTTTGGACAAGTTGGATTATTCG
	2	AGTTTCTTNCTCTAGACTGTTACTTGTATATTTTGGACAAGTTGGATTATTCG
Gst	1	CGCCCCGGTTGGTANTATCTGCTG-AGCATGATGGCAAAGA
	2	ATGGNCCA-TCGCGCCGTTTGGGCGTATCTGNTGTGCCATGACGGCAAAGA
Fer	1	CTCTCCTTAGGTGC-AGTGAGACGAG
	2	GNCTCTCCTTAGGTGC-AGTGAGACGAG
Lectin-1	1	TCNGCGCCGGANANGCGGTCGGAGTAGCTGATGCCTCCGTCGTTTGTCAA TTGGCTTAGAGAGACAGGGCACAAATGTAG
	2	CTTCNGCGCCGGANANGCGGTCGGAGTAGCTGATGCCTCCGTCGTTTGTCAAT TGGCTTAGAGAGACAGGGCACAAATGTAGT
Lectin-4	1	GNNGCCTCTGCGACTTCACTGGCTGCGAGACTAGGGAGGACTTGAAGCCAAT CAACATCAAGGGTTGGTTCTGGTCGAACACCAATCAGAAAATCCC
	2	CGGCCGCCTCTGCGACTTCACTGGCTGCGAGACTAGGGAGGACTTGAAGCCA ATCAACATCAAGGGTTGGTTCTGGTCGAACACCAATCAGAAAATCCC
Mrp4	1	GAGCCAGAAGTTGCGTCTGCTAGTCACGGGTCTATCGACGAGCTTACTAAATT GTAGAANCATTTAAGGCCGTACAAATATCGGGGTCCAT
	2	TGAGCCAGAAGTTGCGTCTGCTAGTCACGGGTCTATCGACGAGCTTACTAAAT TGTAGAAACATTTAAGGCCGTACAAATATCGGGGTCC
Sod	1	AGCGTNNTGCTTATGGG-AGCATCCGCGGATTCTTATAACTTCGTCAGGACC AACCTCCTGTTGGTANGACCCATATCACGGGGCACATCA
	2	AGCGTCCTTCTTATGGG-AGCATCCGCGGATTCTT-AAACTTCGTCAGGACCA ACCTCCTGTTGGTAAGACCCATATCACGGGGCACATCA
Lgbp	1	CTGGCTGTTGGTGNNGT-NATGGTTACTT
	2	CNGGNTGTTGGTGGTGTCAATGGTTACTT
Amy	1	CTCAGAGTTGGTGACATGCGCCGATGTCACGGTCAACGGAGACGGGACAGCG CAGATCACAATTACCAAC
	2	TACTCAGAGTTGGTGACATGCGCCGATGTCACGGTCAACGGAGACGGGACAG CGCAGATCACAATTACCAACG
Hsp704	1	GNACNGGAGACTCTCTGTGTTGAGTT-ANGGANTTGATTAGCGAGCAGAGGG ACCAAGATCTGATGGAGAANGATGAGTTGGACCAATTCATGGACGTTTGCTC G
	2	CTGGNGNNTCTCTGTGTTGAGTTAAAGGANTTGATTAGCGAGCAGAGGGACC AAGATCTGATGGAGGACGATGAGTTGGACCAATTCATGGACGTTTGCNCG
Rad51	1	GCNNTANNTTCTTGTCTATGATTGGNGGATTTTGCTTG-ANGGCCGTCACTGATTGGCGGATGAACTTACGANCTC
	2	GCNNTANNTTCTTGTCTATGATTGGNGGATTTTGCTTGAA--GGCGTCACTGATTGGCGGATGAACTTACGAT
Cat4	1	GCGCCATCGG-AACTTCGCTAAAGTCNACGCGGGCTTCGGACAGGCCATCA GGGAACGCATCGCTAAACT
	2	TCNNNGGNNGCGCNATCGG-AACTTCGCTAAAGTCNACGCGGGCTTCGGAC AGGCCATCAGGGAACGCATCGCTAAACT
Hsp90	1	N
	2	GATCCTAGC-AGCTGGACTCTGGAAAAGANCTTTTCATCAAACTCGAGCCTA ACAAGAACGACCGGACTTTAGCTATTCGGGACAGTGGCAT

Table 3.5: Alignment of transcriptome transcript sequences and those obtained through PCR

Primer set	Sample	Expected product length (nt)	Sequenced product length (nt)	Percent similarity (%)	Gaps (%)
Cnc	1	149	103	99	0
	2	149	103	98.1	0
Gst	1	109	41	65.9	2.4
	2	109	51	76.5	2
Fer	1	81	26	96.2	3.8
	2	81	28	92.9	3.6
Lectin-1	1	122	80	96.2	0
	2	122	83	96.4	0
Lectin-4	1	140	97	97.9	0
	2	140	99	100	0
Mrp4	1	139	91	96.7	0
	2	139	90	97.8	0
Sod	1	134	91	92.3	1.1
	2	134	91	96.7	2.2
Lgbp	1	74	29	86.2	3.4
	2	74	29	93.1	0
Amy	1	118	70	100	0
	2	118	73	100	0
Hsp704	1	143	105	93.3	1
	2	143	102	93.1	0
Rad51	1	115	76	89.5	1.3
	2	115	73	89	2.7
Cat4	1	120	70	97.1	1.4
	2	120	79	88.6	1.3
Hsp90	1	139	1	0	0
	2	139	92	96.7	1.1

GOI primer efficiencies ranged from 82% to 87% (Table 3.6). The significant difference existing between the efficiencies as indicated by two-tailed ANOVA ($P<0.05$) was eliminated by excluding the same reference primers as were excluded in Chapter 2 (*RpL10*, *Gapdh*, and *Syx2*) as well as the GOI primers *Amy* and *Hsp704* (two-tailed ANOVA after removal, $\alpha=0.05$, $F=1.34$ (15, 32), $P=0.24$). By gaining equivalency of efficiencies among the primers, the more streamlined comparative Ct method can be used for expression analysis.

Table 3.6: Amplification efficiencies of GOI primers

Primer set	Efficiency % (mean ± 1 standard deviation
Cnc	85 ± 0.3
Gst	85 ± 0.3
Fer	85 ± 0.5
Lectin-1	85 ± 0.5
Lectin-4	85 ± 0.9
Mrp4	85 ± 1
Sod	87 ± 0.9
Lgbp	87 ± 0.4
Amy	82 ± 0.6
Hsp704	83 ± 0.1
Rad51	86 ± 0.5
Cat4	87 ± 0.9
Hsp90	86 ± 0.4

Gene of interest expression analysis

Significant increases in expression were seen in *Cnc*, *Mrp4*, *Rad51*, and *Hsp90* in response to cadmium exposure ($P<0.05$) (Figure 3.5). Significant increase in expression of *Cnc*, *Mrp4*, and *Hsp90* was seen in response to copper exposure ($P<0.05$) (Figure 3.6). Increased expression of *Cnc*, *Mrp4*, and *Hsp90* was shared across both metals. Expression increase of *Rad51* was specific to cadmium exposure, as was a much larger increase in expression of *Mrp4*.

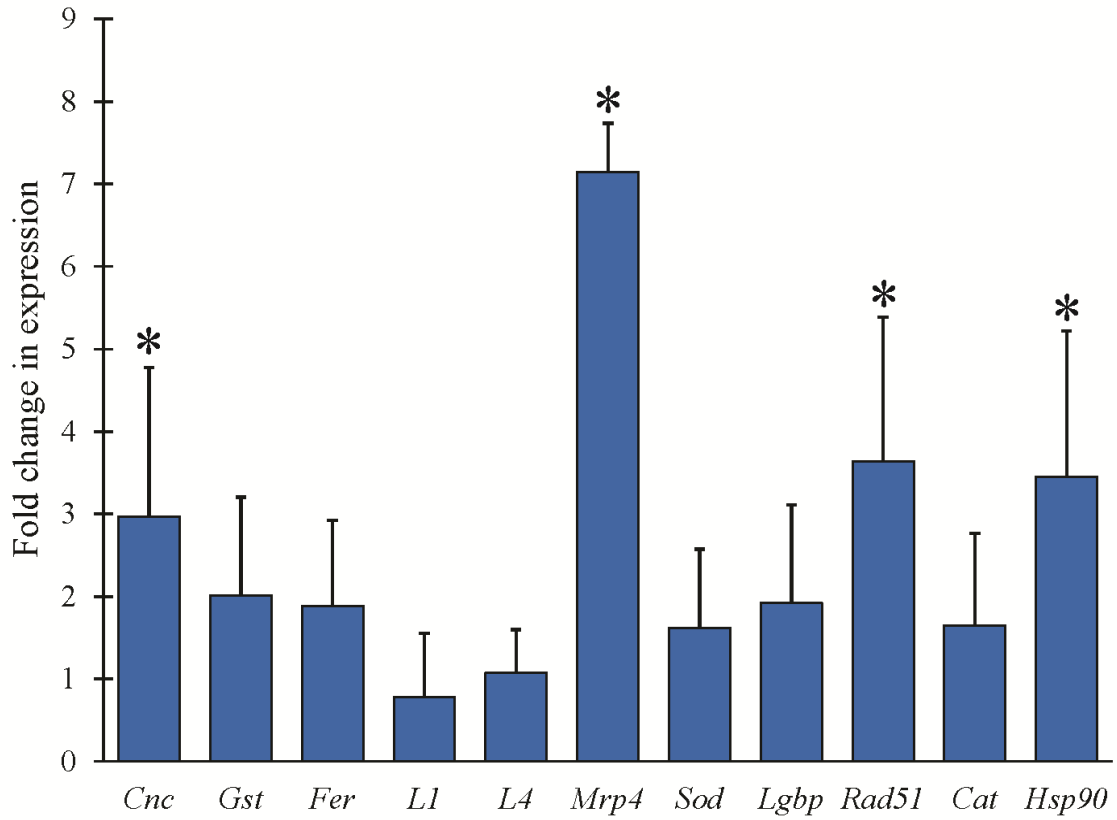

Figure 3.5: Fold change in gene expression of GOI in response to cadmium exposure. Bars are standard errors. Significant differences from control ($P < 0.05$) indicated by *. *L1* and *L4* stand for *Lectin-1* and *Lectin-4*, respectively

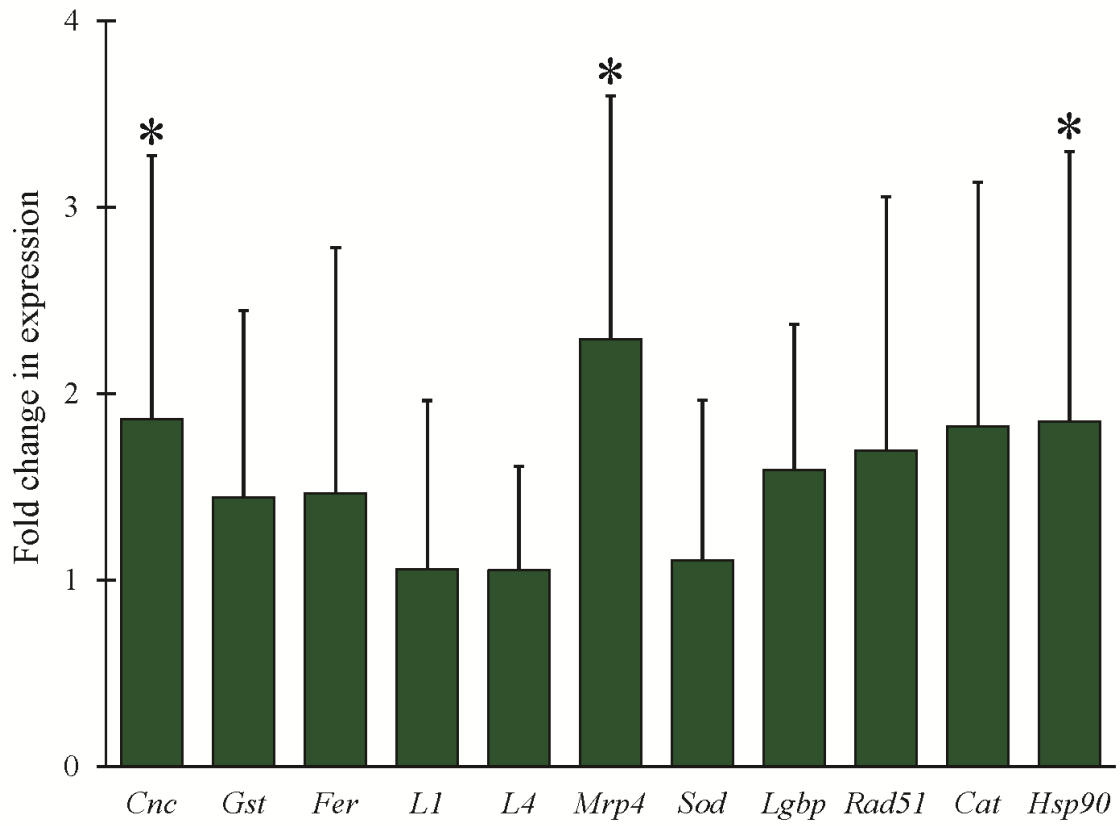

Figure 3.6: Fold change in gene expression of GOI in response to copper exposure. Bars are standard errors. Significant differences from control ($P < 0.05$) indicated by *. *L1* and *L4* stand for *Lectin-1* and *Lectin-4*, respectively

Discussion

Expression of stress-response genes was measured in *H. azteca* after exposure to cadmium and copper to develop molecular biomarkers of exposure for these metals. Cadmium increased the expression of genes for the nuclear receptor *Cnc*, the ABC transporter *Mrp4*, the DNA repair protein *Rad51*, and the heat shock protein *Hsp90*. Copper increased expression *Cnc*, *Mrp4*, and *Hsp90*, but to a much lower degree than cadmium. Increases in the expression of *Cnc*, *Mrp4*, and *Hsp90* were shared between cadmium and copper and could serve as general biomarkers of exposure to many different metals. Increased expression of *Rad51* and a larger increase of *Mrp4* expression

compared to copper were unique to cadmium. These may be unique biomarkers for cadmium at these low concentration levels.

Molecular biomarkers already used for cadmium exposure in freshwater include protein levels and gene expression of metallothioneins, ferritins, GSTs, and amylase (Amiard et al., 2006; Poynton et al., 2007; Soetaert et al., 2007b). Those for copper include metallothioneins, ferritins, lectins, and amylase (Correia et al., 2002; Amiard et al., 2006; Poynton et al., 2007; Poynton et al., 2008a; Poynton et al, 2008b). Significant changes in gene expression are seen in *D. magna* after exposure to LC5 levels of cadmium and copper (18 and 6 μg/L, respectively) (Poynton et al., 2007). Metallothionein protein concentrations in the amphipod *Gammarus locusta* responded to 4 μg/L copper after a minimum exposure time of 6 days (Correia et al., 2002). Expression changes found here in *H. azteca* showed detection sensitivities similar to these previous biomarkers in other organisms. The genes shown to increase in expression to both cadmium and copper (*Cnc*, *Mrp4*, and *Hsp90*) can serve as additional biomarkers of exposure for either of these metals at very low concentrations.

Cap n' collar is a family of proteins that contain basic leucine-zipper domains. Some Cnc proteins are homeotic genes involved in segmentation and appendage development (Mohler et al., 1995; McGinnis et al., 1998; Veraksa et al., 2000; Sharma et al., 2014). One of the best-known cap n' collar proteins is nuclear factor-erythroid 2-related factor (Nrf2). Nrf2 is a transcription factor which responds to many electrophilic chemical stressors including heavy metal salts by inducing a variety of stress-response genes including cytochrome P450s; GSTs; Mrps 2, 3, and 4; metallothionein; ferritin; and

HSPs (Kensler et al., 2007; Maher et al., 2007; Gu and Manautou, 2010; Higgins and Hayes, 2011). This signal cascade begins when the chemical stressor causes Keap1 protein to dissociate from Nrf2, allowing Nrf2 to enter the nucleus and induce expression of many different chemical stress response genes (Itoh et al., 2004; Nguyen et al., 2004). The *Cnc* gene transcript monitored here shows nearly 70% sequence identity with known *Nrf2* transcripts from two insects, a parasitoid wasp *Microplitis demolitor* and the brown marmorated stink bug *Halyomorpha halys*. The increased expression of *Mrp4* and *Hsp90* seen in response to both cadmium and copper in *H. azteca* could be caused, at least in part, by the induction cascade created by Nrf2 (He et al., 2008; Wang and Gallagher, 2013).

Mrp4/ABCC4 (*Mrp4*) is a member of the ABC transporters, a large family of proteins found in the cellular membrane which efflux their transport allocrites out of cells in an ATP-dependent manner. Certain ABCs in the subfamilies ABCB and ABCC are able to transport a large variety of toxic allocrites as a mechanism to detoxify and protect cells. For example some ABCs are implicated in insecticide resistance thanks to this broad-spectrum promiscuous efflux ability, especially members of the subfamilies ABCB, ABCC, and ABCG (Dermauw and Van Leeuwen, 2014). Some ABCs change in expression in response to heavy metal exposure, including the collembolan *Orchesella cincta* and the Asiatic clam *Corbicula fluminea* (Achard et al., 2004; Roelofs et al., 2007). ABCs are also well known for conferring resistance to metals, including cadmium, in plants (e.g., Kim et al., 2007).

Heat shock protein 90 (*Hsp90*) is within the heat shock protein (HSP) family. HSP90 acts as a chaperone to protect protein structures from oxidative damage or to mark damaged proteins for degradation (Parsell and Lindquist, 1993; Becker and Craig, 1994; Kalmar and Greensmith, 2009). Hsp90 is induced by changes in temperature, salinity, oxygen levels, pH, and concentrations of different organic and inorganic chemicals like metals in a variety of vertebrate and invertebrate species, including other crustaceans like the ridgetail white prawn *Exopalaemon carinicauda* and the Zhikong scallop (*Chlamys farreri*) (Feder and Hofmann, 1999; Choi et al., 2006; Gao et al., 2007; Li et al., 2012b). Expression of the *Hsp90* gene is induced by cadmium and copper, as well as by other metals such as selenium, in Pacific oyster (*Crassostrea gigas*) and Pacific abalone (*Haliotis discus hannai*) (Choi et al., 2008; Zhang et al., 2011; Guo and Ki, 2012).

Several interesting responses in *H. azteca* served to discriminate between exposure to cadmium and copper. *Mrp4* responded much more strongly to cadmium (7-fold increase in expression) than it did to copper (2.2-fold). ABCB1 and ABCC 1-4 respond to cadmium exposure; and increased *Mrp4* gene expression and Mrp4 protein titers are associated with cadmium resistance and cadmium stress response, explaining the much stronger response of *Mrp4* to cadmium compared to copper (Szczypka et al., 1994; Oh et al., 2009; Long et al., 2011; Della Torre et al., 2012; Bourdineaud et al., 2015).

Rad51 significantly increased in expression only after exposure to cadmium. Both cadmium and copper can cause DNA damage that would upregulate repair protein genes like *Rad51* (Sagripanti and Kraemer, 1989; Hartwig, 1994; Toyokuni and Sagripanti,

1996; Dally and Hartwig, 1997; Hossain, 2002; Pan and Zhang, 2006). The cadmium-specific response of *Rad51* could indicate that the damage from cadmium is more severe or the transcriptional response to cadmium is longer lasting because cadmium also inhibits DNA repair (Jin et al., 2003). *Lectin-1* was slightly downregulated by exposure to cadmium, though not significantly so. This downregulation could intensify at higher cadmium concentrations or begin at higher copper concentrations, as seen in *D. magna* in response to copper (Poynton et al., 2008a). These immunocompromised organisms would be less able to handle other stressors, especially biological ones.

Some genes (*Fer*, *Gst*, *Sod*, *Cat*) expected to increase in expression, as they have previously in other investigations of metal biomarkers, did not (Poynton et al., 2007). The metal concentrations eliciting expression responses in previous work varies relative to the concentration used here. Cadmium exposure was 20-fold lower (0.06 μg/L) and copper exposure was 1.8-fold lower (2 μg/L) in exposures using *D. magna* (Poynton et al., 2008a). In other studies cadmium exposure was 15-fold higher (18 μg/L) (Poynton et al., 2007) and 1.6-fold higher (2 μg/L) (Lyu et al., 2014). Copper exposure was 1.6-fold higher (6 μg/L) in exposures using *D. magna* (Poynton et al., 2007). First, much of these differences could be attributed to species-specific differences. Second, the concentrations used here perhaps were not sufficiently high to cause changes in expression in genes expected to change in expression. This could provide evidence of a no observed transcriptional effect level (NOTEL) for these individual genes (Lobenhofer et al., 2004; Ankley et al., 2006; Poynton et al., 2008a; Poynton et al., 2008b). The NOTEL is defined as the highest stressor concentration at which no significant change in gene expression occurs. The NOTEL is a suggested way to quickly determine if a substance could have a

toxic effect through a particular molecular pathway of interest, though it has historically been applied at a broader transcriptome scale.

Degrading mRNA could be responsible for the lower overall expression and the higher variability seen in expression values from the copper exposure. mRNA can be extremely transient, with degradation half-life of mRNA ranging from 20 minutes to over 24 hours (Sachs, 1993). The transcriptional response to the low concentrations of metals used here may have occurred quickly. Once homeostasis was reestablished in less than 24 hours, the mRNA would be degraded. Without continuous gene induction until the time of sampling, expression changes would be less likely to be detected. This kind of rapid change in concentration is one disadvantage of monitoring mRNA levels. A time course study of these exposures comparing gene expression at multiple time points would provide data on the optimal timing of sampling for each individual gene, i.e. when each gene peaks in expression following exposure. Finding a compromise between genes in terms of optimal sampling time would allow greater detection of expression changes.

Additional work needed to develop these gene expression biomarkers other than the time course study involve either the metals, the amphipods, or ambient conditions. First the metals should be tested using different concentrations. Developing a dose-response curve for expression for each gene would allow for better estimation of each gene's NOTEL (e.g., Poynton et al., 2008a). Overlapping concentrations between this curve and a mortality-based dose-response curve would create links between these molecular responses and adverse outcomes that could affect an ecosystem. These links form and give weight to Adverse Outcome Pathways (AOPs) (Ankley et al., 2010;

Vinken, 2013). AOPs describe the sequence of events starting at exposure and molecular interaction with a chemical through higher order levels of biological organization, ultimately ending in adverse ecological effects. Changes in gene expression are defined as molecular initiating events within the framework of the AOP (Allen et al., 2014). These events bridge the gap between measured biomarkers and adverse effects at the population and community level (Forbes and Galic, 2016). Detecting molecular initiating events at very low exposure levels is important for risk monitoring and informing risk reduction efforts.

Further work involving *H. azteca* could focus on testing different life stages of the amphipods. Larger *H. azteca* are used here in order to obtain the needed concentration of mRNA for gene expression analysis. Determining how many juvenile amphipods in the sensitive life stage (7 - 14 days old) are needed to conduct these analyses would allow for a better assessment of toxic effects at the most vulnerable point in the amphipod's life. Longer running exposure studies using the sub-lethal concentrations used here should be developed to assess any chronic effects the metals may have across generations. This would help form the links in AOPs between exposure and effect at higher levels of biological organization.

Finally, ambient conditions that could affect gene expression can be explored. Changes in temperature, dissolved oxygen level, salinity, or pH could all potentially affect the organism's response independent of the effect of the stressor of interest, as demonstrated by the range of ambient conditions to which heat shock proteins can respond (Feder and Hofmann, 1999). Many HSPs are involved in physiological

adaptations to extremes in these abiotic parameters, such as diapause and anhydrobiosis (MacRae, 2010; Gusev et al., 2011). All the testing performed and suggested here for cadmium and copper should also be tested again using sediment exposures. As these metals bind readily to sediment and organic material, the exposure of the epibenthic *H. azteca* could be much higher when metals are bound to this material rather than dissolved aqueously.

The work performed here provides several genes that can be used as biomarkers for cadmium and copper at very low concentrations. These biomarkers provide exposure lines of evidence for ERA. However, as previously mentioned, these biomarkers do not provide information on the higher-level effects these metals may have at these exposure concentrations. This is again where AOPs come into play to connect exposure to ecological effects. AOPs can be formed through direct comparison of exposure and effects across biological levels or through modeling of the potential effects that suborganismal changes have on higher levels (Soetaert et al., 2007; Connon et al., 2011; Ananthasubramaniam et al., 2015). However, AOPs always begin with molecular initiating events like gene expression. Changes in gene expression are defined as molecular initiating events within the framework of the AOP (Allen et al., 2014). These molecular initiating events begin the process leading to ecological adverse effects. Detecting molecular initiating events at low exposure levels is important for risk monitoring and informing risk reduction efforts at the population and community levels (Forbes and Galic, 2016). The gene expression biomarkers developed here at low metal concentrations give us the ability to do just that. Examination of genes for potential as biomarkers of exposure and connecting them to effects can be repeated for any chemical,

physical, or biological stressor. To demonstrate the versatility of gene expression biomarkers in *H. azteca*, in Chapter 4 I will determine appropriate biomarker genes for exposure to legacy and current-use insecticides.

Chapter 4: Changes in gene expression as biomarkers of sub-lethal exposure to the insecticides DDT, imidacloprid, and permethrin in *Hyalella azteca*

Abstract

Environmental risk assessment of biocidal chemicals is critical for protection of non-target organisms. Insecticides intended to treat human or crop pests can reach vulnerable aquatic systems through improper handling, improper application, and leaching and runoff. Insecticides impacting aquatic systems include recalcitrant legacy compounds like the organochlorine DDT and more current-use chemicals like pyrethroids and neonicotinoids. Insecticides can partition into sediment and open water, often with a strong preference for one or the other. Organisms that live at the interface of these two compartments are uniquely at risk of exposure to contaminants. The benthic amphipod *Hyalella azteca* was exposed to sublethal concentrations of the insecticides DDT, the pyrethroid permethrin, and the neonicotinoid imidacloprid to assess molecular responses of this potentially vulnerable non-target organism. Expression of detoxification genes including cytochrome P450s and glutathione-*S*-transferase was measured via RT-qPCR after 24 hours of exposure. Expression of the ABC transporter *Mrp4* and the stress-response heat shock protein *Hsp90* increased significantly after DDT exposure. Decreased expression of immune-related genes occurred after permethrin exposure. Imidacloprid exposure did not cause significant changes in expression of any genes, though there was a trend of decreased expression of certain cytochrome P450 genes. Monitoring of the genes that change in expression is a valuable tool for risk assessment

of environmental exposure to these insecticides. Using this narrower set of genes can help streamline further risk studies on sediment exposure to insecticides and mixtures of insecticides. Expression monitoring can also provide a means to connect sublethal molecular responses to ecologically relevant life history changes though Adverse Outcome Pathways, a critical current need in ecotoxicology.

Introduction

Agricultural and residential non-point sources of pollution contribute the vast majority of the insecticides contaminating freshwater systems (Muller et al., 2002; Neumann et al., 2002; Schulz, 2004). As insecticides from these areas enter nearby bodies of water via drift, runoff, or leaching, sensitive non-target arthropod communities can be affected (Fernandez-Alba et al., 2002; Lizotte et al., 2012; Rico and Van den Brink, 2015). The complexity of the effects insecticides may have on these communities is immense. Sensitivities vary by the intrinsic sensitivity of a given species or individual, species-specific life history variables that determine recovery potential, and the toxicokinetic and toxicodynamic variables of the insecticides themselves (Van Straalen, 1994; Escher and Hermens, 2002; Nyman et al., 2014). These non-target communities are also likely exposed simultaneously to a cocktail of discontinued recalcitrant and current-use insecticides with varying modes of action and toxicities (Struger and Fletcher, 2007; Sparks and Nauen, 2015). This makes risk assessment of insecticides critical in order to protect and sustain vulnerable non-target populations and communities and prioritize risk reduction efforts.

The risk of exposure of aquatic non-target organisms to insecticides varies by the insecticide's use and the chemical properties of the insecticide. Insecticides are applied in agriculture in many ways that make runoff and leaching likely, including foliar application, soil injection and drenching, and seed treatments (NASDA, 2014). Spray can drift from application directly into nearby water or can flow in surface runoff or leach from soil, carrying 1 to 10% or more of the amount of insecticide applied to the field (Wauchope, 1978; Schulz, 2004). Some farming operations contribute up to 84% of the total insecticide contamination present in nearby streams (Neumann et al., 2002). Commercial and residential use of insecticides also results in highly contaminated runoff toxic enough to cause adverse effects in aquatic non-target organisms (Weston et al., 2005; Gan et al., 2012). Effects of insecticides in both direct receiving and downstream waters include changes in community composition, direct mortality, and stress-responsive molecular changes (Fulton et al., 1999; Thiere and Schulz, 2004; Pedersen et al., 2006).

Two important, but certainly not the only, chemical properties contributing to insecticide exposure risk in aquatic systems are the octanol-water partition coefficient (K_{ow}, often reported as the log K_{ow}) and the chemical half-life. The log K_{ow} describes the tendency of a chemical to associate with either organic phases (higher values) or with water (lower values). The chemical half-life is the amount of time needed for half of a given amount of chemical to degrade. Half-lives vary by many factors like exposure to light, water, and oxygen. Together these two values play an important role in determining whether or not an organism is at risk of being exposed to a potential toxin in the water. The varied historic and current use and chemical characteristics of the insecticides DDT,

permethrin, and imidacloprid are discussed in more detail to illustrate how these factors affect exposure risk.

DDT was used on a large-scale in agriculture from about 1943 to 1973 in the United States (Smith, 1991; Turusov et al., 2002). DDT has historic and current uses for indoor residential spraying for control of disease-vectoring insects in many developing countries (Curtis and Mnzava, 2000; van den Berg, 2009). DDT and its metabolites are still found in high concentrations and have distributed globally (Bogdal et al., 2013; Octaviani et al., 2015; Booij et al., 2016). Permethrin has current applications in agricultural, commercial/structural, and residential applications, giving it multiple potential non-point sources into aquatic ecosystems (Schulz and Stehle, 2008; Spurlock and Lee, 2008). Imidacloprid and other neonicotinoids were nearly one-third of the monetary value of the 2010 world insecticide market (Casida and Durkin, 2013). In 2010 imidacloprid was the best selling insecticide in the world (Pollack, 2011). Imidacloprid has agricultural, residential, commercial, and veterinary applications (Simon-Delso et al., 2015; van der Sluijs et al., 2015). Imidacloprid is applied in many ways that make it susceptible to runoff including foliar spraying, seed dressing, and soil drenching (Simon-Delso et al., 2015).

The chemical characteristics of these three insecticides give them unique patterns of environmental behavior and exposure potential. Many organochlorines have extremely high log K_{ow} values (e.g., DDT log K_{ow} = 6.91), which means they partition strongly into sediment and organic matter versus water (ATSDR, 2002). Pyrethrins and pyrethroids have varying log K_{ow} values giving them variable behavior. Permethrin has a high log

K_{ow} at 6.5 (ATSDR, 2003). Some neonicotinoids have lower log K_{ow} values, such as imidacloprid at 0.57 (Bonmatin et al., 2015). Low values indicate these insecticides partition into water more readily than DDT or permethrin. The risk of exposure to these insecticides then is generally highest in sediment-dwelling organisms for organochlorines and pyrethroids and highest in open-water organisms for neonicotinoids. Organisms that live at the water-sediment interface therefore are in a uniquely vulnerable position with potential exposures from both water and sediment.

Some insecticides have very long half-lives, such as DDT's soil half-life of 2 to 15 years and aquatic half-life of around 150 years, which contributes to its classification as a persistent organic pollutant (POP) (ATSDR, 2002). Permethrin is relatively recalcitrant to photolytic degradation but is susceptible to aerobic degradation, with a half-life around 2.5 days in aquatic sediment and 39.5 days in soil (Schimmel et al., 1983; Imgrund, 2003). Imidacloprid is highly susceptible to photolysis, with a half-life of 5-43 minutes in organic-poor aquatic systems, but can persist up to 6 months when protected from sunlight (Wamhoff and Schneider, 1999; Liu et al., 2006; Simon-Delso et al., 2015). Insecticides with longer half-lives can build up to large environmental concentrations, increasing exposure risk of non-target organisms.

The patterns of use and the characteristics of DDT, permethrin, and imidacloprid create a high exposure risk of non-target organisms to these insecticides. DDT is no longer used in the United States and other developed countries. However, its tendency to adsorb to sediment and organic matter and extremely long half-life contribute to its continued exposure risk. Permethrin is highly used and partitions to sediment and organic

matter. Sediment-dwelling organisms therefore may be particularly at risk of exposure to permethrin. Imidacloprid, despite its very short half-life in most situations, is produced and used so abundantly that an exposure risk exists from the sheer volume of input of imidacloprid into the environment. The large exposure risks of DDT, permethrin, and imidacloprid create higher risks for adverse effects of these insecticides on non-target organisms.

The effects of the insecticide classes containing DDT, permethrin, and imidacloprid are neurological in nature. Organochlorine insecticides are divided into two groups: the chlorinated alicyclic (e.g., dieldrin, endrin) compounds and DDT-like compounds (DDT and its analogs). Chlorinated alicyclics act as antagonists of GABA-gated chloride channels in postsynaptic neurons, preventing the uptake of chloride ions. No inhibitory postsynaptic potential is produced, allowing action potentials to continue unabated, causing paralysis (Casida and Durkin, 2013). DDT-like compounds also cause paralysis, but via a different mode of action. DDT and its analogs modulate the activity of voltage-gated sodium channels. At first exposure neurons fire spontaneously and rapidly due to increased release of excitatory neurotransmitters, causing tremors. After prolonged exposure the neurons lack the sodium electrochemical gradient needed to continue firing, creating a state of excitatory paralysis (Davies et al., 2007).

Similar to the organochlorines, pyrethroid and pyrethrin insecticides act on voltage-gated sodium channels. Type I pyrethroids and pyrethrins (e.g., permethrin, allethrin) bind to inactive channels and open them to a slowly activating sodium gradient shift, creating repetitive neural firing. The neuron is also shifted into a state of

hyperexcitability due to a change in neural membrane potential, producing an effect called knockdown. Repetitive firing plus creation of the knockdown state leads to incoordination and eventual paralysis (Gammon et al., 1981; Soderlund et al., 2002; Davies et al., 2007). Type II pyrethroids (e.g., deltamethrin) bind sodium channels into an irreversible open state, allowing free flow of sodium ions. The irreversible open state of the channels suppresses all further action potentials (Gammon et al., 1981; Soderlund et al., 2002).

Neonicotinoid insecticides (e.g., imidacloprid, thiacloprid, clothianidin) are agonists of the nicotinic acetylcholine receptor. The neonicotinoid binds to open receptors, holding them in the open conformation. This allows free exchange of sodium and potassium ions, creating an action potential. The overexcitement of neurons by neonicotinoids leads to paralysis (Simon-Delso et al., 2015). Neonicotinoids are arthropod-selective in their toxicity compared to many other insecticides, owing to differences in the structure of the nicotinic acetylcholine receptors between arthropods and other taxa (Nauen et al., 1999; Matsuda et al., 2001; Tomizawa and Casida, 2003; Tomizawa and Casida, 2005).

In this chapter I develop gene expression-based biomarkers in the amphipod *Hyalella azteca* for exposure to DDT, permethrin, and imidacloprid. *H. azteca* is a member of the potentially sensitive crustacean order Amphipoda, living at the sediment-water interface within water bodies often receiving runoff that contains insecticides from agricultural and residential areas (Rico and Van den Brink, 2015). These insecticides pose a great risk to non-target organisms. Sensitive detection methods are needed to

assess exposure risks at contamination levels much lower than those needed to induce visible changes in the organism or the community. Chemical and biochemical methods exist that detect these insecticides at extremely low concentrations (e.g., Mauriz et al., 2007; de Souza Pinheiro and de Andrade, 2009; Feo et al., 2010). However, these methods often require the use of additional hazardous chemicals during processing, expensive input materials, and highly specialized skill sets. Chemical methods also give information with no biological meaning. A molecular method of detection using live organisms provides a compromise between sensitive detection and biological meaning.

Expression of genes for general stress response and detoxification genes will be measured after insecticide exposure. The goal of this study is to identify genes that change in expression in response to these insecticides and that can serve as biomarkers of exposure to these insecticides. Increased expression is predicted in genes for detoxifying proteins like glutathione-*S*-transferases, cytochrome P450s, and ABC transporters. The initial development of biomarkers for individual insecticides lays the foundation for work with insecticide mixtures that more accurately reflect what these amphipods and other non-target arthropods experience in natural settings. Gene expression-based biomarkers will sensitively detect low concentrations of contaminants while providing biological molecular effects that can be connected to higher-level adverse effects on populations and communities.

Materials and Methods

H. azteca were exposed to sub-lethal concentrations of DDT, permethrin, and imidacloprid for 24 hours. RNA was extracted from all samples and reverse transcribed.

Normalization factors (NFs) were developed for each treatment by measuring the expression of candidate reference genes in each sample using RT-qPCR. The expression of the most stable candidate reference genes formed the NFs for each treatment. The expression of stress-response genes of interest was then measured in every sample. This expression was normalized using the NFs. The normalized expression of each gene of interest was compared between the control and each treatment to produce fold changes in expression.

