

Fall 12-2008

**Gene Expression Profiles as Molecular Indicators of Dissolved
Oxygen Stress in Grass Shrimp *Palaemonetes pugio* Holthuis
1949**

Tiandao Li
University of Southern Mississippi

Follow this and additional works at: <https://aquila.usm.edu/dissertations>



Part of the [Aquaculture and Fisheries Commons](#), and the [Marine Biology Commons](#)

Recommended Citation

Li, Tiandao, "Gene Expression Profiles as Molecular Indicators of Dissolved Oxygen Stress in Grass Shrimp *Palaemonetes pugio* Holthuis 1949" (2008). *Dissertations*. 1147.
<https://aquila.usm.edu/dissertations/1147>

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact aquilastaff@usm.edu.

The University of Southern Mississippi

GENE EXPRESSION PROFILES AS MOLECULAR INDICATORS
OF DISSOLVED OXYGEN STRESS IN GRASS SHRIMP
PALAEMONETES PUGIO HOLTHUIS 1949

by

Tiandao Li

Abstract of a Dissertation
Submitted to the Graduate Studies Office
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

December 2008

COPYRIGHT BY

TIANDAO LI

2008

The University of Southern Mississippi

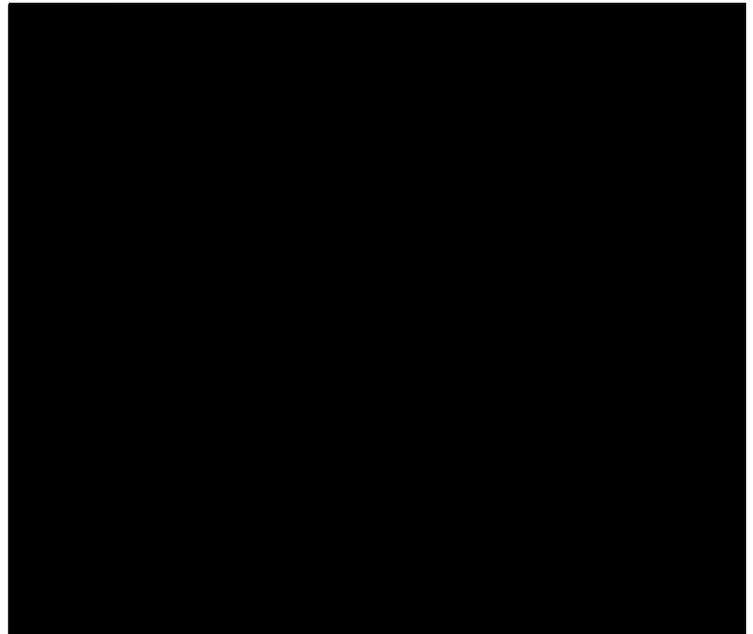
GENE EXPRESSION PROFILES AS MOLECULAR INDICATORS
OF DISSOLVED OXYGEN STRESS IN GRASS SHRIMP
PALAEMONETES PUGIO HOLTHUIS 1949

by

Tiandao Li

A Dissertation
Submitted to the Graduate Studies Office
of The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved:



December 2008

ABSTRACT

GENE EXPRESSION PROFILES AS MOLECULAR INDICATORS OF DISSOLVED OXYGEN STRESS IN GRASS SHRIMP

PALAEMONETES PUGIO HOLTHUIS 1949

by Tiandao Li

December 2008

Occurrence and severity of hypoxia is increasing in coastal and estuarine environments, and recovery of impacted habitats and living resources is slow. Detection of early biological effects of hypoxia is needed for timely remedial action to be taken. The overall objectives of this research was to develop molecular indicators of dissolved oxygen stress to assess the biological impact of hypoxia in coastal estuaries and validate their use through a combination of laboratory and field studies. To achieve these goals, grass shrimp, *Palaemonetes pugio*, oxygen-sensitive and hypoxia-tolerant species abundant in estuarine systems, were exposed to hypoxia under controlled laboratory conditions, and significant changes in gene expression were identified. Grass shrimp were collected from hypoxic field sites to evaluate if these hypoxia-responsive genes can be used as indicators of dissolved oxygen (DO) stress in the aquatic environment.

Hypoxia inducible factor 1 α (HIF-1 α), a key transcription factor that controls a variety of cellular and systemic homeostatic responses to hypoxic stress, was successfully cloned and characterized in crustaceans using RT-PCR and RACE. Grass shrimp HIF-1 α protein shows a high level of conservation with other HIF-1 α proteins in the bHLH, PAS, ODD, and TAD domains. Phylogenetic analysis indicates that grass shrimp and vertebrate HIFs belong to distinct clades within the HIF protein family. HIF mRNA levels were not responsive to chronic or cyclic hypoxia.

Six libraries of expressed sequence tags (ESTs) were constructed by suppression subtractive hybridization (SSH) from the grass shrimp exposed to environmental stress: moderate (DO 2.5 mg/L) and severe (1.5 mg/L) hypoxia, cyclic hypoxia (1.5 \rightarrow 7 mg/L),

contaminant-induced stress (pyrene and copper), and biological stress (molt). Gene Ontology (GO) analysis of libraries showed several genes that were present in only one library suggesting that their expression may be stressor specific. The molting process was accompanied by changes in expression of many genes not found in the hypoxia/copper/pyrene libraries. The resulting annotated transcripts were used to design and construct a cDNA microarray to measure the expression changes in response to hypoxia conditions.

The microarrays were used to examine differentially expressed genes in hypoxic vs. normoxic groups at 6 (H6), 12 (H12), 24 (H24), 48 (H48), 120 (H120), and 240 (H240) hours exposure to chronic hypoxia. Cluster analysis showed two response patterns, composed of an up- (including H6, H24, and H120) and down-regulated (including H12, H48, and H240) dominated cluster. Changes in gene expression are dynamic and transient. There is no differentially expressed gene up- or down-regulated common to all six groups.

Differentially expressed genes were determined in hypoxic vs. normoxic groups after 1, 2, 5 and 10 days exposure to cyclic hypoxia. Sampling on each day was conducted at two different time series, one in the morning (representing low DO, CA) and one in the afternoon (representing high DO, CP). There are distinct differences between the number and identity of specific genes that are significantly down- or up-regulated in shrimp collected at the low DO and high DO points of the cyclic DO cycle. Only a few genes are differentially expressed in grass shrimp exposed to cyclic hypoxia in the field relative to those collected from a normoxic reference site.

In conclusion, grass shrimp HIF is constitutively expressed and not induced by chronic and cyclic hypoxia exposures in both laboratory and field studies. Some differentially expressed genes appear unique at certain time points during laboratory and field exposures. However, changes of significant genes are too dynamic to serve as biomarkers of hypoxia stress in grass shrimp. Gene expression changes of grass shrimp in response to cyclic hypoxia conditions are not only dependent on the duration of exposure but also on the time of day.

ACKNOWLEDGEMENTS

Being at Gulf Coast Research Laboratory of The University of Southern Mississippi has been a truly enriching and fulfilling experience for me. There are many people I would like to thank who have made my life and research here enjoyable.

I would like to express my sincere gratitude to my advisor, Dr. Marius Brouwer, for his overall guidance and support throughout the course of this research. His enthusiasm and passion for doing great research, and interests in so many aspects of biology is unavoidably infectious, and encouraging me to pursue my project. I would also like to thank my committee members, Dr. Erik A. Carlson, Dr. Chet F. Rakocinski, and Dr. Shiao Y. Wang, for willingness to serve on the committee and assistance at various points throughout the project.

I am immensely appreciative of the help from Thea Brouwer throughout the project on many molecular protocols, techniques, and analysis. I would like to thank Steve Manning and Kevin Ryan for their assistance with developing and setting up all grass shrimp collection, culture, and exposure. I am grateful for the support from Nancy Brown-Peterson, Arthur Karels, and Adam Kuhl for field sample collection. I would also like to thank my fellow graduate students and lab technicians for their daily support, general advice and feedback: Rachel Ryan, Jonathan A. Roling, Walter Grater, Christy King, and Laura Hendon, and many others - thanks for being there and making those frustrating days bearable. I would also like to acknowledge the EcoArray Inc. for constructing the libraries.

Finally, I would like to thank my parents for their enormous and unconditional love, encouragement, and support for my life. Special thanks go to my grandparents. Thanks for their understanding that I couldn't stay with them at their last moment. Most importantly, I would like to thank my wife, Bin Zhang, for her love, understanding, and sacrifice for these years. Her unwavering optimism and constant encouragement made this work possible. Little thanks to my daughter, Emma G. Li, for her beaming smile that I get to go home every night, and her little understanding that Daddy can't play with her

sometimes. I always treasure our special moment called “Happy Family”. Words can’t describe the importance of my family, no matter where we go or what we do I know we will always be there for each other.

This dissertation is dedicated to my wife, my parents, and my grandparents for their endless love, patience, and support.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
LIST OF ILLUSTRATIONS	viii
LIST OF TABLES	x
CHAPTER	
I. INTRODUCTION	1
Gulf of Mexico and Hypoxia	1
Effects of Chronic Hypoxia	3
Hypoxia Inducible Factor (HIF)	7
Grass Shrimp	13
Hypothesis and Objectives	15
II. HYPOXIA INDUCIBLE FACTOR, GSHIF, OF THE GRASS SHRIMP	
<i>PALAEMONETES PUGIO</i> HOLTHUIS 1949: MOLECULAR	
CHARACTERIZATION AND RESPONSE TO HYPOXIA	18
Abstract	18
Introduction	19
Materials and Methods	22
Results	27
Discussion	31
Conclusion	35
III. BIOINFORMATIC ANALYSIS OF EXPRESSED SEQUENCE TAGS	
FROM GRASS SHRIMP <i>PALAEMONETES PUGIO</i> EXPOSED TO	
ENVIRONMENTAL STRESSORS	37
Abstract	37
Introduction	38
Materials and Methods	40

Results	47
Discussion	54
IV. CHRONIC HYPOXIA	71
Abstract	71
Introduction	72
Materials and Methods	74
Results	85
Discussion	108
V. CYCLIC HYPOXIA	117
Abstract	117
Introduction	118
Materials and Methods	120
Results	121
Discussion	144
VI. FIELD EXPOSURES	152
Abstract	152
Introduction	152
Materials and Methods	153
Results	154
Discussion	160
VII. SUMMARY	163
REFERENCES	169

LIST OF ILLUSTRATIONS

Figure

1. Amino acid sequence of grass shrimp Trachealess. 29
2. Amino acid sequence of grass shrimp HIF. 29
3. Neighbor joining tree derived from 21 HIF amino acid sequences. 30
4. Changes in grass shrimp HIF expression in response to chronic hypoxia exposure (2.5 ppm) at Day 3, 7, and 14. 31
5. Overview of Clontech PCR-Select cDNA Subtraction Procedure. 44
6. Pipeline of PTA modules. 44
7. Amino acid sequence of grass shrimp hemocyanin. 50
8. Venn diagram of up-regulated transcripts in three DO treatments. 55
9. Venn diagram of down-regulated transcripts in three DO treatments. 55
10. Ratios of background subtracted red (A) or green (B) mean intensities from different slides for self hybridization. 89
11. Individual intensities of red (A) and green (B) channel at different hybridization concentrations. 90
12. Hierarchical clustering using the differentially expressed genes ($p < 0.05$) in hypoxia vs. normoxia at the same time points. 106
13. Venn diagram of chronic hypoxia vs. normoxia at the same time points, down-regulated (A) and up-regulated genes (B) genes. 107
14. Pie chart of KEGG pathways. 109
15. Dissolved oxygen measured during cyclic DO exposure. 123
16. Hierarchical clustering using the differentially expressed genes ($p < 0.05$) in hypoxia vs. normoxia at the same time points. 132
17. Distribution of differentially expressed genes during cyclic (low) exposure into different GO categories. Only GO terms with at least two genes assigned to are shown. 134

18.	Pie chart of KEGG pathways.	145
19.	Map of collection sites. Top: Weeks Bay located near Mobile Bay's east- ern shore. Center: Weeks Creek, cyclic hypoxic site. Bottom: Weeks Bay Mouth, normoxic location.	155
20.	Dissolved Oxygen and Salinity at WBM from September 6 to 13, 2006. . .	157
21.	Dissolved Oxygen and Salinity at WC from September 6 to 11, 2006. . . .	158

LIST OF TABLES

Table

1.	Primers designed for grass shrimp RT-PCR and RACE.	23
2.	Sequencing Summary.	49
3.	Distribution of putative transcripts into different GO categories.	52
4.	Summary of unique genes found only in particular libraries.	56
5.	Primers used for grass shrimp PCR. 11, 47, 52, 59, 124, 173, and 197 are clone names for trachealess and hypoxia-inducible factor 1 alpha.	77
6.	Loop design for comparison of normoxic (N) and chronic hypoxic (H) samples (0, 5, and 10 days; 6, 12, 24, and 48 hours). There were 3 pooled samples for each time point (n=3), for a total of 39 arrays.	81
7.	Summary of qPCR primers.	86
8.	Differentially expressed genes during chronic hypoxia exposures at the same time points.	91
9.	Differentially expressed genes in hypoxia vs. normoxia at the same time points.	93
10.	Distribution of differentially expressed genes into different GO categories. Only GO terms with at least two genes assigned to are shown.	100
11.	Summary of qPCR.	110
12.	Loop design for comparison of normoxic (N) and cyclic hypoxic (H) sam- ples (0, 1, 2, 5 and 10 days). There were 3 pooled samples for each time point (n=3). This design was used twice. Once for samples collected at low DO and once for samples collected at high DO, for a total of 54 arrays.	121
13.	Differentially expressed genes in cyclic hypoxia vs. normoxia at the same time points.	124
14.	Differentially expressed genes during cyclic hypoxia exposures at the same time points ($p < 0.05$).	131

15.	Distribution of differentially expressed genes during cyclic (low) exposure into different GO categories. Only GO terms with at least two genes assigned to are shown.	137
16.	Distribution of differentially expressed genes during cyclic (high) exposure into different GO categories. Only GO terms with at least two genes assigned to are shown.	140
17.	List of differentially expressed genes from field samples.	159

CHAPTER I

INTRODUCTION

Gulf of Mexico and Hypoxia

The Gulf of Mexico, bordered by North, Central and South America, is the ninth largest body of water in the world. The total area of the Gulf of Mexico is approximately 615,000 square miles (1.6 million square km), and almost half of it is shallow intertidal waters. It connects with the Atlantic Ocean via the Florida Strait between the U.S. and Cuba, and with the Caribbean Sea through the Yucatan Channel between Mexico and Cuba. The Gulf's eastern, north, and northwestern shores lie within the states of Alabama, Florida, Louisiana, Mississippi and Texas. This coastline spans 1,680 miles (2,700 km), receiving water from thirty-three major rivers that drain 31 states. Many important industries along the coast include petrochemical processing and storage, shipping, paper manufacture, and tourism. Another important commercial activity is fishing; major catches include various fishes, oysters, crabs, and shrimp (NCAT, 1999).