Pesticide exposure

The insecticides 4,4'-DDT, imidacloprid, and permethrin were chosen for exposure studies to represent the insecticide classes of DDT-like organochlorines, neonicotinoids, and Type I pyrethroids, respectively. Given no available 24-hour LC50 values for these insecticides, 4-day LC50 values were used as the closest available data. 4-day LC50 values for DDT, imidacloprid, and permethrin are 0.17 μg/L, 65.43 μg/L, and 0.021 μg/L, respectively (Lotufo et al., 2000; Stoughton et al., 2008; Andersen et al., 2006; Ding et al., 2012). Exposures were conducted at 20% of the published LC50 values. Exposure concentrations for DDT, imidacloprid, and permethrin were 0.034 μg/L, 13 μg/L, and 0.004 μg/L, respectively. All chemicals were purchased from Sigma-Aldrich as analytical standard grade powders of 100% active ingredient. Stocks of DDT and permethrin were made in 100% ethanol as 1 mg/L and 250 μg/L, respectively. Stock of imidacloprid was made in DI water as 0.4 g/L. The DDT exposure resulted in a final ethanol concentration in the experimental units of 0.0034% ethanol. This is 0.32% of the *H. azteca* LC50 value for ethanol (Bowman et al., 1981). Sufficient 100% ethanol was added to all experimental units including the controls to create a final concentration of

ethanol in each of 0.0034%, similar to the practice in Asselman et al., 2012 and Soetaert et al., 2007a. Low concentrations of all chemicals were used to obtain sublethal effects. The goal was to test gene expression at concentrations at which measurement of mortality or other life history parameters would not differentiate between a control and an exposure. Changes detected in stress response gene expression at these concentrations would serve as more sensitive biomarkers of exposure than life history parameters.

Experimental units consisted of round glass specimen dishes (10 cm diameter, 5.5 cm height, approximately 350 mL maximum volume) with 200 mL of reconstituted fresh water (pH 7.64, conductivity 203 µS/cM) made in accordance to U.S. EPA (2000). Reconstituted water consisted of: 10 L deionized water, 0.5 g $CaSO_4$ (Colorado Scientific), 0.5 g $CaCl_2$ (J. T. Baker), 0.3 g $MgSO_4$ (Sigma-Aldrich), 0.96 g $NaHCO_3$ (Sigma-Aldrich), and 0.04 g KCl (Sigma-Aldrich). Each dish had an air bubbler and a 4" x 4" piece of plastic mesh substrate. Reconstituted water was aerated for 24 hours. DDT, permethrin, imidacloprid, and control exposures had four replicates. Due to a machine error during data acquisition, only three replicates were ultimately used for all treatments. Twenty adult *H. azteca* caught on a number 60 standard sieve were added to each dish (320 total *H. azteca* used) with approximately 2 mg of crushed Tetramin tropical fish food and acclimated for 24 hours.

The appropriate volumes of ethanol and insecticide were added after the 24-hour acclimation period. For the DDT exposure, I removed 6.8 µL of water and added 6.8 µL of 1 mg/L DDT stock. For the imidacloprid exposure, I removed 13.2 µL of water. I then added 6.5 µL of the 0.4 g/L imidacloprid stock and 6.8 µL 100% ethanol. For permethrin

I removed 6.8 µL of water and added 3.2 µL of 250 µg/L permethrin stock and 3.6 µL 100% ethanol. I removed 6.8 µL of water and added 6.8 µL 100% ethanol from the no-insecticide control. All exposures started at 08:00 and ended at 08:00 the following day, with a 16:8 light:dark cycle and light starting at 07:00.

Ten *H. azteca* were collected from each replicate dish by straining through a number 60 standard sieve after 24 hours of exposure. The ten *H. azteca* were placed into microcentrifuge tubes containing 200 µL of RNAlater (Sigma Aldrich) and flash frozen in liquid nitrogen and stored at -80°C until further processing. The total number of surviving *H. azteca* in each replicate was recorded. The proportion of amphipods surviving from each replicate was arcsine square root transformed and analyzed using the non-parametric Kruskal-Wallis analysis of variance in SAS statistical software (v.9.4). RNA extraction and cDNA synthesis was performed as described in Chapter 1. RNA samples were submitted the University of Maryland Sequencing Core for bioanalysis.

Normalization factor analysis

I screened the same candidate reference genes that were tested in Chapter 2 (*αTub*, *Mmp*, *Tbp*, *Ubi*, and *Ubc*). RT-qPCR was performed on a Roche LightCycler 480 Real-Time PCR machine in the Genomics Core at the University of Maryland. Each well of the LightCycler 480 Multiwell 96-well plates (Roche) consisted of 10 µL of 2x LightCycler 480 SYBR Green I Master mix (Roche), 0.5 µL of 10 µM of each primer, 5 µL of 1/10-diluted cDNA, and 4 µL of PCR grade water. Every biological replicate was tested in duplicate. Appropriate no-cDNA –RT controls were tested for all treatments in duplicate. No-template water controls were run with four replicates. The optimized RT-

qPCR protocol used was as follows: 1) a pre-incubation period of 95 °C for 10 min, 2) an amplification period consisting of 30 cycles of 95 °C for 15 s, 61 °C for 30 s, and 72 °C for 30 s, with a single fluorescence measurement performed at the end of each cycle, 3) a single melting curve analysis of 95 °C for 0.05 s followed by an increase of 0.5 °C/cycle from 65 °C to 97 °C, with 5 fluorescence readings/°C, and 4) a cooling period of 40 °C for 1 min. Cycle threshold (Ct) values were determined using fit point analysis on manufacturer provided LightCycler 480 Software release 1.5.0 SP3 (version 1.5.0.39).

The mean of the Ct values of all technical replicates served as the Ct value of each biological replicate. Delta (Ct) values (ΔCt) were calculated using equation (1) for each replicate for each gene, where ε is the efficiency of the primer for the gene being analyzed, $Ct_{gene,min}$ is the minimum Ct value of the gene being analyzed across all replicates, and Ct_x is the Ct value of the biological replicate being analyzed. This served to calibrate all values against the lowest expression value. Results were analyzed for stability of reference gene transcripts using BestKeeper (Pfaffl et al., 2004), GeNorm (Vandesompele et al., 2002), and NormFinder (Andersen et al., 2004). While Ct values were used in BestKeeper, ΔCt values were used in GeNorm and NormFinder. The most stable genes were then used to calculate an NF for each treatment-control pair, i.e. one NF each for the insecticide treatments and one NF for each control-treatment combination. The NFs were calculated as the geometric mean of the ΔCt values of all biological replicates.

$$(1)\ \Delta Ct_x = \varepsilon^{Ct_{gene,min} - Ct_x}$$

The standard error of the mean of the NF (NFsem) was calculated incorporating proper error propagation using equation (2), where $\Delta Ctsem_x$ is the standard error of the mean of the ΔCt value of the first gene in the NF, ΔCt_x is the ΔCt value of the first gene in the NF, $\Delta Ctsem_i$ is the standard error of the mean of the last gene in the NF, and ΔCt_i is the ΔCt value of the last gene in the NF.

$$(2)\ NFsem = NF * \sqrt{\left(\frac{\Delta Ctsem_x}{3 * \Delta Ct_x}\right)^2 + \cdots + \left(\frac{\Delta Ctsem_i}{3 * \Delta Ct_i}\right)_i^2}$$

Gene of interest primer design and testing

I designed additional RT-qPCR primers for cytochrome P450 genes using online design tools from IDT DNA and GenScript (Table 4.1; Appendix D). Cytochrome P450 genes are of interest because of their responses to chemical stressors and their degradation of target chemicals. I tested cytochrome P450 primers using two samples of cDNA from *H. azteca* from the lab population described in Chapter 1. Methods for sample collection, RNA extraction and cDNA synthesis were the same as those found in Chapter 1. Each PCR reaction included: 10 µL of 2x PCR Master Mix (SydLabs), 1 µL cDNA, 1 µL of 100 µM forward primer, 1 µL of 100 µM reverse primer, and 7 µL PCR grade water. PCR was run with the following cycles: 1) 1 cycle of 95ºC for 2 minutes; 2) 25 cycles of 95ºC for 15 seconds, 57ºC for 15 seconds, and 72ºC for 30 seconds; 3) 1 cycle of 72ºC for 7 minutes; 4) held 4ºC indefinitely until PCR products were stored at -20ºC. All reactions were paired with a negative control with 1 µL of water added instead of cDNA. PCR products were run on a 4% agarose gel at 110 V for 40 minutes along with a 10 bp DNA ladder (Life Technologies) and visualized using ethidium bromide and

a UV camera system. Products were sent to GENEWIZ, Inc. for Sanger sequencing using the forward primer of each primer set as the sequencing primer. Experimental PCR product sequences were then aligned to their corresponding expected amplicon sequence (Appendix E) and analyzed for percent similarity using EMBOSS Water (Rice et al., 2000). Note that the expression of genes coding for the actual protein target of the insecticides was not monitored, as there is previous evidence that this may not change in *H. azteca* (Weston et al., 2013).

Table 4.1: Cytochrome P450 GOI transcripts identified in the *H. azteca* transcriptome and their corresponding primers

Transcript name	Abbreviation	Primer set	Forward primer sequence (5' - 3')	Reverse primer sequence (5' - 3')	Amplicon length (nt)
Cytochrome P450 2j2-like	*Cyp2j2*	*Cyp2j2-2*	CATGGGAGGCC TGACTTCTT	CCCGAGGAAAG TCTTACCCA	150
		Cyp2j2-3	TCCCTAGACCC AGACAACCC	CAGCTGCTGAG CATCGAAAT	105
Cytochrome P450 2c9-like isoform x1	*Cyp2c9*	*Cyp2c9-2*	ATGGAGAAGCT CGGCGTAAC	CGATGTAATCG CGAGGATTG	121
		Cyp2c9-3	TTTGGAGTTCC AGCAAGTGG	GTTACGCCGAG CTTCTCCAT	126
Cytochrome P450 2L1-like	*Cyp2L1*	*Cyp2L1*	CACTGCGCTAC CCAAGACAT	GTCCAAGAAGT GCGGAGGAT	141
		Cyp2L1-2	TGGATTTGGAC ATCGGTTTG	GCATTTCCAGC CAATTCCTT	110
Cytochrome P450 3A13-like	*Cyp3A13*	*Cyp3A13-1*	CAGTGCATCCT CTTCATCCTG	GCTCTCGTTAC CCTCAATTCTG	141
		Cyp3A13-2	GTCGAGGCGGC TATATTCGT	GCGTCAGGTCA CGGTACAAC	135
		Cyp3A13-3	GCTACGACACC ACAGCCAAC	CCGCTCTCGTT ACCCTCAAT	118
Cytochrome P450 similar to 6BS1	*Cyp6BS1*	*Cyp6BS1-1*	CGATGTTGCGT CCAGTAAATG	GACGAACACAA AAGCCAATCC	127
		Cyp6BS1-2	TGCAAGTACCT CGACGCTGT	GCACGCTGTAG TCCTTGGTG	94
		Cyp6BS-3	GCCTTCGTGTTT CGATGTTG	TCGTGTCTCTG GCTGTGGAT	113
Cytochrome P450 2u1-like	*Cyp2u1*	*Cyp2u1-1*	CAGACAAGCCT CGATTAGCG	AGCAATGTGTC CTGTGGTATG	142
		Cyp2u1-2	ACAACCCACGC GACATCATA	ACAGCTGCTGA GCATCGAAA	93
		Cyp2u1-3	CTGGACGTGTT GAGGCTGAG	CACGCCATCAG CAATGTGT	113
Cytochrome P450 2j6-like	*Cyp2j6*	*Cyp2j6-1*	ATCGTCTGCCA TCACTCAAC	GGTATTTCATA GCCCTCGTAGG	142
		Cyp2j6-2	GCGGATCGAAT CTGACTTCC	CGAACGGTCAC TTGTCGAAG	126
		Cyp2j6-3	GGCCATCAACC CTGAGATTC	TCCAGTAACAC GGCGTCAAC	123

GOI expression analysis

All experimental replicates were plated with three technical replicates. The RT-qPCR protocol was the same as used above. Threshold cycle (Ct) values were calculated using fit point analysis on LightCycler 480 Software release 1.5.0 SP3 (version 1.5.0.39).

The means of Ct values of all technical replicates were calculated as the Ct values for each biological replicate. Delta (Ct) values (ΔCt) were calculated using equation (3) for each replicate for each gene, where ε is the efficiency of the primer for the gene being analyzed, $Ct_{gene,min}$ is the minimum Ct value of the gene being analyzed across all replicates, and Ct_x is the Ct value of the biological replicate being analyzed. This served to calibrate all values against the lowest expression value.

$$(3)\ \Delta Ct = \varepsilon^{Ct_{gene,min} - Ct_x}$$

Normalized expression levels (NELs) were calculated by dividing each ΔCt value by the corresponding normalization factor (NF) from Chapter 2. The standard error of the mean (sem) of all NELs was calculated incorporating proper error propagation using equation (4), where $NELsem_x$ is the standard error of the mean of the NEL for the gene being analyzed, NEL_x is the NEL of the same gene, NFsem is the standard error of the mean of the NF, $\Delta Ctsem_x$ is the standard error of the mean of the ΔCt value of the same gene, and ΔCt_x is the ΔCt value of the same gene.

$$(4)\ NELsem_x = NEL_x * \sqrt{\left(\frac{NFsem}{NF}\right)^2 + \left(\frac{\Delta Ctsem_x}{\Delta Ct_x}\right)^2}$$

The relative fold change (RFC) in expression of each gene was then calculated relative to the control using equation (5). The standard error of the mean of each RFC (RFCsem) was calculated incorporating proper error propagation using equation (6).

$$(5)\ RFC_{gene} = \frac{NEL_{gene,treatment}}{NEL_{gene,control}}$$

$$(6)\ RFCsem_x = \sqrt{NELsem^2_{x,treatment} + NELsem^2_{x,control}}$$

Relative expression levels (RELs) were calculated for each gene as the NEL of that gene's expression in each individual replicate divided by the geometric mean of the expression of that gene in the control. The geometric mean of the RELs of all replicates of a treatment for a single gene is equal to the RFC of that gene for that treatment. RELs allow for statistical analysis of changes in expression values using all replicates. RELs were analyzed with a Wilcoxon rank-sum test in SAS using the proc npar1way procedure.

Results

H. azteca survival

Mean survival (% ± SD) in the control, DDT, imidacloprid, and permethrin exposures was 97.5 ± 2.9%, 98.8 ± 2.5%, 97.5 ± 5%, and 93.8 ± 4.8%, respectively. No significant effect of any insecticide on survival was found (Kruskal-Wallis analysis of variance, X^2=3.15, df=3, *P*=0.48).

Normalization factor analyses

Extracted samples of RNA were of high concentration and good quality according to bioanalysis (Table 4.2). Descriptive statistics of RT-qPCR results are of the threshold (Ct) values (Tables 4.3 and 4.4). The Ct value is the PCR cycle at which the signal from the sample became higher than the background noise present from sources such as 96-well plate autofluorescence during data acquisition. The standard deviation of the Ct value represents the stability of a gene's expression across samples from the same

treatment. Lower standard deviations show a gene that is similarly expressed across all samples.

Table 4.2: Bioanalysis results of total RNA concentration and quality

Treatment	Replicate	RNA concentration (ng/µL)	RNA integrity number (RIN)
Control	1	1,048	5.9
	2	866	5.7
	3	735	6
	4	820	5.7
DDT	1	704	6.8
	2	1,007	5.9
	3	791	5.5
	4	804	5.3
Imidacloprid	1	952	6.4
	2	944	5.6
	3	622	5.4
	4	963	5.3
Permethrin	1	914	5.8
	2	738	5.6
	3	874	5.4
	4	967	5.1

Table 4.3: Mean Ct values of reference genes *αTub*, *Mmp*, and *Tbp* by treatment. Imid: imidacloprid, Per: permethrin, SD: standard deviation, SE: standard error of the mean

	Gene											
	αTub				*Mmp*				*Tbp*			
Treatment	Control	DDT	Imid	Per	Control	DDT	Imid	Per	Control	DDT	Imid	Per
n	3	3	3	3	3	3	3	3	3	3	3	3
Geometric mean	19.03	19.45	18.78	20.07	23.44	23.26	23.3	24.32	25.19	25.47	26.09	26.33
Arithmetic mean	19.04	19.47	18.79	20.09	23.45	23.28	23.3	24.34	25.2	25.48	26.1	26.35
Min	18.6	18.47	18.58	19.14	22.88	22.05	22.91	23.49	24.71	25.01	25.55	25.64
Max	19.88	20.25	18.97	21.42	24.09	24.49	23.9	25.65	26.1	25.74	26.64	27.49
SD	0.73	0.91	0.2	1.18	0.61	1.22	0.51	1.15	0.78	0.41	0.54	1
SE	0.42	0.53	0.11	0.68	0.35	0.7	0.29	0.66	0.45	0.24	0.31	0.57

Table 4.4: Mean Ct values of reference primers *Ubi* and *Ubc* by treatment. Imid: imidacloprid, Per: permethrin, SD: standard deviation, SE: standard error of the mean

	Gene							
	Ubi				*Ubc*			
Treatment	Control	DDT	Imid	Per	Control	DDT	Imid	Per
n	3	3	3	3	3	3	3	3
Geometric mean	21.99	22.09	22.46	23.07	17.94	18.17	17.87	18.91
Arithmetic mean	22	22.11	22.48	23.09	17.94	18.19	17.87	18.95
Min	21.31	21.07	21.58	22.12	17.68	17.46	17.47	17.66
Max	23.05	23.49	23.5	24.57	18.43	19.37	18.31	20.29
SD	0.925	1.24	0.97	1.3	0.42	1.04	0.42	1.32
SE	0.53	0.72	0.56	0.75	0.24	0.6	0.24	0.76

For the DDT exposure, BestKeeper repeated pair-wise correlation analysis revealed *Ubi* and *Ubc* to have the highest r^2 values when correlated to the BestKeeper index (Table 4.5). GeNorm and NormFinder both found *αTub*, *Ubc*, and *Mmp* to be the most stably expressed genes for the DDT exposure (Table 4.8). For the imidacloprid exposure BestKeeper found *Ubi* and *Ubc* to have the highest r^2 values when correlated to the BestKeeper index (Table 4.6). *Ubc, Mmp*, and *αTub* were most stable under GeNorm and NormFinder analyses (Table 4.8). For the permethrin exposure BestKeeper found *αTub*, *Ubi*, *Ubc*, and *Mmp* to have higher r^2 values when correlated to the BestKeeper index (Table 4.7). *Mmp* and *αTub* were most stable under GeNorm and NormFinder analyses (Table 4.8). All genes ranked by their stability and suitability as reference genes are found in Table 4.9.

The genes chosen to form the NF for each exposure were selected based on the consensus on their stability between all three programs. *αTub*, *Ubc*, and *Mmp* formed the NF for the DDT exposure. *αTub* and *Ubc* formed the NF for the imidacloprid exposure. *αTub* and *Mmp* formed the NF for the permethrin exposure. NFs were calculated for the

exposure and the control separately for each insecticide because the genes comprising the NF differed between the treatments. NFs for each treatment and control, respectively, were 0.6 and 0.65 for DDT, 0.85 and 0.78 for imidacloprid, and 0.46 and 0.77 for permethrin.

Table 4.5: Coefficients of correlation (r values) and *P*-values from the BestKeeper repeated pair-wise correlation analysis for the DDT exposure

Vs.	*αTub*	*Mmp*	*Tbp*	*Ubi*	*Ubc*
Mmp	0.862	-	-	-	-
P-value	0.027	-	-	-	-
Tbp	0.114	0.014	-	-	-
P-value	0.829	0.978	-	-	-
Ubi	0.909	0.941	-0.056	-	-
P-value	0.012	0.005	0.918	-	-
Ubc	0.837	0.877	0.132	0.950	-
P-value	0.038	0.022	0.801	0.004	-
BestKeeper index	0.940	0.942	0.193	0.964	0.957
Coefficient of determination (r^2) vs. BestKeeper	0.88	0.89	0.04	0.93	0.92
P-value	0.005	0.005	0.716	0.002	0.003

Table 4.6: Coefficients of correlation (r values) and *P*-values from the BestKeeper repeated pair-wise correlation analysis for the imidacloprid exposure

Vs.	*αTub*	*Mmp*	*Tbp*	*Ubi*	*Ubc*
Mmp	0.778	-	-	-	-
P-value	0.068	-	-	-	-
Tbp	-0.618	-0.681	-	-	-
P-value	0.192	0.136	-	-	-
Ubi	0.690	0.789	-0.402	-	-
P-value	0.129	0.062	0.429	-	-
Ubc	0.761	0.955	-0.508	0.752	-
P-value	0.079	0.003	0.303	0.085	-
BestKeeper index	0.794	0.855	-0.321	0.925	0.894
Coefficient of determination (r^2) vs. BestKeeper	0.63	0.73	0.10	0.86	0.80
P-value	0.059	0.030	0.534	0.008	0.016

Table 4.7: Coefficients of correlation (r-values) and *P*-values from the BestKeeper repeated pair-wise correlation analysis for the permethrin exposure

Vs.	*αTub*	*Mmp*	*Tbp*	*Ubi*	*Ubc*
Mmp	0.982	-	-	-	-
P-value	0.001	-	-	-	-
Tbp	0.657	0.623	-	-	-
P-value	0.157	0.187	-	-	-
Ubi	0.994	0.988	0.583	-	-
P-value	0.001	0.001	0.223	-	-
Ubc	0.963	0.951	0.725	0.938	-
P-value	0.002	0.004	0.102	0.006	-
BestKeeper index	0.988	0.977	0.751	0.970	0.982
Coefficient of determination (r^2) vs. BestKeeper	0.98	0.96	0.56	0.94	0.96
P-value	0.001	0.001	0.086	0.001	0.001

Table 4.8: Expression stability values from GeNorm and NormFinder analyses

	GeNorm			NormFinder		
Gene	DDT	Imidacloprid	Permethrin	DDT	Imidacloprid	Permethrin
αTub	0.422	0.464	0.289	0.053	0.047	0.028
Mmp	0.422	0.426	0.315	0.031	0.056	0.029
Tbp	0.760	0.890	0.655	0.192	0.225	0.167
Ubi	0.445	0.612	0.348	0.069	0.114	0.048
Ubc	0.397	0.409	0.334	0.053	0.038	0.066

Table 4.9: Ranking of all reference genes by expression stability

	DDT			Imidacloprid			Permethrin		
	BestKeeper	GeNorm	NormFinder	BestKeeper	GeNorm	NormFinder	BestKeeper	GeNorm	NormFinder
Most stable gene	*Ubc*	*Ubc/Mmp*	*Mmp*	*Ubc*	*Ubc/Mmp*	*Ubc*	*Mmp*	*αTub/Mmp*	*αTub*
	Mmp	*αTub*	*Ubc = αTub*	*αTub*	*αTub*	*αTub*	*Tbp*	*Ubc*	*Mmp*
	αTub	*Ubi*	*Ubi*	*Mmp*	*Ubi*	*Mmp*	*αTub*	*Ubi*	*Ubi*
Least stable gene	*Ubi*	*Tbp*	*Tbp*	*Ubi*	*Tbp*	*Ubi*	*Ubc*	*Tbp*	*Ubc*
	Tbp			*Tbp*		*Tbp*	*Ubi*		*Tbp*

Gene of interest primer design

Of all the cytochrome P450 primers tested, *Cyp2j2-2*, *Cyp2j2-3*, *Cyp2c9-2*, *Cyp2L1*, *Cyp3A13*, *Cyp3A13-2*, *Cyp3A13-3*, *Cyp6BS1*, *Cyp6BS1-2*, *Cyp2u1-2*, and *Cyp2j6-2* produced visible bands after electrophoresis of PCR products (Figure 4.1 - bright band in lanes 1 and 14 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Cyp2j2-2*negative control, 3) *Cyp2j2-2* sample 1, 4) *Cyp2j2*-2 sample 2, 5) *Cyp2j2-3* negative control, 6) *Cyp2j2-3* sample 1, 7) *Cyp2j2-3* sample 2, 8) *Cyp2c9-2* negative control, 9) *Cyp2c9-2* sample 1, 10) *Cyp2c9-2* sample 2, 11) *Cyp2c9-3* negative control, 12) *Cyp2c9-3* sample 1, 13) *Cyp2c9-3* samples 2, 14) 10 bp DNA ladder, 15) *Cyp2L1* negative control, 16) *Cyp2L1* sample 1, 17) *Cyp2L1* sample 2, 18) *Cyp2L1-2* negative control, 19) *Cyp2L1-2* sample 1, 20) *Cyp2L1-2* sample 2) (Figure 4.2 - bright band in lanes 1 and 11 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Cyp3A13* negative control, 3) *Cyp3A13* sample 1, 4) *Cyp3A13* sample 2, 5) *Cyp3A13-2* negative control, 6) *Cyp3A13-2* sample 1, 7) *Cyp3A13-2* sample 2, 8) *Cyp3A13-3* negative control, 9) *Cyp3A13-3* sample 1, 10) *Cyp3A13-3* sample 2, 11) 10 bp DNA ladder, 12) *Cyp6BS1* negative control, 13) *Cyp6BS1* sample 1, 14) *Cyp6BS1* sample 2, 15) *Cyp6BS1-2* negative control, 16) *Cyp6BS1-2* sample 1, 17) *Cyp6BS1-2* sample 2, 18) *Cyp6BS1-3* negative control, 19) *Cyp6BS1-3* sample 1, 20) *Cyp6BS1-3* sample 2) (Figure 4.3 - bright band in lanes 1 and 11 is at 100 bp. Lanes: 1) 10 bp DNA ladder, 2) *Cyp2u1* negative control, 3) *Cyp2u1* sample 1, 4) *Cyp2u1* sample 2, 5) *Cyp2u1-2* negative control, 6) *Cyp2u1-2* sample 1, 7) *Cyp2u1-2* sample 2, 8) *Cyp2u1-3* negative control, 9) *Cyp2u1-3* sample 1, 10) *Cyp2u1-3* sample 2, 11) 10 bp DNA ladder, 12) *Cyp2j6* negative control, 13) *Cyp2j6* sample 1, 14)

Cyp2j6 sample 2, 15) *Cyp2j6-2* negative control, 16) *Cyp2j6-2* sample 1, 17) *Cyp2j6-2* sample 2, 18) *Cyp2j6-3* negative control, 19) *Cyp2j6-3* sample 1, 20) *Cyp2j6-3* sample 2).

Only the two samples of the primers producing the strongest bands were chosen for PCR product sequencing (*Cyp2j2-2*, *Cyp2j2-3*, *Cyp2c9-2*, *Cyp2L1*, *Cyp3A13-2*, *Cyp6BS1*, and *Cyp2j6-2*) (Tables 4.10, 4.11). Primer efficiencies were analyzed for only the four primer sets with the highest sequence identity with their expected amplicon sequence (*Cyp2j2-2*, *Cyp2c9-2*, *Cyp6BS1*, and *Cyp2j6-2*) (Table 4.12). Initial statistical analysis of primer efficiencies revealed significant differences. After elimination of *Cyp2c9-2*, no significant difference remained in primer efficiency (two-tailed ANOVA after removal, α=0.05, F=1.39 (df 18, 38), *P*=0.19).

Figure 4.1: PCR products from *H. azteca* cDNA of primers for *Cyp2j2*, *Cyp2c9*, and *Cyp2L1*

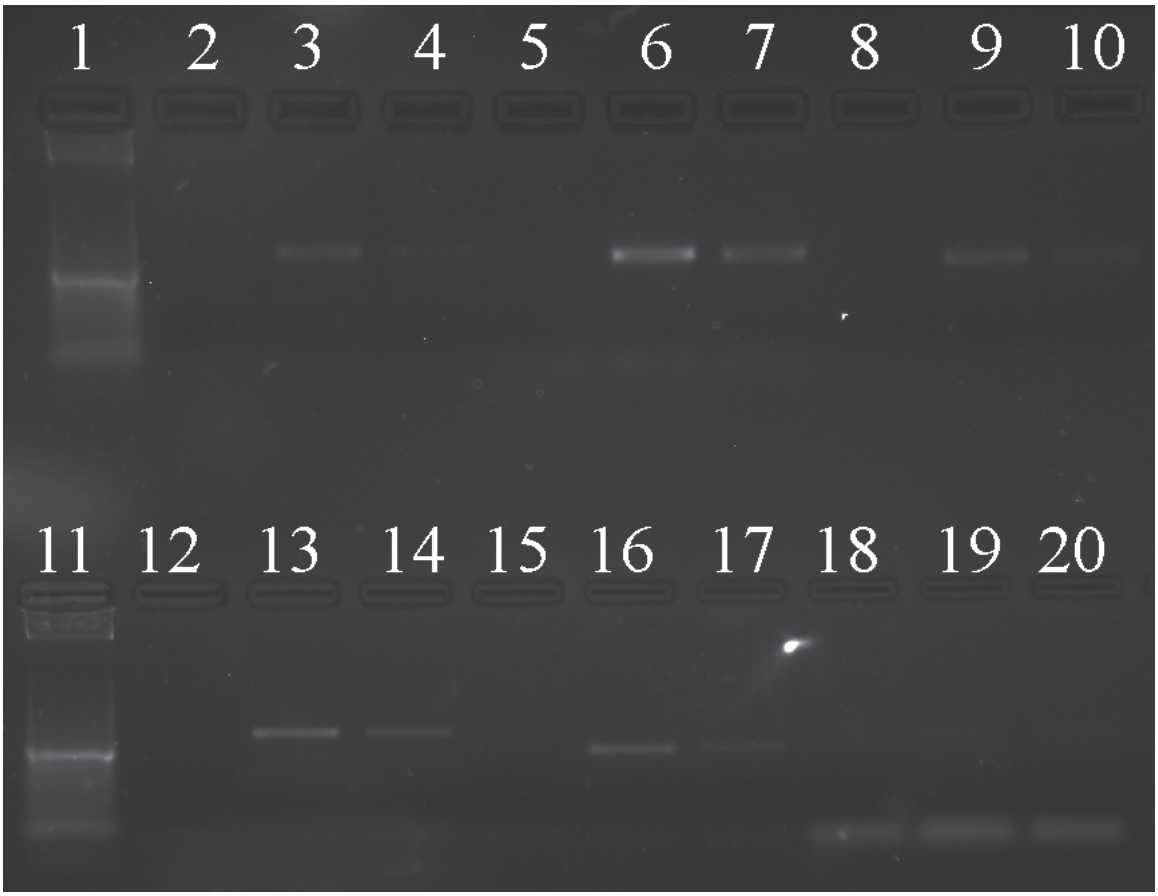

Figure 4.2: PCR products from *H. azteca* cDNA of primers for *Cyp3A13* and *Cyp6BS1*

Figure 4.3: PCR products from *H. azteca* cDNA of primers for *Cyp2u1* and *Cyp2j6*

Table 4.10: Alignment sequences of PCR product sequences and expected amplicon sequences

Primer set	Sample	Product sequence
Cyp2j2-2	1	GCAGCAGTATG--NNNNNGGGAGTGGCTTTAACGGACGGGCTGTTGTGGT CAGAGAACCGACATTTTTTCCTCACTCAACTGCGGAACCTGGGCATGGGT AAGACTTTCCTCGGG
	2	CAGCAGNAT--NNNNNGNGGAGTGGCTTTAACGGACGGGCTGTTGTGGTC AGAGAACCGACATTTTTTCCTCACTCAACTGCGGAACCTGGGCATGGGTA AGACTTTCCTCGGG
Cyp2j2-3	1	CTGCACGCGACTCC-ATGGATTCNGCA---NNGTGGGTTGGCTGGGCATCTG
	2	AAAAAACTGNACGCGACTCC- NTGGNTTCCNCNCNNNTNNCNNTGCTCNNC
Cyp2c9-2	1	ANNTTCTTATGAGCCNNATGAAGAAGATGCATAAAGGAGCACAAGGAGA GCCTCGATGTCAACAATCCTCGCGATTACATCG
	2	CTTATGAGCCACATGAAG- AGTGCATAAAGGAGCACAAGGAGAGCCTCGATG TCAACANNCCTCGCGANTACATCG
Cyp2L1	1	CNNAGGGCACGGTGCTNNTGGGA-TGGCGGGCGGCTGT-TCGGGGACCCN CAGGTCTGG-ATGNNNNCGACGAGTTCTATCCTCCGGAGTTCTTGGNC
	2	GGC-CGGTGC-NNTGGGA-ACGCGGGTGGCTGT-TCGGGCATGCCCTTGG-- GGNATGGAANCGT------TACTTCNAT---------GTCTTGG
Cyp3A13-2	1	TGCTT-NNNAATTATGATAATANCTTATTGNANTNATANTAGGTCTCACC ACAG---ACNATGTGTTGTACCGTGACC
	2	TATTTTGCTT-TTTAATTA---TGATAACATNTTGNACCNCA-----TAG TANT-- AGGCTCCCCACCGGCTTGGTGCCGTGCCC---CGC
Cyp6BS1	1	TNNGATTCA-TATNNNGNACCTCTGATGCGGACTGTCAACCACAGCCAGA GAC-CGAGTAATGGTTTGGCTTTTGTGTTC
	2	TNNCGNACCTCTGATGCGNACTGTCANNCNCAGCCAGAGAC-CGACTTAT GGATTGGCTTTTGTGTTC
Cyp2j6-2	1	ANTNNNCGTTCNCCGNATCTGANNNNNNNTTTCAGATAGCATAGTTCTGT TCCAAAATTTATCGTCCTTCGACAAGTGACCGTT
	2	TTNCGTTCNCCGCATCTGAGGTTAGGTTTCAGATAGCATAGTTCTGTTCC AAAATTTATCGTCCTTCGACAAGTGACCGTTCG

Table 4.11: Alignment statistics of sequenced PCR products and expected amplicon sequences

Primer set	Sample	Expected product length (nt)	Sequenced product length (nt)	Percent similarity (%)	Gaps (%)
Cyp2j2-2	1	150	115	93.9	1.7
	2	150	114	92.1	1.8
Cyp2j2-3	1	105	52	65.4	11.5
	2	105	51	70.6	2
Cyp2c9-2	1	121	82	93.9	1.2
	2	121	76	93.4	1.3
Cyp2L1	1	141	98	78.6	3.1
	2	141	96	58.3	27.1
Cyp3A13-2	1	135	78	67.9	7.7
	2	135	91	59.3	19.8
Cyp6BS1	1	127	80	87.5	2.5
	2	127	83	94	0
Cyp2j6-2	1	126	84	84.5	0
	2	126	83	97.6	0

Table 4.12: Amplification efficiencies of tested cytochrome P450 primer sets

Primer set	Efficiency (% ± SD)
Cyp2j2-2	80.1±1.2
Cyp2c9-2	83.3±1.5
Cyp6BS1	84.7±0.7
Cyp2j6-2	85.4±0.7

Gene of interest analysis

Significant increase in expression of *Lgbp*, *Hsp90*, and *Mrp4* was seen in response to DDT exposure ($P < 0.05$) (Figure 4.4). *Hsp90* and *Mrp4* increased most, by about 4.5-fold. No significant changes in gene expression were seen in response to imidacloprid exposure, though there was a strong trend towards decreased expression of *Cyp2j2* and *Cyp6BS1* (Figure 4.5). Significant expression decrease of *Lectin-1* and increase of *Cyp2j2* were seen in the permethrin exposure ($P < 0.05$) (Figure 4.6). *Gst* and *Lgbp* also decreased in expression strongly, though not significantly.