The Gulf owes its great biodiversity and productivity to abundant nutrients such as nitrogen, phosphorus, and silica, which enhance the growth of marine life. Nutrients provided through upwelling encourage the growth of algae, a basic element of the Gulf food chain. More algae mean more plankton in the euphotic zone, more shrimp and more fish. However, algal production is a good thing only up to a point. The volume of nutrients now delivered to the Gulf due to human-related activities is causing algae to grow too fast, with deadly consequences for these waters. When these plants die, their organic material sinks to the bottom waters and is decomposed by bacteria, consuming oxygen in the process, which lead to hypoxia, and in extreme cases, anoxia (Water Marks, 2004). A number of factors including natural ones can cause oxygen depletion in estuarine ecosystems. The contributions of anthropogenic pollution and eutrophication to this phenomenon are of the major concern.

Nutrients that fuel the Gulf of Mexico hypoxia are mostly derived from fertilizers applied in the watershed of the Mississippi River (Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2001). Generally, excess nutrients lead to increased algal production and availability of organic carbon within an ecosystem, a process known as eutrophication, which is recognized as one of the major water quality concerns in the Gulf of Mexico (EPA, 2003). There are multiple sources of excessive nutrients in watersheds, both point and non-point, and the transport and delivery of these nutrients is a complex process which is controlled by a range of factors. These include chemistry, ecology, hydrology, geomorphology of the various portions of a watershed and that of the receiving system. Both the near-coastal hydrodynamics that generate water column stratification and the nutrients that fuel primary productivity contribute to the formation of hypoxic zones in Gulf of Mexico (Rabalais et al., 2002). Over the last few decades, there have been increases in the frequency, duration, and spatial extent of hypoxic events, which are regarded as one of the major factors responsible for declines in habitat quality and harvestable resources in estuarine ecosystems (Rabalais et al., 1999).

In aquatic ecosystems, hypoxia refers to a depletion of the concentration of dissolved oxygen in the water column from 7mg/L (roughly the maximum solubility of oxygen in estuarine water on an average summer day) to below 2mg/L (NCAT, 2000; NSTC, 2003). Hypoxia can easily occur in semi-closed microtidal shallow waters like coast areas of the Gulf of Mexico. Marine hypoxic areas were first observed in the 1970s. In 2003 146 hypoxic areas in the world oceans were reported (GEO, 2003, 2004). The Gulf of Mexico exhibits the largest area of hypoxia in the United States. Hypoxia has been observed in Gulf of Mexico since routine monitoring since 1980s. In the summer of 2004, the hypoxic zones, “Dead Zone” off the Louisiana Coast, measured up to 8,500 square miles (Water Marks, 2004), larger than the size of Massachusetts. These waters do not carry enough oxygen (<2mg/L) to sustain marine life. Hypoxia begins in late spring, reaches a maximum in midsummer, and disappears in the fall. Fish and other mobile

aquatic species are forced to migrate from hypoxic areas, disrupting their life cycles and increasing their loss to disease and predation. Less mobile species may simply be killed off. Hypoxia is not unique to the Gulf of Mexico; however, the Dead Zone is now considered one of the largest seasonally recurring hypoxic areas in the world (NCAT, 1999).

Effects of Chronic Hypoxia

Effects on Ecosystem and Fisheries

Hypoxia in the Gulf of Mexico has raised considerable concern throughout the United States. This annually enlarging “Dead Zone” is a major threat to the fishing industry and public health because the zone forms in the middle of the most important commercial and recreational fisheries. It is also an indication of the quality of the waters that feed the Gulf, including the Mississippi River Basin. Nitrogen is the most common driver of estuarine eutrophic conditions and currently comes from non-point sources. The nitrogen concentration in Mississippi River Basin water has doubled since the 1950s (EPA, 2003). The growing population in the Mississippi watershed increases activities that can contribute to introducing nitrogen into streams and estuaries (Rabalais et al., 1996).

The Gulf of Mexico Dead Zone is incapable of supporting most marine life, and it has become a serious threat to commercial fishing and recreation industries. The wetlands and marshes of the Gulf of Mexico support a vital ecosystem and one of the world’s greatest fisheries. The fisheries contribute 200,000 jobs and more than \$5 billion to the region annually, while providing up to 30% of the national catch (NCAT, 1999). In 2004 Texas and Louisiana landed 64% of the nation’s shrimp tails (NOAA, 2006). In recent years, hypoxia changed the fisheries significantly. During late spring and early summer, the warm coastal waters and abundant marine vegetation provide juvenile fish, shrimp, and other sea-life with nursery grounds before they move to deeper waters. However, as the Dead Zone grows offshore, it prevents migration, and alters the energy flow of the ecosystem. Hypoxia kills some species and forces others to seek alternative habitats. Even the popu-

lation of mobile species is severely impacted by their inaccessibility to offshore feeding and spawning grounds. At least for the time being, hypoxia appears to promote the in-shore and near-shore fisheries while diminishing the offshore catch. Blocked from moving to deeper waters to complete their life cycles, the species are now caught at a relatively small size, and the market value is lower for small size than for large one. At the same time, commercial fishers move away from the Gulf because of low catches of fish, shellfish, and crustaceans due to hypoxia and anoxia (Turner and Rabalais, 1994).

Hypoxia can profoundly affect the health of an ecosystem and have a variety of impacts on aquatic organisms, including reduced growth rates, decreasing size of reproductive organs, low egg counts, lack of spawning, increased susceptibility to predation, disruption of spawning and recruitment, and in extreme cases, mortality. These changes can further lead to disruption of aquatic food webs due to changes in the relative importance of organisms and pathways of carbon/energy flow, and to large reduction in the abundance, diversity, and harvest of species within affected waters. The negative effects of hypoxia can potentially make both the aquatic population and entire coastal ecosystems more susceptible to additional human and natural stressors (Justic et al., 1995).

Craig et al. (2005) found that hypoxia has resulted in 25% habitat loss for brown shrimp (*Farfantepenaeus aztecus*) on the Louisiana shelf west of the Mississippi delta, and the spatial distribution shift forced them to aggregate on the periphery of the hypoxic zone, where temperatures may be suboptimal for growth. The change in spatial distribution lead to decreases in metabolic scope due to exposure to suboptimal temperatures, low DO concentrations, and possibly lower food supply (Craig and Crowder, 2005), which can result in a long-term decrease in brown shrimp size and weight in the Gulf (Diaz and Solow, 1999).

Effects on Organisms

Many organisms in coastal environments are generally well adapted to hypoxia, however, specific adaptations vary depending on the duration of hypoxia. Some mobile organisms can detect and avoid hypoxic waters. However, such behavior, even though it increases survival, can have substantial costs, in terms of increased stress and energy consumption, potential interference with feeding, decreases in growth rate, and greater susceptibility to predation.

Avoidance is not always possible and animals must rely on physiological mechanisms to take up as much oxygen as possible from ambient environment or switch to anaerobic metabolic pathways to supply energy, or both. Some crustaceans respond to short-term hypoxia by increasing the ventilatory flow of water past the gills to accelerate the diffusion of oxygen into the blood. The process can produce a respiratory alkalosis, which increases the hemolymph pH and thereby giving rise to an adaptive increase in the pH-dependent oxygen affinity of hemocyanins. This short-term adaption to chronic hypoxia only persist for five days, and then returns to the normoxic baseline (deFur and Pease, 1988). Similarly, L-lactate, a by-product of anaerobic metabolism binds to blue crab hemocyanin resulting in increased oxygen affinity (Johnson et al., 1984).

In addition to modulation of oxygen affinity by metabolites, adjustments to prolonged exposure to hypoxia may involve modulation of hemocyanin concentration, structure and function. Long term chronic hypoxia stimulates significant changes in the concentration and structure of hemocyanin molecules of *C. sapidus* (Mangum, 1997) and shrimp *Crangon crangon* (Hagerman, 1986). Hemocyanin concentration of blue crab increases by about 40% to enhance the capacity of the hemolymph to carry oxygen. There are six different types of subunits that make up the large hemocyanin molecule in *C. sapidus* (Mangum and Rainer, 1988). Three of them, subunits 3, 5, and 6, decrease their concentrations in response to chronic hypoxia. The change of hemocyanin structure can increase the oxygen affinity by favoring the more primitive subunits of hemocyanin dur-

ing chronic hypoxia in blue crab (Mangum, 1997). Active species such as crabs and shrimp are also capable of anaerobic metabolism, and utilize it as a survival strategy during hypoxic exposures.

Hypoxia can favor opportunistic species with shorter life cycles. Some algal species that affect human health, previously absent or only present in very small numbers, are now prevalent. For example, the increased incidence of the toxic algae *N. pungens* has been associated with amnesic shellfish poisoning. Human health risks increase when the toxins produced by blooms accumulate in fish and shellfish. Further they may cause problems if airborne toxins from a bloom are inhaled (NSTC, 2003).

Effects of Cyclic (Intermittent) Hypoxia

In addition to chronic hypoxia, oxygen concentrations may vary throughout the day. In most mid-Atlantic estuaries the daily oxygen concentration is associated with time of day (light) and other factors, such as tides. However a strong cyclical pattern of DO occurs in the Gulf of Mexico estuaries. The daily variations of dissolved oxygen range from super-saturation at mid-day due to photosynthesis, to hypoxia at night due to benthic respiration. Estuarine organisms are not only at risk of being subjected to chronic hypoxic conditions, but also face increases in amplitude and frequency of hypoxia/normoxic cycles (Gupta et al., 1996).

Hypoxia and hypoxic/normoxic cycles can cause severe organ damage in mammalian species through the generation of reactive oxygen species (ROS), including superoxide, peroxide, and hydroxyl radicals. During hypoxia, oxidative phosphorylation is inhibited, cellular ATP becomes depleted, and reducing equivalents accumulate in the mitochondrial electron transport chain, a condition known as reductive stress (Dawson et al., 1993; Czyzyk-Krzeska, 1997). The reduced electron carriers are capable of forming the superoxide radical directly from oxygen still available in the tissues or when oxygen is reintroduced. The formation of reactive oxygen species promotes cell killing by oxidation

of intracellular lipids, proteins and DNA, a condition known as oxidative stress. Therefore, cycles of hypoxia and reoxygenation can lead to increased reactive oxygen species and tissue injury (Dawson et al., 1993). Usually the presence of cellular antioxidant defense systems can maintain oxygen radical-induced damage at low levels. These systems include enzymes, such as superoxide dismutase (CuZnSOD in the cytosol, and MnSOD in mitochondria), catalase, glutathione (GSH) peroxidase and GSH reductase, metal-binding proteins such as metallothionein, lipophilic (β -carotene, α -tocopherol) and water soluble (ascorbic acid) vitamins, and the abundant intracellular tripeptide, glutathione. However, when these mechanisms are overwhelmed, oxidative damage ensues. Synthesis of the proteins mentioned above is controlled by the intracellular redox state (Kehrer, 1993; Yu, 1994).

Hypoxia Inducible Factor (HIF)

All organisms possess mechanisms to maintain oxygen homeostasis, which are essential for survival. In a state of hypoxia oxygen demand exceeds supply, and a physiological response is mounted which increases the capacity of blood to carry oxygen to tissues, and alters cellular metabolism, such as facilitating ATP production by anaerobic glycolysis. The hypoxia inducible factor (HIF), conserved during evolution from nematodes to flies to vertebrates, is central to adaptation to low oxygen availability and plays an essential role during evolution (Semenza, 1998). HIF regulates transcription of many genes involved in control of cellular and systemic responses to hypoxia, including breathing, vasodilation, anaerobic metabolism, erythropoiesis and angiogenesis (Bracken et al., 2003). Thus, HIF regulates both short term responses to hypoxia, such as erythropoiesis and glycolysis, and long term responses such as angiogenesis. Therefore, HIF is a master gene that controls oxygen homeostasis during embryonic development and postnatal life in both physiological and pathophysiological processes such as tumor growth and metastasis. The products of HIF target genes are known to promote increased vascularization and

glycolytic metabolism, both of which are essential for solid tumor formation. Constitutive activation of HIF has been correlated with the progression of a variety of human tumors. Prolonged HIF induction leads to the expression of genes affecting the balance between cell death and survival (Ryan et al., 1998).

The discovery of HIF was enabled by the identification of a minimal hypoxia-responsive element (HRE), A/(G)CGTG, in the 5' enhancer region of the erythropoietin gene (EPO), a hormone that stimulates erythrocyte proliferation and undergoes hypoxia-induced transcription (Semenza and Wang, 1992). Subsequent analysis identified HIF as a phosphorylation-dependent protein which binds the major groove of DNA under hypoxic conditions (Bracken et al., 2003). HIF is a heterodimer consisting of one of four hypoxia-regulated α -subunits (HIF-1 α , HIF-2 α , HIF-3 α , and HIF-4 α) and the oxygen-insensitive β -subunit (HIF-1 β , also called aryl hydrocarbon receptor nuclear translocator, or ARNT) (Luo and Shibuya, 2001). The latter is a constitutive nuclear protein which also serves as a binding partner of the dioxin/aryl hydrocarbon receptor (DR/AhR) and hence participates in the cellular response to environmental toxins. In addition, it is an obligate heterodimeric partner for HIF-1 α , and it is also required in multiple signaling pathways (Berra et al., 2001, 2003). HIF-2 α shares 48% amino acid sequence identity and some structural and biochemical similarities with HIF-1 α . However, HIF-2 α has partially different target gene specificity from HIF-1 α , and its expression is limited to the endothelium, lung, and carotid body (Tian et al., 1997). In contrast to HIF-1 α and HIF-2 α , HIF-3 α appears not to be as efficient in mediating the hypoxic response or to act as a negative regulator of hypoxia-inducible gene expression. A splice variant of mouse HIF-3 α , inhibitory PAS protein (IPAS), interacts with the amino-terminal region of HIF-1 α , prevents its DNA binding, and inhibits its activity (Makino et al., 2001). HIF-4 α identified recently in fish shares equal sequence similarity (50%) with HIF-1 α , HIF-2 α , and HIF-3 α from different vertebrate species. Both HIF-1 α and HIF-4 α proteins contain the characteristic bHLH, PAS and ODD (oxygen-dependent degradation) domains typically

found in HIF-1 α proteins (Law et al., 2006).