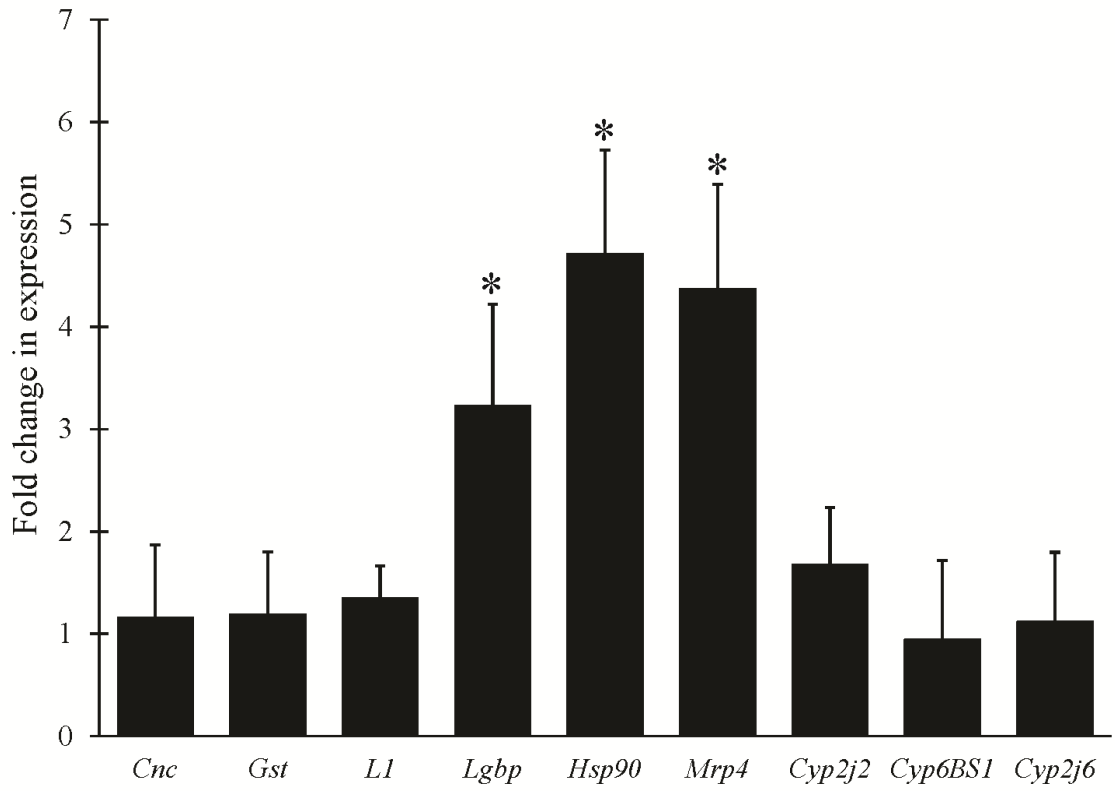

Figure 4.4: Fold change in gene expression of GOI in response to DDT exposure. Bars are standard errors. Significant differences from control ($P < 0.05$) indicated by *. *L1* stands for *Lectin-1*

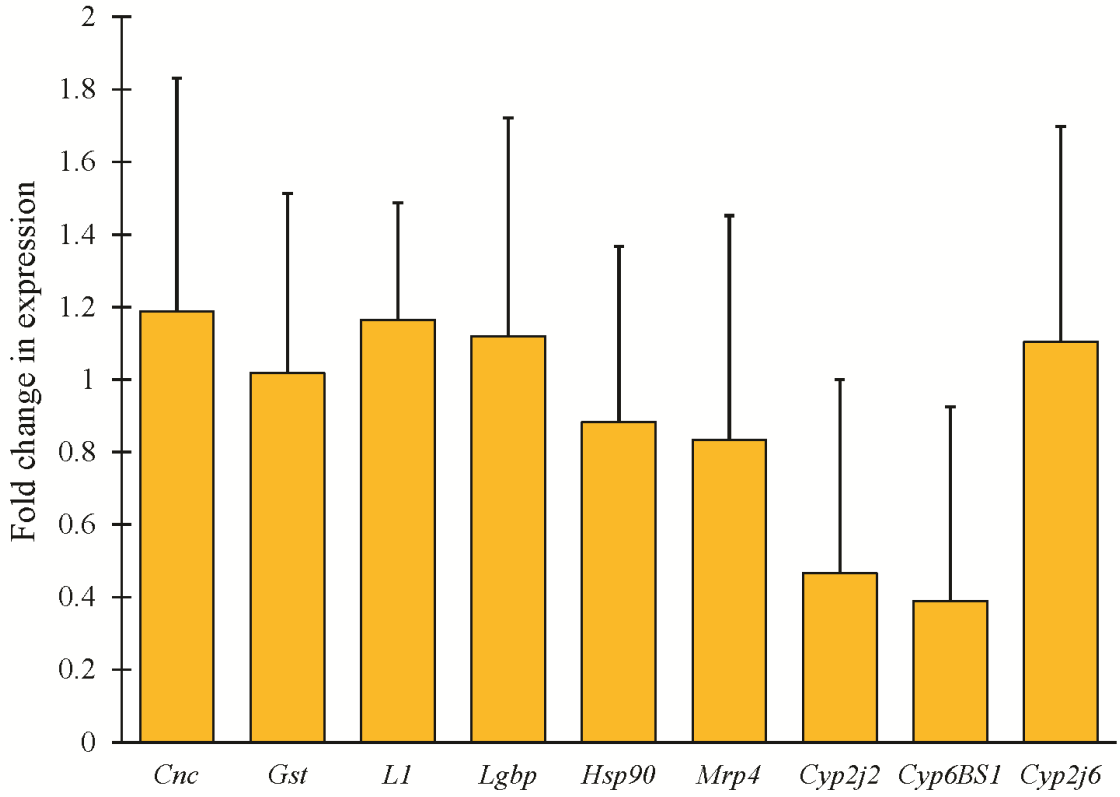

Figure 4.5: Fold change in expression of GOI in response to imidacloprid exposure. Bars are standard errors. *L1* stands for *Lectin-1*

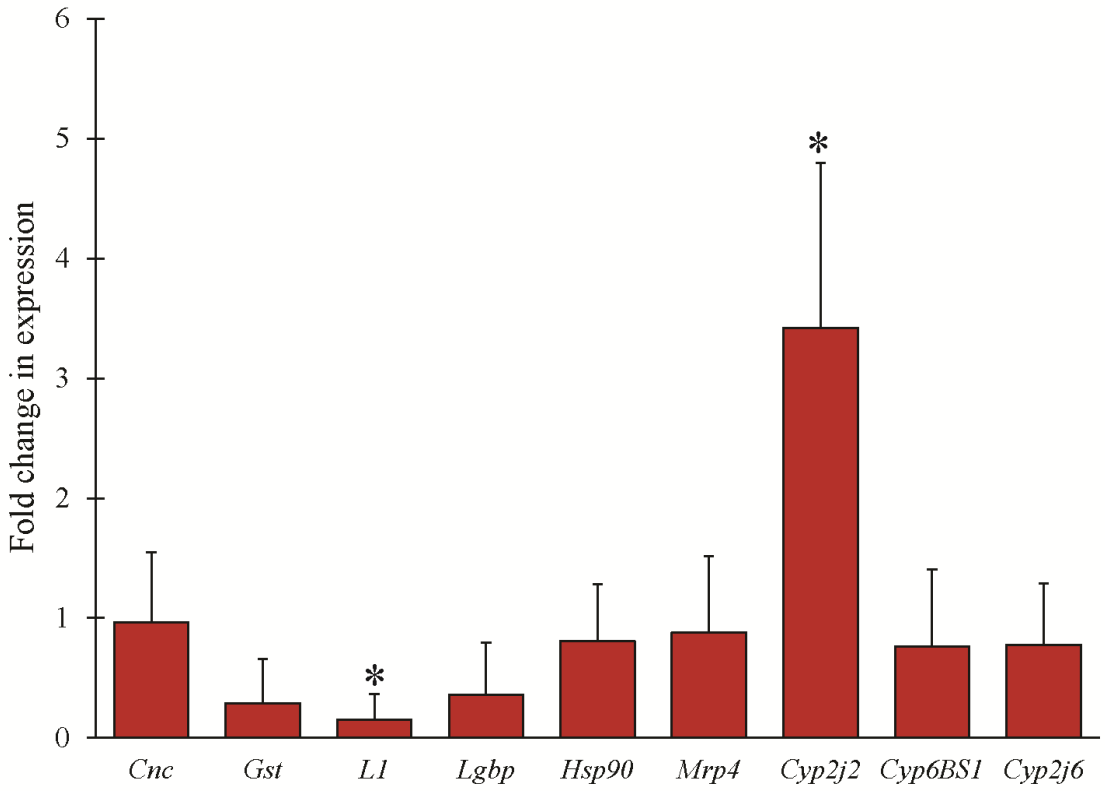

Figure 4.6: Fold change in gene expression of GOI in response to permethrin exposure. Bars are standard errors. Significant differences from control ($P < 0.05$) indicated by *. *L1* stands for *Lectin-1*

Discussion

Expression of stress-response genes was measured in *H. azteca* after exposure to DDT, permethrin, and imidacloprid to develop molecular biomarkers of exposure for these insecticides. DDT elevated the expression of the immune response gene *Lgbp*, the ABC transporter *Mrp4*, and the heat shock protein *Hsp90*. Permethrin exposure increased the expression of the detoxifying cytochrome P450 *Cyp2j2* while decreasing the expression of the immune response gene *Lectin-1*. Imidacloprid did not significantly alter the expression of any genes studied. These changes in gene expression can serve as biomarkers of exposure at low concentrations for DDT and permethrin. The genes

responding to exposure respond to stress and cellular damage that could serve for future connections to adverse effects.

Many molecular biomarkers of insecticide exposure are already in common use. Activities of cholinesterase, acetylcholinesterase, glutathione-*S*-transferase, and catalase are used as biomarkers of organochlorine and organophosphate insecticides, pyrethroids, and neonicotinoids in a variety of aquatic organisms (Jemec et al., 2007; Domingues et al., 2009; Sellami et al., 2014). These molecular responses to insecticides generally fall into two main categories: general stress-response and detoxification. The organochlorine insecticide dieldrin increases expression of the general stress-response heat shock proteins 60 and 70 in several amphipods, including *H. azteca*, (Werner and Nagel, 1997). Dieldrin and chlordane increase expression of certain cytochrome P450s (Coumoul et al., 2002; Dehn et al., 2005). Resistance to DDT is known to involve upregulation of several cytochrome P450s in *Drosophila melanogaster* (Maitra et al., 1996; Daborn et al., 2012; Daborn et al., 2002). These P450s only implicate detoxification genes and proteins in the response to DDT. However, DDT resistance has since been found to be polyfactorial, involving many changes to alter the ability of DDT to penetrate the arthropod cuticle and the organism's ability to both detoxify and excrete the chemical as well (Strycharz et al., 2013). For example, multiple ABC transporters have been associated with resistant *D. melanogaster*, which allow them to more readily excrete DDT or its metabolic products (Pedra et al., 2004; Gellatly et al., 2015). Implication of ABC transporters in DDT resistance is also found in the mosquito *Anopheles arabiensis* and has been expanding to many other arthropods as well (Jones et al., 2012; Dermauw and Van Leeuwen, 2014).

The gene expression responses of *Hsp90* and *Mrp4* to DDT seen here in *H. azteca* are not unexpected given previous knowledge of members of these gene families reacting in other arthropods. *H. azteca* is capable of responding to DDT exposure via general and specific mechanisms to protect from damage and detoxify. What is unique is the upregulation of the immune response-related gene *Lgbp*. No previous work describes a change in immune response genes in arthropods after exposure to DDT, though *Lgbp* is known to increase or decrease in expression in response to other chemical and biological stressors in the environment (Zhang et al., 2010; Chang et al., 2013; Chen et al., 2016).

Environmental exposure of *H. azteca* to pyrethroids is common and chronic enough that some populations have evolved resistance via site mutations in target sodium channels (Weston et al., 2013). Resistance to pyrethroids, including permethrin, is also well known in insects including house flies and mosquitoes, resulting from activity of ABC transporters and cytochrome P450s (Kasai and Scott, 2000; Kasai et al., 2000; Nikou et al., 2003; Muller et al., 2008; Djouaka et al., 2008; Bariami et al., 2012). GST activity increases in response to permethrin exposure in the amphipod *Gammarus pulex* (McLoughlin et al., 2000). This contrasts with the downregulation of the GST measured here in *H. azteca*. The immune response genes *Lgbp* and *Lectin-1* were also downregulated. Many insecticides are known to interact with arthropod immune capacity, usually investigated at only the cellular immune response level (Desneux et al., 2007). Permethrin is highly acutely toxic to *H. azteca* (Andersen et al., 2006; Ding et al., 2012). The immune suppression at sub-lethal permethrin concentrations creates an additional risk for non-target organisms, making them more susceptible to other stressors.

Decreasing environmental concentrations of permethrin to simple a sub-lethal concentration may not be enough to protect non-target populations.

The response to imidacloprid in insects is exemplified by the mechanisms conferring resistance to imidacloprid. Resistance is most often tied to overexpression, either inducible or constitutive, of cytochrome P450 genes. Inducible P450 activity confers resistance to the brown planthopper *Nilaparvata lugens* and to the Asian citrus psyllid *Diaphorina citri* (Puinean et al., 2010a; Tiwari et al., 2011). Constitutive overexpression contributes to resistance in *D. melanogaster*, the green peach aphid *Myzus persicae*, and the tobacco whitefly *Bemisia tabaci* (Daborn et al., 2001; Karunker et al., 2008; Puinean et al., 2010b). Imidacloprid, however, is fairly selective in its toxicity, causing significant mortality in insects while not affecting vertebrates (Song et al., 1997). Non-insect arthropods often suffer chronic, sub-lethal effects from imidacloprid exposure (Lukancic et al., 2010). These chronic effects reduce population growth rates and potentially effect populations and communities (Chen et al., 2010; Van Dijk et al., 2013). Very little is currently known about the molecular responses of non-target non-insect arthropods to imidacloprid. The trend towards decreasing expression of the two cytochrome P450 genes seen here in *H. azteca* indicates that more work is warranted to investigate these effects. The genes studied here are not useful biomarkers for imidacloprid at the imidacloprid concentration examined.

DDT and permethrin had distinct profiles of expression changes. DDT exposure increased expression of *Lgbp*, *Hsp90*, and *Mrp4*. Permethrin decreased expression of *L1* and increased expression of *Cyp2j2*. Imidacloprid showed no significant changes in

expression of any genes monitored here. This allows easy distinction between these insecticides using gene expression profiles. No similarities in expression changes were seen across all three insecticides. Shared increases in the same detoxification genes may occur at higher exposure concentrations, especially between DDT and permethrin, which share the target site of voltage-gated sodium channels.

The gene responses developed here in the model organism *H. azteca* can serve as biomarkers of exposure to DDT and permethrin to help inform ERA analyses. Further studies should focus on the effects of these low insecticide concentrations on ecologically relevant life history parameters like development and fecundity. Connecting exposures and molecular effects to important, larger-scale ecological effects forms an Adverse Outcome Pathway (AOP) (Ankley et al., 2010; Vinken, 2013). AOPs describe the sequence of events starting at exposure and molecular interaction with a chemical through higher order levels of biological organization, ultimately ending in adverse ecological effects. Changes in gene expression are defined as molecular initiating events within the framework of the AOP (Allen et al., 2014). These molecular initiating events begin the process leading to ecological adverse effects. Detecting molecular initiating events at very low exposure levels is important for risk monitoring and informing risk reduction efforts at the population and community levels (Forbes and Galic, 2016). The gene expression biomarkers developed here at low concentrations of insecticides give us the ability to do just that.

Dissertation Summary

My goals in this research were to establish *Hyalella azteca* as a useful molecular model organism for toxicology and risk assessment and to develop molecular biomarkers in *H. azteca* that are more sensitive than their comparable higher-level biomarkers. I sequenced and assembled a transcriptome for *H. azteca* to not only pursue my own research, but also to provide an immense molecular resource for others to pursue research with this amphipod. I identified useful "housekeeping" reference genes and stress-response genes in the transcriptome to use for developing gene expression-based molecular biomarkers. I conducted exposure experiments, challenging *H. azteca* with sub-lethal exposures of the metals cadmium and copper and of the insecticides DDT, permethrin, and imidacloprid. I first measured expression of candidate reference genes using RT-qPCR for all of these treatments and their controls. The programs BestKeeper, GeNorm, and NormFinder determined the reference genes with the most stable expression between controls and treatments. These stable reference genes were used together as normalization factors (NFs). NFs remove outside sources of variation when processing sensitive gene expression data. The genes forming the NFs differed between every treatment, showing how important it is to select reference genes on a by-treatment basis and not by simply selecting from a list of previously used reference genes. I then measured expression of stress-response genes of interest in all treatments and controls. I found a gene expression signature that could be indicative of exposure to metals in general in *H. azteca* (increased expression of *Cnc*, *Mrp4*, and *Hsp90*). I also found that gene expression could differentiate between exposure to cadmium and copper, with *Mrp4*

expressing being much higher after cadmium exposure than with copper and *Rad51* increasing in expression only after cadmium exposure. While there was no gene expression pattern similar to all insecticides, DDT and permethrin exposure produced unique expression profiles. Imidacloprid induced no changes in expression in the genes studied. Finally and importantly, I found no change in survival in any treatment. Changes in gene expression served as more sensitive biomarkers of exposure than mortality. Further studies could compare gene expression sensitivity to other high-level biomarkers like fecundity and behavior. Monitoring expression of these genes during risk assessment will provide a greater ability to determine exposure risk at low stressor concentrations before population and community level adverse effects occur. With these results I have established *H. azteca* as a useful organism for molecular studies in toxicology and for informing ecological risk assessment. Given these data in this dissertation, the now publicly available transcriptome, and the soon to be published *H. azteca* genome, I believe there is a quickly approaching wave of research on molecular applications of this amphipod.

Appendix A: Code used for transcriptome assembly

Fastq screen, example genome indexing for *Daphnia pulex*:

perl bowtie2-build /Volumes/H_AZTECOME/fastq_screen/bowtie2-2.2.3/Dpulex.fasta /Volumes/H_AZTECOME/fastq_screen/bowtie2-2.2.3/Dpulex

Fastq screen, example check of forward and reverse reads against indexed genomes:

perl /Volumes/H_AZTECOME/fastq_screen/fastq_screen_v0.4.4/fastq_screen --paired /Users/ryangott/Desktop/Ha1_Forward/HA_1_GAGTGG_L004_R1_001.fastq /Users/ryangott/Desktop/Ha1_Reverse/HA_1_GAGTGG_L004_R2_001.fastq --aligner bowtie2

Trimmomatic, example trim:

java -jar /Volumes/H_AZTECOME/Trim_out/Trimmomatic-0.32/trimmomatic-0.32.jar PE -phred33 -trimlog /Volumes/H_AZTECOME/Trim_out/HA_1/ha1_001_trimlog.fq /Volumes/H_AZTECOME/Sample_HA_1/Forward/HA_1_GAGTGG_L004_R1_001.fastq.gz /Volumes/H_AZTECOME/Sample_HA_1/Reverse/HA_1_GAGTGG_L004_R2_001.fastq.gz ha1_001_forward_paired_out.fq ha1_001_forward_unpaired_out.fq ha1_001_reverse_paired_out.fq ha1_001_reverse_unpaired_out.fq ILLUMINACLIP:/Volumes/H_AZTECOME/Trim_out/Trimmomatic-0.32/HA1_adapters.fa:2:30:10 LEADING:5 TRAILING:10 SLIDINGWINDOW:4:20 MINLEN:36

BUSCO analysis:

python BUSCO_v1.1b1.py -o HazBUSCOs -in newtranscripts.fa -l arthropoda/ -m trans

Appendix B: Housekeeping gene transcript sequences

Transcript	Length (nt)	Sequence
RpS23	420	TTTGATTAGGGATAGTTTTTGCCAATGCTGAGTCACGAATCGAT TGCTAATGAAAACCAGTCTAGAAGTAGTAGTTACGGTGCCCCG TATGTTTGCGGGTAAAGCCTCTGTCTGATTGAGCTATTTTTCAG AGGTATTGAGGCGAAACAGCCCAACTCGGCCATCAGGAAGTGC GTGCGCGTGCAGCTCATCAAGAACGGCAAGAAGATCTCAGCAT TCGTGCCTCGTGACGGTTGCTTGAATTTCATCGAGGAGAACGAC GAGGTCCTGGTCGCTGGATTTGGTCGTAAAGGTCACGCCGTTGG TGATATTCCCGGTATCCGCTTCAAGGTCGTGAAGGTCGCCAACG TGTCCCTTCTTGCACTTTACAAGGAGAAGAAAGAAAGACCGCG TTCTTAAATAAACGTGACTCGTGAAATT
RpL10	861	CAAAATATTAATTTTTGCTAATTAGCTAGTATGGAATCAAAATA TCCATATATGTAGCTATGTAAGTTAATCCAACTGATCAATATAT CAATTTATTTTTAACTTGTGTAATATGAACTAATGGGTGCTTTGT ATTGTTGCAGAAAGGCATCATGGGGCGGAGACCTGCAAGATGC TATCGCTACTGCAAGAACAAGCCGTACCCCAAAAGTCGGTTCT GTCGTGGTGTGCCTGACCCCAAAATTAGGATTTTTGACTTGGGC CGCAAGAAAGCTGACGTCACTGAGTTCCCGCTCTGTGTCCATCT TGTGTCGGATGAATATGAACAGCTGAGCTCAGAAGCTCTTGAG GCTGCTCGTATCTGTGCCAACAAGTATTTGGTGAAGTACACCGG CAAAGATCAGTTCCACATCAGGATTAGAGTCCACCCCTTCCACG TGATCCGCATCAACAAGATGTTATCGTGCGCTGGGGCTGATAG GCTCCAGACTGGAATGAGAGGCGCTTTTGGCAAGCCGCAGGGC ACAGTTGCTCGTGTCAACATCGGACAGCCCATCATCAGCGTGC GCTCCTACGATAAACATAAGGCTCATGTTGTCGAGGCTCTCAGG AGGGCCAAGTTCAAGTTCCCCGGCCGTCAAAAGATTTTCGTTTC TGGGAAGTGGGGCTTCACGAAGTACGAACGCCCCGACTACGAG AAGATGCTAGCGTCAGGCCATATCAAGTCTGATGGTGTCAATG TCAAGTATATCCCTGAGCACGGTCCACTAATTAAGTGGAAGCG TACGCAGCTCCAGCTTGCTGGCATGGCCTAATGTGCATCCGCGG AAATTTCTGAATAAAGATCACCCGCGCCGTAAT
Act	365	TGCAGATTTTCTTCCCCACGGTGCTGGCCACTGACCGGGAGAGC AGACGGCAGATCTGGGGAGCTGAAGCCCTCAGCCCCGAAGTGC GATCTTCATCAAGCCTGTCGTTCCCCATCAGGCCATCGCACAGG ATATCTAAGTACACTGTCGACTTGAATGCCGTGTCAAGTTTGCT CTCAAAAGTTTTCTCCGAGCTCAAAGTTGACGCCCGGAACTACA ACCTTCAGATATCTGCGCCCAGAATCCTCAATCAGCACACACA GTCTGAGCTCATCAAGATCCTCTTCGAAAAGTTCGGGGTGAAG GGCGTGAACATCACCCACCAGTCCATCCTGGCGCTGTACGCGT ACAACGCCACGTCCGGC
Gapdh	1294	CGATCTCAGAGCTGGCAGTCTAGTCAGAGCATTCGCCGTTGTAC CGTGTAGTCATCACTCCAGCTCAGCTAACATCATGTCGAAGATC GGAATTAACGGATTTGGCCGCATCGGGCGCCTGGTGCTGCGTT GCGCTATTGAGAAGGGAGCTGAAGTGGTGGCGGTGAACGACCC ATTCATCACCTTGGACTACATGGTCTACATGTTCAAGTATGACT CGACTCATGGCACGTTCAAGGGTGACGTGAAGGCTGACGGTGG CAACCTCGTTGTCAACGGCCATGTCATCAAGGTCTTCAATGAAA TGAAGCCTGATAACATTCCATGGGGAAGCGCTGGTGCCGAGTA CATCGTCGAATCCACTGGCGTCTTCACCACTCTTGACAAGGCCA AGGCTCATTTTAGCAACGGCGGCAAAAAAGTAATCATTTCGGC

		CCCATCTGCTGATGCTCCCATGTTCGTCATGGGCGTGAACCACG AGAAATATAGCAAGGACATGACAATTGTGTCGAACGCCTCGTG CACCACCAACTGCCTCGCGCCCGTCGCCAAGGTGCTCAACGAC TCCTTCGGTATTGAGGAGGGACTCATGACGACCGTGCACGCCG TCACCGCTACTCAGAAGACTGTCGATGGTCCCTCGGCCAAGGA CTGGCGCGGTGGCCGGGGCGCTGGCCAGAATATCATCCCTTCA TCCACTGGCGCTGCTAAGGCTGTAGGTAAGGTCATTCCCGAGCT CAACGGCAAGCTCACTGGCATGGCTTTCCGCGTTCCCACGCCCG ATGTCTCTGTCGTCGACTTGACCGTGAGACTTGGCAAGGAGTGC TCCTACGACGAGATCAAGGCCGCTATGAAGGCCGCGTCAGAGG GACCTCTCAAGGGATATCTCGGCTACACCGAAGACGATGTCGT GTCAAGCGATTTCATTGGCGACCACAGGTCATCTATCTTCGATG CCAAGGCCGGCATCCAGCTGAGCAAGACGTTTGTGAAGGTCGT CTCATGGTACGACAACGAGTACGGCTACTCCAACCGTGTCGTTG ATCTTATTAAGCATATGCAGAATGTAGATGCGTAAGGCTGTGA GCCTTCCATGTATTGTTTGTTAGAATCGACTTCATCAAGGATTA CCTTCGCAGTACAGTGGGCTTGAGCGTGATTCATAGCTAGTGAT GGAAATATCTGTTAAACATATATTTTTCCGTTTTCTTGGGACCC CAAGAATTAATTGATTTCATCTTTCGGTCTAACAGACTGTCCTC TGTAAACTGCAATAAACTGGATTAAAATACTG
αTub	1545	TGCTGCATTACGAGGGTCACGAGAAGGCTTACTTGTGAACGCG CGCTCACTTACTTATTCAAGACGTATTCTGCGTCTGTTACTTCAT TGACTCAACTCGGACTTCTGCGTTTCAGTAATCATGCGTGAATG CATCTCGATCCACGTTGGCCAGGCTGGAGTGCAGATGGGCAAT GCCTGCTGGGAGCTGTACTGCCTCGAGCACGGCATCCAGCCCG ATGGCCAGATGCCCTCTGACAAAACCATCGGTGGAGGCGATGA TTCCTTCAACACTTTCTTCTTCGAGACTGGCACGGGAAAACACG TGCCCCGAGCCATCTTCGTTGACCTGGAACCGTCTGTTATAGAT GAGGTCCGGACTGGAGTCTACCGACAACTTTTCCATCCAGAAC AATTGATCACTGGGAAAGAAGATGCTGCTAACAACTATGCACG AGGTCACTATACCATCGGGAAGGAAATCGTTGAAATTGTTTTA GATCGCATTCGTAAAACCGCCGACAACTGCACTGGCCTTCAAG GTTTCTTGGTCTTTCACTCGTTTGGTGGCGGCACTGGTTCCGGTT TCACTTCCCTCTTGATGGAGAGGCTTTCGGTCGACTACGGCAAG AAAAGCAAACTTGAATTTGCCATCTACCCTGCCCCTCAAGTCGC CACTGCCGTCGTCGAACCTTACAACTCGATCCTGACCACACACA CTACTCTGGAACATTCTGACTGCGCCTTCATGGTTGACAACGAA GCAATTTACGATATTTGCCGCCGTAACCTCGATATCGAGCGACC CACTTATACTAATCTGAATCGCATGATCGGTCAGATCGTCTCGT CCATTACCGCTTCCTTGAGGTTCGATGGCGCCCTCAACGTCGAC CTGACTGAGTTCCAAACGAACTTGGTCCCTTACCCGCGTATTCA CTTCCCATTGGTGACATACGCGCCAGTCATCTCTGCCGAGAAAG CCTACCACGAGCAGATTTCAGTGGCCGAGATCACAAACGCTTG TTTTGAGCCCGCCAACCAGATGGTGAAGTGCGACCCGCGTCAT GGAAAGTACATGGCGTGTTGCCTGCTTTTCCGAGGTGACGTCGT GCCCAAGGATGTCAACGCTGCCATCGCTACCATCAAGACCAAG CGGAGCATCCAGTTTGTGGACTGGTGCCCCACGGGCTTCAAGG TGGGCATCAACTACCAGCCACCCACCGTAGTTCCTGGCGGTGA CCTGGCTAAGGTGTCTCGTGCCGTGTGCATGCTTAGCAACACCA CGGCCATCGCTGAAGCCTGGGCTCGTCTGGACCACAAGTTCGA CCTGATGTACGCCAAGCGCGCCTTCGTGCACTGGTACGTGGGC GAGGGCATGGAGGAGGGCGAGTTCACCGAGGCCAGGGAGGAC TTGGCCGCCCTCGAGAAGGACTACGAGGAGGTCGGCGTCGACT CCGCTGAGGATATCATTGGAGAAGGAGAAGAGTATTAATTGAC

		TCATAAATCCTACCACCGGCGTCATGGTATTTTTACCTGCATAT CACGAGCATTAAAACATATCAC
Mmp	1031	CTTCGTCAGATGTCTGCGACAAGATAAATGCGTACCTATTAAAG ACTTTAACAGAATACGAATTTACAAAACTTTTACCTTTCAGCAA ATTTGTCTTTCCTCGTGTACAGGATAAGGACAATCAAAGAAAA ACGAAGAAATCAAACGAGGCGGATCGACAACACTCATTTTCTG ATTGTCGAATGGCCAGAAGGTACGTTCATAGAAAAAGTATGTC TTCACGCCCACTGAAACTGCTGCTGAAATGTTGTAAGGAACTCC GCGCCACACGTCCATGGGTCTTGGGTAATCATTTTCAACTTTGC CGTCACTTCCCAATCTCCAGTACAGATCTCCACTGAATAAATAA GCGTATCCGTTGCGTCCCCATATCATGGCGGCATCTATTTTCGA GACGTTTTTGATGCCAAGGTCTACCAATTTACCTTTCTCGATTG ATTCAGTATTTCGATGTAGCGTGGTCTTAATGAATTTTTCGCCC GAGAAAATAACAAGACTTTTTCCATTGGGTGTTTCAAACGCAG CGTCAATTTTTTCTAGACCTGAAAAAGCTTCCAGACCAAGACGT TCTAATTCTGATGGGTACTCAGAGGTCAACATTGGAAGACGCC ACAAGAATTTATCTTTAAAAAAGTACAGCTCGCCTCTAAAGTA AGTGATTGCATCAAACGATGTGCTACATGCATCGGGGATTTCCA AGCCTGTTGACTTCGGTTTACATGATCCTTGGTTAGGTACAGTT GTGGGTACAGTTGTGCGAGGAACTACGGGAGGTTGATTTGGCT CTCTTGTTGTTGTTGTTGTTGTTGGAGTCGTGGAGGTGCTGGTTG TGGTTGTCGTATCTGGAGTAGTAAACTGAGGCCGATAAGGATTT AATCCAGAATCTGGTACTCCATACAAGGCTTGAATAGCTCGAC GGTCGTCGTCAGGTAACTCAAAAGTTTCCGGATAGCTTTGGTAG TACGGGGCCATCAAAGCCGTGTGCTCGTTGGAGTGTTGCAAAC CCAATGAATGTCCAAGTTCATGAGCCG
Syx	1826	CCAATCTTGTCGGCTCTCTTCAAACCAAAGTCGCTACCATCACG CAGAAGTTCAGAGATGTCCTCGAAGTTAGAACTGAAAACTTGA AAAAGCAAGCGGAGCGTCGTGAGCAGTTCACGGGCGGGACGA GCAGCGGAGCCGTCATGGCTGCTGGCGGTCACCAGGCATCGGT GCTGCTGGCAGACGAGGCTAAAGCCACTGCCGCTGCAAGGAGA ACTGCCAGCAAACCTAATGGCGATGTAGTTCTAGACTTCGATG GCGTAAGCGTTGCTCCAAGCAACAGTGGAGGGGCTTACCAGCA GCAGCTGGCGCTGATGGAGGAACAAGACTCGCTGCTGCAGTCC CGCGCTGAGACCATGAAGACCATCGAGTCAACCATCGTCGAGC TCGGCACGATGTTCACGCAGCTCGCTGCTATGGTCAAGGAGCA AGACGAACTTGTACATCGAATCGATGCCAACGTCGACGATGCA GAGATGAACGTGGAGGCCGCGCATGCGGAACTGCTGAAATATT TCCGATCGGTCTCCTCCAACCGCTGGATGATGTTCAAAGTGTTC GGTATTGTCATCATCTTCTTCATCATTTTCGTTGTCTTCATGGCG TGAGTATCACAATCGATGGATGGTACCACGAAAATGCAAGTTT GTCATTGTTTATGACCACAGATCGAAATGTGAAGTGAACGATT GTTGACGAGTGAAAGGATCTTTCACGTGTAAAATTTTCCCTTTT TTAACGATTTTCCTTCGATTCGAGCCTTGCTTAATGACGGCATA TTTTTATTTTTTTACATAGCTATGAATATTAATTTATTTATATTT AGTTATTTATATTTCCCCTGGAATCAATTATTTTGAGAAATAGT ACATTTCTTCATTTAAATATAACCTCATACTGAACGTTTGATAT CGTATCTATGGATCGATAATATTTCTTTACTTCTTGCTTGCCTTA TCTCCGATTTCCATTTAGAATTTATTTTGTGTGCCCCGCTTCATT AGAACGGTAGTAGCCAGCGAATCGAATTGGACATAGATGAAGT CAAGTCTTCTGAACACCCTCATTTCGAGTATTTAATATATTTAA AGTTATGATTTATTTCGTCAAAATTCCTTCAAATTACGCAGATG ATCAAGTCTATACAATTCATGTCACCGTAGAAATTTATCAGAGT AAGCAGTGAATATAATTGGCTTGTTTTTACCAAATTTTACATAA

		CTGGGCCTATAATTTCAAAATTCGCTAAGTGAACGTAGGTAATT ATGCTCTATTTTCTTGTCTATGATTGGTGGATTTTGCTTGAAAGG CCGTCAATGATTGGCTGATGAACATACGATCTCTTAATTTTTTT GGAATGTATAGGCCCTGTTCTGTGTAAAAACTCAAACGTCCGA ATACGAAGTGACTGCTCACACTTTAACATTTCAATTTTTTATGT ATTACAAAGGCACACTTTGTTATACGATTTCATGTCTCGAAGGG AAACTGTGTATATTGTTCATCGCTTCATACAATTGATAATTTAA CGAGTACTGTACAGCTTTCAAAGAGATATAATAAGCTTATTATG TGTTGAATAACTTTTGTACATGTGTACAGCATTCCACCTAGTAT GCCTAAGTTGGGAAATGAGGTGATTATGAACTGATGATATTGA GTTGCTGTAGACGCTATTGGGCTGACGTAAATGCAGCTGCTCTT CTAACATTGTACACGTTATATATTCGAAGTTTAATATACAGTTT CCAAGGCTTGCTGTTCACTCTCATGAAACTTTGAAAATATACGT AATGTTGATGCATTAACTAACGTGCCAGATATAG
Tbp	1597	GTTATGAACCCGAGTTATTCCCTGGCCTCATTTACCGTATGGTC AAACCCCGCATTGTGCTGCTGATATTTGTCTCTGGCAAGGTCGT CCTGACAGGAGCGAAGGTTAGAGAGCAGATCTACGAGGCATTT GAAAACATCTATCCAATTTTAAAGAGCTTCAAAAAGTAATCCT GGGAGCTATGCCAGCTCTGAGCAACCCCGTACCTGTTCGTCAAT ACTTTAGTTCTTTATTAAGGCATTTAGTTTCGCGTCGTTTGCCTC TACAATCTATGTTAAAATCCTTTACTATCTTAAGATGCGTTGCA ATGCGGACAGTTAGATGCGCTGCGAGGGGCGATGGGTGCATGT GATTTTTCACTGATGAGTTCCATAGCTACAGTATCTAAATAATT AACTACACTTTGAGTATGCTCTCCAGAACTCATGATTTTTATAT TTCGAAATTTAGTAGACCTTTGTTAACTCTAATTTAACTCTAGA TGCGCGAGGAAAGATGTATCCGGATTTACTCAGACAAGCGTTT CATCGGTTGGCTCGTCTGTAATAGAGAAATATTCTCTAAGCCAA GTTCTGGTACCTATATATTCCTGTCTAGACCATTCTTTATGGAGC ATCATGTTATTGAGGTTCTTGTGCGATCATGTGTAGTCTTTTCGT GGAATCAGAATTTTATGCGGACTAAGTTATTTTGAACTTGTTTA TTGTACGTGCTGTTCAACACAAGCATTTCTTCATCACGACTCCT TATACGTTTCCTTTTTTCTGCTTAACAATATTTTTGCTTTAAATC AAGTAAACCTTTCCGTAGTTGAGATACGATTGATTTAAATAATT GTATGGGGATGGAAATTGGGGCACTTTGACATCATACGCTCAG TGTGTTTTTTTTTTTCTGCAAATTTGTTTTTTTTTTTTTGGTTCCT AAGGCATTGCTTGTTACGTTCTTACAACCTATCGTTAATATCTG TTGTCATTGTACTAATTAATTTCTTCCTGACGCGAATGGAGTTG GATAAATGTCACTTATGCCTCACGGCCACGCACTCATAACGAG ATCAGCGTTGTGTTCTGTTGTGCACGTTAAGTCTTGTCCGTAGA CTCGTGCAACAGAGCTTGCTTGTGATGGCCTTTACTTTTGCAGG AGACATTCTGTAGTTTCGCTCATTCAATGAAGAAACTACTGGTT CGTTTAAATTGTTTCATTTTTCGTATACCTAAATACTAGTCACAT ACCAGTAAATGATATGGGGTACTCCTCATTCACTGCAATCAACC ACAGTATCAGCCGTATGCACTTGCCATTCGTTCAGAAATGCTCA CTGGTGACCAATTCTCTCAAATATGCATTACCTTAGTATGTATT AATATGCATACTTCCATGTTTGTGCTTTCTCCAGCAATGGTGAG GGCGCACAGAAGGATGCAGAGCTCGGCTGTTTTATTGATGAGT AGCATGAAGCTTCAGAGATCTCTCCTCGGCATCCCACATAGCG GCGTATCCAGCCATACGGTTGTCTTCAGAGCCATTGATATCATA GTAGAAGTGCAAGTATCGTTAAGAGCATGCATTAGTGTAGTCT GGGACAAGATATTTC
Ubi	3168	GAAGATTTAAACTGTAATTTACACCATTTAATCGCAAAGATGTT CGGATCATCGTGTCCTCATTGCAAGTTGCCGTATGATGAGAAAA AACGGAAACCACTACTGCTGCCTTGTGGTCATGCATCTTGTCGC