In mammalian systems, HIF belongs to a class of transcription factors termed the basic helix-loop-helix/Per-ARNT-Sim (bHLH/PAS) proteins, characterized by two conserved domains, bHLH and PAS. The bHLH/PAS proteins have a stereotypic structure consisting of several domains with diverse levels of conservation. All known members of bHLH/PAS family function as dimers. The bHLH domain is involved in DNA binding and dimerization, and the PAS domain in target gene specificity, transactivation, and dimerization. The bHLH signature domain consists of approximately 15 predominantly basic amino acids responsible for direct DNA binding. This region is adjacent to two amphipathic α helices, separated by a loop of variable length, which forms the primary dimerization interface between family members. The PAS domain, named after the first three identified proteins (Per (period circadian protein), ARNT (aryl hydrocarbon receptor nuclear translocator protein) and Sim (single minded protein)), encompasses 200-300 amino acids containing two loosely conserved hydrophobic regions of approximately 50 amino acids, designated PAS-A and PAS-B. This domain forms a secondary dimerization interface between family members in addition to other roles, such as ligand and chaperone binding in the dioxin receptor (DR) (Isaacs et al., 2002). Despite of not directly binding DNA, the PAS domain has also been reported to confer target gene specificity to the *Drosophila* protein Trachealess (trh) and Single minded (Sim). The functions played by PAS in HIF still remain unknown (Hansson et al., 2002; Berra et al., 2003).

Although oxygen availability regulates multiple steps in HIF-mediated transcriptional activation, the dominant control mechanism occurs through HIF-1 α . The predominant mode of HIF-1 α regulation by hypoxia occurs by post-translational modifications, such as hydroxylation, ubiquitination, and acetylation. In normoxia, the HIF-1 α subunit is constitutively synthesized, but rapidly degraded, mediated by post-translational hydroxylation of conserved proline residues within the oxygen-dependent degradation (ODD) domain, which comprises residues 401-603 in human HIF-1 α (Huang et al., 1998; Mas-

son et al., 2001) and 692-863 in *Drosophila* Sima (HIF-1) (Nambu et al., 1996; Lavista-Llanos et al., 2002). The first proline residues (Pro402 in human HIF-1 α , Pro405 in human HIF-2 α) locates at the N-terminal end of the ODD within a LXXLAP sequence motif, and the second Pro residue (Pro564 in human HIF-1 α , Pro530 in human HIF-2) resides in the C-terminal ODD. A family of iron (II)-dependent prolyl hydroxylase enzymes uses oxygen as a substrate to catalyze the hydroxylation of these critical proline residues. Because oxygen is rate limiting for their activity, these enzymes appear to function as oxygen sensors and provide a direct link between oxygen concentration and the HIF-mediated hypoxic response pathway (Bruick and McKnight, 2002; Bruick, 2003). The hydroxylated proline residues in ODD are recognized by pVHL, the multiprotein product of the von Hippel-Lindau tumor suppressor gene, which functions as an E3 ubiquitin ligase only in the presence of oxygen, and targets HIF-1 α for polyubiquitination and proteasome-dependent degradation (Semenza, 1998), resulting in a half-life of less than 5 min (Erez et al., 2004). Under hypoxic conditions, the proline residues are unmodified and degradation of HIF-1 α is blocked, allowing it to accumulate within the nucleus where, upon binding to HIF-1 β , it recognizes HREs within the promoters of hypoxia-responsive target genes (Huang et al., 1998; Masson et al., 2001; Bruick, 2003).

In addition to proline hydroxylation, an asparagine residue in the C-terminal trans-activation domain (C-TAD) of HIF-1 α is also hydroxylated under normoxic condition blocking its interaction with transcriptional coactivators such as p300, thereby inhibiting transcription of downstream HIF target genes (Lando et al., 2002a,b; Bruick, 2003). Lysine residue (Lys532) located in the ODD domain of HIF-1 α has been reported to be acetylated by an acetyl-transferase named arrest-defective-1 (ARD1), which was originally identified in defective yeast mutants in the mitotic cell cycle. Acetylation of Lys532 favors the interaction of HIF-1 α with pVHL, and thus destabilizes HIF-1 α (Jeong et al., 2002). However, other studies demonstrated that ARD1 had limited, if any, impact on the HIF signaling pathway (Bilton et al., 2005).

Additional mechanisms can influence HIF degradation. One such regulation involves p53. The p53 tumor suppressor is a homotetrameric multifunctional transcription factor induced by DNA damage and cellular stress, including hypoxia. HIF-1 α binds to p53 and the complex may recruit Mdm2 (murine double minute 2), an E3 ubiquitin ligase known to ubiquitinate p53 and mark it for proteasomal degradation. In this manner p53 provides a route for the degradation of HIF-1 α in hypoxic tumor cells. Studies demonstrate p53 primarily mediates slow hypoxic degradation of HIF-1 α that is not hydroxylated at the two proline residues, while VHL mediates rapid normoxic degradation (Hansson et al., 2002).

Modulation of transactivation domain function is a second major mechanism by which HIF-1 α activity is controlled, and the transactivation domains are repressed in normoxia but active under hypoxia. Vertebrate HIF-1 α contains two transactivation domains (TADs) responsible for recruitment of transcriptional coactivators essential for gene expression: N-TAD, the amino-terminal transactivation domain, comprised of amino acid residues 540-580 in mammals, and C-TAD, the carboxyl-terminal transactivation domain, comprised of amino acid residues 786-826 in mammals (Jiang et al., 1997; Pugh et al., 1997; Bruick and McKnight, 2002). These two domains are separated by a region termed the inhibitory domain (ID), which is responsible for normoxic repression of TAD activity.

C-TAD is inactive in normoxia due to hydroxylation of a conserved asparagine residue (N803 in HIF-1 α and N851 in HIF-2 α) via dioxygenase FIH-1 (factor inhibiting HIF) (Lando et al., 2002a,b). However, N-TAD is active in normoxia. Transcriptional activation via N-TAD can lead to robust activation of HIF-1 target genes despite the presence of oxygen through inactivation of pVHL (Berra et al., 2003). N-TAD is sufficient to maintain full-length activity without the C-TAD, whereas the C-TAD alone shows a 25-50% reduction in the absence of N-TAD. Both N-TAD and C-TAD can function independently and the N-TAD likely contributes more than the C-TAD to gene activation. It is N-TAD, not C-TAD, which is responsible for differential transcriptional activity of HIF- α

protein (Hu et al., 2007). N-TAD is highly conserved in vertebrate HIF-1 α and contains the second proline hydroxylation motif in the C-ODD. Regulation of its activity is likely to be a by-product of protein stability (Pugh et al., 1997; Bruick and McKnight, 2002). C-TAD operates independently of the ODD and is able to recruit coactivator complexes such as CBP/p300 only under hypoxic conditions (Kallio et al., 1998; Kung et al., 2000; Bruick, 2003). The molecular event that controls C-TAD activity involves the hydroxylation of an asparagine residue under normoxic conditions. The hydroxylation of asparagine blocks the interaction of C-TAD with the CBP/p300 transcriptional coactivators. Abrogation of asparagine hydroxylation under hypoxic conditions allows for the interaction of C-TAD with CBP/p300 (Lando et al., 2002a,b). The asparaginyl hydroxylase enzyme that catalyzes the reaction belongs to the same family of 2-oxoglutarate/Fe(II) dependent oxygenases as the prolyl hydroxylases that require oxygen for hydroxylation and iron(II) and ascorbate as cofactors (Lando et al., 2002a; Masson and Ratcliffe, 2003). A recent study suggested that both the N-TAD and C-TAD are important for HIF-1 α and HIF-2 α common target genes; however, N-TAD is the principal transactivation domain responsible for target gene specificity of HIF-1 α or HIF-2 α , and C-TAD alone activates some HIF-1 α /HIF-2 α common targets (Hu et al., 2007).

Both prolyl and asparaginyl hydroxylases serve as direct oxygen sensors and must be turned on to fully induce HIF in mammals (Lando et al., 2002a,b). HIF activity is thus subjected to multiple independent levels of regulation responsible for graded responses to subtle changes in oxygen concentration. On the other hand, dependence on two independent regulations ensures that the hypoxic response pathway is tightly controlled.

A third way of controlling HIF-1 α activity may involve the process of nuclear translocation. Under hypoxic conditions HIF-1 α accumulates and translocates to the nucleus and heterodimerises with HIF-1 β . Conversely, HIF-1 α is shuttled back into the cytoplasm during reoxygenation. However, nuclear translocation per se is not sufficient to upregulate gene expression, nor to protect HIF-1 α from degradation (Kallio et al., 1998).

Two nuclear localization sequences (NLSs) are constitutively activated under both normoxia and hypoxia. The N-terminal NLS is situated in the bHLH domain (17-74 in human HIF-1 α), which mediates the nuclear translocation of these proteins, and C-terminal one is located between N-TAD and C-TAD (718-721 in human HIF-1 α). Both NLSs can control the nuclear translocation of HIF-1 α , with the C-terminal NLS being more important in the nuclear translocation under both normoxia, in cells that lack VHL, and hypoxia. The control of nuclear localization is regulated at multiple levels, including stabilization and nuclear import, which may provide a mechanism to activate target genes under normoxia for the development of multiple vascular tumors of VHL syndrome (Luo and Shibuya, 2001).

Grass Shrimp

Grass shrimp can be found along the shores of the Atlantic and Gulf of Mexico of the United States. Their range extends from Maine to Texas. Their basic habitat is the salt marshes and connecting streams with low salinity (Anderson, 1985). There are five species of grass shrimp in the Gulf of Mexico (*Palaemonetes vulgaris*; *pugio*; *intermedius*; *paludosus*; *kadiakensis*), all relatively similar in morphological characteristics and most with over-lapping distribution. Because of their similarities, these species are often misidentified as *P. vulgaris*.

Palaemonetes pugio are among the most widely distributed, abundant, and conspicuous of the shallow water benthic macroinvertebrates in the estuaries of the Atlantic and Gulf Coasts. It is an important food source for many species of commercially important fish that utilize estuaries for refuge or reproduction, and has limited value as fish bait or food for cultured fish or humans (Welsh, 1975). It is also intimately involved in energy and nutrient transport between various estuarine trophic levels: primary producers, decomposers, carnivores, and detritivores (Griffitt et al., 2007).

P. pugio has been recognized as one of the most intensively studied organisms in

ecotoxicology, particularly among estuarine invertebrates. As a useful bioassay test organism (American Public Health Association, 1975), grass shrimp have been shown to can be significantly affected by a number of contaminants, including pesticides, polycyclic aromatic hydrocarbons (PAHs), and metals (Burton and Fisher, 1990; Key and Fulton, 1993; Finley et al., 1998; Key et al., 1998). Some chemical contaminants can also affect reproductive hormone function in grass shrimp (Oberdorster et al., 2000).

Dissolved oxygen regulates the distribution and abundance of grass shrimp (Harper and Reiber, 1999). In Louisiana waters, *P. pugio* are common in DO concentrations of 6 to 11 mg/L. Grass shrimp sometimes climb out of the water during periods of oxygen deficiency, especially during warm summer nights, but such attempts to avoid hypoxia can be effective only for a few hours. Grass shrimp are adapted to the low oxygen environment of the decomposer system to avoid or limit predation and competition (Welsh, 1975). Respiratory studies indicated that *P. pugio* is an oxygen regulator, with its oxygen uptake being fairly constant between pO_2 values of 150 and 50 torr (~ 9 to 3 mg/L). However, grass shrimp is neither a complete oxygen regulator nor a complete oxygen conformer. The critical oxygen pressure at which grass shrimp can no longer regulate their oxygen uptake and become dependent on anaerobic metabolism is between 30 and 35 torr (~ 1.8 mg/L) (Cochran and Burnett, 1996a). They have a limited home range and are hypoxic tolerant, and because of that, must be able to respond to hypoxia found in their immediate environment (Finley et al., 1998; Lee et al., 1998). Also they can be easily maintained in the laboratory (Oberdorster et al., 2000). In aquatic environments, water can become hypoxic as well as hypercapnic, due to increased benthic respiration, resulting in decreases in water pH. The direct result of hypercapnia is acidification of tissues and body fluids, since CO_2 is highly permeable across body surfaces. Cochran and Burnett (1996a) found carbon dioxide increases the hemocyanin oxygen affinity in the grass shrimp. Because of that, moderate hypercapnia induced by hypoxia doesn't affect grass shrimp's ability to regulate oxygen uptake. Taken together, these properties make grass

shrimp an excellent model for the search of molecular indicators of the impact of oxygen stress on estuarine organisms.

Hypothesis and Objectives

Despite the wide-spread occurrence of hypoxia in the world's aquatic ecosystems, a clear view of the impact of hypoxia on living organisms is lacking. In contrast to atmospheric oxygen, the oxygen concentration in aquatic environment varies significantly daily, seasonally, and spatially. Widely divergent organisms have the ability to adapt to variable oxygen concentrations, which suggests that mechanisms of hypoxic sensing and response may have been established early in evolutionary history. Aquatic organisms present us therefore with a unique opportunity to study the evolution, function, and regulation of oxygen dependent genes and their role in environmental adaptation.

The existence of HIF in crustacea and other invertebrates in general (with the exception of *C. elegans* and *Drosophila*) has not been demonstrated. Whether HIF in an aquatic invertebrate, if present, is structurally and functionally equivalent to HIF in vertebrates is unknown. The natural habitat of *Palaemonetes pugio* is often hypoxic, especially in summer, and grass shrimp can adapt to very low environmental oxygen levels. Since HIF has not been characterized in crustacean, this study will clone and sequence grass shrimp HIF-1 α , and test it as a potential molecular indicator of hypoxic stress in laboratory and field studies.