		AACTGTCTATCCGAAATGGAAGACCGAGACGAAATAAGATGCG GAGAAAAGAAGTGCAGAAAATCTTGCACTGATAACGTCTCGGA TCTGCCCATGGTATACGCCCTAATTCCGGCGAGCTCGGGTGCAT CTGGACAAGGTTCTTCCGGCAGGACTGGGCGGGATGTAATTCC CTTGGAGTGCTCTTCATCAATGGCACCGATACCCACAGGGAGA GAAGTCGTTCCTATGGACATGGACGACACAGAATCCCTTGCTG GATCCTCCGTTTCGGCGGTGGTAATTCAGCCTCGTCTGCACACC CCACTTGCGGAGTGTGCATGGATGACTATTGCGACGCCAACCCT CCTCGGCTGCTGACATGTGGCCATACTTTCTGCACACTGTGCCT CCAAGATCTAGAGAAGAAAGATAATTACGCTTGCCCACAGTGC CGCAAAACTCGCAATAAAGGAAGTGTGCCTAAACTTCCTGTAG TCCAAACGCTGATCCACACAAAAAATGCTAAAAATAAAGTGTC CTCCAGGATCTCAACAGCTCAGGTCATGAATAGAGCCCAGGAG CTGCAGCAGGAACACACGCGACGCATGGCTGCGTCAATTGATT CGGACCCAAAAAATATGGAATTGTTTGTCCAGGACATCCGTGG GCAGACGCATACTTTCCGTGTTAACCCAAACGCCGAGACCGTCT TGCAATTGCGGCAGCGACTTCCTGGTCGTCCGGACGACTGTCGC TACCTGACGTTCGAGGGCCGTACTCTGGAAAACCAACGCTATCT CGCGGAATACAACGTTCGTCCGCAATGCACGATCCATGAACGG GGGCGTTTGTCGGGTGGCGAGTTCTAAATTGAGCATTTTGTTGT TGCAGCATAGTTCTAGACGGAAAGGTAGTATTGTTCCAGATGG AGATCCGACCCCTTGATGAAGTGCTTTTAGACTGGAAAATGTTC GAGTCATTATTGGGTAAGCGATGGCTTGAGAAGGAGCTAGGAG AAAAGTTACATGTTAAGTAAAGTTCTATGTCGAGGATGTTTTGC TCGGAGAAGAAGTCTGTGGGATAGCTGAAAGTGAATGGCATCC TATTAGCAAAAGTCAAAAGATACAGGCCGGATATAAAGAATAT GGAACAGATAATATTATATACACTGTTAGCGTTTACTCGAGGAT TTTCTTGTATTACTTTTTTCATTTATTAGGAGGAAATGAGACATT CTTTTACGCATTTCCATATGACCTGTAGCATAAATTATGATTTTA TTTCAAGGGGTTTGTGCCAAGTTTGTACAAGGCATTTTTGAAGT CAAGTTCAATGATTATTTAAAATTCAAAACTATTCTAGCTATTC TGGACGTGTTGATCCCACTGAATTGTTCAACTTGAAACGCATTG CTATAATAATCGCTTGAAGTAAAAATCTTTCATCTCAAGAGTTC TATTATTTGTCATCAAAGGTTGGTACGACCACATATAATTCACA ATTTTGATTGTATACCTGTCAGGATGAAAATGCATATCTTATGA TTTAAGCAGAGGAAAACATTTTTTAATTCGATATCGTACAGGAA AAGCGGACATTTTGATCTAAAAGAAGCAAACTAGTCTACAGTT TTCATTAAATTAAATGACACGCCTACAGGAAAAGGTAAAACTT TCTGTTAGAATTGACACTTTTTGATTAAAAATAATCAAGACCCT AAGCTCCACATGTGCAGTAATTTATGGAGTTGCTTCGTTTCGCT GACTTAGTCAGCCTATCCCGCACACCGACAATAGTATTGAGGA AAATAGAGGCAGAAAGACTCTGCGAGGTCGAAAACCATAACA CGTGATGAAATGTAAAATTTACAACGCATCTTTTCACATAAGAA CTTTTCACTTACATGCGTGCAATATAACAAAAAATATATATAAC TAGTGACTCGTGGGAAACGGTAACTGAATCATTGAATAAAGAT GGCTAAATAAACATTGGAACCTTATATTTAGTATGGCCAAAAT GTATTCTCTTACATCGATACCCTAGAAAAAATTAACAGGATCAG TTCCATCAAAGAATTTAGAGCATATTTTGTAAGAAAAAATAGA TGAGCAAAAAATTATTTAAAGCATGCAGTGAGAGATGAAAGTA AAATTCGGCGACTCAAAGGTTACACAGAAAGACTATTTTACCA TGACAACAGACGAGTGATTTTTTTCCCGTAAACATATCAAACAA ACGTCGAGATAAGATAGTACACCAAACACTCGAATAAGTAGGC GGATGACGCATAGCCGACTGAGAGCTAACAGAACAAGTTTACA TGTCAATCTGTGAGGCCATTTAGTGATAACACGCATGATATTTA TGTTTCAACAAGTGCTAACATTAACGAAGAAGAAAACTACTTT

		AAATTTCTTCCTAATTAATTTCTATATATTAACGTAATATTACGA CTAACTTCAACTAAAGATTAAAGAGTTTCGTTGCATAACAACA AACAACGGGTATTTTACGATAACAACTTCGACTGACACTACAT AATAATATAAACTCGGAAAACTTGATGAGTAAAACCTTAGATC TCAGACAACTACCCTAACTATTGCACGCAAATCTAATAGGTATA GTGGACGATTAAGGTAAAATAAGTACATACTCAGTCCATTCATT AAGGCTTAGAAGTAAGATAAACTTTTCACCACATCCAACACAT CAAATAAGTACTCTTTGTGAAACCAACTTCATTCTCAACTGAGT AACATCGAGCTTCCTCCATGCTTTGATAAACAAGCAACTGATAA ATTACTATAAATTAAAAACCATGAGGAACTTTACGGAGAAAAA GAAAAAATACCCACAAGAATGCTAATAAGTGAAACATATCAGT AAAGTATAAACATTACCACAGATTTGAAGTTTAAAGTTTCAAA GGATATAAGACGTAGAACTTGAATACCAAATCTATGAGTCAGA ACATCATAACACCATACGCTGAGATACGTGGTGGCAACAGGCA AGAAATATTAAATTCCACCTCAGGTCTAATTAAAG
Ubc	1567	TAATTGTTTGCAGGAATTGCAAGACTTGGGTCGTGACCCTCCAG CCCAGTGCTCAGCAGGTCCAGTAGGGGAGGACATGTTCCACTG GCAGGCCACCATCATGGGTCCAGCTGACAGTCCCTTCCAAGGG GGCGTTTACTTCCTCACCATCCACTTCCCCACAGACTACCCCTT CAAACCCCCCAAGGTGGCATTCACAACTCGCATATATCACCCC AACATAAACAGTAACGGTTCAATATGTCTGGATATTCTTCGTGG TCAATGGAGTCCTGCATTAACTATTTCTAAAGTATTGCTGTCGA TCTGCTCTCTGCTGACGGACCCGAACCCGGAAGATCCTCTGGTG CCTGAGATCGCGCGGATGTACAAGTCTGACCGTGAGAAGTACA ACGAATTGGCCAGAGAATGGACGCGCAAGTATGCCATGTGATC TTCCTAACGCGCTCTGGGGTGACTGGCCGCGCCGCTGCTGCTTG TGTGTCACTCCGCCCTCACAGCACCGCCACCACCACCACCACCA CGTCCTCCTCTTCTACCATTCTTGATCTGCTTCTCTTCATGCCTC TAAGTCCATTATAGTTATTTCTGTGGTCGCAGAGCTATCCTGTT GATGTCGTTGCGCAAAGAGCTGTCGTGGCAATGCCAGAGCTGC GCATTTGCACCCATTAGGTCTTTCCTTGCCTAACTGAATGCTAC GTCCGAAAATAATTTGGTATGGGCCGGGGATGTTAAGTGGTCC CAGAACTCGCTGGGTACGCTCAGGTTCTTATTGAATGCGCGTGA ATAAATTGTTTTTACCTCGTGTATTTTAATTGGTATAGGCTTGTC AGATTAGAATTTTCTTAAATTTGGGTGTGGGTGTAAAACTCTGC TCAATATTCAACTTCTCCCTCCGAGTCTTGAGATTTCACGAACT CCCACGAACCGCGCGTTTGAATGGTGTTCAGTGCTTCTGACCGC AACATTTGCTGTATTCCCAGCAATTCGCTGTCTCTGAACTTCGA ATATCGTTCTATGGACGCGTGTGTTACCGTTCCTAATATTTCCCC TCGCAGCTCCGTCAGCTAACATGATCCTAAGTCTGTACATGATG CTGAGAGCCTTCATTATTTGAGCGACTGGCCCTTGTTTGTTCCT ACGTCCATTTTAATGTGTGGCGGTAACTGATCGGAGGATTATAA TAGAGCACGAAGAAGCAGTTAACGTGGTGGTTGCTTGATCAGT TATTCTATTGTCACTCTCTTTATCGCCTGGGCCTCTAGACGCGTT GATCATCGCATATTGTATAGGATCCAATAATCTGCTTAAGAGTT AGTCATTATTAAATTGATCATATTTATGAATGATTTTGCCTGCTC ATCTGGGATTGACGGCATCGTTTAGTTCTAATTGATTTTTTTTGT ATTCTATAATGATTCACAATGCTCTCATACGAACTGTACATTTT TTTTATCCCTTTCAATTGTTGTATTGACTGCCAAATTCAGTTTCT AAAGCACGAAGATGGTTTATTATTTTTTTCAATTTAAAAAATGT TTTTCTTTGCGATTATACTGGAATAAAT

Appendix C: Housekeeping gene transcript amplicon sequences

Primer set	Amplicon length (nt)	Sequence
RpS23	127	CAATGCTGAGTCACGAATCGATTGCTAATGAAAACCAGTCTAGAAG TAGTAGTTACGGTGCCCCGTATGTTTGCGGGTAAAGCCTCTGTCTGA TTGAGCTATTTTTCAGAGGTATTGAGGCGAAACA
RpL10	132	GATTTTCGTTTCTGGGAAGTGGGGCTTCACGAAGTACGAACGCCCC GACTACGAGAAGATGCTAGCGTCAGGCCATATCAAGTCTGATGGTG TCAATGTCAAGTATATCCCTGAGCACGGTCCACTAATTAA
Act	140	GCCATCGCACAGGATATCTAAGTACACTGTCGACTTGAATGCCGTG TCAAGTTTGCTCTCAAAAGTTTTCTCCGAGCTCAAAGTTGACGCCCG GAACTACAACCTTCAGATATCTGCGCCCAGAATCCTCAATCAGCAC A
Act2	130	AAAGTTGACGCCCGGAACTACAACCTTCAGATATCTGCGCCCAGAA TCCTCAATCAGCACACACAGTCTGAGCTCATCAAGATCCTCTTCGA AAAGTTCGGGGTGAAGGGCGTGAACATCACCCACCAGT
Act3	126	TTCATCAAGCCTGTCGTTCCCCATCAGGCCATCGCACAGGATATCTA AGTACACTGTCGACTTGAATGCCGTGTCAAGTTTGCTCTCAAAAGTT TTCTCCGAGCTCAAAGTTGACGCCCGGAACTA
Gapdh	150	GGATTACCTTCGCAGTACAGTGGGCTTGAGCGTGATTCATAGCTAG TGATGGAAATATCTGTTAAACATATATTTTTCCGTTTTCTTGGGACC CCAAGAATTAATTGATTTCATCTTTCGGTCTAACAGACTGTCCTCTG TAAACTGCAA
αTub	149	CCTACCACGAGCAGATTTCAGTGGCCGAGATCACAAACGCTTGTTT TGAGCCCGCCAACCAGATGGTGAAGTGCGACCCGCGTCATGGAAA GTACATGGCGTGTTGCCTGCTTTTCCGAGGTGACGTCGTGCCCAAG GATGTCAACGCT
Mmp	146	TTGGTTAGGTACAGTTGTGGGTACAGTTGTGCGAGGAACTACGGGA GGTTGATTTGGCTCTCTTGTTGTTGTTGTTGTTGGAGTCGTGGA GGTGCTGGTTGTGGTTGTCGTATCTGGAGTAGTAAACTGAGGCCGA TAAGGAT
Syx	76	CAGCAAACCTAATGGCGATGTAGTTCTAGACTTCGATGGCGTAAGC GTTGCTCCAAGCAACAGTGGAGGGGCTTAC
Syx2	92	ACTTCGATGGCGTAAGCGTTGCTCCAAGCAACAGTGGAGGGGCTTA CCAGCAGCAGCTGGCGCTGATGGAGGAACAAGACTCGCTGCTGCA G
Syx3	91	CGCTGAGACCATGAAGACCATCGAGTCAACCATCGTCGAGCTCGGC ACGATGTTCACGCAGCTCGCTGCTATGGTCAAGGAGCAAGACGAA
Syx4	132	AGGAGAACTGCCAGCAAACCTAATGGCGATGTAGTTCTAGACTTCG ATGGCGTAAGCGTTGCTCCAAGCAACAGTGGAGGGGCTTACCAGCA GCAGCTGGCGCTGATGGAGGAACAAGACTCGCTGCTGCAG
Tbp	149	CCATGTTTGTGCTTTCTCCAGCAATGGTGAGGGCGCACAGAAGGAT GCAGAGCTCGGCTGTTTTATTGATGAGTAGCATGAAGCTTCAGAGA TCTCTCCTCGGCATCCCACATAGCGGCGTATCCAGCCATACGGTTGT CTTCAGAGCC
Ubi	148	CATGGTATACGCCCTAATTCCGGCGAGCTCGGGTGCATCTGGACAA GGTTCTTCCGGCAGGACTGGGCGGGATGTAATTCCCTTGGAGTGCT CTTCATCAATGGCACCGATACCCACAGGGAGAGAAGTCGTTCCTAT GGACATGGAC
Ubc	119	TGGGTGTGGGTGTAAAACTCTGCTCAATATTCAACTTCTCCCTCCGA GTCTTGAGATTTCACGAACTCCCACGAACCGCGCGTTTGAATGGTG

		TTCAGTGCTTCTGACCGCAACATTTG

Appendix D: Gene of interest transcript sequences

Transcript	Length (nt)	Sequence
Keap	1396	CCAAGTTTATTCAATTGGATCGAAACTTTCTACTGTATCGATGGG ACTTTATCCATCAGAGCCACCCACTGCCCACAGTTCTCCCATATG CTTGACAAAGCCATGATAACTTCTCGCTTGAGACATGTTGTTCAG TTCAGTTAGGGTCTGGGTTTTAACGTTCAAGAGATATAAATCCTT GCTGACATTAGGACTTTCTTCGCCAAGAAGCGAGCCTCCAAACGC CTTTCGTCCTCCGCACATATATACTTTATTTCCGAAAACAGCGAGT CCCAAAAATATTGTTGATTGGGGTAGCATTGCTTCCCATGTTAGC CTCTCACTTTTTTTGGCTCTAATATCTATTACTTGGATGGTGCGAC CTAAGCTATGAAAGTTTTCGTTCATGAGATTCAGGTTGACTCCTCC GACTTACAACATGAAATCGGGTCTCAGCTTCGTCCTGTTTACTTG AGACTTTGACGTGAACTTGGAGTGCAGTGAACGCTGGAGCTTCGT AATGGCCAGGGACCGAGTGACATCTGTTTCGATCGCAGACTGCCT TCTTAGAGCGTTCCTTATTGCCAATATTGAGTCGTTTTCTATTTAA GCCTCCGCTATCTTTTCTTGGAGCAGTTCCGCAGCGTAGTTTTCAA TATTTTCCAGGTTTCCATCAATTTCTTCCTCTGATTCCTGTTTCCAA AGTCCCCAAGCAAGAAGTAGCTCGTAGATAATGAGATGGAGCCC ATAATTTTGATGCGCAACAATTTAGTTCATTACGTATTCATCGAA GTATTCTTTATCCTCCGGGTAATCGAAGCTATCAGTTCGCACACA AGCCAGAATGCTATCTAAGACTTGTGTTTGGTAGGACTAGTGAAT TTCGTTAATTTTCTACAAAATTTTCTCCCGGTTCTTCCACTTCGCC AACGTCCCAAAAGATTCTTGTTCTGCTTCATCCTCTGTCGATATCT AAAGAATGTTAGTCGTCGGATGATCGTTCCTAGGCGGGATCTATA ATTTCCAATCATTTGCTAATGATTTTGAATGAACCTTGGGCCGATT TTGCATATTTTCTTCAGCTGACCGAATGTTTTCTCGGTTAAAAAGT TTATCTTTTTCTGACTGGTCGACTGCAGGACATTTCAGGAAATGCT TGAAGGCCGCAGTTACCGTAGAGGTCCTCATCTGCAGGAACAGGT TTCCCGTTGAAGCAGCCTCGCCCACGGATTCCATTACTATTGACTT TGTTGCTTGTTTGCAATTGCGAAGTTCTTTGAGAGCGAATATTAA ATTTTGCTTAATAAAATGGAAAGCTAAAAGCTTTTAATCGCCAAC AGGGGATGTCGTCCGGGATGGAAGTAAAAATCCTCCTGAAGAAC TGATTTCCTTTTTTGAAACGCTGAAAGAATTTATC
Cnc	2128	GTCGTATAGGGTTGAGCTCTCTCTATTTTCCGTGTATCGATGTACA TTAGTGTTGTTATTTCACCCTAAATCATTTCAAAAGTCTAGCTGCC ATTAATTTCAATTAAAAATGGGTGGATCAGATAATACAGAGCCTT CCGGAGCCAAGGGTGCTGGCACCTTGAGAAAACAGTTCAGCTTTC TTGCGATTTATAGCAAACTTATTCCAGCTCTATCTGAAGTATCATC TTCACTGGAGCACATCCATCATAACCACACATACCACATGTCACC TGAGGGGCCATCTGGACTGCCTCGGCCATCACAGAGAGACCCTA ATAAAAGTAAAGCCCGTAAGCCTGAACCCGAACGTGCCTTAACC CGTGATGAAAAACGAGCTCGAGCTCTGAACCTGCCCATCTCATGT GATGACATAATCCACTTACCAATGGATGAGTTCAATGAAAGGATC TCTAAGTATGACCTCACCGAGAATCAGCTCTCACTCATCAGAGAT ATCAGAAGAAGAGGGAAGAATAAGGTGGCTGCTCAGAATTGCCG TAAACGTAAATTGGATCAGATCATGCATTTAGCTGAGGAAGTTAA AGTGATACAGAGTCGCAAGAATGAACTCATATCCCAGTATGAGTT TCTAAGTGGGGAGAGACTTCGAGTCAAGCATAAATTCTCTCTTCT TTATAGACACATATTCCAGCACCTGCGAGATAGCGAGGGTCAGCC ATATTCGCCACACGACTATAACCTGCAACAGTCAGCTGACGGCAG TGTACTCCTGGTGCCCAAGTCCGCTCCTCACCTGCCCCTGGACCCT

		GGTCTGCCTCCTATTCCAACCAATTCCAGTAAACAGGATGACGAC CCTAGCAAAGGCCGGCCATCTCATCAGCAACCTCCTCACCAACAT CACAATCCTCATTGACGATATCATTAATGCTCTTGCTAACAACATT AGAATTAATATACTTACGAGTTGTGCTATTTTTTAAAAGACTTCA ACATTATTCGTTTTGTTGGTCAGTTTAGTCAAGTATTTCCAATTAG ATCGTCCAATGTTGCAATTCCGACGGAGTGGTGTCATTTTCCTCCT GAGACCATTTGGCTTTATTCCAGTAGCTCTCTGGCTTTATATAAGG TGGTATTATGTTTATTATTTTTGTGCATTGAGTTCTGATTAAGTAA GGCTGTATCAGCCCGATGCTCGTGAGGTACGTAAGAAACTTTCAA GTTCATTTTCGCGCCTTTTACATTGCATCAGCTCTGCGGCCAATTG TTAACTTCTTGTTACAGTTATTAGTTATTTAAAAGTGTGCATGACA ACTTCATCTGTATAGTGACCTGGTACCATTACAATTGCGAGTCTTT ATGTTTTAGAATAATGACGAAAAGAAGCTACGCAGCTTCTCGATA TTGTACATTTAATCATATGTATGTTTCTTGCAGCATGTCGGAACAT GTGTTTCTTCGTGTCGGCGAACTGTATCTTTATGACAAGGTGTTCC TGACTATCAGAATTGGTTTTCTTTAGCAATTCTAAAGCGAGTTCA AATATCAGGAGCATCATATTGATCCACTTATGACTTATATAATAC AAATGCTATCTGTGGAGGTGGCTGTGCATAGTGTAAATATTGAAG TAGTTTCTTCCTCTAGACTGTTACTTGTATATCTTGGACAAGTTGG ATTATGTGAACGTGCCTTCATGCGAGGTGATTGCCCGCTGCTCTT GATGGTCTCTTCGTGGAATATTAGCTTAATTTTCCTTGAAAATGTT CTCCCTATATGATATCTGACACTCAGCTTAAGAAAATAACTTGGA ATTGGAAACCCGCTGGTGTCTAGCATTTTTTCAAAATTATTTTAAA ATAATTTCTCAACTTTCCCAATGATGTTATATGGACTTTTCTCCAT TTTCAGATAACGAAAAATTGAGTATTTGATGTATTAAGTACTACA ATTTTTGTAACATTTTTTGTCATTGTTCGAAATCTGTGATAAATAG GAAGTTTTGTGCTCAATGGTTCAGTTTGATCTATCAGCTTTACTCT TTACTTCGTGCCTGATTGCTCATTTTCGTGTCTTTTTTTTTATGTTC GAGTCTCACGCATTAACACAATTAAAAAAAAAAAACTTC
Mrp4	3234	TTCGACACGCATCCAATCGGACAAGTTTTGAACAGGTTTAGCAAA GACCTGGGCCAGATAGACGATCTCTTGCCCTCTTGTTTTTGGGATT TCTTTGAAATCTCGCTCAATATTGTGGGCATCATAACTGTCATTGC CTCTGTGAACCCTTACCTGCTCATACCCACCTTCATTCTGGCCGGC ATATTCTGGTACATCAGAAGGTTCTACCTTGGCTCTGCCAGAGAC ATCAAACGGCTTGAGGGCATCACCCGAAGTCCAGTTTTCTCGCAC CTGAGCACCTCGCTGCACGGCCTCACCACCATCAGGGCCTTCAAA GCACAGGAATCGTTCGCCAAGGACTTCGATGACATTCAAGACGTA CATTCCGCAGCCTGGTTCTTGTTTATCTGCACGACGCGCTGGTTTG GCATCTACTTGGATTGCTTATCGTGCCTTTACATCGCCATCGTTAC GTACAGCTTCTTGGGCAACAGCCATTCTCTGGGAGGAGACGTCGG GCTGGCGATATCGTCAGCCATGAGCTTGTCAGGAATGTTTCAGTG GGGAGTGCGTCAGTCAGCCGAGGTTGAAAACCAGATGACGTCAG TCGAGCGCGTGCTGGAGTTCAGTAAACTGGAACCAGAGGCGCCA CTCACCACTGAAGAAGACAAGAAGCTGGACCCGCAGTGGCCGCG TTCTGGCAGCATTGAGTTCATTGACGTGAGCCTTAAGTACGACGA AGCCAAACCTCCCGTCTTGAAGAAGCTCAACTTCGCTGTGTCGTC TGGGGAGAAGATCGGCATCGTGGGACGTACAGGAGCCGGAAAGT CTTCCCTCATTTCGTGCCTCTTCCGCCTGACGGAGCCGACCGGCCA AATCATCATCGACGGCGCTGACGTCAGCAAGTTGGGCCTCCACAC GCTACGTAAAAACATCTCTATCATTCCCCAGGATCCAACTCTTTTC AATGATACATTCCGCAAAAATCTGGATCCATTCTCTCAGCATTCA GACGAAGAACTTTGGCTTGCGCTCGAAGAAGTGCAGCTGAAGGA AGCAGTGCAAGAGACCCCCGGGGGGCTGGAGCACGTGGTGTCGG AGGGCGGCAGTAACCTGAGCGTGGGGCAGCGTCAGCTCGTGTGC

		CTCGCCAGGGCCATCATCTCGCATAACCCCATCCTGCTCATGGAT GAGGCCACCGCCAACGTCGACCCCAAGACGGACGAGCTCATCCA GACGACCATCCGAGACAAGTTCCGGTCGTGTACGGTGCTCACGGT CGCCCACAGACTCCACACCATCATGGACAGCACCAGAGTCATGGT GCTCAGCGCCGGCCAGCTCAAAGAGTTCGACGCCCCACACTTGCT ACTCAAGGACAAGAAATCTGTGTTCTCTGAGCTGGTGAGCCAGGC AGGTACATCCACTGCATCACACCTAAGGAACCTGGCGTTAACGGC CTTCTCGATGGCCAGCGCCGACGATGGCAAGAGATCGAGCAGCA GCAGCACCAGCCCAACAGGCCGTAGCCCTGTTAGCGTCGCCAGTG ATCTTAGTCCTGTAAGCTTAATAACTGCAAAAGATGTGGACATGA CCCATATTAAAACCACGTCAGAAAGAGGAAGCGATAAAGACTCA GAGGCCATAGCGCCATCTAGCCCAAGAATTGACCCAGAAGTTGC GTCTGCCAGTCACGGGTCTATCGACGAGCTTACTAAATTGTAGAA ACATTTAAGGCCGTACAAATATCGGGGTCCATTCATCTTTGGTGT AATGATTGTTGATATTTTTCTTACCATGATAAATTCCCGTCCATTT AATGGCTGGTATTTTTAATAGTTAATCGCTTTTACTTCTTATTACT TTGATACTGACGATGTGAAAATGAAGGATGAGAAGTCGGCAATT ATATTGCAACACCTTTGTGAATCTTATTGGATGTTAAAGTTTCAGT TTTACGTTAAACTAAGTCTCCATTACATTTTACCGTAAAATAAAG GAAATTTTACCATATGCCTGATTCACTGGTTTTGGTACCAGTGGA CGTTGAAGACTTCAGGGACGTGATCAGTTTCGCGACGTGTTAATT ATGTGGTCAAATTGAGACATGATGGAATTCCCTGCGTAGCGCTTA TTATTGCTATAATTTTGAGTTGGAAAATGCCTACTTTCCTACATGA GTCAGCTTGCTGTCTTACGCGTTGAGCATTCTACGTCAAGCTTTCT ACTTTTTGCGATTCTCTAATAATCTTGTCAAGTAAAATAACTATTC ATCACTATGAATTGTATTAGGTAATTCATAGGTGCCAATGTTTCG GAATTGGACCTGAGAATAGCCTCGCTTACGATTATTTGTAAAAAT TTCTTGATACAATTAAGAATCGGTTTAAAATACAAATATGTGCTA GTGGACCTGGATGAAACCTAGCGCCTTCATATTTGATTAAAGGAA CGAATATATTGTTGAACATAAATTTCGTCCTCATTTTCCATATAAA TTGATATCACAAAGGACGAAGAAATAAGTATGATTAGGTTAATAT GAAACTTTAATGCAGCAATCATTTATGTCAGCCCGACCTGAATAT TTACAACAACAATTTAACAGTTGCGACTATACTTGTCATTGCTCG GGAAAGAAAGCGAATAGCGCCATTCACGAGACAATTCGCTGCTC AAGGCTTTGGGTTCTTCCCGACACGATTACTCTACCTTCAGAGAA CTATGTAATGCTAAAAGCTCTTCAACTATTTCCCTTAACAATGAG CTTAATAATATATAGCTTTACATCTTAAATCTCTGCATACATGCAA CTGCAGCGTTACGAATGCATGAAGACTTGCCAGGGATCTAGTCGT CGAGTTTATCAAGTCTCTCCTTGTGTCGCTTGAGGGCCTCTTGGTG GGCTTTGATCTGCTCTTCGTGGAACTTCACCTCCTTCTCCAGACCG TCCTTCATGTCCTTCAGCTGCTTGGCTTGCTGAACGAAACAAAGC AAATTAGATTCGACTAAATACGCGATCTTCCTGAGCATCTGATTT GGTTATTAGTACGCCCGTGCTGAATTTTCAGTGACCGGTAGAATC GATAGAAAATCTTCAACAGAATGATATACTTATCACTAACTGTAT CAAGTAGCATCAATGATGGCATCGTTTCCATCATCATGTGATCAT GATTTCAAGAGACACAAGTCTCCTTACAAAAAAACGGTAGTCATG AAACACTTTTTTCAATCACAATTTTGAGAAAAAAC
Mrp5	909	CAGCACATGGGATATCAACAGCAACAACACGAGTCCTGTTGGCA TCAACGTGACATTCGATAGCAGCAACAACATCGCAGACAACCCA CACCTCAACTTCTACCTGAGCGTCTACGGCAGCACCGTGATCTTC ATCCTAGCGACGTCGCTGCTGCGTGGAGTGATCTTCATGAAGACC AGCATCCACGCGTCCACGCGCCTTCACGACCAAGTCCTCAACGTC GTCTTCCGAAGTCCCTTAACCTTTTTCGACACGACTCCCGTGGGTC GCATCCTAAACATTTTCTCAAGGGACATGGACGAAGTGGATGTTC

		GACTACCTATGAGCCTGGAGATGTTCCTGCAGAACATTTGGTTGG TCAGCTGTTCGCTGCTGCTTGTATGCCTCGTCTTCTGGCAGTTCCT GCCAATGTTAGTGATCCTCGCCATCATCTTCATGGTCATCAGAAA CATCTTTAGAATCGGCATAAGGGACTTCAAACGCCTTGAGAACGT CTCCCGGTCCCCGCTCTACAGCCACGTTAGTACGACACTTTCAGG ACTTGACACCATCCACGCCTACGGTAAACAGGACCTCTTCTCTAA CAAGTTCTACTACCTCTTCGATGAAAGTTCGACGGCTTTTTACCTT TTCAACTGCGCCATGCGGTGGCTGGCAATAAGACTGGACATGTTG GCCCTCGGGGTGACCACATGCACCGCTATTCTCACTGTCCTCCTG CGTGGGCAAGTGGCGGCTGCGTTCGCGGGTCTTGCACTGGCTTAC TCCTCCCAGCTCTCCGGTATATTCCAGTACACTGTACGACTCAGC ACAGAAACCGAGGCCAGATTCACTTCAGTTGAAAGAATTCAGTCT TACATAGAAACCCTTGAGAGTGAAGGGCAGCACTCGGAGGTGAT ACCTCCGGC
Sod	1246	CTGAAGGAAGGCCTCCGAAGTACGTCTGACATCATCCAGCTGGTG CTCTACCCATCTGACAATGAGAGCAACAGCCCAAAAGAAGCAAG CGTCCTTCTTATGGGAAGCATCCGCGGATTCTTAAAACTTCGTCA GGACCAACCACCTGTTGGTAAGACCCATATCACGGGGCACATCAC GGGTCTGACACCCGGCCCCCACGGGTTCCATATCCACCAGACAGG CAACCTCACCGAGGGCTGCAAGTCAACCGGGGGCCACTATAACC CTTTCGAGCGTGTGCACGGTGCTCCCCATGACCACGTTCGCCATG TTGGTGACTTGGGTAACATCGTGGCAAATGAGAGAGGCGAGGCT CTTGTGGACATCACGGACCACCAGGTGACTTTGGTGGGACCCAAC AGCGTCGTGGGACGAGCCTTCGTCGTGCACGAGGGTCGCGACGA CCTGGGACAAGGAGGACACGAGACCAGCAAGACCACCGGCAACG CTGGCGGCAGAGTTGGATGCGGAGTCATTGCTAGAGCTTGAACTT CAGACGTCGAGGATTTAAAAACTATTGACGATAATTTTGAATTAG GTATATTTTTTGACGACAAAAGGTTTAGGTTTTGACGACAGTATT GGGCCACTCATTCTGGCATAGAATCGAGAATGACCTAATATAAAA TAATATGAAATGACTTGAGAAAGAAACAACAATCAATAGTTGAA TTGATAATGCCGTGTTCGACTGTTGGCTCCATAAGACTAACCGAA GCAAGTTACCAACCAATGCGACCAGTTGATTTTTAGCACAGCTTG TTGTTGATTAGATTCGAAACATCAAAGTAATTCCATGGTTTGGTA TTATGCAAGTCACAGTGGCATCAGTGTTATCTGCTCCACTCAACTT AGCCGTTTCTCTCATCAAATGAAGTTGATATTTGATTAAAAATTAT TTGAGATATGTTCTTCTTTAAGACTCAGTAGCAAGTTTGTGAGAA GATAAATAAAATTATAAGTGTCTAAAATCCTTGAGTGAGATGGTG CGATCGTTGAGTGAAAAATTTACCTCGCTGTGCAATGGACTGACT GAAATGAAAGCTCATATGAAAAATACAGGGAAAAAATGGACGTA AGAAACTAGAGTTGATTTATCCGAAGTCCCCGCTTCGAGCGAAAA TTCGATCACGGGACGCCGATATTGCGCTATCGTCACTGACTGCGC ATGCAGCATCGTACGCAGTTCAAGGTCGCTTTTCTG
Hsp70	1505	CTCACCTTGGAGGGGAAGATTTTGACAATCGCCTCGTCAATCACT TCGTTGAAGAATTTAAGAGGAAATATAAGAAAGATATCAAAGAT AATAAGCGAACTTTGCGACGACTGAGAACTGCGTGTGAACGAGC CAAGAGGACCCTATCCTCATCTGCACAAGCTAGCATTGAGATTGA TTCGCTTTACGAGGGCATCGACTTCTACACATCCATCACTCGCGC CCGATTTGAAGAAATGTGCTCTGACCTCTTCAGAGGAACCCTGGA TCCTGTGGAGAAGTCGCTTCGAGATGCCAAAATGGATAAAGGAC ATATCCACGAGATTGTATTGGTTGGAGGTTCAACTCGTATTCCTA AAATTCAAAAACTACTACAAGACTTCTTCAACGGCAAGGAGTTGA ACAAATCAATCAACCCTGACGAAGCTGTTGCCTATGGTGCTGCAG TTCAGGCTGCCATTCTAACCGGTGACAAATCTGAAGCCGTTCAAG ACCTTCTCCTCCTCGATGTCGCCCCACTATCCATGGGCATTGAGAC

		AGCTGGTGGAGTTATGACTGCCCTCATCAAGCGGAACACTACCAT TCCCACTAAACACTCGCAGATCTTCACCACCTACTCCGACAACCA GCCCGGTGTGCTCATCCAGGTTTACGAAGGCGAGCGCGCCATGAC AAAAGACAACAACCTCCTCGGCAAGTTCGAGCTCACAGGCATCC CTCCTGCTCCTCGGGGCGTGCCTCAGATTGAGGTCACGTTTGATA TCGACGCCAATGGCATCTTGAACGTGTCTGCTGTCGACAAGTCCA CCGGCAAGGCAAACAAGATTACCATCACCAACGACAAAGGTCGC CTCAGCAAGGAGGACATCGAGAAGATGGTACAGGATGCCGAAAA GTACAAGGCAGAGGATGATAAGCAAAAGGAGAAGATTGCCGCCA AGAACAGTCTTGAGTCATACTGCTTCAACATGAAATCTACTCTCG AAGATGATAAACTCAAAGACAAGGTGCCAGAAGAGGATCGCAAA AAGGCTCTAGATGCTTGCAGTGAAGCCATCAAGTGGCTGGATGCA AATCAGCTCGCAGATAAGGAAGAGTTTGAGTTCAAACAGAAAGA GGTCGAGAAGATTTGTTCTCCTGTAATTACCAAGCTCTACGGAGG TGCAGGTGGAATGCCCGGTGGCATGCCAGGTGGTATGCCAGGAG CGGCATCTGGTGCCGGTGGTCCAACGATAGAAGAGGTTGATTAAT AAGTGACTTTTCTGTCATTCGTTGTGCCTTTATTGATTGTAGAAAT TTTGTTTTGTTCTGAGTAAAAGCGTGTGTCTCTGTTTTATATGAGC TAATTACCATGTGTAATGAGCATGTGGGACGCACGAGATTTATTT CTTAAGTGTGCCTCTTGAAGTGATGTACATTACTATACTAGTCGTG AGAATCTGTGCGGGCATTTATTTTCTTTACCTTGTTCTTAACTTTA GTTGTAAGGTTTAGTTCTAATTTCT
Hsp704	1283	CTTCAAGTCATCCCTCAATCGAGCTCGTTTGATGGAATTGTGTTCT GACCTCTTCAAGAAAACCATGGAGCCGGTTGATGCAGCTTTGACG TCGGCAAAAATGACTATTGATGACATCGACGAGATTGTTCTAGCT GGCGGCTCTACCAGAATTCCGAAAATACAGGAGCTTCTAAAGGA AAAGTTTAATAACCGCAAGCTCAATCAGTCGATCAATCCTGATGA AGCTGTTGCTTACGGAGCTGCTCTTCACGCAGCTTCTAAAACTGG TGACAAGACAATTGAGAACATAAAGTTCGTAGATGTCACCCCAAT GACACTTGGGATAAATGTTCTTGGAAACTTGATGAGCCGAATTAT TCCACGCAACTCCAAACTTCCTGCGACTAGGAAGAAGAAGTTCGT CACCACTGTTGACTATCAGACAGGTGTGGATATCAAAGTGTACGA AGGAGAACGACTCAGTAGTGACGAGAACAACTTGCTGGGTGAGT TCTCTTTGGAAAATATCCAACCAGCTAAGAAAGGTGTCCCACAAA TAGAGGTAACTTTCGTTGTCGATAAGAATGGAATCTTGACTGTGA CTGCTGAGGACACAGCCACCGGCAGCAACGAAGAGATCGTTATC CGCTGCAACAAAGGCCGACTGCCGAAGCATGAAGTCGAGCGAAT GATCACCGAAGCGAAAGAATTCCAGAAAGATGACGAAAAGAGG GAGAAGCTTACAGCTGCTCGACATGAACTGGAGACTCTCTGTGTT GAGTTAAAGGACTTGATTAGCGAGCAGAGGGACCAAGATCTGCT GGAGAACGATGAGTTGGACCAATTCATGGACGTTTGCTCGGAAA CTGAAGAGTGGCTTGCCGAAAATGAACAGGCAACGGAAGAAGAG TACCGTTCCCGTTACGATGCTCTTCTTGAAGTCAAAGAGCGCATG CTCAGTAACTAAATTTTTAGCCATAGCTTGGTTTTCAATGCCTTCA TTTTCAATATGAAATATCAACAAAAAACTACAGTCATTTGTCGTG TGTTTTAGAATTCTTTGGTTATGTTTTTCATTATTGTATTCATGACT TGAATTGCAATAACATTCTGCATTTGTGATTAGTGCTAATGCCTCT TGCTTGGAACTCTAATTGGGACTTGAGTTTATACTAGTGCGTAAT AAAAAAGGCTCCTTGATACCAGGTATACTGGGTTTTGGGTGCTTT TAAATTTGATCCTCCTATCAAAATGTTGCGAGTCTCAATTGCCATT GAATGATTCTAATTCCAATATTCAGA
Hsp90	2606	TCAGACTTGAGCTAGCAATAACAGTGAACAGTGCAGTCAGGCAA GACGCTGAATCCGCGATATCCAGGTTTGCTTCACAGCCACCAGCT TGCACCTTCTATAACACTGCCGCTATTTCAAGATGCCCGAAGAAA

		CTGTAATGGAGACGGAGGGCGTCGAGACGTTCGCCTTCCAGGCTG AGATCGCCCAGCTGATGTCCCTCATCATCAACACATTCTACAGCA ACAAGGAAATCTTCCTGCGTGAATTAATCTCAAACTCGTCTGATG CTCTGGACAAGATCAGATATGAATCCCTTACAGATCCTAGCAAGC TGGACTCTGGAAAAGATCTTTTCATCAAACTCGAGCCTAACAAGA ACGACCGGACTTTAACTATTCGGGACAGTGGCATTGGTATGACGA AGGCTGACCTCATCAACAACCTGGGCACCATTGCCAAGTCCGGCA CCAAGGCCTTCATGGAGGCCCTCAACGCTGGCGCCGACATCTCTA TGATCGGTCAGTTTGGTGTGGGCTTCTACTCCGCCTACCTCATCGC TGACAAGGTCACCGTCGTCTCTAAGCATAACGACGATGAGCAATA CATCTGGGAGTCTTCTGCTGGAGGCTCGTTCACAGTGAAGGCAGA CAGCGGCGAGCCTATTGGTCGCGGCACCAAGATTATCCTGCACCT TAAGGAGGATCAGACTGAATACCTCGAGGAGCGTCGTCTCAAGG AAATAGTCAAGAAGCACTCGCAGTTTATTGGTTATCCTATCCGAC TGTTGGTCGAGAAGGAGCGCGACAAGGAAGTGTCTGATGATGAG GAGGAAGAAGAGAAGAAAACGGACGAAGATAAAGAGAAGGACG ATGACGAGGAAAAGCCAGAAGTAGAAGATGTGGGCGAAGATGA AGACGCAGATAAGAAAGACAAGAAGAAAAAGAAGAAAATCAAG GAAAAATACACGGATGAGGAAGAACTCAACAAGACGAAGCCCCT GTGGACTCGCAATCCCGATGATATATCTAAGGAAGAATACGGAG AATTTTACAAGAGCCTCACCAATGACTGGGAAGATCATCTGGCCG TAAAGCACTTCAGTGTGGAGGGTCAGCTAGAGTTCCGTGCTCTGC TTTTCGTGCCCAAACGAGCTCCGTTTGATCTATTCGAAAACAAGA AGCAGAAGAACAAGATTAAGCTGTACGTGCGACGTGTCTTCATCA TGGAGAACTGCGAAGATCTTATTCCCGAGTATCTGAACTTCATCA ACGGTGTAGTTGACTCAGAAGACCTGCCTCTTAACATTTCTCGTG AGATGCTGCAACAGAACAAGATCTTGAAGGTAATTCGAAAGAAT ATTGTAAAGAAAGTTTTGGAACTTCTGGACGAGCTCACCGAAGAT AAGGAAAGCTTCAAGACGTTCTACACGCACTTCTCGAAGAATCTG AAGCTTGGCATCCACGAGGACTCGACCAATCGCAAGAAATTGGC AGACTTCCTGCGCTATCATACGTCTGCCAGCGGTGATGACATGTG TAGCTTGAAGGACTACGTGTCCCGAATGCCTGAAAAACAGAAGC ACATCTATTACATCACCGGCGAAAGCAAGGAGAGCGTTGCGAAC TCAGCGTTCGTGGAGCGCGTGAAGAAGCGCGGTCTTGAGGTGATC TACATGGTGGATCCCATCGATGAGTACTGCGTGCAGCAGCTGAAA GAGTACGACGGCAAGCAGCTGGTGTCGGTGACCAAGGAGGGTCT TGAGCTTCCTGAAGATGAGGAAGAAAAGAAGGCATTCGAAGAAA AGAAGAGCAAGCTTGAGAACCTGTGCAAGGTAATGAAGGACGTA CTGGACAAGAAAGTAGAGAAAGTGGTTGTGTCGAACCGTCTGGT GTCATCACCCTGTTGCATCGTTACTTCTCAGTACGGTTGGACGGCT AACATGGAGCGTATCATGAAGGCACAAGCCCTGCGAGACACGTC GACAATGGGCTACATGGCAGCTAAGAAGCACCTCGAAATTAATC CAGACCACAGCATCGTCGAAGCTCTGCGCGTCAAGGTAGACGCT GACAAAAATGACAAGTCTGTCAAGGATCTCGTAATGCTGCTGTTC GAGACCTCGCTTCTGTCCTCTGGCTTCTCACTTGAGGATCCCACCG TGCACGCTTCTCGCATATACAGGATGATAAAGCTCGGCTTGGGAG TCGAGGATGAAGAAGAATCAGGAGACGGCGAGGCCCCAATGAGC GAAGAGATGCCTCCCTTGGAGGGCGACGAGCCTGATGATTCTCGA ATGGAGGAAGTAGATTAAACTTGGCGATTTTCGTTACAGTTGTTC AATTTTGGCATTGTTCATAGTTCATCGGCTCATTGTCATTTATTTT GGCAATTGTTAATTGTTCATTTCATTTAGTGTTAGTTGAGCATTGT CCTATTGCGCGAAGTAGGTTTAAGACTGAATCTAGTGACTTGGTT AAGGTTCGTCTTATCTAGTTATAGTTGATAGTTCTTCGTTTCGTGG CTGTTAATACAATAGCTCGATGATAGTCATGTATTCCATTCTTTTT TAGTTGGAAGAATCTTGTAGTGAAAGGAAAGTTATTTTATTTTTTC

		ATAAATTGGAA
Fer	914	GATCTAGCTACTGAAATTTTATGATGTTAGAGATGCAATATTTCT GAATAAATATGACTTTTACAGCCCTGGAGTCATTATTATGAGCGA GAAGATGTTGCTCTGCATGGACTGGCCAAGTTCTTCCGGAAGAAC AGTGATGAAGAAAGAGAGCACGCTCAGAAATTCATGAAGTATCA AAATAGTCGTGGTGGTCGCATCGTACTGCAAGCGATTGCTGCCCC CTCGTTGCAGGAGTGGGGCTCAGCCTTGGACGGCCTGCAAGCGGC CCTTGACCTCGAGAAACAAGTCAACCAGTCACTGCTTGACATTCA CGTGCTGGCGAGCACCCACTCAGATCCTCACCTCACCAACTTCCT CGAGGGAGAGTTCCTAGAAGAACAGGTTGAAGCTATCAAGGAGC TCGCGGATATGATCACACGCCTCAACCGCGCCGGCCCCACAGGCC TTGGAGAATACATCTTCGACAAGGAGCTTCGCGACTGATATCTCT ACTGAATACTTCCTTTCTAAAATCCTGGCCTCTATATACTTGCGGT TATCACAAGCCTCTCCTTAGGTGCAAGTGAGACGAGGAATGGATC TGTGGCTCGTTCGACCAGCATCTGAATGGAAGTAATGTTGCATAA ACGAACCTTCCCAAAATTTACATCTTATCCTCGGAATGCAGATTTT TATTTTAGTTTACGTTGGATCTGAGAAGTTGCCTATATATCCTGTT AAATTTTTCAGAAGATTTGGTTTTTCATTCCTTTGAAAGGACTTGA ATCTATTAATTCATTTCAAGCTAGTCAATGACGTCGTACAGACAC TTGACACAAGTTGCAGCACCGTTCCTTCGACCTAAGATTAGATGG GTTCGCGTAATTGTGTTATAAGTTTTCTACTCTTGTGTTCGTGAAT ATACTAAATAC
Cat	2993	TGCCGACTGGTTGCCAAAACAGAAGGTTGTGTAGTTGCGCCGCTT ATAATTTTTGCTCGTTATCAACATCACCATGTCGCGAGATATTGCT GCCGATCAGCTCAAGCTATTCAAGAGCTCTGCAACTGAAAAGGA ATCATGCCCCGGATCGCTCACGACGAGCTCCGGCTGCCCAGTGGG CGACAAGCTCAACTCCCTGAGCATCGGCCCGCGCGGTCCTCTGCT GCTGGAGGACCTGCAGTTCTTGGACGAGATGGCCCACTTTGATCG GGAGCGCATCCCCGAGCGCGTGGTGCACGCCAAGGGCGGAGGTG CGTTCGGAACGCTGGAGATCACGCACGACATAACCAAGTACTGC AAGGCTGCCGTGTTCAGTGAAGTGGGCAAGAAGACGCCCCTCGC CATCAGGTTCTCCACCGTTGGCGGTGAGAGCGGATCAGCCGACAC TGCCAGGGATCCTCGTGGATTTGCGATAAAATTCTACACTGAGCA AGGTAACTGGGACCTCGTGGGGAACAACACTCCGATCTTCTTCAT TCGGGATCCTGTGCTCTTCCCTTCCTTCATCCACACGCAGAAGAG AAACCCTGCCACGCACCTTAAGGACCCTGATATGATGTGGGACTT CATCAGCCTGCGACCAGAGACCACTCACCAGGTGTGCTTCTTGTT CTCTGACCGCGGCACCCCCGATGGCTTCCGCTTCATGAACGGCTA CGGCTCACACACCTTCAAACTCGTCAACAAAGAGGGCCAGCCCGT CTACTGCAAGTTCCATTTCAAGACCGATCAGGGCATCAAGGAACT TGTCGGCGGAGCGCGCGACCGAGCTGGCAGGCACCGACCCCGAC TACTCCATCCGTGACCTGTACAACTCCATCGCCGCGGGCCAGTAT CCCTCATGGACTATGTACATCCAGGTGATGACATACGAACAAGCT GAAAAGTACCGCTACAATCCGTTCGACGTCACCAAAGTTTGGCCG CACAAGGAATTCCCTCTCATTCCGGTTGGAAAAGTTAAGCTGGAT CGTAACCCCAAGAACTATTTCGCCGAGGTGGAGCAGATCGCGTTC TCGCCGGCTCACATGGTGCCCGGTATTGAAGCGTCGCCAGACAAG ATGCTGCAGGGCCGCCTCTTCTCGTACACCGACACTCACCGCCAC CGCCTCGGGCCAAACTACCAACAGATACCCGTCAACTGCCCCTAC CGTACCAAGGCTGCTAACTACCAGCGAGATGGGCTCATGTGTGTC ACTGACAACCAGGCTGGGGCACCAAACTACTACCCCAACAGCTTC TCTGGCCCCGCTGACTGCAAGAGCTACAACGGCTGTAAAGCAAA GTACTCTGGTGATATCGAACGCTTCAATTCTTCGGATGAAGATAA CTTTTCCCAGTGCACTGACTTCTACAGAAACGTTTTATCTGAAGCC

		GAGCGCGGCCGCCTAGCCGACAACATAGCAGGCCACTTGGTTTTC GCCATGGACTTCATCCAGGACCGCGCCATCGGAAACTTCGCTAAA GTCGACGCGGGCTTCGGACAGGCCATCAGGGAACGCATCGCTAA ACTCAAGACCGCCAAACTTTAGAAATTCCACCATCTAATTCCTGA AAATTATTGCGAAGGCTATGATTTTTGGCTGGTCTATTCATGACAT TCAAGATTATTAATATGAATATGCTCTCTTCACTCTTATTTTTGAT CTGTGATGGTTTCGGTGGTGTTGTAATTGAACGTTTTAGTTGTAAA GCAACACGGACGATGACGAAAGTGCCCGTCTCTGGTATTTGTGTC TGAAACATTTTCTTTGCGCTCATAATATTCCATGCGTATTTTGAAA GCCTTACATTATTTTATATTAAATTATCATGGATATTATTAACTTG TAATCTACAGAAATCGGGTTTGACTAGACTCACGGATCTGAAACG AGTGGCTTATTTAGTATCGTCGTTAGCATCAGTTATGTCCGTACTA CTTTTGCTGGAAGCATAATTCTCGATGATTACTTCAGCTATTTCAG CGTAACGGTGCCCAGCATCAATGAATCACATTCCGCAGAGTTGTC TAAATGATGGTCGACTAGTTTATATCAAGATTATTTTATTATATGA TTTATTGCTCACTCTTAACTACAGAATCCCTTAGCAAAATCTCAAC GAACTTAACATTCAACAAAGCAGAGTAATTTTTGTCAATGGATTT GTCGGAGTTACCCTATTTTTAAGGACGTTTTGCTGTGAAATGTTCG TTTAGTGATGGCAATTTACGTCAGCGTATTATTCACTAAATCGTGC CTTTTGTTCCTATTGTGGGAAATGCATTGATCTGGCTGAATGATTT CGTTTGAAATAATGCCTAACTTCAATATCATCTCAAAATTGAAGA GGTATATAAGTTCATGAGAACTTATCGCTCATGTGAATCGGCAAT AATATGTTGAGCTAAGCGTTTATTGCGTCGCTGGTTTTCCTAAGTT GAGTGTTTAAAATTTTTTAATTTTTTGGTAAATTTTAAGCGAATGA ACTAAAGTTGTCATAAAGTCTTCTCGGTGTGGCTTCATTGTTTTTA CAAGAATTTAATTTTCGAATATCTTGTTGTTTGTTTCAGAGTTTGC TTGTTTGAGAGTATAACCTGCATTTAATATTTTTCTATTAACTCGT TGCTGGTTGTTTGAACTAGCTACTTTGTACTGCTTGGAGCGCGATT CCGGTTTCATTGAAGGTTTATGTAGTATTTTTTTCTGTCTCGTATT ACGGGTTAGATAATGTTGCCACTGTTCTAATGCTCGTTCTATTATT ATATCGTGGAGTTGTTGGATCTTTGTGCGGTACATACATTTCAATG ACCCTTTATCTATTATGCTTCAGTGAGTAATAAATCGGTAAAATA GCCTGCCTTTTATTCTATTTCTGGAAAAGTAAAAATTTACTTTTGA TTAATTTCATAAGCAAGCGTATTTTTAACTGTATCTTTGGACTTAG ATTCA
Amy	1067	CCAGTGCAGTAACGGATGGGTTTGTGAACACCGTTGGCCGTCTAT CACAAAAATGGGTTCGTTTCCGAAACGCCGTGGCTGGAGCGCTTA CGGCAGACAACAAATATAACGACGGCAACGTTGTGGCCTTCTCGC GAGGAAACCTGGGCTTCTTCGCCATGATCAAGCAAGGCACCTTGT CTATGAACCTGCAAACTGGAATGCCAGCTGGGACATACTCAGAGT TGGTGACATGCGCCGATGTCACGGTCAACGGAGACGGGACAGCG CAGATCACAATTACCAACGAGGAAGAACCAATCTTCGCTATCTGC CAAGGCTGCAGTTGTGATGACGCACCTATTGTGACCGCCACGCAG GGACCAGATATAACTCCAACCACCCCTGTGCCCACGACTCCTGGT CCCACGGAGCCCAGCATCACGGACGGAGTTCACCGAACTGTCATC TTCGTTCAGAAGCAAACAAACGACGGACAAGACCTCTTCGTCCGG GGCGGCATCGACGCAACCGTTAGGCCAGGTTGCACCGACGACGT CACCACGGACCCATGCGCGATCGACCACACCGTGAACTCATTGGG CACCACAACGCACTACGACAAGTACAACGCATGGAGTGAAGGCG ACACCAAACTTGACTGGTTCGGTGCTGAGGCTGGCCAAGGTACCT ACGGGGGGCAGGCTGCATCTGGAACTCCTCTTGCCTGGACCAGCA GCAACCCCAGCAACCCGGGATACCAACCGCTCAACAAGTACGGC GACCATTACTGGATGCTGGACGTAGACATGGACTGCAGTCAGACC GAGCAGGGCTGGTTCGAGTTGAAGGCGTTCGTCTCCAACTCTGGT

		GACGGATGGGAGGGCGACATTGCGCAGGCTGGCAACTGCAAGGG CGCGGCAGGCGGCAGTGCTCCGTATGCTACGAAGAACCACGTCG GCCGCTGCGGGTACCTCAACGTCTTCAGTTTCGACTCGGGCTCTT GCACCGTCGAAATATTGGCTTAATTCTTGGAAATTTTCCACGTTGT AGCAATCAACAATAAAGTATACACAAAATTAAAAATT
Lgbp	714	CTTCCTTCGCCGACAACTTCCACACGTGGCGCATGGACTGGACGC CCGACTCCATCACGTATTATCTTGACGACAAGCAAGTGCTGCGTG TGGACCCTGGCAAGAGCTTCTGGGACTTCGGCGGTTTAGGGAACA GCGGCCGCGGCAACCCATGGAGCGGCGGCTCCAAGATGGCACCC TTTGATCGTGAATTCTACCTCATCCTGAACCTGGCTGTTGGTGGTG TCAATGGTTACTTCCCAGACGGTCTCTCTGGCAGTTCCCCCAAGC CGTGGTCCAACCATCAACCCGATGCTCCTGCGGCGTTCTGGAAAG CTAAGAACCAGTGGCTGCCATCATGGCAGAACGGAAAGCCCGGC ATCAGCGAGTCTGCAGCCCTTCAAGTCGACTACATCAAAGTCTGG AAGCTGTCTTAAGAATGTACACAATAGAACAGTTAACCATAAATT GTTTCGCGCCAACGTTCAATTTAATTTACCGTGTCTTATAGCACAG GCAGTAAATACATGAACAAAATCAGTTCCAAAATAGCGCGTTCAT TAACAGTTCATTAACATATTCAGAGTACTCGTAATTTCAGAGTAA GTTTTTCGCTTTGTCCTGTAGTAAATTTTTCATATTGACTTTACTGC ATTTATACAGATTTATTTGGCGACACTTTTTGTTGGCAGTATCATG TCAAAACATTTGTAAGCAATAAATTGTTGAGCAGCA
Lectin-1	1787	TGACAGCTTGCTCGGCGGAAGACTTCGCGTGCAGCAACGGCCAGT GCATTTCCATCACAAACGTGTGCGATGGAGTGCAACATTGTTCTG ACAATTCAGATGAAACGAGGTGTGCGCAAATTGACATGGAAGAG TGCCCTCCCCTGTACACTAACGTGTCAGGGCGATGTCTGCATATC GACATTGTGACGGGAACTTGGTTCGAAATGCACTATCGTTGCAGG GAGATCGGCGGAGTAATGGTAGAAGTTGACGACGCCAACTTCTTC TACGATCTCGTCAACGTATTGAAGCAACAATTTGCTTCGAATACC AAAAATTTCTGGATTGGAGGCCAAAAACCAGCAGGGTTGGAATC GTCTTACGTCTGGAACACAACAGGAACCTTCATTGCTATGGGAAG TCCTTTTTGGGCCGTGTACGATTCCGGGTCGTCGACTGGTGATTAT TCCCAGGTAACTAATCTAACAAATCAAAAATCTCAACAGACTTAG TCAAAAAGTAATACCAGCAGCGAGAAATTTAAATTGAAGTGAAG AACTTAAGAATGTTCAACTTTAAACTGATCGAACTAGATTTGGTC AGATCAGAGTCACAACTCAGAGGCCTACAACGACGTACTTAACA AAATTTTTCGTTTTCTAAGCGTTTCGGGTCAACAGCGTAAGAAAG CAACCGATTTGGCGAACTACTATTCATTACAATCAGAGCCATGAT ATACAGGTGATAAAGTTGACAATATAATCAGAGAAAACTAAATT TAGCATACGGGGATTGTGACGGGAATCAGATTAGGACCTTGCTAA CGTACTGCTTAAATCACACAGGAGCCGATGGAGTCCGCGGAGGC TCAGTACGTTTACCTCAACAAAGACAGGTTCCTGTACTTCAGCGC CGGAGATGCGGTCGGAGTAGCTGATGCCTCCGTCGTTTGTCAATT GGCTTAGAGAGACAGGGCACAAATGTAGTCGCATTTCAACGAAC TTGACCACATGAGAATCAGCAATGCTAATCAAACTTTTAGTACGT ATTTCGATTGTAGCAAAATTAATCAACATAAAGAAATTAAATGCA TATAACGTAACGCCTATATTGTGTCCGAACTATTATACACTAAAA AAGTAATTTCCTTGATCCTTTCCAAACTATGTAGCCGTGATGTCAG TAAAAAAAGTGATTCGAACGGAATGAGGCTAACACCAGGGGATT TACAACGGAAACGGATGTGAAGGGAATTTACTAATACGTAAATTT CCTCTCTACTCTTGTACGCGAAATGCAATCTAATGAAAACCATTA CTGGCGCTTCCGATCAAATCCTTAATCTACATAGAAAAAATCGAC AGAAATGACAGGATTGTGCTAGTGTTGCCATCAGCAGCTGGGATT GCCACATCAATTAATACTAAATACCAAACATGTATTCTACGTTAG AGTGGTCGGAAATACGGAATAGTAACCAGTCGGAAATAACCAAG

		GTTGCGATAACTTATAGTCTTAGGCACAGCTGTGCACATAATTTTT TTAATATACAAAATACTGAAAAATATTTACTGAAAAAAAAATGAT TTCTGTAGTTTCCACTTATTGTTACGTATTTGTAAAAAAATACTCC TTTAAGGCATTGTTCCAAAAGGGTACACTAGATCTATGGTTATTT GCCTTTACTATTCAGCTTCTTTTTTCTTTTAATGGCAGCCTCAACA GGTATTTTGCAGGCAACTTTAGGACCGTTGATAACGCAGATGACG TTGACCCGATCGTTGTAGGGAAAATTTGGCAGCAGC
Lectin-4	927	CTAAGGTATACCAGTTGTATGTAGCGACCAGCGAAAACACGTCG GCAGAAGTGAAAGTGAGCAATGAAGTACTTAGCACTCCTAACCC TGCTGGTCATCGCTGGCGTGTGCGTGGCGCAGCAGCAAGATCAAG CCGCCCAGCAAGGAACTCTCCTTGCTCTTCCAGAGCTCGAGCTTT GCCAGACCCGACCAAAGCACTGGCAGTTCAACAATCACTGGTACT ACTTCTCGTGGGACGATCCCTCTAACATCGAGACGGATCCTGCGA CTGGCCAGCCTGCGGGCAAGAAGGTCAACTGGCTCGATGCCAGA AACGAGTGCAGGAAGCGGTGCATGGACGCCGTCGGCATGGAGAC CTTGCAAGAGAACAACATGATCTTCAAGTTCCTTGAGAGCCGTAA CGTGACTTACATCTGGACCTCCGGCCGCCTCTGCGACTTCACTGG CTGCGAGACTAGGGAGGACTTGAAGCCAATCAACATCAAGGGTT GGTTCTGGTCGAACACCAATCAGAAAATCCCCCCAACCAACGCA GTCATTCCCACTTGGAACATCCAGCCCTGGAGTGCTGCAGGGCAC TTCGGTGGGCCACAGCCCGACAACGCTGAGTTCCAAATCAACAAT ACCACGGAGTCGTGCTTGGGCGTGCTCAACAACATCTACGATGAC GGCATCAAATGGCATGATATTGCATGCTACCACAAAAAGCCAGT GATCTGTGAAGACTCTGATGTTCTAATCCAGTACATCGAAGCCGT CCACAATGTTAAGCTGTAGATTTGTTGGCTAGTAATCGTCGAAAT TCTGAACTATAAGATGAATTTGCCAGTCGGAACACTTGAAATAAA TTATTGGCAAGGTCAAGCTGACAGTCCGGGAGACCCGCACTGAA ACGTTTAACTTGTTGGTTTGTTCCTTCTGAATTTT
Gst	1345	CAAGAAGATCCTGGGAGCGGAATCATTCTTCAACTCGAGTCTGGC CGCTGGCGGTTATCGCCGGTCGGCTGCTTCAGTTGCTCGCTGACG CTCGGCGTGTGTTGACGACATCGGGAAAAGTCTGTTATCAGATGG ATTTCTACTACCACCCCGTCTCCGCTCCGTGCCGGTCTTGCCAGCT CACGGCTAAGGCTCTCGGGCTTAAACTCAACCTCAAGTTCTTGGA TCTGATGAAGGGAGAGCAGAACAACCCAGAGTTCCTTGCCATCA ATCCAGAGCATTGTGTACCTACCCTGGTTGACGGCGACTTCAAGC TATGGGAAAGTCGCGCCATTTGCGCGTATCTCGTGAACCAGTACG GCAAAGACGACAGCCTGTACCCTAAGGACCCAAAGACGCGGGCC GTCGTAGACCGAATGCTTTATTTCGACATGGGCACCCTTTACAAG AGGTTCGGCGAGTACGTGTATCCCACGGTATTTCACGGCGCCCCT GTGGACAAGAAAAAACTGGAACCCATCATGGAGGCCCTGGGCTG GCTGAACGGCTACCTCGAGGGAGGGGACTACCTAGCGGGAGCTT CCCCTACCATCGCCGACTTCGCTGTTGTGGCGACCGTGGCAACTT TCGTGGCCTCGGGAATCGACATCTCGGGACAACCCAAAGTGGTG GCTTGGTACGCCAGGTGCAAAGCCAACTTCGCAGGTTCAGACGA GAACGAAGAAGGGGCCCAGCAGTTCGGCGGTTGGGTCAAGGCGA AAATGTGAAAATTCAATTCAAAATCCTATTATCCACTATCCTTTTT TGGAAGCGCATGTTTTTCCTGTCATCTTGAAGTGCCGAGAAAGCG CTATCCCTCAAATTTCTGCATTCTACTATCCGAGTCATCTCTCTGA ATTCACATGCGGCTAACTTATGAAATTGAATGAAAATTTGTGCAG TTCTCCATTTTAATTTTGCCCTTGAAGTTCAAGCTATTTCCAAACC GAATCATTCTTGATTGTCGCCCAAGTACTGTTGGGCGAAGCATCC TTATCTGGCGTTATGAACGTGGCAGCCCCAGGCCACGATGGATCG CAGCCTCTGATAATTATCGCTCATCCGCCCTCATTTGGGCCACCTG ATGCACGATGCTGTTGCTTGCTTCTCACTGTCGGTAGTGGTGCAG

		AGGCAGTCGATCAGCGATTGTGTCCAGGCGCATGTCCAATCGTCT CTTGAAGGTCTCGAGGGTGAACTCACCCTCTCTTAGAGGGTGAAT CAGTTCACACACTTTCAGAGTTAATACACAGGTTTCAGAATTGAT TTACTGAACACAGTTCAGAATTAATAAATTCTCTACCCTCAG
Cyp2c9	620	GGGGAGATCTGGCACGGCGGTCGCCGCTTCACAATCCGCCATCTG CGCGACATGGGCATGGGCAGGTCATCTCTGGAGTCGCGCATCGCT TTGGAAGCGCACAATCTTGTCGAAAACTTCAAAAAAGTAACCGA CGTTGCGACCGAAGTACCGAGGTCTTTATCCATCGCTATACTCAA CATCATCTGGGAACTCTGCGCAGGAAAGCGACACGACTTGGACG ATGAGAAGGCTTTGGAGTTCCAGCAAGTGGTGGTCGACATTTTCC ACGATTTCCAGGGCAACGTTGTCATTTTCGACACATTCCCATTCCT GCTTCCCATAACACCGCGTTTCGTCATGGAGAAGCTCGGCGTAAC AAAATATATGGAGGACTGCGATATTCTTATGAGCCACATGAAGA AATGCATAAAGGAGCACAAGGAGAGCCTCGATGTCAACAATCCT CGCGATTACATCGACGCGTTCCTCATCGAAATGCAGCTGGAAAAG AATGCCGGCAGCGCTATATTCACCGAGGAGAATCTGATGGTAAG CTTGGGTGACATTTTCACGGCCGCATCCGAGATCACGTCCATTAC ACTGCGCTGGCTCATCCTCTACATGATCAAGCATCCAGA
Cyp18a1	2818	GGAAGTTCGCCATTTGCTGGACTGCGTCCACAAGACGGGAGGGG AAGCGGTCAGCATGGATCAGTTCCTGACGGCCCTCTCCACGAACG TCATCACATCTTTCTTATTATCAACTCGGTATGATGCCTACGATCC TATCTTCCTGAAACACAAAGAAATCAACGGAGAAGGATTCAAAC TCTTTGCTAAGTTGGACATGGTGAACTATATCTGGGCTTTGAAGTT CCTGCCTGGGGTCGTGAAGGACATGAAGAAGCTGTGCGCCAATC ACGCCGAGTCGATGGATTTCTTCCGAAACTTGATCCAGGAGCGCA GGAAAACTTACATACCCGGGGACGACAAGGACGTACTGGACTAT TACTTGTCCGTCGAGTACGAGCGAGAGACGCCCATCCTCAATTTC GATGAGCAAGTCGCTCAGGTGCTCTGCGACTTGGTGAGCGCAGGC AGCGAGAGCGTCAAGACGACCGCAAGTTGGGTGCTCGCTTACCTC CTCCACCACGACCACGTCGTCCACAAGATGCGCTGCGAGATCGAC GAACACGTCGGCAGGGATCGGCTCCCAAATATGGCCGACCAAAC CCGACTAATTTATTGCCAGGCCGTTATAAACGAAGTCATGCGCAT TGCCAACGTTGTCCCTATTGCTCCGAACCACGCAGCGGAAGCTGA TTTCAATTTACATGGATATTTTATCCCAAAAGGAAGTTCTGTCGTG GCTCTTCTCTATGCCTGTCACATGAGCCCGAAGTATTGGGATGAA CCGCAAGAATTTCGGCCTGAAAGATTCATCAACGAGCAAGGGCG GCTAACGTGCCCTCAGGCTTTCATACCTTTCGGGAAAGGACAACG GAATTGTTTGGGTGACACTATAGCTCAGTCGGAAATATTCCTTCTT CTCACGGCCCTCATTCAAAATTTTACTCTAGAAAACCCTATTGGTT GTCCTTTGCCTTCATTGGAGGGCGTCCAAGGCGCAACATTGTGTC CTCAGGACTTCAAGGTAAGCATGAAATCGCGTTCAAGTAAAACTG TCACCAGCGATCCGCGGCCTCCGCCCGCGACCACAATTCTCATGC ATGGTTAAATGATCGTCACACGTTCATAATGCCAACTACTAATTA CCTGATTATGTCAGCTCATTACCTGTTCATTGCTTTCAAGAAACAT TTTGATCGTAGCATTAATTCTGTTATCTTCATTTTTTAGAAGTTTA ATTGTTGATTTTGTTGAAACCCTTGAAAGAGAATCCCTTTTCGTCG CCATCACGTGCGGCTCGCGTTGTGGCGACCAAATGAGGCACATTG TGTTATTATCAACTGAATCAAATTGTTGAGATGTGTTATCGTCGA GGACAAATACTCTTCAATTAGAAGCTCATCTATTTATAAAAAAAA TGTTTGTACGGCTATTGGACTATGTCAGTTTAATTCTCTAATCTAT AATATAAGCTGAGACCAATATTGTAAAGATGACCAAACCCATGC GGGGAAATGAAATCAGGCAAGCCCATAGCCGAAGAAAAGATTAT GGCTTTCAGTGATACAATTAAATCAACTTGGGCATCATAGTCAGC TCACAATGTTGATGATGGTGATCACATGCAAGCTCTCGAACACAC

		CGAACATTACGAATACACTGAGCGTCCGGAACACCTCACTATAGG CCGTGAGCTTCGAGACTTTATGTTGTGTTTTGAATAGTTTCGTGAC GTATTAAAGTTCTTGATCGTTTGCTTATGCATTTGTGTTGCTAGAA ATTTGTGTCGGAAAACAAGAGGTGCAAATCAGTGTTGTTGAGAAT CACATTTTATCTTGCTTATGGTATTACATTTCGTTTGATTACATTTT TGAGTCCGCTTTAAGATGTACCATAATCACGCTACATCTAGCGCT AGTTCAAGCTAGATAATGACTAATCCCACAAATTATAAATAAGTG CCAATGTAGCTTCGACGAGTAAAAATTCCCAATTGTGTAAGTTAC CAACTATATTGTAAATAGTTTATCTAAGTATTATATTTTTAATAAC AATGTGAAGACTATATCGTCAACGCAAGCTCGGTGAAGTAAATA CTATCATGGCTAAATTTAAATATGTTACGGGTACTCTTTGTTAAGT AATAATCATATTATCGTAGCAAGTTATGGAATTTCATAACTACGC AATCGGCTACCTTATATTGACCAAGTTAATTATAAACAGCGCACG TAAAAGCTTTTATCCGTCTCAGAAATTCAAAATTTTCATATCAACT TCATGGTATCCAATAATTTCCAGCCGTTTTTGCACACCATTCCTGA AACAAGCATTCGTTGTTCTAGGCAGTGGCGCAAAATTACAGTAAA TTAGTGCTGCTTACAAACTGAACTATATTTGGAAAGAACTTTAGT TTTTTGATCAATTGTGTTGTTATATTTTGAAGTAACTACGATACTG CTTGATTACCTTGGGCCGTACAATCAAAACTCGCTTAAGATATCG CAAAACTCGCGGCCAAAATCTCGTATGTCATTGGTGGGTTCAACA GGAAAAGCGGTCGTTGATTGGTGGATGAAAAGTTTATTTCTTAGC TATTTTTGATGGTACGGCACCATGGTCAACCGTGAATTTTATTTCT CTCAAATGTAGCCAGCCGATGCGCGCTTTAACGTAATGTACGCGC TTTAATGTAATGTATTATGTAGAGAAGCTCATGTGATATGTATGC TACGTACGTTTAAGATTTTATCTTACGCTCATCTCTGAATGGCATT ATGTTCAGACAACTTTTGTATACGCGAGAAGTCAAGTCCAAGTTA AATAAATAGCATTTCAA
Cyp2j2	1497	AAAACTTCCAGGCTCTAAAGTTCCAGGCTCTGTTCCAGGCTCAGC CCCCATGTCTTTAAAATCTTGGGCTTAGCCACCTCTTTCCTTAGTA CCTGGCGCTTAGCCCCTCTTTCTTCAAAACCTGGAGGTTTAGGCCT CACTGCTCTTGAAAACTGATGCTTAGCCCCTCTGTATTTAAAACCC GAGAGGTTTAGCCCACAATGCCCACCTGTTCTTGAAAGCTGGGGC CTAGCCCACCTGTCCCTGAGAACTGGGGCTGAGCCCTTCTGTCTTT AAAACCGAGAGGATTAGCCCACCTGTTTTAAGTGATTTTGAAAGC TGGGGGTTTAGCCCCATGTCCTTAAGCCTAGTGATACCCCGGTAA GCACTATTGGTATCCTCTTTAACTCATTGAATCAGGCACAAAGCC TATGCTTTCTCTAGCCTCTAGATTTATTTAGCAGAATCTTGAGCTC TTAGGTTAATCAATTCACTCTCATATAACTCATGGATACGATTTTA CTGTACAGTTATATTATGTTATACTAGTTGAGGTTAGCATCATTCC CAGTAAATAGTTTTCAAGAAACGAATAATCAAATGCAGCATGCG GAAGGTTCTTTCTAACACGTCATCTTCACAGCACGAAATGGGGCG CTTATCCTATCGTCGTTCTTACCGAATGCAGTCTCATCAGGAAAG CCTTTCAACATCCAAGCTTTCATGGGAGGCCTGACTTCTTGTCATT CAGAATAGTGCAGCAGTATGAAAAGAGGGGAGTGGCTTTAACGG ACGGGCTGTTGTGGTCAGAGAACCGACATTTTTTCCTCACTCAAC TGCGGAACCTGGGCATGGGTAAGACTTTCCTCGGGGACGCCATAG AAGATGAGGCGTCGCTGCTGATGGACCATCTCGAGAAGAACTGC CTCGATAAACCTGTAGTCATAGACAGCGCTCTCAACACGGCAATT ATCAACGTACTATGGCTGATGTTTACAAGTAAGCGATACGACGTG GGTGACCCTGAAATAGTGCAATATGTTCAAGTCCTGGATATAATG CTGCGCGTTACTCGGAAGAAAATCTTTCTGCTTGATGTTATTCCTG TGTTGCCAAAAATCCTACCTACGTTCATCCTCAACTCGTTATTCGA CTTCGATGTGGGAGAAAATTACTTTAAAGAGCTGACCAAAATGTC TGAGAATAAAATTCGAGAACATTTGACATCCCTAGACCCAGACA