Several hypoxia responsive genes have been identified and characterized in *P. pugio*. These include cadmium metallothionein (AY935987), mitochondrial superoxide dismutase (AY935986), HSP70 (AY935982), cytosolic manganese superoxide dismutase (AY211084), hemocyanin (AY935988), and hypoxia-inducible factor 1 alpha (AY655698) (Brouwer et al., 2007; Li and Brouwer, 2007). In addition a custom cDNA macroarray with a limited number of 78 clones from a hypoxia-responsive SSH cDNA library has been developed (Brouwer et al., 2007). However, the relationship between a stressor and

a given effect may involve hundreds of genes simultaneously. High-throughput genomic technologies such as microarrays can perform expression profiling of a large number of genes in a single experiment, and are therefore an ideal tool for the exploration of genetic regulatory pattern changes induced by hypoxia. For this study, six EST libraries have been generated from grass shrimp exposed to environmental stress: moderate (DO 2.5 mg/L) and severe (1.5 mg/L) hypoxia, cyclic hypoxia (1.5→7 mg/L), contaminant-induced stress (pyrene and copper), and biological stress (molt). Using cDNA microarrays constructed from these cDNA libraries, genes responsive to different environmental stressors may be identified, and gene expression profiles determined in specified conditions can be used as molecular indicators of contaminant impacts on grass shrimp in coastal waters. Therefore, the major objective of this research is to develop molecular indicators of DO stress to assess the biological impact of hypoxia in coastal estuaries and validate their use through a combination of laboratory and field studies. The remaining challenge is to determine if the molecular indicators can be linked to ecologically-important parameters, such as reproductive fitness, which, if impaired, may have population-level impacts.

Hypotheses

1. HIF is present in grass shrimp under hypoxic conditions.
2. HIF expression can be used as a molecular indicator of chronic and cyclic hypoxia exposure in grass shrimp in both laboratory and field studies.
3. Gene expression profiles of hypoxia-responsive genes can be used as indicators of chronic and cyclic hypoxic stress in grass shrimp from both laboratory and field studies.

Objectives

1. Clone and sequence grass shrimp HIF.

2. Perform sequence analysis of ESTs from SSH libraries.
3. Determine if HIF can be used as an indicator of chronic and cyclic hypoxia exposure by analyzing HIF expression level using microarrays and real-time quantitative PCR.
4. Determine if expression of hypoxia responsive genes can be used as an indicator of chronic and cyclic hypoxia exposure in the laboratory and field using microarrays.
5. Validate microarray results using qPCR of selected genes.

CHAPTER II

HYPOXIA INDUCIBLE FACTOR, GSHIF, OF THE GRASS SHRIMP
PALAEMONETES PUGIO HOLTHUIS 1949: MOLECULAR CHARACTERIZATION
AND RESPONSE TO HYPOXIA

Abstract

Hypoxia inducible factor 1 α (HIF-1 α) is a key transcription factor that controls a variety of cellular and systemic homeostatic responses to hypoxic stress. Expression and function of HIF-1 α have not been studied in crustaceans, which experience wide fluctuations of oxygen tensions in their aquatic environment. Here we show that a HIF-1 α homolog, gsHIF, is present in the hypoxia-tolerant grass shrimp *Palaemonetes pugio*. Using RT-PCR and 3' and 5' RACE, we cloned a full-length gsHIF cDNA (3822bp) with an open reading frame encoding a 1057 amino acid protein. Similar to vertebrate HIF-1 α , gsHIF has one basic helix-loop-helix (bHLH) domain, two PAS domains, an oxygen-dependent degradation domain (ODD) with two proline hydroxylation motifs, and a C-terminal transactivation domain (C-TAD) with an asparagine hydroxylation motif. In addition to these conserved sequences, gsHIF has a unique 230 amino acid sequence (aa 790-1020) not found in any vertebrate HIF proteins. Phylogenetic analysis indicates that grass shrimp and vertebrate HIFs belong to distinct clades within the HIF protein family. Expression analysis shows that gsHIF is constitutively expressed under normoxic (7.5 ppm DO), moderate (2.5 ppm DO) and severe (1.5 ppm DO) hypoxic conditions. In addition to gsHIF, we cloned a fragment of a second bHLH-PAS transcription factor from the grass shrimp, which had one bHLH and two PAS domains, and an overall 68% amino acid sequence homology with *Apis mellifera* trachealess protein.

Keywords - *Palaemonetes pugio*; hypoxia; hypoxia-inducible factor; gene expression; phylogenetic classification; trachealess protein.

Introduction

All organisms possess mechanisms to maintain oxygen homeostasis which are essential for survival. In a state of hypoxia, when oxygen demand exceeds supply, a physiological response is mounted which increases the capacity of blood to carry oxygen to tissues, and alters cellular metabolism, such as facilitating ATP production by anaerobic glycolysis. The hypoxia inducible factor (HIF), conserved during evolution from worms to flies to vertebrates, is central to adaptation to low oxygen availability (Semenza, 1998). HIF regulates the transcription of many genes involved in control of cellular and short-term and long-term systemic responses to hypoxia, including glycolysis, erythropoiesis, breathing, vasodilatation, and angiogenesis. HIF controls oxygen homeostasis during embryonic development and postnatal life in physiological processes and also in pathophysiological processes such as tumor growth and metastasis (Ryan et al., 1998).

The discovery of HIF was enabled by the identification of a minimal hypoxia-responsive element (HRE), A/(G)CGTG, in the 5' enhancer region of the erythropoietin gene. Subsequent analysis identified HIF as a phosphorylation-dependent protein which binds the major groove of DNA under hypoxic conditions (Bracken et al., 2003). HIF is a heterodimer consisting of one of four hypoxia-regulated α -subunits (HIF-1 α , HIF-2 α , HIF-3 α , and HIF-4 α) and the oxygen-insensitive HIF-1 β subunit. The latter is a constitutive nuclear protein which also serves as a binding partner (so called Arnt or aryl hydrocarbon receptor nuclear translocator) of the dioxin/aryl hydrocarbon receptor (DR/AhR) and hence participates in the cellular response to environmental toxins. In addition, HIF-1 β /Arnt is required in multiple signaling pathways (Berra et al., 2001, 2003).

In mammals, HIF belongs to a class of transcription factors termed the basic helix-loop-helix/Per-Arnt-Sim (bHLH/PAS) proteins, characterized by two conserved domains, bHLH and PAS. The PAS domain was named after the first three proteins in which it was identified: Per (period circadian protein), Arnt (Ah receptor nuclear translocator protein) and Sim (single-minded protein). All known members of bHLH/PAS family function as

dimers. The bHLH domain is involved in DNA binding and dimerization, and the PAS domain in target gene specificity, transactivation, and dimerization. The bHLH signature domain consists of approximately 60 amino acids with two functionally distinct regions. The basic region, located at the N-terminal end of the domain, is involved in DNA binding and consists of 15 amino acids with a high number of basic residues. The HLH region, at the C-terminal end, functions as a dimerization domain and is constituted mainly of hydrophobic residues that form two amphipathic helices separated by a loop region of variable sequence and length. The PAS domain encompasses 200-300 amino acids containing two loosely conserved hydrophobic regions of approximately 50 amino acids, designated PAS-A and PAS-B. This domain forms a secondary dimerization interface between family members in addition to other roles, such as ligand and chaperone binding in the dioxin receptor (DR) (Isaacs et al., 2002). Despite of not directly binding DNA, the PAS domain has also been reported to confer target gene specificity to the *Drosophila* protein Trachealess (trh) and Single minded (Sim). The functions played by PAS in HIF still remain unknown (Hansson et al., 2002; Berra et al., 2003).

Although oxygen availability regulates multiple steps in HIF-mediated transcriptional activation, the dominant control mechanism occurs through HIF-1 α . In normoxia, HIF-1 α subunit is constitutively synthesized, and two proline residues in the so-called HIF oxygen-dependent degradation domain (ODD) are hydroxylated by HIF prolyl hydroxylase. This modification targets HIF for rapid degradation by the ubiquitin-proteasome pathway, resulting in a half-life of less than 5 min (Erez et al., 2004). Under hypoxic conditions, the proline residues are unmodified and degradation of HIF-1 α is blocked, allowing it to accumulate within the nucleus where, upon binding to HIF-1 β , it recognizes HREs within the promoters of hypoxia-responsive target genes (Huang et al., 1998; Makino et al., 2001; Bruick, 2003). In addition to proline hydroxylation, an asparagine residue in the C-terminal transactivation domain (C-TAD) of HIF-1 α is also hydroxylated under normoxic conditions blocking its interaction with transcriptional coactivators such

as p300, thereby inhibiting transcription of downstream HIF target genes (Lando et al., 2002a,b; Bruick, 2003).

Most of our knowledge of molecular responses to hypoxia comes from in vitro studies of terrestrial mammalian systems. Relatively little is known about molecular responses of aquatic organisms to hypoxia. The oxygen content in aquatic environment varies markedly daily, seasonally, and spatially. Due to the low oxygen content of water, increased respiration by benthic (micro)organisms stimulated by excess organic material, a condition known as eutrophication, can cause hypoxia or anoxia (Rabalais et al., 2002). It is thus not surprising that environmental oxygen levels play a significant role in the evolution of aquatic animals. They have developed various physiological and biochemical adaptations to enable survival in hypoxic and anoxic environments, including air breathing organs, specialized metabolic pathways enabling long-term anoxic survival, and modifications of hemoglobin molecules to optimize oxygen transport. At the same time, they present a unique opportunity to study the evolution, function, and regulation of oxygen dependent genes and their role in the environmental adaptation (Soitamo et al., 2001). Since widely divergent organisms have the ability to adapt to variable oxygen concentrations, mechanisms of hypoxic sensing and response may have been established early in evolutionary history.

The first full-length cDNA (3605 bp) of HIF-1 α in fish was cloned from rainbow trout *Oncorhynchus mykiss*, which encodes a protein sequence of 766 amino acids that shows a 61% similarity to human and mouse HIF-1 α (Soitamo et al., 2001). To date ~20 additional HIF-1/2/3/4 α cDNAs have been reported in several fish species (Law, 2002; Powell and Hahn, 2002; Strausberg et al., 2002; Law et al., 2006), but HIF-1(-2, -3 or -4) α has not yet been characterized in crustacea and the mechanisms and conditions by which HIF regulation occurs in hypoxic shrimp have not been elucidated. Here, we present the sequence of HIF cDNA from the grass shrimp *Palaemonetes pugio* and its response to moderate and severe hypoxia. This marsh-resident species is commonly found

in habitats that are often hypoxic, and can adapt to very low oxygen levels found in their immediate environment (Finley et al., 1998; Lee et al., 1998). In general, grass shrimp has been shown to be uniquely physiologically adapted to stressful tidal marsh habitats (Welsh, 1975). The use of a non-model, yet commonly occurring resident species in these studies allows laboratory results to be more easily related and applied to field measurements.

Materials and Methods

Extraction of Total RNA and Cloning of Full-Length cDNA

Total RNA was isolated from grass shrimp, *Palaemonetes pugio*, hepatopancreas using Stat-60 (Tel-Test, Friendswood, TX, USA) according to the manufacturer's protocol. After precipitation, RNA was stored in RNA Secure (Ambion). Single-stranded cDNA was generated from mRNA by reverse transcription (RT) with Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen) and oligo dT primers (5'-TTTTTTTTTTTTTTVN-3').

Block Maker (Henikoff et al., 1995) was used to find several highly conserved blocks in a set of HIF protein sequences from different species (*Danio rerio* AAH46875, *Mus musculus* AAC53455, *Rattus norvegicus* AAD24413, *Xenopus laevis* CAB96628, *Homo sapiens* AAC50152, *Drosophila melanogaster* AAC47303, *Caenorhabditis elegans* AAK62778, *Oncorhynchus mykiss* AAK30364, and *Gallus gallus* BAA34234), where a block is an aligned array of amino acid sequence segments without gaps that represents a highly conserved region of homologous proteins. The resulting blocks in the Blocks Database format were imported into the CODEHOP program (Rose et al., 1998) to design the forward and reverse primers for PCR amplification, HIFF1, HIFF2, HIFR1, and HIFR2 (Table 1).

PCR conditions for the primary PCR were 94°C for 2 min for 1 cycle, 35 cycles of 94°C for 30 sec, 45°C for 30 sec, and 68°C for 2 min, followed by 68°C for 7 min.

Secondary PCR cycle conditions were 94°C for 2 min for 1 cycle, 35 cycles of 94°C for 30 sec, 50°C for 30 sec, and 68°C for 1 min, followed by 68°C for 7 min. The PCR reactions using AccuPrime™ SuperMix I (Invitrogen) gave two amplification products, ~700 (hif700) and ~900bp (hif900). Bands were visualized by ethidium bromide on a 1% agarose gel, purified with QIAquick Gel Extraction Kit Protocol (Qiagen), and T/A cloned using pGEM®-T Easy Vector System (Promega), and sequenced. The ~900 bp and ~700 bp PCR products were identified as fragments of Tracheless and HIF-1 α (see Results).

Table 1: Primers designed for grass shrimp RT-PCR and RACE.

Primer Name	Nucleotide Sequences
HIFF1	5'-GGG CGG AAG GAG AAG TCC MGN GAY GCN GC-3'
HIFF2	5'-ACA CAA CGT CAC CAC CCA CYT NGA YAA RGC-3'
HIFR1	5'-GGT ACT GGC CGG TCG TCM CYT GNC CYT T-3'
HIFR2	5'-GGT CCG CAG CTG CAA ATC RTC RTC NAT-3'
900F1	5'-CAG AGG AAG GTC AAG CAG GGT CAC A-3'
900F2	5'-GCT GCC ACT TCA AGA GTT CGG GAT ATA GAG-3'
900R1	5'-CCA TCG TCA TCC CGG ACA TCG TA-3'
900R2	5'-CGC GGT CTC ATG TTG CCT CCT TTA-3'
700F1	5'-GCT TGT GAA GGG CGA GGA CGA GT-3'
700F2	5'-GCC CTG GAC TCG GAA CTC ATC AAA G-3'
700R1	5'-AAT GAT GTC GCC TTC GGT AGA GAG CAC-3'
700R2	5'-GTT AGA CAA GCA TGG CAG AGG GC-3'
F1	5'-CCA GGA AGT AGC CCA GAA TAT GAC G-3'
F2	5'-ACG CAC ACC AGA GCC ACC TAA AGC-3'
F3	5'-GCG GTA AAG ATG GAG ATG ATG GAG-3'
F4	5'-GTA GCT CAC CTC TCC AAG ATC ACC A-3'
T1	5'-GAC TAC ACA CCA GAT GAA CTG CAA GG-3'
T2	5'-GTC CCT CTA CCC CTT GTG TCA CG-3'

RACE was performed using the SMART-RACE kit (BD Bioscience) to generate the 3' and 5' ends of the Tracheless and HIF cDNA. Gene-specific primers were designed based on the initial cDNA sequences of hif700 and hif900. Primary and secondary

primers for 3' and 5' ends of hif900 were 900F1 and 900F2, and 900R1 and 900R2, respectively. For hif700, the primary and secondary primers for 3' and 5' ends were 700F1 and 700F2, and 700R1 and 700R2, respectively (Table 1). Touchdown PCR conditions for the primary PCR were 5 cycles of 94°C for 30 sec and 72°C for 3 min, 5 cycles of 94°C for 30 sec, 70°C for 30 sec, and 72°C for 3 min, 35 cycles of 94°C for 30 sec, 65°C for 30 sec, and 72°C for 3 min, followed by 72°C for 5 min. Secondary PCR cycle conditions were 40 cycles of 94°C for 30 sec, 68°C for 30 sec, and 72°C for 3 min, followed by 72°C for 5 min. The PCR products were purified and cloned and sequenced as described above.