		ACCCACGCGACATCATAGATTACTACCTCCTGGAAAAAACTGAAC GCGACTCCAATGGCTTCCTCACTCATTTCGATGCTCAGCAGCTGT GCGGCATGATGAATGACGCATTTATAGCTGGCTTGGACACGACCG CCTCCACCCTGCGCTGGATGATGCTCTACCTAACCATCTATCCTGA CGTTAGCATGCCTTCTTGCGCTTCACTGCTGACAATACAAATAGA GTTTTATCACCAGTGAAGAGGCGTATCTCATCGTGTGAATCATTTT CAAC
Cyp	2239	CGCAATTGTGCATCACTGTCACCATAAGAAGTTCTTCCACAAGCA ACCTTAATGTTCGTCTCATATGTTCTAGATGACAGTCTACCTAATT ATTTTCCTGTATTTAAAACGTCAGTCAAAAGCGCAATTTCAAGCT GTGCTTGCGTTGCAGATGCTGGCCATGGGCTTCAACGCCGTGATA TCCCGCATGGAGGAATACTTTCCTCGGCCCCTCGAGTTCCTGCCG GAGCGGTGGCAGCGCGACCGCCCCTACGGCGCCATCCATCCCTTC GCGTCGCTGCCCTTTTCCCACGGCACCAGGATGTGCATCGGCAAG CGTCTGGCCGAGCAGGAGATATACACCTTCCTCATCAGGATGCTG CAGAAGTACACGGTGACGTTCGAGTCTGACGAAGAGATGGAGGT GAAGACGCAGCTTGTCCTCAAGCCCGCCACTCAGCTCGCCTTTAA GTTCCACCCGAGGAACTGATTCGCCCTCTCGAATTGCGTGCCAAT CTTGGCTTACGTTTCGTGGCAAGCTGTCACCAACGCGTCCCCGTA TCGTGCTGTCAACAGCGCCTCATGCTACAGACAAACCACAAGTGC CTTCGTCTTTCAACATCAACTGTTCCTGCAATGCACTGCCTACCTG CGTTGCACGAGCGGACGGCACCTAATAATCTTGTTATATATAATC AATAAAAGGCGACCCGTTTAATTGGATAAATAAATCTTTGTATGG GTCAAAAATTAGTCTTCAGATATTAATATTGCAAATTTTTGATATC ATCTTGGAGCATTGGCATTGGTTGTGAGGTGATAAGCCGGAGAGC CGGTGACTTTATGGGAATATATGTAATCGTCATTCTTATCCCGGC AGGTAGGCTTCCTGTATTATATTAATGAGCAAATCTTAGTTATTCT TCGCAGTACAAACGATTTTATTAACTAATTTTAGCTCAAACATAA TTAGCCTATGCCCTGAGAGTGCGAGTGGCACACAAACAACTGAAT TCTCTCATTATGCAAGTCGACAAAACTGGACCATGACCGCTATTC GACTATAATATTATTTGTCACATTTTAAATTATGTTAATTATTTGA GAATATTCACGTGTAGGGCTCCCTGGAGATTATTTTGCCGTGCCC CAAGCAGTTGGTAGTGCAATAAATTTATCTGCCGCATTTGTCATG CCTGATCACTAGTTCTTGCTGAGCTGCCTTGTTCTGTTCATGCTCT GGGGCAAATTTACGTCGTCGTTCCTAACCGGAATGAATCATCTGT AAAATTATTAATAGAATTGACAATTAGTAGTGGTAGTAGGAAAA GGTTTGCTAGCGCTTCAGATAGCTTATCTATTTCACGCCGGAGCC CTGCCATAATAATTTACTTATAGTAACTTGCTATGAACGGTGGTT AGTATGGGTCCTAACGATGCCTGTCCGGCGCGCTTCTGTAAGCGT GCCCAGAAAAGAGGCTTGCTATGTTTTCTTCTTATCCTAAATTAA ATTAGAATTTACATTTTTAATATGTTTAGTATTTTCTATCTAATAG TTATCTTTGTATGGACTCGGAAGACTCGCCCCACTGCCTGATATA ATCAACTGAAGGAGACTGAACTGAAGAATGTAAAGCTTAACTAT ATATATATATAATATATACTACTATTTGATATGTAAGAGTAACCA TCATATGCTATTGTTATGTTAGAAATTAAGCTGTGTTCCAGCGTCG AACGTTACTTTTACCAATATTTTATATTCACACCAAAAAAGCATTT CTGAAATGACTTCTCTTGTTGTGTCGTTGAACATTTATTGTTAATC CATCTATATATTTTATTAGTGGGTAAACGGCGGTATTTTGTAATCT GTATCAGTGCACTAAATTCAAACGAGTTATCGTAACAAAATATAA GTTTTTATTATTGCCTTTGCTATAATTTTACAGCCTACGCGTGTTA CCGACTACCGATTGTTCTGGGTTGGCAAGTTGTGTTGGTTGTTACT GGAAGATCTTTTCCTTACGTCAGCAATCTGCTGTCTTGCTAACCGA ATGGCTGCTCATGCTGCGCTCCAGGTGCGAATTGCTTTGAATGCT TTCGGTGCTACTTCCGCTGCTCGTAAATTAATAAGTTAACGATAA

		AATTATCTACCACGAAGTCTGATGAATATTTTCTGGTGTCTTGAA ACGCTTTGCAGCATTTATCTCGTCCACTGCTGCCCAGAAACACCA AATGGTTTTTCCTCGCACTTGCG
Cyp2L1	1695	CGTCTCCGTTGTGCCGCTGCTGGTCCCGCACTGCGCTACCCAAGA CATCGAAGTGAACGGCTTCCATGTACCCAAGGGCACGGTGCTAAT GGGACACGCGGGCAGCTGTCACGGGGACCCGCAGGTCTGGCAGC GTCCCGACGAGTTCTATCCTCCGCACTTCTTGGACGAAAACGGCA AATTCTGTCCCAAGAAAAACGGCTTCATCCCTTTCTCGATTGGTC GGCGCGTGTGCCCCGGAGAGAACTTGGCCCGGCAGACGATGTTC CTGTTCGGCAGCGCGCTGCTGCAAACCTTCATGTTCGAGGCCCCC GAGGGCGAGGTACTGAGCACTCAACGCGACCCTGCCGAGCGCAT GATTATCATCCCCAAGCCCTTCAGGGTCATCATGAGGCAGCGCAG TTAGGCTGTGTTGTTACAGTGCAGTATCGCAATGCAGGATTGCAG TGCAGTGTGACAGTGTAGTGTGACAGAGTAGTGCGACAGAGTAG TGTGACATAGTAGTGTGACATTGTATTGTGACAGTGCAGTGTGAC ATTGCTCTATGACAGTGTAGTTTTGTCACGGAGCCACTCCCACTG AGCGTAACGTGAGCCTATATACTATCACTAATAATCCAGCCATGG ATCGTTACATTTCTTGCAGAATACAAGGGAATAAGGACGATTTAA CACAAAACTTTCATTCATCCTACGACGATAAGATAAACAGAGTGT TTGTGAGTTGAAGTTGGGTCAAGACTTCGGAGGTCTCTCTGTATG AGATAAAACTTTTACAGACACTAAACGTACACAATGGGTCATGCC TAAAACTGTATCATAAAAACTAGAAAGACCTCTTTACATTTATAG CTATACATGTTGACATCACACTAATATATTACAGGTCACAATAAA ACGCAACGATATTGAACCTATCTAAAAATATTCAAGTATTCCGTG AACAGTACACATTCCGTGCATAGTACACAGTCCCGTATGGTGGCG CAGGTCGAGTGATCCTATCGAGCTACTATTCCGCAATACAGAGTG GATCCGGCAATATCGCATGGACATCCGAAGCCATATGACTAAAA AAGTGGTCAATATAAGAACATCTATAAAAGGTTCATTTCCCTCCT CTGCAGTGCGACGTTTCTCCCAACCCTACCACCTATACCACTGCCT TCAAAAGTCTGGATTTGGACATCGGTTTGCCATCCTCCTTTACAAC CTTACCCTCCTCCCAAGAAGTTAATGTACACTCAGGTAGGCCGCG CTTTGAGAAAGGAATTGGCTGGAAATGCCTTCAAAAGAAATGTG AATACGGGAAGAATGGTGAAGGATGCACATTCTCATACCCTAGA ATGTGCTATAAATTCTTTGACTATGGAAGCGGTCATCCCAGAGGG TGCAAAAGAGGTAGCGATTGCCTCTTATCATCCTACGCTGTGCTG GAGTATCATGAATAGAGAGAATTGCACCAAAGAAAACTGTCGTT TCCACCACATAAAATCAAGAACAAAAAGGGGAGAAAATGTCAAT AGAGCGAGTACCAATGTTATGCCCTCACAGTAAATTGCGATGCCC ACCGCACTGCACGACAGCGTTGCCAACTGCTATTTTTTCATACAA GATTTCACAAATCTGGTATTTTGAGACGATGTCTGGTATTGGTAT ATTTTATCTGGTATTATTACCTGAATCTGGAATATTT
Cyp3A13	2113	AGTTCGCCGAGGGGACTTTTTGGACCTGCTGCTGGATGCCAGGGA GTTGACGAAGGAACAGGAGGGCGAGGTCGATCCTAAGGGGGTGA AAAAAGGCGGATTGAAGGACTCGACCATCGTGGCGCAGTGCATC CTCTTCATCCTGGCGGGCTACGACACCACAGCCAACACCCTCTGC TTCGCCGCAGCGCTCCTCGCCCGTAACCCTCAATGTCAGGCTCGT CTTAGGGCGGAGATCGCGGCCAGAATTGAGGGTAACGAGAGCGG CAAACTCACCTACCAGGACGTCATGGAGTGCAAGTACCTCGACGC TGTACTTAGTGAAACGCTGCGCATGTATCCTCCGGCTCATCTGTTG GAGAGGAAATGCACCAAGGACTACAGCGTGCCCGACTCGGACGT GAAGATCACGCAGGGCATGTTCGTCACCGTGCCCGTGTACTCCCT GCACCGCGACGAGCGCTACTTCCCCGACCCCGACACCTTCCTACC CGACCGCTTCATGCCCGCTACTAAGAACGACATCCCGTCGGGGGT GTACCTTCCCTTCGGGGCTGGCCCCCGTATGTGCATCGCGGAACG

		CTTCGCTCGCATGGAAGCACGACTCGCACTCGCTGACCTCATCAT GAACTTTGACCTCTCAGTCGAGCCGGGGTCAGAGGAGCTCAAAGT CTCTAAGGAGTTCGGCATCCTGCGACCTGATCCTGCGAGCCTGAA CCTCATCTTGAGGGACGCGGCATCGAACGCGGCGGGTGCCACAG GCGCTGCCTCGGATGCCACAGGCGCTGCCTAGGTGCCACAGGCGC TCCCTAGCATATTCTGAGTTCTACAAAAATTCTTGATTTTCGCCTC TGTCTTGATTAGTTTCTTGTTGCGTCCAATAAATGTTAGTAACCAA GTCAGAAATATATTTGATTCAAAATGCCGGACCTCTGATGCGGAC TATCATCCACAGCCAGAGACACGAGTAATTGATAGGCTTTTGTGT TCGTCGAGGCGGCTATATTCGTCATGGGTGTAGTGATTTCTCAGA TATATTATGGTGCTTCATTAATTATTTTGATAACCTATTGGACCAC AGCATGTTGCACTTCAGAGTACATGTGTTGTACCGTGACCTGACG CTGAGAGGCTTGCTGCGAGTAGCGTGAATCCATTGGCACCGCCCC ACGTATGGGGCCCAGACAGTTACATGAATACTAAGAATGTAACG GTTGATGGTTGTGTCTAGGCCAATAGGTAATTAAATAGGCACTGC AGAGAGCCCTCGTATCAGTCATTCCAACAGTGGCTGGTTACTGTT GAGCAAAGATGCAGATGAAATTGATATCTAATTAACTGAAACCC ACCAAACAAGCATTGAATTATCTTCTTGTACGGTAATGGTAAATA AAAAACATGGGACTTAATAGTTCTGTTTTAAAGACGTTCCTCCAT AATTCTTAAGGTACGACAATTTCTTCCTTAGCGACCATCACAATT GGATGACGGGCAGACGCTCGATACACGAGCGGCCCGTAGGACCG TTAACTGTACATCTTGTGATTAGTCTACACATGTCTTTGATGTTTA ATTAACACATTAATTATTAAGATCTCATGACAACCGTCATTTCAG CGACGTTATTGTTGGCATCACTCGGCAGAATCGCTAATTCCGATC GAGTGCCAAGATTAAGAAGAGATAAGAAAAGTTAGTAAGACGAA TAGTGTAGATTACAAGACTAGAAAAAAAGACTGAATAGAGCATG TAGCGCAGGAAAAAAAAGTTTACATTTTCCAAAAAGATATCCGG CATTTCCAAGATTTGCTGTGTGCTTTGACCCAAACGTGTTTAAAA GCGTTTTGTGCTCACCAAACCAAACGCAGGCAGAACGATGAGAC GAATCTTATGAACAGCAGAACACGTAAGCTGCCACGACAGGCGC ACACGCAATAATTTACTACTATAAAATTAACAGCGGAATACATTC ACAGAATTCTTCAAAAATTGGTACGAACTTACAGACAAGTACTAA AATCATCCGCAAATTACTATAAAAACTAATTAGGCTTAGTTAATA AATATTCATTTAGATTAGAAGCTAATCAATAGAAATGATGCCCAC AACGCG
Cyp6BS1	1134	GGCAAACTCACCTACCAGGACGTCATGGAGTGCAAGTACCTCGA CGCTGTACTCAGTGAAACGCTGCGCATGTATCCTCCGGCTCATCT GTTGGAGAGAAAATGCACCAAGGACTACAGCGTGCCCGACTCGG ACGTGAAGATCACGCAGGGTATGTTCGTGACCGTGCCCGTGTACT CCCTGCACCGCGACGAGCGCTACTTCCCCGACCCCGACACCTTCC TACCCGACCGCTTCATGCCCGCTACTAAGAACGACATCCCGTCCG GGGTGTACCTTCCCTTCGGGGCTGGCCCCCGTATGTGCATCGCGG AACGCTTCGCTCGCATGGAGGCACGACTCGCACTCGCTGACCTCA TCATGAACTTTGACCTCTCAGTCGAGCCGGGGTCAGAGGAGCTCA AAGTTGCCAAGTCGTTCGGCATCCTGCGACCTGATCCTGCGAGCC TGAACCTCATCTTGAGAGACGCGGCATCCAGCGCGGCCGGTGCCA CAGGCGCTGCCTAGCATATTCTGAGTTCTACCTAAATTCTTGATTT TCGCCTTCGTGTTTCGATGTTGCGTCCAGTAAATGTTAGTTACAAA GTCTGAAATACATTTGATTCATTATGCCGCACCTCTGATGCGGAC TGTCATCCACAGCCAGAGACACGAGTAATGGATTGGCTTTTGTGT TCGTCGAGGCGGCTATATTCGTCATGGGTGAAGTGATTACTCGGT TATTTTGCTTTATTAATTATGATAACATATTGGACCACATAGTATG TGGTACTACAAACTGAGTAGTACCGTGACCTGACGCTGCAGGCTT GCTGCGACTATCGTGAATCCATGAGTGGCATCCTGCGTATGGGGC

		CCCGACAGTTCATGAATACAAAGACGTAAAGCTTGATGGTTGTGT CTAGAAGGCTAATAGAATTATATAGGCTGTGCAGAGAGCCCTCGT ATTAGTCTTTTCAGCAGTGGCTGGTTAATGTTGAGAAAAGATGCA GATGAAATTGATATCTGATAAACTGAAACCCACCAAACAACTATT GAATTATCTTCTTGTACGGTAATGGTAAATAAAAAGCATGGGACG TAATAGTTCTGTTTTAAAGACGTTCCTCCATAATTCTTAGGGTACG ACAATTT
Cyp2u1	851	CTGTACGTTAGTCGGAAGAAAATCTTTCTGCTTGACATCATTCCA GTGTTGCCGAAAATCCTACCTATGTTCATTCTCAACGCGTTCTTCG ACTTCAATGTGGGAGAAAATTACTTTCAAGAGCTGACCAAAATGT CTGAGAATAAAATTCGAGAACATTTGACATCCCTGGACCCAGACA ACCCACGCGACATCATAGATTACTACCTCCTGGAAAAAACTGAAC GCGACTCCAATGGCTTCCTCACTCATTTCGATGCTCAGCAGCTGT GCGGCATGATGAATGACACATTTTTGGCTGGCTTGGACACGACCG CCTCCACCCTGCGCTGGATGATGCTCTACCTAACCATCTATCCTGA CGTACAGGCAAAGGCTCATCAATGCCTGGACGAGGCAGTGCCTA GAACTCGGCTTCCTTGCCTGGCAGACAAGCCTCGATTAGCGTATG TAGAAGCGCTGGTGCTGGACGTGTTGAGGCTGAGCTCTTTGGCGC CTCTGTTGATGTTTCATCGCGCCACAGATGACGTCGTATTCGAGG GCTATACCATACCACAGGACACATTGCTGATGGCGTGCAGCGAA ATTTGCCACAAAAATAAGTCCTTCTGGCAGAAACCTGACCAACTT TACCCTGAACATTTCCTGGACGAAAATGGGAAGCTTGACAGCGAT AAGGAAAATTTCATCCCCTTCTCAGTCGGCCGACGCCAATGTGTG GGGGAGTCTTTGGCCCGCATGCAACTTTTCCTCTTCACGGCGGCC ATCTTGCACCGATTCAAGCTCGAAGCTGCGACGACCACTTCGCTG AAATCAGAATCTGATCCTGCCACGACCATTTTCAACAGGCC
Cyp2j6	2428	CTTGACCCCAACAACCCTCGTGACGTCATCGACCAGCTGCTGCTG GAGCGTGCCCACAGAGACGTCCCTGAATACCTCACTGCTGAAGAT GAATTGGATTTCAAATGCATTATGGGGGACATGTTTTTCGCCGGT TCAGAGACTACCACAACAACTCTTCGATGGATGATACTGTACATG GCCATCAACCCTGAGATTCAAGCTAAAATTCACAAATGTTTGGAC CAAGTTGTACCTAAAGATCGTCTGCCATCACTCAACGACAGAAAT CAACTGCAGTACGTTGACGCCGTGTTACTGGAGGTCCTGAGAGTG AGTTCACTTGCACCCATTTTGCTGCACGCAACTACAGCTGACGTG ACCTACGAGGGCTATGAAATACCCAAGGGCACGGTGGTTCTTGGT TGTGCCGAAATGTGCCATCGCGATCCCGCTTATTGGAAGCATCCA GACCAGTTCTATCCGGAACACTTCCTGGACAACGAAGGGAAACTC GACGGGAAAAAAGAAGGATTTATACCTTTCTCAATCGGGCGCCG CCAGTGTCTCGGGGAGTCTCTGGCTCGCATGGAGCTCTACCTCTT CTCGACGGCCATTTTGCAGCGCTTCACCATAGAGCCACCGGAGGG CGTCACGCTCTCTACGGAAACTGATCCGTCGCAGATGATGTTTCG TGTACCCAAGCCTTATGAAGTGGTCCTGAAGAGGAGAATGTAGTC CAGACATTGCGTCTCTTCCAGCCGATTTGACTGAAGTTATACTTTT TCATGTCTGCTGCGATTGCGCTATGCGCTGTTCGAAAAACCACAA AAAAATTGTTGGACCGACGTTTTAGTGGCCGGACGTCTTGTGAAA TTCGACCTATATATATCCGGTTCGGGTTAACTCCCAACAAAACGG TGAAATGGTTTCATTTCGCGAAACGGCTAACCGTTAAAAACGACG CTTCAACCGTTTTTTCGAGGGAGCGCAGCCGTGCACCACCAATTT TGTGCACCAATAAGCCCTGAATCGAACGCCTATGATCTCTCCGGA CATCCACCAATCAGCCGTGAACCACTCTGCTTTTTCTTCCTCCGGC CTCGCATAGAACCTGCCCTGCACCACCAGGCCGCGGATCGAATCT GACTTCCGGGCATCGTTAAATTACGTAATTTTACGTTCACCGCATC TGAGGTTAGGTTTCAGATAGCATAGTTCTGTTCCAAAATTTATCGT CCTTCGACAAGTGACCGTTCGACCGAGGGCTGGTTCTACCCACGG

		CCTAGATACGGAAAGCTTGAAGTGTTGAAAGTTACGTTCCGGAAG GCAGGAATCATGATGGTCCATTACATCCAGGGCAGTGGTGGTCTT CGGCAGGCGACCCTTTTTCGATCATTGATATACTCTTCAATCCTCG GGCTGGGGAAAACCTGTGGTGCCAATTCAGCTATTTTTCCGCTAG ATCTAGCGTATTTTCAACGCATTTGGCGGAAAAAGTGTATGTAAA ATGTAGGTAATGCGATCTTTTTAGATACTAACACTTTTTAGCTATA CTAACACTGAATGATTTGGCAACACTGGGAAAACAGGTGCGTGA TTGAAAGCCTCGAATAGAAATTTAGCAAATCTGTATGAACCTGTA AAGACTCCAATTCATAACTCGCGCTCTCATTTCTTGTTAACGTCCT AATTATTTTTATTAAATCCCCAAGTTCGATGAACTGTTTTTGTTGC TTCATGTCCTTATTAGCTAAAATCCAACTATTTACTTCAATGTTCA TTATTAGGCAATAGTAGAATATTAAATCTAAAATATTATCAATTT TGAAAAATGAACAATAAGCGTGTAATTTCTGAATAATAAAAATTC AGCAAACGACAAAATACTTTTAGGTTCTAATAGGGTCATTCAAAA ATTGTGAATTCTGTTTAGAGTGCGTAAGCGGGCAACATTAGCCAT GCACTATAGGGCGGCTTAGCTTTATTTTCTGGCCAGCTTCGTCGTG CGAACAACTGCTAGTAGTAAGCTGCCATGTTTTCACTACAATTCA ATGCTTGCATTGTATAGATAGTTAAAAATTTCGAAGTGTAACTGT CAAAACTGAATAACCGCCTGCTCTCAATATTCTTTCAACATTGTTT TGTTTTGTTGTGTCTTTGCTTTGCTATGTCATGTCTTTATTTTATCG ACCATCACCCACACATGTGTTGGTCAACTGCATCCTTCTTCCATCA TCAATAAATTTGACTCGGTTCTCAACACGCCTTACTTAACCGTTGA CTACATAGTTACCCTAAATAATTGTGAAAGGGGCTGCCGTTTGTA ACTTCGAAATCCTATCTCAAACCACTTTAGCACCCTCTTTCCTCAA ACCTGATAAAATTAGAGGGACTAGTTTTGCCTCGGCAAAGTGCTT AGGACTTATTAATCCTAATTGTTTATTTG
Rad50	1532	AGACCTGCTCGTTAAGCTCCGCAACACGAGGAACGATCTCGCCGC CAAGCAAATCCAAATGAAGACTAAGCAACAGGCGGTGGAGAAGA AGCGTGAGGAAGTGAAGCGTCTGGAAGAGGAGAGCTGCAAGGA GCGGAAGAGCAGAAACGAGCGTGAGGAGCTTGTGGGCCCCGCCA GGACTGCCGTCAAGGAAGAGCAGGCCAAGAAAGATGCCGCCGTG GCACGACGACATCTGCTCGACAGAGAAGCTCGAACTAAGATTGA ATGTGTTCGTGAGCTCCTGAGACCAATAGAGACCCTGCAAGCGAG TATTAGGAATTTTATGTCAAGCAACAGTGAGTTGAAAATCCTTGC GAGCAAACAAGAAGTACAGAAGTTGCGTCTCGAGGAGCAGGAGA AGCAGAAACTGAAGGAGCAGCACGAGACGAGCCTCAAGAAGCTC GAGAAAGAGCTTGTAAACGAACGCGAGTATGAACGCAGTCTTAA TACTGAAAAGAAGATACGGGAACGAAAGCGAGAGGTCGAAGAC ATACGGGAGGAGGTGATGCTTCTGAAAGCTAAAATTGCTTCCTCG CAACAGGAGTCCTTGACGTTGGAGAAAAGTCGTCTCTCCAGAGA GTTTCACGAGCTCGAGAAAGAGTTGAACTTGAGTCGCGGTCGCTC CGATGAGGTGAACAAGAGAGTCCGTGAGCTTCAGCGAGACTTGC AGCTTCCTCAGTTTAAGGAAGCCGTGCGCAAGTATCGCGAGAAA ATGCTACTTATCATGTGTACAAGGACCGCGTGTGAGGACCTGAAC AAATACTACAAGGCTCTCGACACAGCCATCATGAGGTTCCATGCC GAGAAGATGACAGCCATCAACAGGATCATTCACCAGCTCTGGAA CTCCACCTACAAAGGCAACGATATCGACAAGATTGAGATCAAAA CGGATGAAAATGAATCTAAAGGTGCGGACAAGCGGCGCACTTAC AACTACCGCGTGGTGATGCACAAGCGCGAAGTGGAGCTCGACAT GCGCGGCCGCTGCTCGGCCGGTCAGAAGGTGCTCGCCTCTCTCAT CATCAGGATGGCTTTAGCAGAGACTTTCAGCAAGAACTGCGGCAT ACTTGCGCTAGATGAACCAACGACTAACTTGGACCGCGATAACAT CCGATCTCTGGTGCGGTCTCTGGTCGGGATCGTCAATCAGCGCTC TCAGCAACGCAACTTCCAGCTCATCGTCATCACCCACGACGACGA

		ATTCCTCAGTGGGCTGGCTGAAGTCGACCAAATTCAAGATTATTA TAACGTGCGGCGCGATGACCGTGGGTTCTCATACGTTCATCGTGT GGAATTGGACCAGACGCGACGATAACGTTAGCTTGAAGAAGAAA AATTCTTTTGGACGAGTTCTTCAGTTCCATAGCGTCATCTTTCACG CACTCTTCTAGTAAAAATTATTTTTTCATTCCATATGGACACAATC CTAAAGTCGGCGAAAAGTAAAGAAATGTTACGTAAATTTTGTGTT AAGATTCAATTCTTCCTTG
Rad51	3529	CTTCACTTGATGAGGTTTTGGGTGGTGGTGTCCTTCTAGGTGCCCT CACTGAAGTTGTTGGCACTCCAGGAGCTGGCAAGACCCAACTCTG TTTACAGTTGAGTGTCAGTGTTCAGCTGCCGAAGTGGTGCCATGG TATGGAGGGAGAAGCTGTATTCATAGATGCTGAGGGCAGCTTCAT TGCATCCAGACTACTTGAAATTGCTGACTCAGCTGTATCTCACTGT GATCAACTCCGTTCATCTGATGCTTATCAGAATCAAGACAAGGAA GAAAATGGGTCCTCCTTTACCTCAGAGTCCATACTGAAGAATGTA CATTATTTCAGATGCCTGTCACAACTCAGTGTGGCTTCCTGCGTTA AGCTGTTGCCTGAATTTATTTCCAAACATCCTCGGGTGAGGATCG TCATCATTGACTCCATCGCATTCCACCTTCGTCACGACACCCACG ACAACAGGTGTAGAAAAAGTGATCCGACTTTCGTTTCGTTTCAAA ATCTCTTAATAGTTCAGCTAAGAATGGATCACCACTTTGTCTCTAG AATAATTAATTCATTTTTGAATGTAATAATTTGAATAATCTTCTGC AAGCGTGTCCATTCTTCTATGCTGTAGAAGAAGTTTCTCCTGAAC CGATTAACTGCAGAAATCACGTAACTTTGATTGCAGGATTCGAAC AAGCCAGTTGTGCTCCATTACTCAAGATCTTATCCAACTAGCAAC GAATCATCGAGTCGCGGTTGGTAATTTCACCGTTCGTTATTGAGG AGTTACTTCATCTGTCGGATTCCAACTATCTTCTTTGCGTTTATTC AACTGAAGCCTTACATATGGATTGCAAAACTTTTGGTGGTTGTCA CCAATCAAATGACGACTCGCATCGCACATATTGACTCTAAACTCA TGAAAAATTCCGAGTGTCTGAATGACCGCGTATCGGACGATGAA AAACGTCGGCGCACTGCCACCCCTCATGAGGAAGAGATGACTGA AGGCAGAGATCTCAATTCGAGCCCAAACAAGGAGAATCTAAAAC GTAAATTGAGCGGTGCGCGGAGTGACACAATCTCTGTCAGCGAGT CTGTGACTACTAGGCGTGAAACGAACGCCACGTTGACGCCCGCCC TGGGGGAGAGTTGGGGGCACTGCGCTACTATTCGTCTGTTGCTTC GCTGGCAGCACGGAAGAAGGAATGCAACGCTTCTCAAGTCTCCC CATAAAGCCACTGCTTCCTGTTCTTATGCGGTCACAGCTGCAGGG ATAAGAGGACTATTAGAAACTTGAGTTCAGGTGCGCGCTCATGAC TTTTATTTCAACAAACCACAAAACTTTATTTCATCACAGCACCAA AAAATTTTTAGTACTGCAATAGCATATCTAGATGTAGGTATTTTC AAACTTGATCATGAGCGAACAAATAAAAATGTGCCTATCAACTA GGGATAGTTTATTGTACTAGACTGACTAACTAACGAGTCCCTCAT CATAGAGCATATATTTTAAACTAAAATCTAGCTCTAGTAACGTTC ATGGATGCAAAAACCTTGAGATGCTACATCAGGTTACTCCTGTAC AGTTTTCTTACTATGTATCGTTCAAATCAGTGGCTGATTTACACCT TAAACATTTCATGCTATGAGCTTAATGGACTTACAGCAATCCAAA CAATAAACGTAGATTAACAACGTCGCAAACTGATAAAATTTTAAT CTTCAAGAGCCGGGTCACATACTCTGCCTGCATTAATTTCAGATTT AGAAAAAATCTGAACTGGGGCGATACTCCCAAAACTCGCAAAGT GAACTTAAGTAATTACGCTATATTTTCTTGTCTATGATTGGCGGAT TTTGCTTGAAAGGCCGTCACTGATTGGCGGATGAACTTACGATCT CGCAACTTTTTTTGAAAGTATAGGCCCGGAAGCTTAACTATTACA ATGATACACATCGAAAGTGTCAATGTGAACAAACGTGGACTGTTA GGCGGCCTTCACATTAGCTTGGGAACTCTAGCAAATACGTAACAG ATTTACTCACGTACAGAAACTAGTCCTTCCCATGGTATCGTTTCAT GTAACTCAAGTATTGAACAAAGTAAAGTTATTTTTCATTCGTCCTT

		TCATTGACCTGCTACTTGTTGTCAAAATTTGTGGAACGCCCCGGCT ATTTTTTTCTCAAGCATCAAGTGATCATAACAAGAAGTCGAAGCT AGGGATACTTCCTATCATCCTATGTTATTAACAACTGAAAAAGTA ATAAATTGTTTTTTCGAAACGACTCACGTGTAGAAAGCACTAGAC CAACTGATTTCATATTTTAATTTTATTAAAATAACATTTACACGAC AAATTTAGCTCCACACTCTCCAACATCAAACCTTTCAGAATTTCG GACAGATTTACTTTTAATTATTGAGTTTAATTTTGAAATTTGAGAG ATTATATTGAGATTGATTGTAGAAAAATTGACAGCAATTCTTATT ATCGGTAACCTAAAATTGTAAAAAATTCATTAAAATTGGATGTGT GGTCACTTACTGCCCCACGCCGGGACAAGCTCTTCGTCACTATAC TCATTCGAGTCTTCACTATCGCTACGAGGGTTTTTTTTTACGTTTTT GTCACTGAATTTTGTAACTTTGCGTTCAGGACTTGAAGGTTCATCT TCCTCGGATGACGAGTCAGATGCGCCAGAAGAATTGTTTCTAACG CTTTGATGGCTCAAATTTGTCACTTTGCGTTCGGGACTTGCAGGTT TATCTTCTTCGGATGACAAGTCAGATGAATTCCTTGAAAGGGAGA TGCGTCTTTCAGCGATGTCTTTATCTCTCCCTGAATTACATAGTTT CTTTTGGGCCAAAGCAAGATTACCTGAGCCATTCACTGACTTATT TTCCAGCTTCCTTTTCTTGTTCTGCTTCGAATTTGAGGTCGACATT ACACCGAAAGAAGACTTGTTCAATCCAGATTCTGAACGTCCACTG GGCGCAGATTCATTCTGTATCGCAGTAGCTAACGCTTTCCTTGGA CTTGACGCCGACATCAAGCCAAGTGTAGACTTGCTCAATCCAGCT TTTGAACGTTCGCTGGACGCAGATTCATTTTTTGTCATCGTAAATG ACGTCTTCCTTGGACTTGATGCTGAAATTAAGCCAAGCGTAGACT TGCTCAATCCTAATTTAGAACGTTGGCAAGATACAGAATCATTTT CTGTTCTAGTAGGCAACTCCTCGTCCGAATCATCTTCCGAGTCCCT ATCTGAGTATGAAGTAGGTGACGAGGTACGCGAGAGGGAGGGAA TTACCTTGGTTTGCGCAGGTGGTGTCAATCCTTGATCCGTGTCCTC AGAATCCGACTCTTTCACCGAACTGCCGCTACTGCTGTTGTTGCTC CGATTGATTGCCTGCTCACTTTTCTTGGACGGAGAAGCTAATCGC GGCGATTTGTTTGAACTCATCAAACGTTTGTGCTCCTTTCCGGCGA GTTTCCGACTATTACCAGCAGCAAGCGGTGTAGACGTGCATGGCT G

Appendix E: Gene of interest transcript amplicon sequences

Primer set	Amplicon length (nt)	Amplicon sequence
Keap	137	GCCTCCGCTATCTTTTCTTGGAGCAGTTCCGCAGCGTAGTTTTC AATATTTTCCAGGTTTCCATCAATTTCTTCCTCTGATTCCTGTTT CCAAAGTCCCCAAGCAAGAAGTAGCTCGTAGATAATGAGATG GAGCCC
Keap2	122	CCTCCGGGTAATCGAAGCTATCAGTTCGCACACAAGCCAGAAT GCTATCTAAGACTTGTGTTTGGTAGGACTAGTGAATTTCGTTA ATTTTCTACAAAATTTTCTCCCGGTTCTTCCACTTC
Keap3	87	CTGGTCGACTGCAGGACATTTCAGGAAATGCTTGAAGGCCGCA GTTACCGTAGAGGTCCTCATCTGCAGGAACAGGTTTCCCGTTG A
Keap4	90	AACAGGTTTCCCGTTGAAGCAGCCTCGCCCACGGATTCCATTA CTATTGACTTTGTTGCTTGTTTGCAATTGCGAAGTTCTTTGAGA GCG
Cnc	149	AAATGCTATCTGTGGAGGTGGCTGTGCATAGTGTAAATATTGA AGTAGTTTCTTCCTCTAGACTGTTACTTGTATATCTTGGACAAG TTGGATTATGTGAACGTGCCTTCATGCGAGGTGATTGCCCGCT GCTCTTGATGGTCTCTTCG
Mrp4	139	GGAAGCGATAAAGACTCAGAGGCCATAGCGCCATCTAGCCCA AGAATTGACCCAGAAGTTGCGTCTGCCAGTCACGGGTCTATCG ACGAGCTTACTAAATTGTAGAAACATTTAAGGCCGTACAAATA TCGGGGTCCAT
Mrp5	141	TTTCTCAAGGGACATGGACGAAGTGGATGTTCGACTACCTATG AGCCTGGAGATGTTCCTGCAGAACATTTGGTTGGTCAGCTGTT CGCTGCTGCTTGTATGCCTCGTCTTCTGGCAGTTCCTGCCAATG TTAGTGATCCT
Mrp5-2	104	CCTTCACGACCAAGTCCTCAACGTCGTCTTCCGAAGTCCCTTA ACCTTTTTCGACACGACTCCCGTGGGTCGCATCCTAAACATTTT CTCAAGGGACATGGACG
Mrp5-3	93	GGACATGGACGAAGTGGATGTTCGACTACCTATGAGCCTGGA GATGTTCCTGCAGAACATTTGGTTGGTCAGCTGTTCGCTGCTGC TTGTATG
Sod	134	GCTCTACCCATCTGACAATGAGAGCAACAGCCCAAAAGAAGC AAGCGTCCTTCTTATGGGAAGCATCCGCGGATTCTTAAAACTT CGTCAGGACCAACCACCTGTTGGTAAGACCCATATCACGGGGC ACATCA
Hsp70	77	ACCATTCCCACTAAACACTCGCAGATCTTCACCACCTACTCCG ACAACCAGCCCGGTGTGCTCATCCAGGTTTACGA
Hsp704	143	ATGACGAAAAGAGGGAGAAGCTTACAGCTGCTCGACATGAAC TGGAGACTCTCTGTGTTGAGTTAAAGGACTTGATTAGCGAGCA GAGGGACCAAGATCTGCTGGAGAACGATGAGTTGGACCAATT CATGGACGTTTGCTCG
Hsp90	139	AAACTCGTCTGATGCTCTGGACAAGATCAGATATGAATCCCTT ACAGATCCTAGCAAGCTGGACTCTGGAAAAGATCTTTTCATCA AACTCGAGCCTAACAAGAACGACCGGACTTTAACTATTCGGGA CAGTGGCATT
Fer	81	TCCTTTCTAAAATCCTGGCCTCTATATACTTGCGGTTATCACAA GCCTCTCCTTAGGTGCAAGTGAGACGAGGAATGGATC
Cat	83	CACACCTTCAAACTCGTCAACAAAGAGGGCCAGCCCGTCTACT