Because the sequencing reactions gave ambiguous results for the 3' and 5' ends, additional sets of primary and nested primers were designed for trachealess and HIF, to get the full-length cDNA sequences. For trachealess the primary and secondary primers were T1 and T2. For HIF the first set of primary and secondary primers was F1 and F2, and the second set of primers was F3 and F4 (Table 1). The PCR cycling parameters for the primary and secondary PCR were the same as above. The product bands were isolated from the gel, purified and sequenced.

All primers were synthesized by Invitrogen (Carlsbad, CA, USA). DNA sequences were determined by the University of Maine Sequencing Center (Orono, ME, USA) and in house using a Beckman CEQ 8000 Genetic Analysis System. Sequence analyses and homology searches were performed using the online BLAST suite of programs (NCBI). Conserved domains in the sequences were identified using the NCBI Conserved Domains database and by alignment of grass shrimp HIF with HIF-1 α from organisms with known HIF domain structures.

Hypoxia Exposures

Grass shrimp were collected in the vicinity of Ocean Springs, Mississippi, in Davis Bayou using dip nets. Adult females and males were segregated by sex and maintained in the laboratory at 15 psu and $27 \pm 1^\circ\text{C}$ for 7 to 30 days prior to experimentation.

During acclimation and experimentation periods, grass shrimp were fed brine shrimp nauplii once daily and commercial flake food once daily. During all acclimation and experimentation periods, shrimp were held in artificial seawater (Fritz Super Salt, Fritz Industries, Mesquite, TX, USA) diluted to 15 psu.

Exposures were conducted in an intermittent flow-through system described by Manning et al. (1999). Normoxic (7.5 ppm DO) and hypoxic (2.5 ppm and 1.5 ppm DO) conditions within the treatment aquaria were established and maintained as described before (Brouwer et al., 2004, 2005, 2007; Brown-Peterson et al., 2005). The flow-through test system provided 1L every 20 minutes (resulting in 3 complete volume additions/day) to each of the 35L test aquaria using a separate water delivery partitioner for each of the normoxic and hypoxic treatments. Oxygen levels were controlled by bubbling nitrogen into a holding tank which gravity fed to the partitioner used to deliver flow-through hypoxic seawater. A 24 hour timer was used to activate a solenoid valve which controlled nitrogen introduction into the holding tank at intervals that maintained oxygen in the holding tank at a level which resulted in the desired oxygen concentration when introduced into the test aquaria. An additional partitioner provided flow-through normoxic seawater, and normoxic conditions were maintained by gently bubbling oxygen into the cells of the water partitioner prior to delivery of water to the individual aquaria. Female grass shrimp were housed individually in retention chambers constructed from 10 cm Petri dish bottoms with a 15 cm high collar of 500 μ m nylon mesh placed into 35L flow-through glass aquaria in a water bath held at $27 \pm 1^\circ\text{C}$. In all experiments, oxygen was monitored continuously in one hypoxic flow-through aquaria, and DO, temperature and salinity were measured in all flow-through aquaria once or twice daily using a YSI Model 600XLM data sonde. After 3, 7 and 14 days of exposure, 10 shrimp per treatment were sacrificed and weighed (40.2-41.7 mg). The thorax/hepatopancreas of the shrimp was stored in 1 ml RNAlater at -20°C for nucleic acid extraction.

Phylogenetic Analysis

A multiple alignment of 21 HIF-1/2/3/4 α amino acid sequences was performed with ClustalX 1.83 (Thompson et al., 1994, 1997; Jeanmougin et al., 1998). The aligned sequences were then used to calculate distances between pairs of protein sequences. The neighbor-joining method (Saitou and Nei, 1987) was applied to the distance matrix to calculate a tree using 1000 bootstrap trials to derive confidence values for the groupings in the tree. Corrections for multiple substitutions were made as described by Kimura (1983). Alignment positions where any of the sequences had gaps were excluded from the analysis. An additional tree was constructed using the maximum parsimony method implemented by Felsenstein's PHYLIP package v3.6 (Felsenstein, 1985). The aligned sequences in PHYLIP format were imported into the SEQBOOT algorithm to generate 500 data sets by bootstrap resampling (Felsenstein, 1989). The multiple data sets were used to calculate most parsimonious trees with PROTPARS. The resulting tree output file was used as input in the program CONSENSE that calculates a majority rule consensus tree with confidence intervals. *Caenorhabditis elegans* HIF-1 α was used as outgroup.

HIF Expression Analysis

HIF mRNA levels were determined using custom cDNA macroarrays printed with cDNA from 78 clones from a hypoxia-responsive suppression subtractive hybridization (SSH) cDNA library (Brouwer et al., 2005, 2007), and with cDNA from the ~700 bp gsHIF clone. gsHIF PCR products with a final concentration of 100 ng/ μ L were robotically spotted in duplicate onto neutral nylon membranes using 100 nL pins as described by Larkin et al. (2003).

Total RNA was extracted from normoxic (7.5 ppm DO) and hypoxic (2.5 ppm and 1.5 ppm DO) grass shrimp as described above. Five to eight individual shrimp, out of ten sampled, in each treatment group gave sufficient amounts of RNA for reverse transcription and labeling. Radiolabeled probes were generated by random primer labeling

of DNase-treated (DNA-free, Ambion) total RNA with [α - 33 P]dATP (2'-deoxyadenosine 5'-triphosphate). Hybridization and wash steps were performed as previously described (Larkin et al., 2003).

The membranes were exposed to a phosphor screen (Molecular Dynamics) at room temperature for 48 hr. The blots were quantitatively evaluated using a Typhoon 8600 imaging system (Molecular Dynamics). For each cDNA clone the general background of each membrane was subtracted from the average value of the duplicate spots on the membrane. Intensity values for all genes were transformed to the log base 2 and normalized to the median array intensity. HIF expression data were analyzed by conducting a one-way ANOVA across time for both DO regimens. Data were tested for homogeneity of variance and normality of distribution using SigmaStat 3.11 (SYSTAT Software, Inc. San Jose, CA, USA). Where normality test failed a Kruskal-Wallis one-way ANOVA on ranks was performed.

Results

Cloning and Sequencing of Trachealess from Grass Shrimp

Using total RNA extracted from the hepatopancreas of grass shrimp and degenerate primers designed by CODEHOP, two initial PCR products of \sim 900bp and \sim 700bp were obtained. Sequencing showed the \sim 900 bp product to be 870 bp long, with an open reading frame of 290 amino acids without start and stop codons, with a 78% sequence identity to trachealess protein from *Tribolium castaneum* (GenBank Accession no: XP_967112). 3' and 5' RACE were performed in an attempt to obtain the complete trachealess coding sequence. 3' RACE products formed one contiguous sequence with the 870 bp product with a 327 bp overlap. All 3' RACE products had a TGA stop codon in the same position. All 5' RACE products had an identical 252 bp sequence at the 3' end which formed one contiguous sequence with the 870 bp product with a 120 bp overlap. The 5' end sequences of the 5' RACE clones were ambiguous and there was no clear translation start.

Hence we show only the 84 amino acids, corresponding to the conceptual translation of the 252 bp that are identical in all 5' RACE products (Figure 1). The 501 amino acid sequence shown in this figure represents therefore a fragment of trachealess protein, with part of the N-terminal sequence (~45 amino acids) missing. It should be noted that trachealess proteins that show greatest homology with grass shrimp trachealess, *Tribolium castaneum* (67%), and *Bombyx mori* (65%; GenBank Accession no. BAA22946) trachealess proteins (Matsunami et al., 1999) are 834 and 849 amino acids long, respectively.

Cloning and Sequencing of HIF from Grass Shrimp

Sequencing showed the ~700 bp PCR product to be 673 bp long. BLASTX search revealed 49% identity with HIF-1 protein from *Tribolium castaneum* (GenBank Accession No. XP_967427). The complete HIF sequence was obtained using SMART 5' and 3' RACE. However, the first 3' RACE product did not give the full-length product. Use of new primer sets (F1/F2 and F3/F4, Table 1) designed from the sequence of the previously found 3' RACE product resulted in the full-length 3' sequence. The complete gsHIF sequence (GenBank Accession no. AY655698) was 3288 bp long, with an open reading frame encoding a 1057 amino acid protein with the initiation methionine at position 135 bp and stop codon TAG at 3308 bp and a molecular weight of 114.67 kDa (Figure 2). There is an overall 46% homology with *Tribolium castaneum* HIF-1 protein. The amino acid sequence of grass shrimp HIF is the second longest in size compared to HIF-1 α of most vertebrates (~800 AA) and *Drosophila* (1507 AA).

Molecular Phylogenetic Analyses

Neighbor joining and maximum parsimony analysis produced trees of the same topology with similar high bootstrap scores. Both methods revealed 2 well-supported clades. (1) Invertebrate (arthropod) HIFs (*Palaemonetes*, *Drosophila* and *Tribolium*) with 93% bootstrap support in the neighbor joining tree (Figure 3) and 87% in the maximum

```

1 XXXFAPEQW RQLPDGSILE LRKEKSRDAA RSRGKENYE FYELAKMLPL PPAITSOLDK
61 ASIIRLTISY LKLRDFTLHG DPPWPRDHS GTKNLKGGNM RPRTMSCMTM DIFET
121
181 LATGQPLPSP SSLGSEEGQA GSQGTMNPDV ATCMAVTSTS QHKGYDRAFC IRMKSTLTKR
241 GCHFSSCYR VVLILGHLRP QYVFSHRKS APTLMGLVAL AIALPPPSVH EVRLESDMPV
301 TRITDFRIA HCEPKVADLL DYTPDELOGR SLYPLCHGQD VDKLRKTHVD LIEKGGVMSP
361 YFRLLNKTGG YTNMQTCATV VINNKNGDEQ NIICVNYIIS RTQYDTLVMD QTQLDPALAN
421 MKRDDLDYTN PATPEPSEGS CGVVGEESSS SPRSTANTTG GGGSIGTRS PPTGGPPEQI
481 STPLRGTPSV PLDGAQIKTE VITL

```

XXX indicate N-terminal sequence upstream from FAPE is ambiguous. No distinct translation start is found.

20-72: bHLH Domain; **116-177**: PAS Domain; **296-395**: PAS Domain.

Figure 1: Amino acid sequence of grass shrimp Tracheales.

```

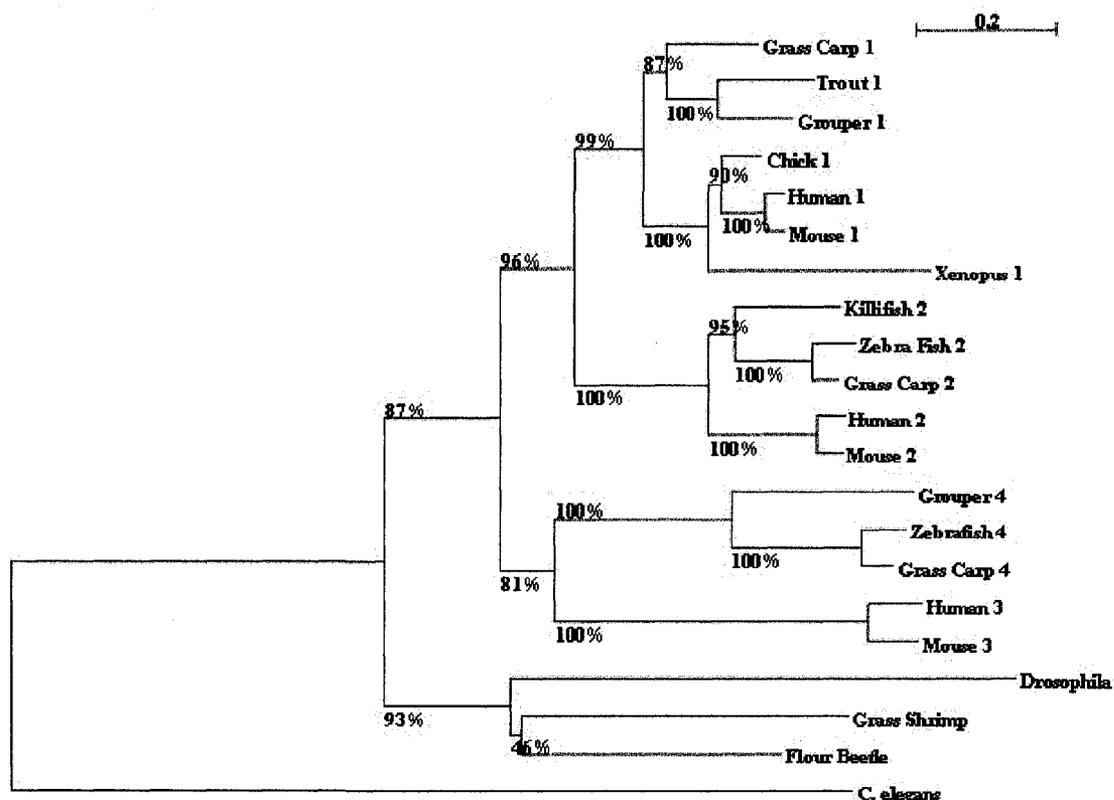
1 MPHEKKEGVA TGKSEAKEKR RNSEKRKEKS RDAARCRRGK ESEIFTELAN ALPLPARTIS
61 MLDKASVMRL TIAFLKTRAL CHACLTKKSD MGECKKLDLE MDSLEMKALD GFLLVSTEG
121 DIIYSSENTA TELGLPQVDV MCOCLEVEYTH PGDHEEARAL VSAKGFPPQEP RHAFRLKCT
181 LIANGRSVNL KSASYKVVVV SGELVNGEDE SWLVALGTFV FIPSNIIEEPL DKQTFVSKHS
241 LDMKFTYVDD NVSEFCGHAS SELLGRSLYE MHHALDSELI KDAYKTLRIK GOVETGRYRF
301 LAREGGYVWV VTQATLIHGP KDHKPQYVVC LNYVVSGVES PGEVLSEQQF LNCSKNNNNK
361 STKSLDSAVP VALSVPTSAT QKVTQPKPEV QPKVNVGVFA SSPVSRVIPA PAPPLQPPTP
421 VAATSKIEFP RTEEMNKOYL IEPEDQPYGV ELKDEPDDL LT HLAPSGGDTG VPLEVPIFKP
481 EDDVFTQIP ISYTDGALFT SPTAIPENIL EPGSSPEYDE YEVTDRKAC NRESGGKIIV
541 NNINGCSPSS ICSSPGSACG LRTPEPPKAL ISQAAVFOSS PGSKPSHRRV VESNRPIGAT
601 ESLEFQLNET ARESYVNIEL EVENQXNDLD EFUMRAPFIP LSNELMLNQ DDEMWGAQPC
661 SLPMGKRNSK YTSLLNGDED SSLAQLLRDR DPPIAGCGEP KHLERHPGD PRCSTYQQGK
721 FLDGGGSFVO PNQVLPCHEG CKDCDDGGGD LVEDDPPQVY MHETVEPPPE LINVESRQHE
781 LISAKRQHSF NSSPLIGHEK LGSLEYENRQ RSEPLQDHQQ PSPQRSPEGS QRQVPMQSH
841 HSGIRELTFE YAPTMQQLLI SKEPITVRCG RGGVCGSCGG GGGGLSAPLS LHKGDSVER
901 NLLNLNGEIE DASREGEVQV FAAPIRLTQD RMTANLLADD GSSHLSYPKL RILTGGGSEF
961 MQAGQHSLSKI SSCFTKCGGT RCGDDNGSNS ACGGGKSTFG ESFLQRGRRQ DPLLLVDPDL
1021 TIPSLSELSE LDFEVNAPAN IGNLQCADL LMALDQA