		GCAAGTTCCATTTCAAGACCGATCAGGGCATCAAGGAACT
Cat2	104	CGTTCGACGTCACCAAAGTTTGGCCGCACAAGGAATTCCCTCT CATTCCGGTTGGAAAAGTTAAGCTGGATCGTAACCCCAAGAAC TATTTCGCCGAGGTGGAG
Cat3	134	TCTTCATTCGGGATCCTGTGCTCTTCCCTTCCTTCATCCACACG CAGAAGAGAAACCCTGCCACGCACCTTAAGGACCCTGATATG ATGTGGGACTTCATCAGCCTGCGACCAGAGACCACTCACCAGG TGTGC
Cat4	120	CGACAACATAGCAGGCCACTTGGTTTTCGCCATGGACTTCATC CAGGACCGCGCCATCGGAAACTTCGCTAAAGTCGACGCGGGC TTCGGACAGGCCATCAGGGAACGCATCGCTAAACT
Amy	118	GGCACCTTGTCTATGAACCTGCAAACTGGAATGCCAGCTGGGA CATACTCAGAGTTGGTGACATGCGCCGATGTCACGGTCAACGG AGACGGGACAGCGCAGATCACAATTACCAACG
Lgbp	74	AGATGGCACCCTTTGATCGTGAATTCTACCTCATCCTGAACCT GGCTGTTGGTGGTGTCAATGGTTACTTCCCA
Lectin-1	122	GGCTCAGTACGTTTACCTCAACAAAGACAGGTTCCTGTACTTC AGCGCCGGAGATGCGGTCGGAGTAGCTGATGCCTCCGTCGTTT GTCAATTGGCTTAGAGAGACAGGGCACAAATGTAGT
Lectin-4	140	AAGTTCCTTGAGAGCCGTAACGTGACTTACATCTGGACCTCCG GCCGCCTCTGCGACTTCACTGGCTGCGAGACTAGGGAGGACTT GAAGCCAATCAACATCAAGGGTTGGTTCTGGTCGAACACCAAT CAGAAAATCCC
Gst	109	TCCTTGCCATCAATCCAGAGCATTGTGTACCTACCCTGGTTGAC GGCGACTTCAAGCTATGGGAAAGTCGCGCCATTTGCGCGTATC TCGTGAACCAGTACGGCAAAGA
Cyp2c9	104	CGACATTTTCCACGATTTCCAGGGCAACGTTGTCATTTTCGACA CATTCCCATTCCTGCTTCCCATAACACCGCGTTTCGTCATGGAG AAGCTCGGCGTAACAA
Cyp2c9-2	121	ATGGAGAAGCTCGGCGTAACAAAATATATGGAGGACTGCGAT ATTCTTATGAGCCACATGAAGAAATGCATAAAGGAGCACAAG GAGAGCCTCGATGTCAACAATCCTCGCGATTACATCG
Cyp2c9-3		TTTGGAGTTCCAGCAAGTGGTGGTCGACATTTTCCACGATTTCC AGGGCAACGTTGTCATTTTCGACACATTCCCATTCCTGCTTCCC ATAACACCGCGTTTCGTCATGGAGAAGCTCGGCGTAAC
Cyp18a1	72	AAGTTCTGTCGTGGCTCTTCTCTATGCCTGTCACATGAGCCCGA AGTATTGGGATGAACCGCAAGAATTTCG
Cyp2j2	149	CATCTTCACAGCACGAAATGGGGCGCTTATCCTATCGTCGTTC TTACCGAATGCAGTCTCATCAGGAAAGCCTTTCAACATCCAAG CTTTCATGGGAGGCCTGACTTCTTGTCATTCAGAATAGTGCAG CAGTATGAAAAGAGGGGAGT
Cyp2j2-2	150	CATGGGAGGCCTGACTTCTTGTCATTCAGAATAGTGCAGCAGT ATGAAAAGAGGGGAGTGGCTTTAACGGACGGGCTGTTGTGGT CAGAGAACCGACATTTTTTCCTCACTCAACTGCGGAACCTGGG CATGGGTAAGACTTTCCTCGGG
Cyp2j2-3	105	TCCCTAGACCCAGACAACCCACGCGACATCATAGATTACTACC TCCTGGAAAAAACTGAACGCGACTCCAATGGCTTCCTCACTCA TTTCGATGCTCAGCAGCTG
Cyp	72	GAAGAGATGGAGGTGAAGACGCAGCTTGTCCTCAAGCCCGCC ACTCAGCTCGCCTTTAAGTTCCACCCGAGG
Cyp2L1	141	CACTGCGCTACCCAAGACATCGAAGTGAACGGCTTCCATGTAC CCAAGGGCACGGTGCTAATGGGACACGCGGGCAGCTGTCACG GGGACCCGCAGGTCTGGCAGCGTCCCGACGAGTTCTATCCTCC GCACTTCTTGGAC

Cyp2L1-2	110	TGGATTTGGACATCGGTTTGCCATCCTCCTTTACAACCTTACCC TCCTCCCAAGAAGTTAATGTACACTCAGGTAGGCCGCGCTTTG AGAAAGGAATTGGCTGGAAATGC
Cyp3A13	141	CAGTGCATCCTCTTCATCCTGGCGGGCTACGACACCACAGCCA ACACCCTCTGCTTCGCCGCAGCGCTCCTCGCCCGTAACCCTCA ATGTCAGGCTCGTCTTAGGGCGGAGATCGCGGCCAGAATTGAG GGTAACGAGAGC
Cyp3A13-2	135	GTCGAGGCGGCTATATTCGTCATGGGTGTAGTGATTTCTCAGA TATATTATGGTGCTTCATTAATTATTTTGATAACCTATTGGACC ACAGCATGTTGCACTTCAGAGTACATGTGTTGTACCGTGACCT GACGC
Cyp3A13-3	118	GCTACGACACCACAGCCAACACCCTCTGCTTCGCCGCAGCGCT CCTCGCCCGTAACCCTCAATGTCAGGCTCGTCTTAGGGCGGAG ATCGCGGCCAGAATTGAGGGTAACGAGAGCGG
Cyp6BS1	127	CGATGTTGCGTCCAGTAAATGTTAGTTACAAAGTCTGAAATAC ATTTGATTCATTATGCCGCACCTCTGATGCGGACTGTCATCCAC AGCCAGAGACACGAGTAATGGATTGGCTTTTGTGTTCGTC
Cyp6BS1-2	94	TGCAAGTACCTCGACGCTGTACTCAGTGAAACGCTGCGCATGT ATCCTCCGGCTCATCTGTTGGAGAGAAAATGCACCAAGGACTA CAGCGTGC
Cyp6BS1-3	113	GCCTTCGTGTTTCGATGTTGCGTCCAGTAAATGTTAGTTACAAA GTCTGAAATACATTTGATTCATTATGCCGCACCTCTGATGCGG ACTGTCATCCACAGCCAGAGACACGA
Cyp2u1	142	CAGACAAGCCTCGATTAGCGTATGTAGAAGCGCTGGTGCTGGA CGTGTTGAGGCTGAGCTCTTTGGCGCCTCTGTTGATGTTTCATC GCGCCACAGATGACGTCGTATTCGAGGGCTATACCATACCACA GGACACATTGCT
Cyp2u1-2	93	ACAACCCACGCGACATCATAGATTACTACCTCCTGGAAAAAAC TGAACGCGACTCCAATGGCTTCCTCACTCATTTCGATGCTCAG CAGCTGT
Cyp2u1-3	113	CTGGACGTGTTGAGGCTGAGCTCTTTGGCGCCTCTGTTGATGTT TCATCGCGCCACAGATGACGTCGTATTCGAGGGCTATACCATA CCACAGGACACATTGCTGATGGCGTG
Cyp2j6	142	ATCGTCTGCCATCACTCAACGACAGAAATCAACTGCAGTACGT TGACGCCGTGTTACTGGAGGTCCTGAGAGTGAGTTCACTTGCA CCCATTTTGCTGCACGCAACTACAGCTGACGTGACCTACGAGG GCTATGAAATACC
Cyp2j6-2	126	GCGGATCGAATCTGACTTCCGGGCATCGTTAAATTACGTAATT TTACGTTCACCGCATCTGAGGTTAGGTTTCAGATAGCATAGTT CTGTTCCAAAATTTATCGTCCTTCGACAAGTGACCGTTCG
Cyp2j6-3	123	GGCCATCAACCCTGAGATTCAAGCTAAAATTCACAAATGTTTG GACCAAGTTGTACCTAAAGATCGTCTGCCATCACTCAACGACA GAAATCAACTGCAGTACGTTGACGCCGTGTTACTGGA
Rad50	72	CTCCGATGAGGTGAACAAGAGAGTCCGTGAGCTTCAGCGAGA CTTGCAGCTTCCTCAGTTTAAGGAAGCCGT
Rad51	115	GATACTCCCAAAACTCGCAAAGTGAACTTAAGTAATTACGCTA TATTTTCTTGTCTATGATTGGCGGATTTTGCTTGAAAGGCCGTC ACTGATTGGCGGATGAACTTACGATCTC

Bibliography

Achard, M., M. Baudrimont, A. Boudou, and J. P. Bourdineaud, 2004. Induction of a multixenobiotic resistance protein (MXR) in the Asiatic clam *Corbicula fluminea* after heavy metal exposure. Aquat. Toxicol. 67: 347-357.

Allen, T. E. H., J. M. Goodman, S. Gutsell, and P. J. Russell, 2014. Defining molecular initiating events in the adverse outcome pathway framework for risk assessment. Chem. Res. Toxicol. 27: 2100-2112.

Altenburger, R., M. Nendza, and G. Schuurmann, 2003. Mixture toxicity and its modeling by quantitative structure-activity relationships. Environ. Toxicol. Chem. 22 (8): 1900-1915.

Amiard, J.-C., C. Amiard-Triquet, S. Barka, J. Pellerin, and P. S. Rainbow, 2006. Metallothioneins in aquatic invertebrates: Their role in metal detoxification and their use as biomarkers. Aquat. Toxicol. 76: 160-202.

Ananthasubramaniam, B., E. McCauley, K. A. Gust, A. J. Kennedy, E. B. Muller, E. J. Perkins, and R. M. Nisbet, 2015. Relating suborganismal processes to ecotoxicological and population level endpoints using a bioenergetic model. Ecol. Appl. 25: 1691-1710.

Andersen, B. S., B. M. Phillips, J. W. Hunt, V. Connor, N. Richard, and R. S. Tjeerdema, 2006. Identifying primary stressors impacting macroinvertebrates in the Salinas River (California, USA): Relative effects of pesticides and suspended particles. Environ. Pollut. 141: 402-408.

Andersen, C. L., J. L. Jensen, and T. F. Orntoft, 2004. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res. 64: 5245-5250.

Andrews, S., 2014a. FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Andrews, S., 2014b. FastQ Screen. Babraham Bioinformatics. <http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/>.

Ankley, G. T., G. P. Daston, S. J. Degitz, N. D. Denslow, R. A. Hoke, S. W. Kennedy, A. L. Miracle, E. J. Perkins, J. Snape, D. E. Tillitt, C. R. Tyler, and D. Versteeg, 2006. Toxicogenomics in regulatory ecotoxicology. Environ. Sci. Technol. 40: 4055-4065.

Ankley, G. T., R. S. Bennett, R. J. Erickson, D. J. Hoff, M. W. Hornung, R. D. Johnson, D. R. Mount, J. W. Nichols, C. L. Russom, P. K. Schmieder, J. A. Serrano, J. E. Tietge, and D. L. Villeneuve, 2010. Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environ. Toxicol. Chem. 29: 730-741.

Araya, M. T., A. Siah, D. Mateo, F. Markham, P. McKenna, G. Johnson, and F. C. J., Berthe, 2008. Selection and evaluation of housekeeping genes for haemocytes of soft-shell clams (*Mya arenaria*) challenged with *Vibrio splendidus*. J. Invert. Pathol. 99: 326-331.

Asselman, J., S. P. Glaholt, Z. Smith, G. Smagghe, C. R. Janssen, J. K. Colbourne, J. R. Shaw, and K. A. C. De Schamphelaere, 2012. Functional characterization of four metallothionein genes in *Daphnia pulex* exposed to environmental stressors. Aquat. Toxicol. 110-111: 54-65.

ATSDR, 2002. Toxicological profile for DDT, DDE, and DDD. Atlanta, GA. Department of Health and Human Services, Public Health Service.

ATSDR, 2003. Toxicological profile for pyrethrins and pyrethroids. Atlanta, GA. Department of Health and Human Services, Public Health Service.

ATSDR, 2004. Toxicological profile for copper. Atlanta, GA. Department of Health and Human Services, Public Health Service.

ATSDR, 2012. Toxicological profile for cadmium. Atlanta, GA. U.S. Department of Health and Human Services, Public Health Service.

Bariami, V., C. M. Jones, R. Poupardin, J. Vontas, and H. Ranson, 2012. Gene amplification, ABC transporters, and cytochrome P450s: Unraveling the molecular basis of pyrethroid resistance in the dengue vector, *Aedes aegypti*. PLoS Negl. Trop. Dis. 6: e1692.

Barwick, M. and W. Maher, 2003. Biotransference and biomagnification of selenium, copper, cadmium, zinc, arsenic, and lead in a temperate seagrass ecosystem from Lake Macquarie Estuary, NSW, Australia. Mar. Environ. Res. 56, 471-502.

Becker, J. and E. A. Craig, 1994. Heat-shock proteins as molecular chaperones. Eur. J. Biochem. 219: 11-23.

Beliaeff, B. and T. Burgeot, 2002. Integrated biomarker response: A useful tool for ecological risk assessment. Environ. Toxicol. Chem. 21 (6): 1316-1322.

Bogdal, C., E. Abad, M. Abalos, B. van Bavel, J. Hagberg, M. Scheringer, and H. Fiedler, 2013. Worldwide distribution of persistent organic pollutants in air, including results of air monitoring by passive air sampling in five continents. Trends Analyt. Chem. 46: 150-161.

Bolger, A. M., M. Lohse, and B. Usadel, 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30 (15): 2114-2120.

Bonizzoni, M., Y. Afrane, W. A. Dunn, F. K. Atieli, G. Zhou, D. Zhong, J. Li, A. Githeko, and G. Yan, 2012. Comparative transcriptome analysis of deltamethrin-resistant and –susceptible *Anopheles gambiae* mosquitoes from Kenya by RNA-seq. PLOS One 7 (9): e44607.

Bonmatin, J.-M., C. Giorio, V. Girolami, D. Goulson, D. P. Kreutzweiser, C. Krupke, M. Liess, E. Long, M. Marzaro, E. A. D. Mitchell, D. A. Noome, N. Simon-Delso, and A. Tapparo, 2015. Environmental fater and exposure; neonicotinoids and fipronil. Environ. Sci. Pollut. Res. 22: 35-67.

Booij, P., I. Holoubek, J. Klanova, J. Kohoutek, A. Dvorska, K. Magulova, S. Al-Zadjali, and P. Cupr, 2016. Current implications of past DDT indoor spraying in Oman. Sci. Total Environ. 550: 231-240.

Borgmann, U., W. P. Norwood, and C. Clarke, 1993. Accumulation, regulation, and toxicity of copper, zinc, lead, and mercury in *Hyalella azteca*. Hydrobiologia 259: 79-89.

Bourdineaud, J.-P., S. Cambier, S. Preveral, P. Catty, and C. Forestier, 2015. Partial inventory of ABCB and ABCC transporter genes responding to cadmium and zinc contamination in zebrafish *Danio rerio*. Environ. Anal. Toxicol. 5: 1000276.

Bustin, S. A., 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. J. Molec. Endocrinol. 29: 23-39.

Bustin, S. A. and T. Nolan, 2004. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J. Biomol. Tech. 15: 155-166.

Bustin, S. A., V. Benes, J. A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M. W. Pfaffl, G. L. Shipley, J. Vandesompele, and C. T. Wittwer, 2009. The MIQE guidelines: Minimum Information for Publication of Quantitative Real-time PCR Experiments. Clin. Chem. 55: 611-622.

Calza, S. and Y. Pawitan, 2010. Normalization of gene-expression microarray data, in: Fenyo, D. (Ed.), Computational Biology. Humana Press, New York City, pp. 37-52.

Camacho, C., T. Madden, N. Ma, T. Tao, R. Agarwala, and A. Morgulis, 2015. BLAST command line applications user manual. Bethesda, MD: National Center for Biotechnology Information.

Carpi, A., 1997. Mercury from combustion sources: A review of the chemical species emitted and their transport in the atmosphere. Water Air Soil Pollut. 98: 241-254.

Casida, J. E. and K. A. Durkin, 2013. Neuroactive insecticides: Targets, selectivity, resistance, and secondary effects. Annu. Rev. Entomol. 58: 99-117.

Chang, C.-C., A. Rahmawaty, and Z.-W. Chang, 2013. Molecular and immunological responses of the giant freshwater prawn, *Macrobrachium rosenbergii*, to the organophosphorus insecticide, trichlorfon. Aquat. Toxicol. 130-131: 18-26.

Chen, X. D., E. Culbert, V. Hebert, and J. D. Stark, 2010. Mixture effects of the nonylphenyl polyethoxylate, R-11 and the insecticide imidacloprid on population growth rate and other parameters of the crustacean, *Ceriodaphnia dubia*. Ecotoxicol. Environ. Safety 73: 132-137.

Chen, Y.-Y., J.-C. Chen, Y.-H. Kuo, Y.-C. Lin, Y.-H. Chang, H.-Y. Gong, and C.-L. Huang, 2016. Lipopolysaccharide and β-1,3-glucan binding protein (LGBP) bind to seaweed polysaccharides and activate the prophenoloxidase system in white shrimp *Litopenaeus vannamei*. Dev. Comp. Immunol. 55: 144-151.

Choi, C. Y., B. H. Min, N. N. Kim, S. H. Cho, and Y. J. Chang, 2006. Expression of HSP90, HSP70 mRNA and change of plasma cortisol and glucose during water temperature rising in freshwater adapted black porgy, *Acanthopagrus schlegeli*. J. Aquac. 19: 315-322.

Choi, Y. K., P. G. Jo, and C. Y. Choi, 2008. Cadmium affects the expression of heat shock protein 90 and metallothionein mRNA in the Pacific oyster, *Crassostrea gigas*. Comp. Biochem. Physiol. C 147: 286-292.

Cherry, D. S., J. H. Van Hassel, J. L. Farris, D. J. Soucek, and R. J. Neves, 2002. Site-specific derivation of the acute copper criteria for the Clinch River, Virginia. Hum. Ecol. Risk Assess. 8: 591-601.

Colas, F., A. Vigneron, V. Felten, and S. Devin, 2014. The contribution of a niche-based approach to ecological risk assessment: Using macroinvertebrate species under multiple stressors. Environ. Pollut. 185: 24-34.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles, 2005. Blast2GO: A universal tool for annotation, visualization, and analysis in functional genomics research. Bioinformatics 21 (18): 3674-3676.

Connon, R., H. L. Hooper, R. M. Sibly, F. Lim, L. Heckmann, D. J. Moore, H. Watanabe, A. Soetaert, K. Cook, S. J. Maund, T. H. Hutchinson, J. Moggs, W. de Coen, T. Iguchi, and A. Callaghan, 2008. Linking molecular and population stress responses in *Daphnia magna* exposed to cadmium. Environ. Sci. Technol. 42: 2181-2188.

Connon, R. E., S. Beggel, L. S. D'Abronzo, J. P. Geist, J. Pfeiff, A. V. Loguinov, C. D. Vulpe, and I. Werner, 2011. Linking molecular biomarkers with higher level condition indicators to identify effects of copper exposures on the endangered delta smelt (*Hypomesus transpacificus*). Environ. Toxicol. Chem. 30: 290-300.

Cooper, W. E., 1965. Dynamics and production of a natural population of a fresh-water amphipod, *Hyalella azteca*. Ecol. Monogr. 35: 377-394.

Correia, A. D., G. Lima, M. H. Costa, and D. R. Livingstone, 2002. Studies on biomarkers of copper exposure and toxicity in the marine amphipod *Gammarus locusta* (Crustacea): I. Induction of metallothionein and lipid peroxidation. Biomarkers 7 (5): 422-437.

Costa, P. M., S. Caeiro, C. Vale, T. A. DelValls, and M. H. Costa, 2012. Can the integration of multiple biomarkers and sediment geochemistry aid solving the complexity of sediment risk assessment? A case study with a benthic fish. Environ. Pollut. 161: 107-120.

Coumoul, X., M. Diry, and R. Barouki, 2002. PXR-dependent induction of human *CYP3S4* gene expression by organochlorine pesticides. Biochem. Pharmacol. 64: 1513-1519.

Croteau, M.-N., S. N. Luoma, and A. R. Stewart, 2005. Trophic transfer of metals along freshwater food webs: Evidence of cadmium biomagnification in nature. Limnol. Oceanogr. 50: 1511-1519.

Curtis, C. F. and A. E. P. Mnzava, 2000. Comparison of house spraying and insecticide-treated nets for malaria control. Bull. World Health Organ. 78: 1389-1400.

Daborn, P., S. Boundy, J. Yen, B. Pittendrigh, and R. ffrench-Constant, 2001. DDT resistance in *Drosophila* correlates with *Cyp6g1* over-expression and confers cross-resistance to the neonicotinoid imidacloprid. Mol. Genet. Genomics 266: 556-563.

Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil, S. Jeffers, N. Tijet, T. Perry, D. Heckel, P. Batterham, R. Feyereisen, T. G. Wilson, and R. H. ffrench-Constant, 2002. A single P450 allele associated with insecticide resistance in *Drosophila*. Science 297: 2253-2256.

Dally, H. and A. Hartwig, 1997. Induction and repair inhibition of oxidative DNA damage by nickel(II) and cadmium(II) in mammalian cells. Carcinogenesis 18: 1021-1026.

Davies, T. G. E., L. M. Field, P. N. R. Usherwood, and M. S. Williamson, 2007. DDT, pyrethrins, pyrethroids, and insect sodium channels. IUBMB Life 59: 151-162.

De Souza Pinheiro, A. and J. B. de Andrade, 2009. Development, validation and application of a SDME/BC-FID methodology for the multiresidue determination of organophosphate and pyrethroid pesticides in water. Talanta 79: 1354-1359.

Dehn, P. F., A. Allen-Mocherie, J. Karek, and A. Thenappan, 2005. Organochlorine insecticides: Impacts on human HepG2 cytochrome P4501A, 2B activities and glutathione levels. Toxciol. in Vitro 19: 261-273.

Della Torre, C., R. Zaja, J. Loncar, T. Smital, S. Focardi, and I. Corsi, 2012. Interaction of ABC transport proteins with toxic metals at the level of the gene and transport activity in the PLHC-1 fish cell line. Chem. Biol. Interact. 198: 9-17.

Dermauw, W. and T. Van Leeuwen, 2014. The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance. Insect Biochem. Mol. Biol. 45: 89-110.

Desneux, N., A. Decourtye, and J.-M. Delpuech, 2007. The sublethal effects of pesticides on beneficial arthropods. Annu. Rev. Entomol. 52: 81-106.

Dheda, K., J. F., Huggett, S. A. Bustin, M. A. Johnson, G. Rook, and A. Zumla, 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. BioTechniques 37: 112-119.

Ding, Y., P. F. Landrum, J. You, A. D. Harwood, and M. J. Lydy, 2012. Use of solid phase microextraction to estimate toxicity: Relating fiber concentrations to toxicity – Part I. Environ. Toxicol. Chem. 31: 2159-2167.

Djouaka, R. F., A. A. Bakare, O. N. Coulibaly, M. C. Akogbeta, H. Ranson, J. Hemingway, and C. Strode, 2008. Expression of the cytochrome P450s, *CYP6P3* and *CYP6M2* are significantly elevated in multiple pyrethroid resistant populations of *Anopheles gambiae s.s* from Southern Benin and Nigeria. BMC Genomics 9: 538.

Domingues, I., A. R. Agra, K. Monaghan, A. M. V. M. Soares, A. J. A. Nogueira, 2010. Cholinesterase and glutathione-*S*-transferase activities in freshwater invertebrates as biomarkers to assess pesticide contamination. Environ. Toxicol. Chem. 29: 5-18.

Dondero, F. M. Banni, A. Negi, L. Boatti, A. Dagnino, and A. Viarengo, 2011. Interactions of a pesticide/heavy metal mixture in marine bivalves: A transcriptomic assessment. BMC Genomics 12: 195.

Draghici, S., P. Khatri, A. C. Elklund, and Z. Szallasi, 2006. Reliability and reproducibility issues in DNA microarray measurements. Trends Genet. 22: 101-109.

Du, Y., L. Zhang, F. Xu, B. Huang, G. Zhang, and L. Li, 2013. Validation of housekeeping genes as internal controls for studying gene expression during Pacific oyster (*Crassostrea gigas*) development by quantitative real-time PCR. Fish Shellfish Immunol. 34: 939-945.

Duan, Y., Guttman, S.I., and Oris, J.T., 1997. Genetic differentiation among laboratory populations of *Hyalella azteca*: Implications for toxicology. Environ. Toxicol. Chem. 16: 691-695.

Duan, Y., Guttman, S.I., and Oris, J.T., 2000. Genetic structure and relationships among populations of *Hyalella azteca* and *H. Montezuma* (Crustacea: Amphipoda). J. North Am. Benthol. Soc. 19: 308-320.

Elinder, C. G., 1992. Cadmium as an environmental hazard. IARC Sci. Publ. 118: 123-132.

Environment Canada, 2013. Biological test method: Test for survival and growth in sediment and water using the freshwater amphipod *Hyalella azteca*, second ed. EPS 1/RM/33. Ottawa, Canada.

Escher, B. I. and J. L. M. Hermens, 2002. Modes of action in ecotoxicology: Their role in body burdens, species sensitivity, QSARs, and mixture effects. Environ. Sci. Technol. 36: 4201-4217.

Evers, D. C., K. A. Williams, M. W. Meyer, A. M. Scheuhammer, N. Schoch, A. T. Gilbert, L. Siegel, R. J. Taylor, R. Poppenga, and C. R. Perkins, 2011. Spatial gradients of methylmercury for breeding common loons in the Laurentian Great Lakes region. Ecotoxicol. 20: 1609-1625.

Fan, W.-H., G. Tang, C.-M. Zhao, Y. Duan, and R. Zhang, 2008. Metal accumulation and biomarker responses in *Daphnia magna* following cadmium and zinc exposure. Environ. Toxicol. Chem. 28: 305-310.

Feder, M. E. and G. E. Hofmann, 1999. Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annu. Rev. Physiol. 61: 243-282.

Feldmannova, M., K. Hilscherova, B. Marsalek, and L. Blaha, 2006. Effects of N-heterocyclic polyaromatic hydrocarbons on survival, reproduction, and biochemical parameters in *Daphnia magna*. Environ. Toxicol. 21: 425-431.

Feldmesser, E., S. Rosenwasser, A. Vardi, and S. Ben-Dor, 2014. Improving transcriptome construction in non-model organisms: Integrating manual and automated gene definition in *Emiliania huxleyi*. BMC Genomics 15: 148.

Feo, M. L., E. Eljarrat, and D. Barcelo, 2010. A rapid and sensitive analytical method for the determination of 14 pyrethroids in water samples. J. Chromatogr. A 1217: 2248-2253.

Fernandez-Alba, A. R., M. D. Hernando Guil, G. D. Lopez, and Y. Chisti, 2002. Comparative evaluation of the effects of pesticides in acute toxicity luminescence bioassays. Anal. Chim. Acta 451: 195-202.

Forbes, V. E., A. Palmqvist, and L. Bach, 2006. The use and misuse of biomarkers in ecotoxicology. Environ. Toxicol. Chem. 25: 272-280.

Forbes, V. E. and N. Galic, 2016. Next-generation ecological risk assessment: Predicting risk from molecular initiation to ecosystem service delivery. Environ. Int. 91: 215-219.

Fulton, M. H., D. W. Moore, E. F. Wirth, G. T. Chandler, P. B. Key, J. W. Daugomah, E. D. Strozier, J. Devane, J. R. Clark, M. A. Lewis, D. B. Finley, W. Ellenberg, K. J. Karnaky Jr., and G. I Scott, 1999. Assessment of risk reduction strategies for the management of agricultural nonpoint source pesticide runoff in estuarine ecosystems. Toxicol. Ind. Health 15: 201-214.

Gammon, D. W., M. A. Brown, and J. E. Casida, 1981. Two classes of pyrethroid action in the cockroach. Pestic. Biochem. Physiol. 15: 181-191.

Gan, J., S. Bondarenko, L. Oki, D. Haver, and J. X. Li, 2012. Occurrence of fipronil and its biologically active derivatives in urban residential runoff. Environ. Sci. Technol. 46: 1489-1495.

Gao, Q., L. Song, D. Ni, L. Wu, H. Zhang, and Y. Chang, 2007. cDNA cloning and mRNA expression of heat shock protein 90 gene in the haemocytes of Zhikong scallop *Chlamys farreri*. Comp. Biochem. Physiol. B. 147: 704-715.

Geiser, D. M., J. I. Pitt, and J. W. Taylor, 1998. Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. Proc. Natl. Acad. Sci. 95: 388-393.

Gellatly, K. J., K. S. Yoon, J. J. Doherty, W. Sun, B. R. Pittendrigh, and J. M. Clark, 2015. RNAi validation of resistance genes and their interactions in the highly DDT-resistant *91-R* strain of *Drosophila melanogaster*. Pestic. Biochem. Physiol. 121: 107-115.

Georgopoulos, P. G., A. Roy, M. J. Yonone-Lioy, R. E. Opiekun, and P. J. Lioy, 2001. Environmental copper: Its dynamics and human exposure issues. J. Toxicol. Environ. Health Part B 4: 341-394.

Goff, S. A., M. Vaughn, S. McKay, E. Lyons, A. E. Stapleton, D. Gessler, N. Matasci, L. Wang, M. Hanlon, A. Lenards, A. Muir, N. Merchant, S. Lowry, S. Mock, M. Helmke, A. Kubach, M. Narro, N. Hopkins, D. Micklos, U. Hilgert, M. Gonzales, C. Jordan, E. Skidmore, R. Dooley, J. Cazes, R. McLay, Z. Lu, S. Pasternak, L. Koesterke, W. H. Piel, R. Grene, C. Noutsos, K. Gendler, X. Feng, C. Tang, M. Lent, S. Kim, K. Kvilekval, B. S. Manjunath, V. Tannen, A. Stamatakis, M. Sanderson, S. M. Welch, K. A. Cranston, P. Soltis, D. Soltis, B. O'Meara, C. Ane, T. Brutnell, D. J. Kleibenstein, J. W. White, J. Leebens-Mack, M. J. Donoghue, E. P. Spalding, T. J. Vision, C. R. Myers, D. Lowenthal, B. J. Enquist, B. Boyle, A. Akoglu, G. Andrews, S. Ram, D. Ware, L. Stein, and D. Stanzione, 2011. The iPlant collaborative: Cyberinfrastructure for plant biology. Front. Plant Sci. 2: 34.

Gomez-Olivan, L. M., N. Neri-Cruz, M. Galar-Martinez, P. Vieyra-Reyes, S. Garcia-Medina, C. Razo-Estrada, O. Dublan-Garcia, and A. Y. Corral-Avitia, 2012. Assessing the oxidative stress induced by paracetamol spiked in artificial sediment on *Hyalella azteca*. Water Air Soil Pollut. 223: 5097-5104.

Goncalves, C., M. Martins, M. H. Costa, S. Caeiro, and P. M. Costa, 2013. Ecological risk assessment of impacted estuarine areas: Integrative histological and biochemical endpoints in wild Senegalese sole. Ecotoxicol. Environ. Safety 95: 202-211.

Gonzalez, E. R., and L. Watling, 2002. Redescription of *Hyalella azteca* from its type locality, Vera Cruz, Mexico (Amphipoda: Hyalellidae). J. Crustacean Biol. 22: 173-183.

Gu, X. and J. E. Manautou, 2010. Regulation of hepatic ABCC transporters by xenobiotics and in disease states. Drug Metab. Rev. 42: 482-538.

Guan, R. and W. X. Wang, 2006. Multiphase biokinetic modeling of cadmium accumulation in *Daphnia magna* from dietary and aqueous sources. Environ. Toxicol. Chem. 25: 2840-2846.

Guo, R. and J.-S. Ki, 2012. Differential transcription of heat shock protein 90 (*HSP90*) in the dinoflagellate *Prorocentrum minimum* by copper and endocrine-disrupting chemicals. Ecotoxicol. 21: 1448-1457.

Gusev, O., R. Cornette, T. Kikawada, and T. Okuda, 2011. Expression of heat shock protein-coding genes associated with anhydrobiosis in an African chironomid *Polypedilum vanderpanki*. Cell Stress Chaperones 16: 81-90.

Handy, R. D., T. S. Galloway, and M. H. Depledge, 2003. A proposal for the use of biomarkers for the assessment of chronic pollution and in regulatory toxicology. Ecotoxicol. 12: 331-343.

Hargrave, B. T., 1970. The utilization of benthic microflora by *Hyalella azteca* (Amphipoda). J. Anim. Ecol. 39: 427-437.

Harris, J. M., 1993. The presence, nature, and role of gut microflora in aquatic invertebrates: A synthesis. Microb. Ecol. 25: 195-231.

Hartwig, A., 1994. Role of DNA repair inhibition in lead- and cadmium-induced genotoxicity: A review. Environ. Health. Perspect. 102: 45-50.

Hauton, C., T. Tyrrell, and J. Williams, 2009. The subtle effects of sea water acidification on the amphipod *Gammarus locusta*. Biogeosci. 6: 1479-1489.

He, X., M. G. Chen, and Q. Ma, 2008. Activation of Nrf2 in defense against cadmium-induced oxidative stress. Chem. Res. Toxicol. 21: 1375-1383.

Heckmann, L., R. Connon, T. H. Hutchinson, S. J. Maund, R. B. Sibly, and A. Callaghan, 2006. Expression of target and reference genes in *Daphnia magna* exposed to ibuprofen. BMC Genomics 7: 175.

Heid, C. A., J. Stevens, K. J. Livak, and P. M. Williams, 1996. Real time quantitative PCR. Genome Res. 6: 986-994.

Higgins, L. G. and J. D. Hayes, 2011. The cap'n'collar transcription factor Nrf2 mediates both intrinsic resistance to environmental stressors and an adaptive response elicited by chemopreventative agents that determine susceptibility to electrophilic xenobiotics. Chem. Biol. Interact. 192: 37-45.

Hoke, R. A., P. A. Kosian, A. M. cotter, F. M. van Dermeiden, G. T. Ankley, G. L. Phipps, and D. J. Durhan, 1995. Check studies with *Hyalella azteca* and *Chrionomus tentans* in support of the development of a sediment quality criterion for dieldrin. Envrion. Toxicol. Chem. 14: 435-443.

Hook, S. E., H. L. Osborn, L. A. Golding, D. A. Spadaro, and S. L. Simpson, 2014. Dissolved and particulate copper exposure induces differing gene expression profiles and mechanisms of toxicity in the deposit feeding amphipod *Melita plumulosa*. Environ. Sci. Technol. 48: 3504-3512.

Hossain, Z. F. H., 2002. Studies on the interaction between Cd2+ ions and DNA. J. Inorg. Biochem. 90: 85-96.

Hou, D., J. He, C. Lu, L. Ren, Q. Fan, J. Wang, and Z. Xie, 2013. Distribution characteristics and potential ecological risk assessment of heavy metals (Cu, Pb, Zn, Cd) in water and sediments from Lake Dalinouer, China. Ecotoxicol. Environ. Safety 93: 135-144.

Huggett, D. B., D. Schlenk, and B. R. Griffin, 2001. Toxicity of copper in an oxic stream sediment receiving aquaculture effluent. Chemosphere 44: 361-367.

Huggett, J., K. Dheda, S. Bustin, and A. Zumla, 2005. Real-time RT-PCR normalization; strategies and considerations. Genes Immun. 6: 279-284.

Imgrund, H., 2003. Environmental fate of permethrin. Environmental Monitoring Branch, California Department of Pesticide Regulation.

Itoh, D., K. I. Tong, and M. Yamamoto, 2004. Molecular mechanism activating Nrf2-Keap1 pathway in regulation of adaptive response to electrophiles. Free Radical Biol. Med. 36: 1208-1213.

Jemec, A., T. Tisler, D. Drobne, K. Sepcic, D. Fournier, and P. Trebse, 2007. Comparative toxicity of imidacloprid, of its commercial liquid formulation, and of diazinon to a non-target arthropod, the microcrustacean *Daphnia magna*. Chemosphere 68: 1408-1418.

Jin, H. Y., A. B. Clark, R. J. C. Slebos, H. Al-Refai, J. A. Taylor, T. A. Kunkel, M. A. Resnick, and D. A. Gordenin, 2003. Cadmium is a mutagen that acts by inhibiting mismatch repair. Nat. Genet. 34: 326-329.

Jin, Y., R. Chen, L. Sun, W. Wang, L. Zhou, W. Liu, and Z. Fu, 2009. Enantioselective induction of estrogen-responsive gene expression by permethrin enantiomers in embryo-larval zebrafish. Chemosphere 74: 1238-1244.

Jones, C. M., H. K. Toe, A. Sanou, M. Namountougou, A. Hughes, A. Diabate, R. Dabire, F. Simard, and H. Ranson, 2012. Additonal selection for insecticide resitance in urban malaria vectors: DDT resistance in *Anopheles arabiensis* from Bobo-Dioulasso, Burkina Faso. PloS ONE 7.

Kalmar, B. and L. Greensmith, 2009. Induction of heat shock proteins for protection against oxidative stress. Adv. Drug Deliv. Rev. 61: 310-318.

Kanehisa, M. and S. Goto, 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28 (1): 27-30.

Karr, J. R. (1981). Assessment of biotic integrity using fish communities. Fisheries 6 (6): 21-27.

Karunker, I., J. Benting, B. Lueke, T. Ponge, R. Nauen, E. Roditakis, J. Vontas, K. Gorman, I. Denholm, and S. Morin, 2008. Over-expression of cytochrome *CYP6CM1* is associated with high resistance to imidacloprid in the B and Q biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). Insect Biochem. Mol. Biol. 38: 634-644.

Kasai, S. and J. G. Scott, 2000. Overexpression of cytochrome P450 CYP6D1 is associated with monooxygenase-mediated pyrethroid resistance in house flies from Georgia. Pestic. Biochem. Physiol. 68: 34-41.

Kasai, S., I. S. Weerashinghe, T. Shono, and M. Yamakawa, 2000. Molecular cloning, nucleotide sequence, and gene expression of a cytochrome P450 (CYP6F1) from the pyrethroid-resistant mosquito, *Culex quinquefasciatus* Say. Insect Biochem. Mol. Biol. 30: 163-171.

Kemble, N. E., E. L. Brunson, T. J. Canfield, F. J. Dwyer, and C. G. Ingersoll, 1998. Assessing sediment toxicity from navigational pools of the upper Mississippi River using a 28-d *Hyalella azteca* test. Arch. Environ. Contam. Toxicol. 35: 181-190.

Kensler, T. W., N. Wakabayashi, and S. Biswal, 2007. Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. Annu. Rev. Pharmacol. Toxicol. 47: 89-116.

Ki, J., S. Raisuddin, K. Lee, D. Hwang, J. Han, J. Rhee, I. Kim, H. G. Park, J. Ryu, and J. Lee, 2009. Gene expression profiling of copper-induced responses in the intertidal copepod *Tigriopus japonicas* using a 6K oligochip microarray. Aquat. Toxicol. 93: 177-187.

Kim, D.-Y., L. Bovet, M. Maeshima, E. Martinoia, and Y. Lee, 2007. The ABC transporter AtPDR8 is a cadmium extrusion pump conferring heavy metal resistance. Plant J. 50: 207-218.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg, 2013.TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions, and gene fusions. Genome Biol. 14: R36.

Kozera, B. and M. Rapacz, 2013. Reference genes in real-time PCR. J. Appl. Genetics 54: 391-406.