```

24-76: bHLH Domain; **101-161**: PAS Domain; **234-321**: PAS Domain; **459-464**: proline hydroxylation motif in N-terminal oxygen-dependent degradation domain (N-ODD); **634-640**: proline hydroxylation motif in C-terminal C-ODD; **1025-1046**: C-terminal transactivation domain (C-TAD); **1034-1038**: asparagine hydroxylation motif.

Figure 2: Amino acid sequence of grass shrimp HIF.

parsimony tree (data not shown), and (2) vertebrate HIF-1/2/3/4 α with 87 (neighbor joining) and 82% (maximum parsimony) bootstrap support. Two subclades comprised of HIF-1/2 α (96% bootstrap support) and HIF-3/4 α (81% bootstrap support) were present in the vertebrate HIF protein family (Figure 3).



Bootstrap values on each branch indicate the percentage of trees (1000 replicates) in which that branch is present. 1, 2, 3 and 4 represent HIF-1 α , 2 α , 3 α and 4 α , respectively.

GenBank accession numbers for each species from top (Grass Carp 1) to bottom (*C. elegans*) are: AAR95697; AAK30364; AAW29027; BAA34234; AAF20149; NP_034561; CAB96628; AAL95711; ABD33838; AAT76668; AAC51212; NP_034267; AAW29028; AAQ94179; AAR95698; AAD22668; AAC72734; AAC47303; AAT72404; XP_967427; CAA19521.

Figure 3: Neighbor joining tree derived from 21 HIF amino acid sequences.

In Vivo Expression and Response Pattern of HIF to Severe and Moderate Hypoxia

To study the expression of gsHIF under normoxic and hypoxic conditions, grass shrimp were exposed to normoxic, moderate hypoxic (2.5 ppm), and severe hypoxic (1.5ppm) conditions and shrimp were sampled on day 3, 7, and 14. The measured DO values were 7.71 ± 1.61 ppm and 2.47 ± 0.50 ppm for the moderate hypoxia exposure, and 7.50 ± 0.53 ppm and 1.55 ± 0.23 ppm for the severe hypoxia exposure. The normalized normoxic and hypoxic expression levels of gsHIF, in response to 2.5 ppm DO are shown in Figure 4. One-way ANOVA across time showed no statistical difference between groups ($p = 0.065$). Similarly, using Kruskal-Wallis ANOVA on ranks, no difference between groups was observed in the 1.5 ppm DO exposure ($p = 0.587$).

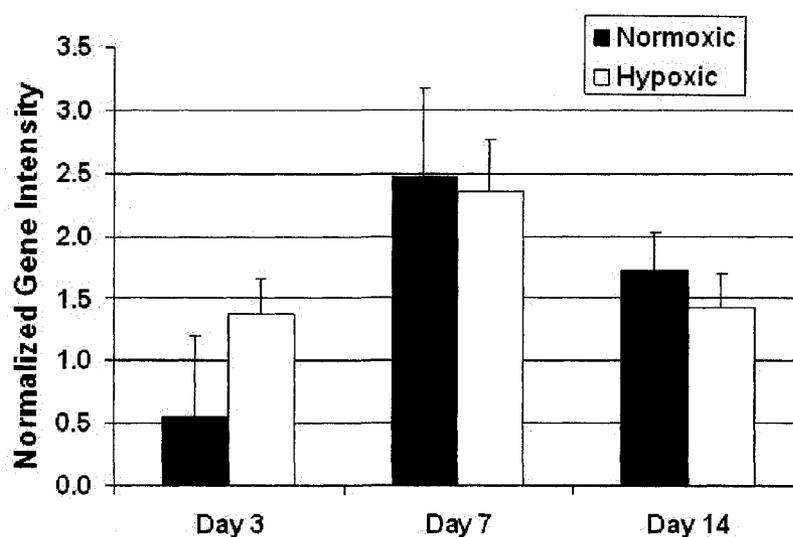


Figure 4: Changes in grass shrimp HIF expression in response to chronic hypoxia exposure (2.5 ppm) at Day 3, 7, and 14.

Discussion

Like HIF-1 α , the tracheless protein (TRH) is a member of the bHLH-PAS family of transcription factors, which shows high identity with HIF-1 α in the bHLH domain

(Isaac and Andrew, 1996). TRH is required for tube formation and is a key regulator of salivary gland and tracheal development in *Drosophila*. Growth of tracheal tubes, which comprise the oxygen delivery system in insects, is stimulated by hypoxia and involves the nitric oxide/cyclic GMP (NO/cGMP) signal transduction pathway (Wingrove and O'Farrell, 1999). NO, hypoxia and HIF-1 α are functionally linked. The gene of the inducible form of nitric oxide synthase (iNOS) is upregulated by HIF and NO impairs normoxic degradation of HIF-1 α by inhibition of prolyl hydroxylases (Melillo et al., 1995; Metzzen et al., 2003). *Drosophila* larvae exhibit rapid NO/cGMP-mediated responses to hypoxia including behavioral changes allowing larvae to escape local hypoxia (Wingrove and O'Farrell, 1999). Grass shrimp show similar types of behavior by climbing out of water during periods of oxygen deficiency (Anderson, 1985). It will be interesting to determine whether hypoxia-induced behavioral responses in grass shrimp share a regulatory mechanism with the HIF and NO/cGMP system as found in *Drosophila*.

The fragment of grass shrimp trachealess protein obtained in this study shows a high level of conservation with other trachealess proteins in the bHLH/PAS regions. Grass shrimp trachealess bHLH domain (20-72) shows 98% identity with bHLH from *Drosophila melanogaster* (GenBank Accession no. AAA96754) and *Bombyx mori* (BAA22946) trachealess. The PAS-A (116-177) domain is most similar to that of *Drosophila* (88%) and the PAS-B (296-395) domain is most similar to that of the red flour beetle *Tribolium castaneum* (GenBank Accession no: XP_967112) (78%). The transcriptional activation domain in the C-terminal portion of the *Drosophila/Bombyx/Tribolium* molecule (Isaac and Andrew, 1996) is not found in the 501 amino acid sequence of the grass shrimp trachealess protein.

HIF protein of the grass shrimp shows a high level of conservation with other HIF-1 α proteins in the bHLH/PAS regions, with amino acid sequences 23-76, 109-179, and 234-321 corresponding to bHLH, PAS-A, and PAS-B domains respectively. In general, the bHLH domain is involved in DNA binding and dimerization, and the PAS do-

main in target gene specificity, transactivation, and dimerization. In addition to the bHLH/PAS domains, HIF-1 α proteins contain domains and sequence motifs that are involved in nuclear translocation, transactivation and post-translational modifications, which control HIF protein stability and transcriptional activity.

HIF-1 α is rapidly degraded under normoxic conditions, mediated by post-translational hydroxylation of conserved proline residues within a polypeptide segment known as the oxygen-dependent degradation domain (ODD), which comprises residues 401-603 in human HIF-1 α (Huang et al., 1998; Masson et al., 2001) and 692-863 in *Drosophila* Sima (HIF-1) (Nambu et al., 1996; Lavista-Llanos et al., 2002). One of the proline residues subjected to hydroxylation in vertebrate HIF-1 α resides at the N-terminal end of the ODD within a LXXLAP sequence motif (residues 397-402 in human HIF-1 α), which is conserved in grass shrimp (residues 459-464, LTHLAP). A second Pro residue (Pro-564) resides in the C-terminal ODD (residues 561-567, MLAPYIP in human HIF-1 α), with corresponding sequences in *Drosophila* present in residues 847-853 (MRAPYIP) and in grass shrimp in residues 634-640 (MRAPFIP). The hydroxylated proline residues in ODD are recognized by pVHL, the product of the von Hippel-Lindau tumor suppressor gene, which functions as an E3 ubiquitin ligase and targets HIF-1 α for polyubiquitination and proteasome-dependent degradation (Semenza, 1998). The prolyl hydroxylase enzymes that catalyze the hydroxylation of these critical proline residues use oxygen as a substrate. Because oxygen is rate limiting for their activity, these enzymes appear to function as oxygen sensors and provide a direct link between oxygen concentration and the HIF-mediated hypoxic response pathway (Bruick and McKnight, 2002; Bruick, 2003). Accordingly, HIF-1 α protein increases exponentially in human HeLa cells exposed to decreasing oxygen concentrations, with a half-maximal response between 1.5 and 2% O₂ and a maximal response at 0.5% (Jiang et al., 1996). In rainbow trout and chinook salmon cells, maximum accumulation of HIF-1 α occurs at much higher oxygen levels (5% O₂), a typical oxygen tension of venous blood in normoxic animals, suggesting a role

for oxygen-dependent gene regulation not only during environmental hypoxia, but also in the normal physiology of these fish (Soitamo et al., 2001). Whether this difference in HIF-1 α stabilization applies to terrestrial and aquatic organisms in general remains to be determined.

Under hypoxic conditions proline residues are unmodified and HIF-1 α accumulates and translocates to the nucleus, a step which is mediated by a bipartite nuclear localization signal (NLS) in the C-terminus of the human protein (Luo and Shibuya, 2001). Grass shrimp HIF does not have this signal sequence, but has two potential NLSs between aa 4-20 and 25-41, which have the characteristic bipartite NLS structure consisting of two adjacent basic domains separated by a 10 amino acid spacer sequence. Interestingly, human HIF-1 α contains a similar bipartite NLS between aa 17-33, which mediates nuclear import of a GFP-HIF-1 α /1-74 chimeric protein (Kallio et al., 1998).

Modulation of transactivation domain function is a second major mechanism by which HIF-1 α activity is controlled. Vertebrate HIF-1 α contains two transactivation domains (TADs) responsible for recruitment of transcriptional coactivators essential for gene expression: N-TAD, the amino-terminal transactivation domain, comprised of amino acid residues ~540-580 in mammals, and C-TAD, the carboxyl-terminal transactivation domain, comprised of amino acid residues 786-826 in mammals (Jiang et al., 1997; Pugh et al., 1997; Bruick and McKnight, 2002). N-TAD is highly conserved in vertebrate HIF-1 α and contains the second proline hydroxylation motif in the C-ODD. Regulation of its activity is likely to be a by-product of protein stability (Pugh et al., 1997; Bruick and McKnight, 2002). Grass shrimp HIF appears to lack N-TAD. However, as discussed above, the conserved proline hydroxylation motif still exists.

C-TAD operates independently of the ODD and is able to recruit coactivator complexes such as CBP/p300 only under hypoxic conditions (Kallio et al., 1998; Kung et al., 2000; Bruick, 2003). The molecular event that controls C-TAD activity involves the hydroxylation under normoxic conditions of a conserved Asn residue in aa 801-805 (EV-

NAP) of human HIF-1 α . This Asn hydroxylation motif is conserved in grass shrimp HIF in aa 1034-1038. The hydroxylation of Asn blocks the interaction of C-TAD with the p300/CBP transcriptional coactivators. Abrogation of Asn hydroxylation under hypoxic conditions allows for the interaction of C-TAD with CBP/p300 (Lando et al., 2002a,b). The asparaginyl hydroxylase enzyme that catalyzes the reaction belongs to the same family of 2-oxoglutarate/Fe(II) dependent oxygenases as the prolyl hydroxylases (Lando et al., 2002a; Masson and Ratcliffe, 2003). Both prolyl and asparaginyl hydroxylases serve as direct oxygen sensors and must be turned on to fully induce HIF in mammals (Lando et al., 2002a,b). HIF activity is thus subjected to multiple independent levels of regulation responsible for graded responses to subtle changes in oxygen concentration.

Grass shrimp HIF mRNA levels are not noticeably affected by hypoxia. This observation, combined with the conservation of the ODD, suggests that gsHIF protein levels may be controlled at the (post)translational level as found for vertebrate HIF-1 α and -2 α . Recently 2 HIF isoforms have been identified whose induction appears to occur at the transcriptional level: HIF-3 α in mammals (Heidbreder et al., 2003) and HIF-4 α in fish (Law et al., 2006). Molecular phylogeny analysis, using neighbor joining and maximum parsimony methods, shows that grass shrimp HIF clusters firmly with HIFs from other invertebrates, which are distinct from the vertebrate HIF-1/2/3/4 α family. The transcriptionally controlled mammal HIF-3 α and fish HIF-4 α proteins form a distinct subclade within the vertebrate HIF (1/2/3/4 α) clade. These results suggest that the gene duplication giving rise to the invertebrate and vertebrate HIFs preceded the duplication resulting in the post-translationally and transcriptionally controlled forms of vertebrate HIF. Whether invertebrates also have HIFs that are under transcriptional control is unknown at present.

Conclusion

As a first step towards understanding the molecular mechanisms that underlie the adaptation of hypoxia-tolerant crustacea to low dissolved oxygen concentrations, we

have cloned an HIF-1 α homolog from the grass shrimp *Palaemonetes pugio*. The encoded amino acid sequence shows high level of conservation with vertebrate HIF-1 α in the bHLH, PAS-A, PAS-B, ODD (with the two proline hydroxylation motifs) and C-TAD (with the asparagine hydroxylation motif) domains. Conservation of important structural motifs suggests that the function, stability and transactivation of grass shrimp HIF are controlled by similar molecular mechanisms as the vertebrate HIF-1 α proteins. However, grass shrimp HIF contains a large polypeptide sequence (aa 790-1020) which has no matching sequences in GenBank. Whether this region conveys unique functional properties to grass shrimp HIF remains to be determined. Similar to what is found for vertebrate HIF-1 α , grass shrimp HIF is constitutively expressed and not induced to an appreciable extent by hypoxia.

CHAPTER III

BIOINFORMATIC ANALYSIS OF EXPRESSED SEQUENCE TAGS FROM GRASS SHRIMP *PALAEMONETES PUGIO* EXPOSED TO ENVIRONMENTAL STRESSORS

Abstract

Six libraries of expressed sequence tags (ESTs) were constructed by suppression subtractive hybridization (SSH) from the grass shrimp, *Palaemonetes pugio*, exposed to environmental stress: moderate (DO 2.5 mg/L) and severe (1.5 mg/L) hypoxia, cyclic hypoxia (1.5→7 mg/L), contaminant-induced stress (pyrene and copper), and biological stress (molt). A total of 1553 ESTs were clustered and assembled using Paracel Transcript Assembler software. The resulting 661 potential transcripts included 181 contigs and 480 singlets. All assembled sequences were annotated by BLAST searches against the public protein database. Gene Ontology (GO) terms for each sequence were provided using Goblet software. A total of 312 assembled transcripts matched a protein with an E-value less than 1E-5. 18% of the most similar matches were from different crustaceans. Large proportions of sequences had no significant BLAST hits (52.8%) or GO terms (64%). GO analysis by libraries showed several genes that were present in only one library suggesting that their expression may be stressor specific. Up-regulation of muscle proteins and GSH-peroxidase appeared specific for chronic (1.5 mg/L) and cyclic hypoxia exposures, respectively. Several genes involved in sulfur redox and (homo)cysteine metabolism were all down-regulated in response to cyclic hypoxia. Up-regulation of cytochrome c oxidase subunit I and down-regulation of vitellogenin was a common response to chronic (1.5 mg/L and 2.5 mg/L) and cyclic DO exposures. The molting process was accompanied by changes in expression of many genes not found in the hypoxia/copper/pyrene libraries. The cDNA clones and sequence information can be used for future functional analysis and for construction of microarrays for monitoring of environmental stressors in coastal waters using wild or caged grass shrimp.

Keywords - *Palaemonetes pugio*; grass shrimp; crustacean; copper, pyrene; hypoxia; gene expression; annotation.

Introduction

Grass shrimp, *Palaemonetes pugio*, are among the most widely distributed, abundant, and conspicuous of the shallow water benthic macroinvertebrates and can be found in the salt marshes along the shores of the Atlantic and Gulf of Mexico of the United States (Anderson, 1985). Although grass shrimp have only limited value as fish bait or food for cultured fish, their ecological importance is unquestioned. Grass shrimp have been extensively documented as prey of fishes and other carnivores and they transport energy and nutrients among various estuarine trophic levels (Griffitt et al., 2007).

P. pugio has also been recognized as a model species in bioassays for assessment of water quality (American Public Health Association, 1975) and has been used in reproductive studies in the laboratory (Oberdorster et al., 2000). Much information has been published about mortality and sub-lethal effects of various toxic chemicals on grass shrimp. In bioassay studies, anthropogenic contaminants found in estuaries, including pesticides, polycyclic aromatic hydrocarbons (PAHs), and metals have been shown to impact grass shrimp growth, size, reproductive capacity, molting, and survival (Burton and Fisher, 1990; Key and Fulton, 1993; McKenney et al., 1998; Lee et al., 2000). Dissolved oxygen regulates the distribution and abundance of grass shrimp in estuarine habitats (Harper and Reiber, 1999). They have a limited home range and are hypoxic tolerant, and because of that, can respond to hypoxia found in their immediate environment (Finley et al., 1998; Lee et al., 1998).

In recent years, genomic approaches have been increasingly applied in the field of toxicology. Such toxicogenomics studies can enhance our understanding of the mechanisms that underlie toxic effects of chemicals on living tissues of various organisms and may help to identify gene expression profiles that may serve as biomarkers of ex-

posure (Calzolari et al., 2007). However, this approach is hampered by the very limited genomic information that is available for non-model organisms that play an important role in ecosystem function. For example, despite the ecological and toxicological significance of grass shrimp, little is known about its genome. Among nine grass shrimp mRNA sequences submitted to GenBank, only two represent complete coding sequences, cadmium metallothionein 1 mRNA (348bp, AY935987), and hypoxia-inducible factor 1 alpha mRNA (3822bp, AY655698) (Brouwer et al., 2007; Li and Brouwer, 2007). Another 42 expressed sequence tags (ESTs) from suppressive subtractive hybridization (SSH) libraries prepared from grass shrimp exposed to three different xenobiotics have been deposited in GenBank (Griffitt et al., 2006).

Clearly an alternative approach is needed to obtain the transcriptome of an important species without full knowledge of the organism's genome. One method that can be applied to acquire information on the transcriptome is sequencing of ESTs from cDNA libraries. An expressed sequence tag (EST) is a short and partial sequence of a transcribed nucleotide sequence produced by sequencing of a cloned mRNA from cDNA libraries. The resulting ESTs represent portions of expressed genes, and the most highly expressed genes will be represented many times by identical or nearly identical clones in the libraries. This form of analysis has been employed in analyzing invertebrate ESTs from cDNA libraries from aquatic organisms, such as water flea *Daphnia magna* (Watanabe et al., 2005), blue crab *Callinectes sapidus* (Coblentz et al., 2006), and copepod *Calanus finmarchicus* (Hansen et al., 2007).

EST sequences can be clustered and assembled into overlapping contiguous sequences that represent unique transcripts. In this way, if the genome of the organism that originated the EST has been sequenced one can align the EST sequence to that genome. In this respect, ESTs become a tool to refine gene discovery, and the experimental conditions in which those ESTs are obtained may provide information on the potential function of the corresponding gene. Moreover, ESTs can be used to construct DNA microarrays

that then can be employed to determine gene expression in specified conditions.

Previous studies of grass shrimp in our laboratory have used a gene-by-gene approach for the study of significantly expressed genes during hypoxic exposures. Several hypoxia responsive genes were identified and characterized from grass shrimp. These include cadmium metallothionein (AY935987), mitochondrial superoxide dismutase (AY935986), HSP70 (AY935982), cytosolic manganese superoxide dismutase (AY211084), hemocyanin (AY935988), and hypoxia inducible factor 1 α (AY655698) (Brouwer et al., 2007; Li and Brouwer, 2007). This gene-by-gene approach proved to be successful and necessary for early studies. However, the advent of various transcriptomic profiling techniques provides new rapid and accurate means for discoveries of genes and patterns of genes induction involved in responses to environmental stress.

This paper reports EST sequencing, computerized clustering, assembly, and annotation of the sequences from six cDNA libraries from grass shrimp exposed to a variety of natural and anthropogenic stressors: chronic hypoxia, cyclic hypoxia, molt, copper, and pyrene. This project will improve our understanding of the Crustacean stress response, may enhance the discoveries of new genes and novel functions or pathways, and provide the foundation for future studies with microarrays constructed from these sequences.

Materials and Methods

Grass Shrimp

Grass shrimp were collected in the vicinity of Ocean Springs, Mississippi in Davis Bayou using dip nets. Adult females and males were segregated by sex and maintained in the laboratory at 15 psu and $27 \pm 1^\circ\text{C}$ for 7 to 30 days prior to experimentation. During acclimation and experimentation periods, grass shrimp were fed brine shrimp nauplii once daily and commercial flake food once daily. During all acclimation and experimentation periods, shrimp were held in artificial seawater (Fritz Super Salt, Fritz Industries, Mesquite, TX, USA) diluted to 15 psu.

Laboratory Exposures

Exposures were conducted in an intermittent flow-through system described by Manning et al. (1999). Normoxic (DO 7.5 mg/L) and hypoxic (DO 1.5 mg/L and 2.5 mg/L for chronic DO, and 1.5-7 mg/L for cyclic DO) conditions within the treatment aquaria were established and maintained as described before (Brouwer et al., 2004, 2005, 2007; Brown-Peterson et al., 2005). The flow through test system provided 1L every 20 minutes (resulting in 3 complete volume additions/day) to each of the 35L test aquaria using a separate water delivery partitioner for each of the normoxic and hypoxic treatments. Oxygen levels were controlled by bubbling nitrogen into a holding tank which gravity fed to the partitioner used to deliver flow-through hypoxic seawater. A 24 hour timer was used to activate a solenoid valve which controlled nitrogen introduction into the holding tank at intervals that maintained oxygen in the holding tank at a level which resulted in the desired oxygen concentration when introduced into the test aquaria. An additional partitioner provided flow-through normoxic seawater, and normoxic conditions were maintained by gently bubbling oxygen into the cells of the water partitioner prior to delivery of water to the individual aquaria. Grass shrimp were housed individually in retention chambers constructed from 10 cm Petri dish bottoms with a 15 cm high collar of 500 μm nylon mesh placed into 35L flow-through glass aquaria in a water bath held at $27 \pm 1^\circ\text{C}$. A 16 hour-light and 8 hour-dark photoperiod was maintained. In all experiments, oxygen was monitored continuously in one hypoxic flow-through aquaria, and DO, temperature and salinity were measured in all flow-through aquaria once or twice daily using a YSI Model 600XLM data sonde. After 3 and 5 days of exposure at 2.5 mg/L DO the thorax/hepatopancreas of 10 shrimp per treatment was removed and stored in 1 mL RNALater (Ambion Inc. Austin, TX, USA) at -20°C . Hepatopancreas from shrimp exposed to 1.5 mg/L DO and cyclic DO was archived in RNALater for RNA extraction after 3 days, and 3 and 5 days, respectively.

Copper exposures consisted of a seawater control and three 96-hr copper treat-

ments (800, 240 and 100 $\mu\text{g/L}$) in a flow through system described above. Copper treatments were generated by injection of a copper stock solution (copper chloride dissolved in distilled water) into dilution water. Appropriate volumes of copper stock solution were delivered to the exposure system by precision syringe pumps and injected into a treatment splitter box/mixing chamber. Immediately after injection of stock, 2 L of diluent water from a water partitioner was delivered to each splitter box/mixing chambers to produce the desired concentrations. Each treatment splitter box/mixing chamber delivered appropriately diluted copper to duplicate treatment aquaria through calibrated delivery lines. Copper concentrations in the treatment aquaria were measured by atomic absorption spectroscopy at 24, 48, 72, and 96 hours. At the end of the 96 hour exposure hepatopancreas from all surviving shrimp was removed and stored in 1 mL RNALater.

Pyrene exposures consisted of a seawater control, a solvent control and three pyrene treatments (50, 15 and 4.5 $\mu\text{g/L}$) in a flow through system described above. Pyrene treatments were generated by injection of the compound dissolved in triethylene glycol (TEG) into dilution water. Appropriate volumes of pyrene stock solution were delivered to the exposure system by precision syringe pumps and injected into a treatment splitter box/mixing chamber. Immediately after injection of stock, 2 L of diluent water from a water partitioner was delivered to each splitter box/mixing chambers to produce the desired concentrations. Each treatment splitter box/mixing chamber delivered appropriately diluted pyrene to duplicate treatment aquaria through calibrated delivery lines. Pyrene concentrations in the treatment aquaria were measured at 24, 48, 72, and 96 hours using reverse phase HPLC with fluorescence detection (Oberdorster et al., 2000). At the end of the 96 hour exposure hepatopancreas was dissected from all surviving shrimp and stored in 1 mL RNALater.

To examine effects of molt cycle on gene expression, hepatopancreas tissues were dissected from shrimp at 1, 3, and 5 days after molting and stored in 1 mL RNALater at -20°C .

Suppression Subtractive Hybridization

Hepatopancreas tissues in RNAlater were shipped on dry ice to EcoArray Inc. (Alachua, FL, USA) for suppression subtractive hybridizations (SSH). Subtractive hybridizations were performed in both directions in order to obtain both up-regulated genes (subtraction run in forward direction; transcripts expressed at higher level in exposed specimen) and down-regulated genes (subtraction run in reverse direction; transcripts expressed at higher level in the untreated specimen). In the forward direction, the treated tissue that contains up-regulated gene transcripts is referred to as the “tester”, while the untreated specimen is called the “driver”. For a reverse subtraction, the “tester” and “driver” designation for the tissues is switched. Subtractive hybridizations were constructed using the Clontech (Palo Alto, CA, USA) suppression subtractive hybridization kit following the manufacturer’s recommendations as briefly outlined below (Figure 5).

Total RNA was extracted using phenol:chloroform STAT-60 (Tel-Test, Friendswood, TX, USA) and homogenization. RNA from shrimp exposed for 3 and 5 days to hypoxia (DO 2.5 mg/L and cyclic DO) and RNA from shrimp exposed to different copper and pyrene concentrations was pooled. Poly-A⁺ mRNA was then isolated using the Oligotex mRNA Midi Kit (Qiagen, Valencia, CA, USA) and converted to cDNA. The tester and driver cDNAs were digested with Rsa I, a restriction enzyme that yields blunt-ended cDNA fragments. The tester cDNA pool was then divided into two portions, each of which was ligated with a different cDNA adapter sequence. Two sequential hybridizations were then performed. For the first hybridization, an excess of driver was added to each of the tester pools, the samples were heat denatured and then allowed to anneal to each other, resulting in the generation of several different hybrid sequences of cDNA.