Kramer, V. J., M. A. Etterson, M. Hecker, C. A. Murphy, G. Roesijadi, D. J. Spade, J. A. Spromberg, M. Wang, and G. T. Ankley, 2011. Adverse outcome pathways and ecological risk assessment: Bridging to population-level effects. Environ. Toxicol. Chem. 30: 64-76.

Langmead, B. and S. L. Salzberg, 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357-359.

Le, T., E. Lim, N. Hong, S. Lee, Y. S. Shim, J. R. Hwang, Y. Kim, and J. Min, 2013. Proteomic analysis in *Daphnia magna* exposed to As(III), As(V), and Cd heavy metals and their binary mixtures for screening potential biomarkers. Chemosphere 93: 2341-2348.

Lee, K., S. Raisuddin, J. Rhee, D. Hwang, I. T. Yu, Y. Lee, H. G. Park, and J. Lee, 2008. Expression of glutathione S-transferase (GST) genes in the marine copepod *Tigriopus japonicas* exposed to trace metals. Aquat. Toxicol. 89: 158-166.

Leet, J. K., S. Sassman, J. J. Amberg, A. W. Olmstead, L. S. Lee, G. T. Ankley, and M. S. Sepulveda, 2015. Environmental hormones and their impacts on sex differentiation in fathead minnows. Aquat. Toxicol. 158: 98-107.

Lenz, P. H., V. Roncalli, R. P. Hassett, L. Wu, M. C. Cieslak, D. K. Hartline, A. E. Christie, 2014. *De novo* assembly of the transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) – The dominant zooplankter of the North Atlantic Ocean. PLOS One 9 (2): e88589.

Li, C., S. Weng, Y. Chen, X. Yu, L. Lu, H. Zhang, J. He, and X. Xu, 2012a. Analysis of *Litopenaeus vannamei* transcriptome using the next-generation DNA sequencing technique. PLOS One 7 (10): e47442.

Li, J., J. Han, P. Chen, Z. Chang, Y. He, P. Liu, Q. Wang, and J. Li, 2012b. Cloning of a heat shock protein 90 (HSP90) gene and expression analysis in the ridgetail white prawn *Exopalaemon carinicauda*. Fish Shellfish Immunol. 32: 1191-1197.

Lindel, T., H. Hoffman, M. Hochgurtel, and J. R. Pawlik, 2000. Structure-activity relationship of inhibition of fish feeding by sponge-derived and synthetic pyrrole-imidazole alkaloids. J. Chem. Ecol. 26 (6): 1477-1496.

Liping, Y., N. Xuanxuan, C. Leilei, Z. Linbao, Z. Jianmin, L. Xiaoli, and W. Huifeng, 2014. Expression profiles of two small heat shock proteins and antioxidant enzyme activity in *Mytilus galloprovincialis* exposed to cadmium at environmentally relevant concentrations. Chin. J. Oceanol. Limnol. 32: 334-343.

Liu, W., W. Zheng, Y. Ma, and K. K. Liu, 2006. Sorption and degradation of imidacloprid in soil and water. J. Environ. Sci. Health Part B 41: 623-634.

Lizotte Jr., R. E., F. D. Shields Jr., J .N. Murdock, and S. S. Knight, 2012. Responses of *Hyalella azteca* and phytoplankton to a simulated agricultural runoff event in a managed backwater wetland. Chemosphere 87: 684-691.

Lobenhofer, E. K., X. Cui, L. Bennett, P. L. Cable, B. A. Merrick, G. A. Churchill, and C. A. Afshari, 2004. Exploration of low-dose estrogen effects: Identification of no observed transcriptional effect level (NOTEL). Toxicol. Pathol. 32: 482-492.

Long, Y., Q. Li, Y. Wang, and Z. Cui, 2011. MRP proteins as potential mediators of heavy metal resistance in zebrafish cells. Comp. Biochem. Physiol. Part C 153: 310-317.

Lord, J. C., K. Hartzer, M. Toutges, and B. Oppert, 2010. Evaluation of quantitative PCR reference genes for gene expression studies in *Tribolium castaneum* after fungal challenge. J. Microbiol. Methods 80: 219-221.

Lotufo, G. R., P. F. Landrum, M. L. Gedeon, E. A. Tigue, and L. R. Herche, 2000. Comparative toxicity and toxicokinetics of DDT and its major metabolites in freshwater amphipods. Envrion. Toxicol. Chem. 19: 368-379.

Lowry, J. K. and A. A. Myers, 2013. A phylogeny and classification of the Senticaudata subord. nov. (Crustacea: Amphipoda). Zootaxa 3610 (1): 1-80.

Lukancic, S., U. Zibrat, T. Mezek, A. Jerebic, T. Simcic, and A. Brancelj, 2010. Effects of exposing two non-target crustacean species, *Asellus aquaticus* L., and *Gammarus fossarum* Koch., to atrazine and imidacloprid. Bull. Environ. Contam. Toxicol. 84: 85-90.

Lundstedt, S., P. A., White, C. L. Lemieux, K. D. Lynes, I. B. Lambert, L. Oberg, P. Haglund, and M. Tysklind, 2007. Sources, fate, and toxic hazards of oxygenated polycyclic aromatic hydrocarbons (PAHs) at PAH- contaminated sites. Ambio 36 (6): 475-485.

Lyu, K., X. Zhu, R. Chen, Y. Chen, and Z. Yang, 2014. Molecular cloning of manganese superoxide dismutase gene in the cladoceran *Daphnia magna*: Effects of

microcystin, nitrite, and cadmium on gene expression profiles. Aquat. Toxicol. 148: 55-64.

Ma, Z. K. Chen, Z. Yuan, J. Bi, and L. Huang, 2013. Ecological risk assessment of heavy metals in surface sediments of six major Chinese freshwater lakes. J. Environ. Qual. 42: 341-350.

MacDonald, D. D., C. G. Ingersoll, D. E. Smorong, J. A. Sinclair, R. Lindskoog, N. Wang, C. Severn, R. Gouguet, J. Meyer, and J. Field, 2011. Baseline ecological risk assessment of the Calcasieu estuary, Louisiana: Part 2. An evaluation of the predictive ability of effects-based sediment-quality guidelines. Arch. Environ. Contam. Toxicol. 61: 14-28.

MacRae, T. H., 2010. Gene expression, metabolic regulation and stress tolerance during diapause. Cell. Mol. Life Sci. 67: 2405-2424.

Machado, J. and M. Lopes-Lima, 2011. Calcification mechanism in freshwater mussels: Potential targets for cadmium. Toxicol. Environ. Chem. 93: 1778-1787.

Maher, J. M., M. Z. Dieter, L. M. Aleksunes, A. L. Slitt, G. Guo, Y. Tanaka, G. L. Scheffer, J. Y. Chan, J. E. Manautou, Y. Chen, T. P. Dalton, M. Yamamoto, and C. D. Klaassen, 2007. Oxidative and electrophilic stress induces multidrug resistance-associated protein transporters via the nuclear factor-E2-related factor-2 transcriptional pathway. Hepatol. 46: 1597-1610.

Maitra, S., S. M. Dombrowski, L. C. Waters, and R. Ganguly, 1996. Three second chromosome-linked clustered *Cyp6* genes show differential constitutive and

barbital-induced expression in DDT-resistance and susceptible strains of *Drosophila melanogaster*. Gene 180: 165-171.

Major, K., D. J. Soucek, R. Giordano, M. J. Wetzel, F. Soto-Adames, 2013. The common ecotoxicology laboratory strain of *Hyalella azteca* is genetically distinct from most wild strains samples in Eastern North America. Environ. Toxicol. Chem. 32: 2637-2647.

Mantione, K. J., R. M. Kream, H. Kuzelova, R. Ptacek, J. Raboch, J. M. Samuel, and G. B. Stefano, 2014. Comparing bioinformatics gene expression profiling methods: Microarray and RNA-seq. Med. Sci. Monit. Basic Res. 20: 138-141.

Marioni, J. C., C. E. Mason, S. M. Mane, M. Stephens, and Y. Gilad, 2008. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 18: 1509-1517.

Mastin, B. J. and J. H. Rodgers, Jr., 2000. Toxicity and bioavailability of copper herbicides (Clerigate, Cutrine-Plus, and copper sulfate) to freshwater animals. Arch. Environ. Contam. Toxicol. 39: 445-451.

Mathews, T. and N. S. Fisher, 2008. Evaluating the trophic transfer of cadmium, polonium, and methylmercury in an estuarine food chain. Environ. Toxicol. Chem. 27: 1093-1101.

Matsuda, K., S. D. Buckingham, D. Kleier, J. J. Rauh, M. Grauso, and D. B. Sattelle, 2001. Neonicotinoids: Insecticides acting on insect nicotinic acetylcholine receptors. Trends Pharmacol. Sci. 22: 573-580.

Mauriz, E., A. Calle, J. J. Manclus, A. Montoya, A. Hildebrandt, D. Barcelo, and L. M. Lechuga, 2007. Optical immunosensors for fast and sensitive detection of DDT and related compounds in river water samples. Biosens. Bioelectron. 22: 1410-1418.

McGinnis, N., E. Ragnhildstveit, A. Veraksa, and W. McGinnis, 1998. A cap 'n' collar protein isoform contains a selective Hox repressor function. Development 125: 4553-4564.

McLoughlin, N., D. Yin, L. Maltby, R. M. Wood, and H. Yu, 2000. Evaluation of sensitivity and specificity of two crustacean biochemical biomarkers. Environ. Toxicol. Chem. 19: 2085-2092.

McNulty, E. W. F. J. Dwyer, M. R. Ellersieck, I. E. Geer, C. G. Ingersoll, and C. F. Rabeni, 1999. Evaluation of the ability of reference-toxicity tests to identify stress in laboratory populations of the amphipod *Hyalella azteca*. Environ. Toxicol. Chem. 18: 544-548.

Meij, R., 1991. The fate of mercury in coal-fired power plants and the influence of wet flue-gas desulphurization. Water Air Soil Pollut. 56: 21-33.

Miao, J., L. Chi, L. Pan, and Y. Song, 2015. Generally detected genes in comparative transcriptomics in bivalves: Toward the identification of molecular markers of cellular stress response. Envrion. Toxicol. Pharmacol. 39: 475-481.

Mistry, J., R. D. Finn, S. R. Eddy, A. Bateman, and M. Punta, 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41 (12): e121.

Mittapalli, O., X. Bai, P. Mamidala, S. P., Rajarapu, P. Bonello, and D. A. Herms, 2010. Tissue-specific transcriptomes of the exotic invasive insect pest emerald ash borer. PLoS One 5: e13708.

Mohler, J., J. P. Mahaffey, E. Deutsch, and K. Vani, 1995. Control of *Drosophila* head segment identity by the bZIP homeotic gene *cnc*. Development 121: 237-247.

Muller, K., M. Bach, H. Hartmann, M. Spiteller, and H.-G. Frede, 2002. Point- and nonpoint-source pesticide contamination in the Zwester Ohm Catchment, Germany. J. Environ. Qual. 31: 309-318.

Muller, P., E. Warr, B. J. Stevenson, P. M. Pignatelli, J. C. Morgan, A. Steven, A. E. Yawson, S. N. Mitchell, H. Ranson, J. Hemingway, M. J. I. Paine, and M. J. Donnelly, 2008. Field-caught permethrin-resistant *Anopheles gambiae* overexpress CYP6P3, a P450 that metabolises pyrethroids. PLoS Genetics 4: e1000286.

NASDA, 2014. National Pesticide Applicator Certification Core Manual, 2nd ed. National Association of State Departments of Agriculture Research Foundation, Washington, DC.

Nauen, R., U. Ebbinghaus, and K. Tietjen, 1999. Ligands of the nicotinic acetylcholine receptor as insecticides. J. Pest. Sci. 55: 608-610.

Nebeker, A. V., G. S. Schuytema, W. L. Griffis, J. A. Barbitta, and L. A. Carey, 1989. Effect of sediment organic carbon on survival of *Hyalella azteca* exposed to DDT and endrin. Environ. Toxicol. Chem. 8: 705-718.

Nett, M., H. Ikeda, and B. S. Moore, 2009. Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat. Prod. Rep. 26: 1362-1384.

Neumann, M., R. Schulz, K. Schafer, W. Muller, W. Mannheller, and M. Liess, 2002. The significance of entry routes as point and non-point sources of pesticides in small streams. Water Res. 36: 835-842.

Nguyen, T., C. S. Yang, and C. B. Pickett, 2004. The pathways and molecular mechanisms regulating Nrf2 activation in response to chemical stress. Free Radical Biol. Med. 37: 433-441.

Nikou, D., H. Ranson, and J. Hemingway, 2003. An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistance strain of the malaria vector, *Anopheles gambiae*. Gene 318: 91-102.

Nordberg, G. F., 2009. Historical perspectives on cadmium toxicology. Toxicol. Appl. Pharmacol. 238: 192-200.

Nussey, G., J. H. J. van Vuren, and H. H. du Preez, 2000. Bioaccumulation of chromium, manganese, nickel, and lead in the tissues of the moggle, *Labeo umbratus* (Cyprinidae), from Witbank Dam, Mpumalanga. Water S. A. 26: 269-284.

Nyman, A.-M., K. Schirmer, and R. Ashauer, 2014. Importance of toxicokinetics for interspecies variation in sensitivity to chemicals. Environ. Sci. Technol. 48: 5946-5954.

Octaviani, M., I. Stemmler, G. Lammel, and H. F. Graf, 2015. Atmospheric transport of persistent organic pollutants to and from the Arctic under present-day and future climate. Environ. Sci. Technol. 49: 3593-3602.

Oh, S.-H., S.-Y. Lee, C.-H. Choi, S.-H. Lee, and S.-C. Lim, 2009. Cadmium adaptation is regulated by multidrug resistance-associated protein-mediated Akt pathway and metallothionein induction. Arch. Pharm. Res. 32: 883-891.

Oppenheim, S. J., R. H. Baker, S. Simon, and R. DeSalle, 2015. We can't all be supermodels: The value of comparative transcriptomics to the study of non-model insects. Insect Mol. Biol. 24 (2): 139-154.

Pan, J., J. A. Plant, N. Voulvoulis, C. J. Oates, and C. Ihlenfeld, 2010. Cadmium levels in Europe: Implications for human health. Environ. Geochem. Health 32: 1-12.

Pan, L. and H. Zhang, 2006. Metallothionein, antioxidant enzymes, and DNA strand breaks as biomarkers of Cd exposure in a marine crab, *Charybdis japonica*. Comp. Biochem. Physiol. Part C 144: 67-75.

Parsell, D. A. and S. Lindquist, 1993. The function of heat-shock proteins in stress tolerance: Degradation and reactivation of damaged proteins. Annu. Rev. Genet. 27: 437-496.

Pauchet, Y., P. Wilkinson, M. van Munster, S. Augustin, D. Pauron, and R. H. ffrench-Constant, 2009. Pyrosequencing of the midgut transcriptome of the poplar leaf beetle *Chrysomela tremulae* reveals new gene families in Coleoptera. Insect Biochem. Mol. Biol. 39: 403-413.

Pedersen, J. A., M. A. Yeager, and I. H. Suffet, 2006. Organophosphorus insecticides in agricultural and residential runoff: Field observations and implications for total maximum daily load development. Environ. Sci. Technol. 40: 2120-2127.

Peralta-Videa, J. R., M. L. Lopez, M. Narayan, G. Saupe, and J. Gardea-Torresdey, 2009. The biochemistry of environmental heavy metal uptake by plants: Implications for the food chain. Int. J. Biochem. Cell Biol. 41: 1665-1677.

Pfaffl, M. W., A. Tichopad, C. Prgomet, and T. P. Neuvians, 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. Biotechnol. Lett. 26: 509-515.

Pinot, F., S. E. Kreps, M. Bachelet, P. Hainaut, M. Bakonyi, and B. S. Polla, 2000. Cadmium in the environment: Sources, mechanisms of biotoxicity, and biomarkers. Rev. Environ. Health. 15: 299-323.

Pollack, P., 2011. Fine chemicals: The industry and the business, 2nd ed. John Wiley & Sons, Inc., New York.

Poynton, H. University of Massachusetts Boston, Boston, MS, U.S.A., personal communication.

Poynton, H. C., J. R. Varshavsky, B. Chang, G. Cavigionlio, S. Chan, P. S. Holman, A. V. Loguinov, D. J. Bauer, K. Komachi, E. C. Theil, E. J. Perkins, O. Hughes, and C. D. Vulpe, 2007. *Daphnia magna* ecotoxicogenomics provides mechanistic insights into metal toxicity. Environ. Sci. Technol. 41: 1044-1050.

Poynton, H. C., A. V. Loguinov, J. R. Varshavsky, S. Chan, E. J. Perkins, and C. D. Vulpe, 2008a. Gene expression profiling in *Daphnia magna* Part I: Concentration-

dependent profiles provide support for the no observed transcriptional effect level. Environ. Sci. Technol. 42: 6250-6256.

Poynton, H. C., R. Zuzow, A. V. Loguinov, E. J. Perkins, and C. D. Vulpe, 2008b. Gene expression profiling in *Daphnia magna*, Part II: Validation of a copper specific gene expression signature with effluent from two copper mines in California. Environ. Sci. Technol. 42: 6257-6263.

Poynton, H. C., J. M. Lazorchak, C. A. Impellitteri, B. Blalock, M. E. Smith, K. Struewing, J. Unrine, and D. Roose, 2013. Toxicity and transcriptomic analysis in *Hyalella azteca* suggests increased exposure and susceptibility of epibenthic organisms to zinc oxide nanoparticles. Environ. Sci. Technol. 47: 9453-9460.

Poynton, H. C., J. M. Lazorchak, C. A. Impellitteri, M. E. Smith, K. Rogers, M. Patra, K. A. Hammer, H. J. Allen, and C. D. Vulpe, 2011. Differential gene expression in *Daphnia magna* suggests distinct modes of action and bioavailability for ZnO nanoparticles and Zn ions. Environ. Sci. Technol. 45: 762-768.

Protopopova, M. V., V. V. Pavlichenko, R. Menzel, A. Putschew, T. Luckenbach, and C. E. W. Steinberg, 2014. Contrasting cellular stress responses of Baikalian and Palearctic amphipods upon exposure to humic substances: Environmental implications. Environ. Sci. Pollut. Res. 21: 14124-14137.

Puinean, A. M., I. Denholm, N. S. Millar, R. Nauen, and M. S. Williamson, 2010a. Characterisation of imidacloprid resistance mechanisms in the brown planthopper, *Nilaparvata lugens* Stal (Hemiptera: Delphacidae). Pestic. Biochem. Physiol. 97: 129-132.

Puinean, A. M., S. P. Foster, L. Oliphant, I. Denholm, L. M. Field, N. S. Millar, M. S. Williamson, and C. Bass, 2010b. Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid *Myzus persicae*. PLoS Genetics 6: e1000999.

Radonic, A., S. Thulke, I. M. Mackay, O. Landt, W. Siegert, and A. Nitsche, 2004. Guideline to reference gene selection for quantitative real-time PCR. Biochem. Biophys, Res. Comm. 313: 856-862.

Rice, P., I. Longden, and A. Bleasby, 2000. EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 16 (6): 276-277.

Rico, A. and P. J. Van den Brink, 2015. Evaluative aquatic invertebrate vulnerability to insecticides based on intrinsic sensitivity, biological traits, and toxic mode of action. Environ. Toxicol. Chem. 34: 1907-1917.

Roelofs, D., J. Marien, and N. M. van Straalen, 2007. Differential gene expression profiles associated with heavy metal tolerance in the soil insect *Orchesella cincta*. Insect Biochem. Mol. Biol. 37: 287-295.

Rosner, U., 1998. Effects of historical mining activities on surface water and groundwater- An example from northwest Arizona. Environ. Geol. 33: 224-230.

Sabater, S., H. Guasch, M. Ricart, A. Romani, G. Vidal, C. Klunder, and M. Schmitt-Jansen, 2007. Monitoring the effect of chemicals on biological communities: The biofilm as an interface. Anal. Bioanal. Chem. 387: 1425-1434.

Sachs, A. B., 1993. Messenger RNA degradation in eukaryotes. Cell 74: 413-421.

Sagripanti, J.-L. and K. H. Kraemer, 1989. Site-specific oxidative DNA damage at polyguanosines produced by copper plus hydrogen peroxide. J. Biol. Chem. 264: 1729-1734.

Samson, S. M., J. E. Dotzlaf, M. L. Slisz, G. W. Becker, R. M. van Frank, L. E. Veal, W. Yeh, J. R. Miller, S. W. Queener, and T. D. Ingolia, 1987. Cloning and expression of the fungal expandase/hydroxylase gene involved in cephalosporin biosynthesis. Biotechnol. 5: 1207-1214.

Saussure, H., 1958. Memoire sur divers crustaces nouveaux des Antilles et du Mexique. Memoires de la Societe de physique et d'histoire naturelle de Geneve 14: 417-496.

Scharlaken, B., D. C. de Graaf, K. Goossens, M. Brunain, L. J. Peelman, and F. J. Jacobs, 2008. Reference gene selection for insect expression studies using quantitative real-time PCR: The head of the honeybee, *Apis mellifera*, after a bacterial challenge. J. Insect Sci. 8: 33.

Schimmel, S. C., R. L. Garnas, J. M. Patrick Jr., and J. C. Moore, 1983. Acute toxicity, bioconcentration, and persistence of AC 222,705, benthiocarb, chlorpyrifos, fenvalerate, methyl parathion, and permethrin in the estuarine environment. J. Agric. Food Chem. 31: 104-113.

Schulz, R., 2004. Studies on Exposure, effects, and risk mitigation of aquatic nonpoint-source insecticide pollution: A review. J. Environ. Qual. 33: 419-448.

Schulz, R. and S. Stehle, 2008. Synthetic pyrethroids in agricultural surface waters: Exposure, effects, and risk mitigation, in: Gan, J., F. Spurlock, P. Hendley, and D.

P Weston (Eds.), Synthetic pyrethroids: Occurrence and behavior in aquatic environments. American Chemical Society, Washington, DC, pp. 171-201.

Sellami, B., A. Khazri, H. Louati, F. Boufahja, M. Dellali, D. Sheehan, P. Aissa, M. Ridha Driss, E. Mahmoudi, and H. Beyrem, 2014. Effects of permethrin on biomarkers in Mediterranean clams (*Ruditapes decussatus*). Bull. Environ. Contam. Toxicol. 92: 574-578.

Sharma, P. P., T. Gupta, E. E. Schwager, W. C. Wheeler, and C. G. Extavour, 2014. Subdivision of arthropod *cap-n-collar* expression domains is restricted to Mandibulata. EvoDevo 5.

Shaw, J. R., J. K. Colbourne, J. C. Davey, S. P. Glaholt, T. H. Hampton, C. Y. Chen, C. L. Folt, and J. W. Hamilton, 2007. Gene response profiles for *Daphnia pulex* exposed to the environmental stressor cadmium reveals novel crustacean metallothioneins. BMC Genomics 8: 477.

Shi, X., W. Guo, P. Wan, L. Zhou, X. Ren, T. Ahmat, K. Fu, and G. Li, 2013. Validation of reference genes for expression analysis by quantitative real-time PCR in *Leptinotarsa decemlineata* (Say). BMC Research Notes 6: 93.

Shuhaimi-Othman, M., and D. Pascoe, 2007. Bioconcentration and depuration of copper, cadmium, and zinc mixtures by the freshwater amphipod *Hyalella azteca*. Ecotoxicol. Environ. Safety 66: 29-35.

Simao, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov, 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31: 3210-3212.

Simon-Delso, N., V. Amaral-Rogers, L. P. Belzunces, J. M. Bonmatin, M. Chagnon, C. Downs, L. Furlan, D. W. Gibbons, C. Giorio, V. Girolami, D. Goulson, D. P. Kreutzweiser, C. H. Krupke, M. Liess, E. Long, M. McField, P. Mineau, E. A. D. Mitchell, C. A. Morrissey, D. A. Noome, L. Pisa, J. Settele, J. D. Stark, A. Tapparo, H. Van Dyck, J. Van Praagh, J. P. Van der Sluijs, P. R. Whitehorn, and M. Wiemers, 2015. Systemic insecticdes (neonicotinoids and fipronil): Trends, uses, and mode of action and metabolites. Environ. Sci. Pollut. Res. 22: 5-34.

Smith, A. G., 1991. Chlorinated hydrocarbon insecticides., in: Hayes, W. J. and E. R. Laws (Eds.), Handbook of Pesticides Toxicology. Academic Press Inc., San Diego/New York, pp. 731–915.

Soderlund, D. M., J. M. Clark, L. P. Sheets, L. S. Ullin, V. J. Piccirillo, D. Sargent, J. T. Stevens, and M. L. Weiner, 2002. Mechanisms of pyrethroid neurotoxicity: Implications for cumulative risk assessment. Toxicol. 171: 3-59.

Soetaert, A., K. van der Ven, L. N. Moens, T. VAndenbrouck, P. van Remortel, and W. M. De Coen, 2007a. *Daphnia magna* and ecotoxicogenomics: Gen expression profiles of the anti-ecdysteroidal fungicide fenarimol using energy-, molting-, and life stage-related cDNA libraries. Chemosphere 67: 60-71.

Soetaert, A., T. Vandenbrouck, K. van der Ven, M. Maras, P. van Remortel, R. Blust, and W. M. De Coen, 2007b. Molecular responses during cadmium-induced stress in *Daphnia magna*: Integration of differential gene expression with higher-level effects. Aquat. Toxicol. 83: 212-222.

Song, M. Y., J. D. Stark, and J. J. Brown, 1997. Comparative toxicity of four insecticides, including imidacloprid and tebufenozide, to four aquatic arthropods. Environ. Toxicol. Chem. 16: 2494-2500.

Spanier, K. I., F. Leese, C. Mayer, J. K. Colbourne, D. Gilbert, M. E. Pfrender, and R. Tollrian, 2010. Predator-induced defences in *Daphnia pulex*: Selection and evaluation of internal reference genes for gene expression studies with real-time PCR. BMC Molec. Biol. 11: 50.

Sparks, T. C. and R. Nauen, 2015. IRAC: Mode of action classification and insecticide resistance management. Pestic. Biochem. Physiol. 121: 122-128.

Spurlock, F. and M. Lee, 2008. Synthetic pyrethroid use patterns, properties, and environmental effects, in: Gan, J., F. Spurlock, P. Hendley, and D. P Weston (Eds.), Synthetic pyrethroids: Occurrence and behavior in aquatic environments. American Chemical Society, Washington, DC, pp. 3-25..

Stohs, S. J. and D. Bagchi, 1995. Oxidative mechanisms in the toxicity of metal ions. Free Radical Bio. Med. 18: 321-336.

Stohs, S. J., D. Bagchi, E. Hassoun, and M. Bagchi, 2000. Oxidative mechanisms in the toxicity of chromium and cadmium ions. J. Environ. Pathol. Toxicol. Oncol. 19: 201-213.

Stoughton, S. J., K. Liber, J. Culp, and A. Cessna, 2008. Acute and chronic toxicity of imidacloprid to the aquatic invertebrates *Chironomus tentans* and *Hyalella azteca*

under constant- and pulse-exposure conditions. Arch. Environ. Contam. Toxicol. 54: 662-673.

Strong, Jr., D. R, 1972. Life history variation among populations of an amphipod (*Hyalella azteca*). Ecol. 53: 1103-1111.

Struger, J. and T. Fletcher, 2007. Occurrence of lawn care and agricultural pesticides in the Don River and Humber River Watersheds (1998-2002). J. Great Lakes Res. 33: 887-905.

Strycharz, J. P., A. Lao, H. Li, X. Qiu, S. H. Lee, W. Sun, K. S. Yoon, J. J. Doherty, B. R. Pittendrigh, and J. M. Clark, 2013. Resistance in the highly DDT-resistance *91-R* strain of *Drosophila melanogaster* involves decreased penetration, increased metabolism, and direct excretion. Pestic. Biochem. Physiol. 107: 207-217.

Stuhlbacher, A. and L. Maltby, 1992. Cadmium resistance in *Gammarus pulex* (L.). Arch. Environ. Contam. Toxicol. 22: 319-324.

Suedel, B. C., E. Deaver, and J. H. Rodgers, Jr., 1996. Experimental factors that may affect toxicity of aqueous and sediment-bound copper to freshwater organisms. Arch. Environ. Contam. Toxicol. 30: 40-46.

Szczypka, M. S., J. A. Wemmie, W. S. Moye-Rowley, and D. J. Thiele, 1994. A yeast metal resistance protein similar to human cystic fibrosis transmembrane conductance regulator (CFTR) and multidrug resistance-associated protein. J. Biol. Chem. 269: 22853-22857.

Thiere, G. and R. Schulz, 2004. Runoff-related agricultural impact in relation to macroinvertebrate communities of the Lourens River, South Africa. Water Res. 38: 3092-3102.

Tiwari, S., A. D. Gondhalekar, R. S. Mann, M. E. Scharf, and L. L. Stelinski, 2011. Characterization of five *CYP4* genes from Asian citrus psyllid and their expression levels in *Candidatus* Liberibacter asiaticus-infected and uninfected psyllids. Insect Mol. Biol. 20: 733-744.

Tomizawa, M. and J. E. Casida, 2003. Selective toxicity of neonicotinoids attributable to specificity of insect and mammalian nicotinic receptors. Annu. Rev. Entomol. 48: 339-364.

Tomizawa, M. and J. E. Casida, 2005. Neonicotinoid insecticide toxicology: Mechanisms of selective action. Annu. Rev. Pharmacol. Toxicol. 45: 247-268.

Toyokuni, S. and J.-L. Sagripanti, 1996. Association between 8-hydroxy-2'-deoxyguanosine formation and DNA strand breaks mediated by copper and iron. Free Radical Biol. Med. 20: 859-864.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter, 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7 (3): 562-578.

Turusov, V., V. Rakitsky, and L. Tomatis, 2002. Dichlorodiphenyltrichloroethane (DDT): Ubiquity, persistence, and risks. Environ. Health Persp. 110: 125-128.

U.S. EPA, 1991. Title 40 CFR Parts 141, 142, and 143 – National Primary Drinking Water Regulations; Final Rule. Washington, DC: U.S. Government Printing Office.

U.S. EPA, 1998. Guidelines for Ecological Risk Assessment. Washington, DC: Risk Assessment Forum, U.S. Environmental Protection Agency. EPA/630/R-95/002F. Federal Register 63(93): 26846-26924.

U.S. EPA, 2000. Methods for measuring the toxicity and bioaccumulation of sediment-associated contaminants with freshwater invertebrates, second ed. EPA 600/R-99/064. Washington, DC.

U.S. EPA, 2002. Title XIV of the Public Health Service Act, The Safe Drinking Water Act. Washington, DC: U.S. Government Printing Office.

U.S. EPA, 2012. Title 40 CFR Appendix A to Part 423 – 126 Priority Pollutants. Washington, DC: U.S. Government Printing Office.

Van den Berg, H., 2009. Global status of DDT and its alternatives for use in vector control to prevent disease. Environ. Health Persp. 117: 1656-1663.

Van der Sluijs, J. P., V. Amaral-Rogers, L. P. Belzunces, M. F. I. J. Bijleveld van Lexmond, J.-M. Bonmatin, M. Chagnon, C. A. Downs, L. Furlan, D. W. Gibbons, C. Giorio, V. Girolami, D. Goulson, D. P. Kreutzweiser, C. Krupke, M. Liess, E. Long, M. McField, P. Mineau, E. A. D. Mitchell, C. A. Morrissey, D. A. Noome, L. Pisa, J. Settele, N. Simon-Delso, J. D. Stark, A. Tapparo, H. Van Dyck, J. van Praagh, P. R. Whitehorn, and M. Wiemers, 2015. Conclusions of the Worldwide

Integrated Assessment on the risks of neonicotinoids and fipronil to biodiversity and ecosystem functioning. Environ. Sci. Pollut. Res. 22: 148-154.

Van Dijk, T. C., M. A. Van Staalduinen, and J. P. Van der Sluijs, 2013. Macro-invertebrate decline in surface water polluted with imidacloprid. PLoS ONE 8: e62374.

Van Straalen, N. M., 1994. Biodiversity of ecotoxicological responses in animals. Neth. J. Zool. 441/442: 112-129.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman, 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3 (7): research0034-research0034.11.

Veraksa, A., N. McGinnis, X. Li, J. Mohler, and W. McGinnis, 2000. Cap 'n' collar B cooperates with a small Maf subunit to specify pharyngeal development and suppress Deformed homeotic function in the *Drosophila* head. Development 127: 4023-2037.

Verbost, P. M., G. Flik, R. A. C. Lock, and S. E. W. Bonga, 1987. Cadmium inhibition of Ca-2+ uptake in rainbow trout gills. Am. J. Phys. 253: R216-R221.

Verbost, P. M., G. Flik, P. K. T. Pang, R. A. C. Lock, and S. E. W. Bonga, 1988. Cadmium inhibits plasma-membrane calcium transport. J. Mem. Biol. 102: 97-104.

Verbost, P. M., G. Flik, P. K. T. Pang, R. A. C. Lock, and S. E. W. Bonga, 1989. Cadmium inhibition of the erythrocyte Ca-2+ pump – A molecular interpretation. J. Biol. Chem. 264: 5613-5615.

Vinken, M., 2013. The adverse outcome pathway concept: A pragmatic tool in toxicology. Toxicol. 312: 158-165.

Vinodhini, R. and M. Narayanan, 2008. Bioaccumulation of heavy metals in organs of fresh water fish *Cyprinus carpio* (Common carp). Int. J. Environ. Sci. Tech. 5: 179-182.

Wamhoff, H. and V. Schneider, 1999. Photodegradation of imidacloprid. J. Agric. Food Chem. 47: 1730-1734.

Wan, Q., I. Whang, C. Y. Choi, J. lee, and J. Lee, 2011. Validation of housekeeping genes as internal controls for studying biomarkers of endocrine-disrupting chemicals in disk abalone by real-time PCR. Comp. Biochem. Physiol. C 153: 259-268.

Wang, L. and E. P. Gallagher, 2013. Role of Nrf2 antioxidant defense in mitigating cadmium-induced oxidative stress in the olfactory system of zebrafish. Toxicol. Appl. Pharmacol. 266: 177-186.

Wang, Z., M. Gerstein, and M. Snyder, 2009. RNA-seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet. 10: 57-63.

Watanabe, H., K. Kobayashi, Y. Kato, S. Oda, R. Abe, N. Tatarazako, and T. Iguchi, 2008. Transcriptome profiling in crustaceans as a tool for ecotoxicogenomics: *Daphnia magna* DNA microarray. Cell Biol. Toxicol. 24: 641-647.

Watras, C. J., R. C. Back, S. Halvorsen, R. J. M. Hudson, K. A. Morrison, and S. P. Wente, 1998. Sci. Total Environ. 219: 183-208.

Wauchope, R. D., 1978. The pesticide content of surface water draining from agricultural fields - A review. J. Environ. Qual. 7: 459-472.

Wellborn, G. A., R. Cothran, and S. Bartholf, 2005. Life history and allozyme diversification in regional ecomorphs of the *Hyalella azteca* (Crustacea: Amphipoda) species complex. Biol. J. Linn. Soc. 84: 161-175.

Werner, I., and R. Nagel, 1997. Stress proteins HSP60 and HSP70 in three species of amphipods exposed to cadmium, diazinon, dieldrin, and fluoranthene. Environ. Toxicol. Chem. 16 (11): 2393-2403.

Wester, R. C., H. W. Smith, and H. Chasis, 1992. *In vitro* percutaneous absorption of cadmium from water and soil into human skin. Fundam. Appl. Toxicol. 19: 1-5.

Weston, D. P., R. W. Holmes, J. You, and M. J. Lydy, 2005. Aquatic toxicity due to residential use of pyrethroid insecticides. Environ. Sci. Technol. 39: 9778-9784.

Weston, D. P., H. C. Poynton, G. A. Wellborn, M. J. Lydy, B. J. Blalock, M. S. Sepulveda, and J. K Colbourne, 2013. Multiple origins of pyrethroid insecticide resistance across the species complex of a nontarget aquatic crustacean, *Hyalella azteca*. Proc. Natl. Acad. Sci. 110 (41): 16532-16537.

Witt, J. D. S. and Hebert, P. D. N., 2000. Cryptic species diversity and evolution in the amphipod genus *Hyalella* within central glaciated North America: A molecular phylogenetic approach. Can. J. Fish Aquat. Sci. 57: 687-698.

Zeng, V., K. E. Villanueva, B. S. Ewen-Campen, F. Alwes, W. E. Browne, and C. G. Extavour, 2011. *De novo* assembly and characterization of a maternal and developmental transcriptome for the emerging model crustacean *Parthyale hawaiensis*. BMC Genomics 12: 581.

Zhang, D., J. Ma, J. Jiang, L. Qiu, C. Zhu, T. Su, Y. Li, K. Wu, and S. Jiang, 2010. Molecular characterization and expression analysis of lipopolysaccharide and β-

1,3-glucan-binding protein (LGBP) from pearl oyster *Pinctada fucata*. Mol. Biol. Rep. 37: 3335-3343.

Zhang, W., C. Wu, K. Mai, Q. Chen, and W. Xu, 2011. Molecular cloning, characterization, and expression analysis of heat shock protein 90 from Pacific abalone, *Haliotis discus hannai* Ino in response to dietary selenium. Fish Shellfish. Immunol. 30: 280-286.

Zhang, Y., Y. Liu, J. Zhang, Y. Guo, and E. Ma, 2015. Molecular cloning and mRNA expression of heat shock protein genes and their response to cadmium stress in the grasshopper *Oxya chinensis*. PLoS One: e0131244.

Zhang, Z.-Y., X.-Y. Y., D.-L. Wang, H.-J. Yan, and X.-J. Liu, 2010. Acute toxicity to zebrafish of two organophosphates and four pyrethroids and their binary mixtures. Pest Manag. Sci. 66: 84-89.

Zhao, S. and R. D. Fernald, 2005. Comprehensive algorithm for quantitative real-time polymerase chain reaction. J. Comput. Biol. 12: 1047-1064.