For the second hybridization, the two different tester pools are mixed together in the presence of an excess of driver without denaturing, and new hybrids are formed. The ends of the differentially expressed cDNA sequences are then filled in by DNA polymerase and two rounds of PCR are performed to enrich for these cDNA clones. During

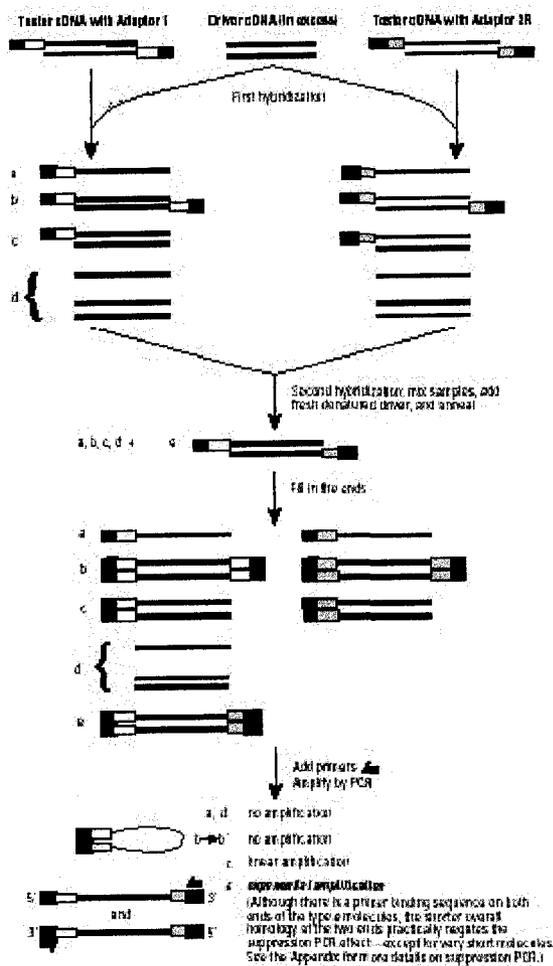


Figure 5: Overview of Clontech PCR-Select cDNA Subtraction Procedure.

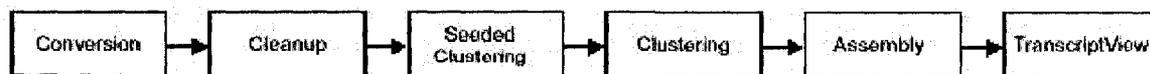


Figure 6: Pipeline of PTA modules.

the PCR reaction, the non-differentially regulated, hybrid clones are either not amplified or do not exhibit exponential amplification. Thus, only differentially regulated genes are enriched by this process. The resultant pool of cDNA clones obtained from the PCR reaction were then shotgun ligated into the pGEM-T Easy cloning vector (Promega, Madison, WI, USA) and then transformed into DH5 α cells and plated onto Luria-Bertani (LB) agar plates containing ampicillin and oxacillin (100 μ g/mL each). Recombinant colonies were then picked from the plates and sequenced in a 96-well high throughput format using standard methods (ICBR Sequencing Core Facility, University of Florida).

Assembly

The resultant expressed sequence tags (ESTs) were assembled using Paracel TranscriptAssembler package (PTA v 3.0.0) with Paracel-optimized version of CAP4 algorithms (Huang and Madan, 1999) that have been extensively modified to account for the peculiarities of EST datasets while retaining important information about possible alternative splice forms (Paracel, 2002). PTA takes EST data as input and passes the data through a series of custom steps to automatically clean, cluster, and assemble the sequences, group ESTs based on measured pairwise similarity, and reconstruct EST datasets into virtual transcripts (Figure 6).

Libraries of EST sequences are seldom perfect. Often they contain contaminants that don't represent the source mRNA. The Cleanup stage in the PTA pipeline removes contaminant sequences and selectively mask regions of ESTs to convert input ESTs into high quality sequences. Sequences that matched with *E. coli* DNA and grass shrimp RNA (mainly 16S mitochondrial rRNA sequences) were removed from further processing. Clone vector contaminants, such as poly-A/T tails and vector, were replaced with a masking character (typically "X"). This process applies when only a section, not the whole sequence, is judged to be useless for further analysis. Nine grass shrimp nucleotide sequences from GenBank (AY211084 [cytosolic MnSOD], AY655698 [HIF-1 α], AY935982

[HSP70], AY935983 [ribosomal protein S20], AY935984 [ribosomal protein S14], AY935986 [mitochondrial MnSOD], AY935987 [metallothionein 1], AY935988 [hemocyanin], and AY935989 [actin]) were used for seeded clustering of the EST sequences. Overlapping sequences in highly similar clusters were grouped into a single contiguous transcript (contigs). The output was a set of contigs and singlets (single sequence cluster) generated for each cluster.

Annotation

All ESTs included in the assembly, whether singlets or contigs, were annotated by sequence similarity comparison against the non-redundant protein database (nr), nucleotide database (nt), and Swiss-Prot database with the BLAST algorithm (BLASTX and BLASTN) (Altschul et al., 1997).

For each putative transcript, the five most similar entries in the databases were recorded. Assignment of putative transcripts with the protein database entries required maximal E-values of $1E-5$. Any sequences failing to match the protein database entries were successively searched with the nucleotide database with E-values less than $1E-10$ as recognized putative identities. The exact choice of most related sequences in each group of alignments (clusters) depended not only on the best hit values but also on the detailed information of matched sequences (Liu et al., 2006), such as having at least 50+ consecutive amino acids with 50%+ similarity, position of sequence in the assembled consensus sequence (contig), and analysis of potential protein conserved domains and functional sites.

To identify the function of each sequence, the matched known genes of annotated ESTs were classified into different functional categories according to Gene Ontology (GO) (Ashburner et al., 2000). The sequence analysis in terms of GO annotation using the GOblet software (<http://goblet.molgen.mpg.de/>) was conducted on all of the assembled sequences. The GOblet software (Hennig et al., 2003; Groth et al., 2004) took ESTs in

FASTA format and performed similarity searches against GO-annotated databases for various model organisms using BLAST. GO summary of trees and tables were constructed based on all GO terms of the respective hits.

Pathway analysis was carried out according to KEGG mapping (Kanehisa et al., 2006). EC numbers were assigned to sequences by BLASTX search against protein database (nr) with a cut off value of $1E-5$. The sequences were mapped to KEGG biochemical pathways according to EC distribution in the pathway database.

The sequence analysis was carried out using EMBOSS (Rice et al., 2000). BLAST report parsing, and features and annotations assignments were carried out using BioPerl (Stajich et al., 2002), and statistical analysis was performed using R (R Development Core Team, 2008). A web site was developed (<http://orca.st.usm.edu/~litd/gs/>) where the user may download the information of the grass shrimp ESTs, including singlets and contigs, uncover which ESTs belong to which contigs, and obtain and search the BLAST annotations against different databases.

Results

Water Quality Parameters

The average DO levels during chronic DO exposures were 1.55 ± 0.23 , and 2.47 ± 0.50 (mg/L), respectively. The measured copper and pyrene concentrations were 122 ± 7 , 237 ± 6 , and 719 ± 22 ($\mu\text{g/L}$), and 4.14 ± 1.26 , 13.69 ± 3.56 and 44.15 ± 9.08 ($\mu\text{g/L}$), respectively. Over the 96 hour exposure, temperature, DO, pH, and salinity recorded as follows: 26.51 ± 0.76 ($^{\circ}\text{C}$), 7.30 ± 0.21 (mg/L DO), 8.39 ± 0.04 (pH), and 15.41 ± 0.30 (‰) for copper, and 27.00 ± 0.22 ($^{\circ}\text{C}$), 6.96 ± 0.23 (mg/L DO), 8.41 ± 0.03 (pH), and 15.36 ± 0.28 (‰) for pyrene.

Libraries and Sequencing

In order to generate a diverse set of sequences from grass shrimp, six independent cDNA libraries were constructed from mRNA extracted from grass shrimp exposed to chronic DO (1.5 and 2.5 mg/L), cyclic DO (1.5-7 mg/L), copper and pyrene, and from post molt shrimp. A total of 1553 clones, ranging from 51 base pairs (bp) to 834 bp, were sequenced. The average length of the sequenced ESTs was 490 bp with a median of 513 bp and a mode of 480 bp. The average insert size for the six libraries was 561, 579, 512, 450, 478, and 409 bp for the chronic DO (1.5 and 2.5 mg/L), cyclic DO, post molt, copper exposure, and pyrene exposure, respectively (Table 2).

Assembly

The first phase of assembly was performed by the seeded clustering module (Figure 6), which compared the input 1553 ESTs data against 9 well-characterized seed sequences in GenBank. 17 sequences matched one of three mRNA sequences (AY655698 [HIF-1 α], AY935983 [ribosomal protein S20], and AY935988 [hemocyanin]), and they were grouped into a cluster with that mRNA sequence and were removed from further consideration in the pipeline. Of these 17, 14 sequences matched AY935988 (*Palaemonetes pugio* hemocyanin mRNA). Further analysis of these 14 sequences extended 173 amino acids of AY935988 to 311 amino acids long (Figure 7). The amino acid sequence represented a fragment of hemocyanin protein, with parts of hemocyanin_C domain (Ig-like domain), and hemocyanin_M domain (copper containing domain). BLASTX search against non-redundant protein database (nr) revealed 76% identity with hemocyanin subunit L from *Marsupenaeus japonicus* (ABR14693).

Table 2: Sequencing Summary.

Library	Sequences	Name ^d	Mean Length (bp)
Chronic DO (1.5mg/L), up-regulated	45	18A (E-H)	684
Chronic DO (1.5mg/L), down-regulated	113	18A (A-D) and 18F	512
Chronic DO (2.5mg/L), up-regulated	94	16B	592
Chronic DO (2.5mg/L), down-regulated	91	16A	565
Cyclic DO, up-regulated	167 ^a	35A and 45A	535
Cyclic DO, down-regulated	165 ^b	35B, 45B/Rev/Rev2	490
Post Molt, up-regulated	190	37A/AR/C	398
Post Molt, down-regulated	168	37B/D	510
Copper Exposure, up-regulated	148	39A/C	471
Copper Exposure, down-regulated	154	39B/D	485
Pyrene Exposure, up-regulated	66	41A/C	395
Pyrene Exposure, down-regulated	152	41B/D	415
Total	1553 ^c		490

^a 96 sequences corresponded to 16S mitochondrial rRNA. ^b 8 sequences corresponded to 16S rRNA and 19 to 28S rRNA. ^c contains 145 rRNA and 45 Coliphage sequences. These sequences were removed during the first step of the Paracel assembly process.

^d EST (clone) identification numbers (see <http://orca.st.usm.edu/~litd/gs/> for more detail).

```

1  DFPFWWRDSYSHKLDKRGEMFFWVHNQLAVRFDAERISNYLEPVEELHWEKPIHDFGAFPH
61  ASYKYGGAFPSRPDDIEFEDVDGVARVRDMVIYESRIRDAIAHGVTIKEDGTHIDIMNER
121  GVDVLGDVIESSLYSPNVQYYGALHNTAHIMLGROTDPHGKYNMPPGVMEHFETATRDPG
181  FFRLHKYMDNIFREHKDSLPSYTFEDLDFKGVKVAIDGKLETYFEDFEYSLVNAVDD
241  TEDIPDVEIDTYVPRLNHKEFSYSIDIKNDKGADTLATIRIYIWPBKDYNGVEFNFDGGR
301  WQAIELDRFWVKLSAGNNHIV

```

1-194 Hemocyanin_M, copper containing domain

200-321 Hemocyanin_C Ig-like domain

Figure 7: Amino acid sequence of grass shrimp hemocyanin.

The second phase of assembly (Figure 6), clustering module, sorted the remaining ESTs into clusters, or groups of closely related sequences, based on local similarity score. Initially each EST was placed in its own cluster and each sequence was compared against all other sequences in both 3' and 5' orientations for pairwise comparison. A greedy clustering algorithm combined with pairwise comparison results iterated through all pairwise comparison hits, and joined the two sequences that generated the hits into a single cluster. It produced 444 singlets and 165 clusters.

The third phase of assembly operated on the clusters created during the first two phases (Figure 6). The sequences in each cluster were assembled into a single contiguous sequence (contigs), and there were often multiple contigs generated per cluster. Singlets, representing transcripts sequenced only once, are ESTs that don't fit into clusters, or if in clusters, don't form any part of contigs. Simple clusters are those that assemble into a single contig. Complex clusters assemble into some combination of contigs and/or singlets and so contain ESTs probably representing more than one closely related transcript (Coblentz et al., 2006). Of the 165 clusters, 28 were considered as complex clusters because they contained more than one putative transcript (contigs or singlets). There were 81 singlets in clusters. Based on the criteria listed in Materials and Methods, the total number of putative transcripts selected for further annotations was 661, of which, the majority (399+81, 72.6%) were singlets, which probably represented the rarer transcripts, and the others (181) were contigs. The number of ESTs per contig ranged from 2 to 67,

with a mean of 7, and a median of 4. The length of the contigs ranged from 124 to 1138 bp, with an average length of 492 bp (see <http://orca.st.usm.edu/~ltd/gs/> for further detail). The number of putative transcripts (singlets and contigs) among the six libraries was 52, 82, 118, 157, 147, and 105 for chronic DO (1.5 and 2.5 mg/L), cyclic DO, post molt, copper exposure, and pyrene exposure, respectively. Of the 480 singlets, 34, 59, 90, 111, 108, and 78 sequences came from the above six libraries, respectively.

Annotation

Each of 661 putative transcripts was compared against NCBI non-redundant protein database (nr) using BLASTX. Of those 513 (77.6%) matched with at least one significantly similar sequence. Of these, 376 were singlets and 139 were contigs. Of the assembled transcripts, 5.1% (26/513), 28.3% (145/513), and 27.5% (141/513) matched a protein with an E-value less than 1E-75, between 1E-75 and 1E-25, and between 1E-25 and 1E-5, respectively. The most similar matches for each of the grass shrimp transcripts were also sorted by organism. Similarities to Arthropoda and chordate genes were identified in a large proportion of the annotations because the public protein database includes a number of completely or nearly completely sequenced genomes of species in these phyla. 17% (87/512) of the most similar matches with chordate sequences were identified in the model organisms *Homo*, *Mus*, and *Rattus*. 43% (220/513) of the most similar matches were from the phylum Arthropoda, with insects (121: *Tribolium*, *Apis*, *Aedes*, and *Drosophila*,) and crustaceans (94: *Litopenaeus*, *Macrobrachium*, *Marsupenaeus*, *Pacifastacus*, *Pandalus*, and *Penaeus*) being by far the most common arthropod classes.

The BLAST results for grass shrimp transcripts provide some insight into the nature of the most abundant, differentially expressed, mRNAs. The three contigs that are made up of about 10 ESTs did not match any proteins in the NCBI database. Another six most commonly sequenced transcripts (16, 31, 10, 10, 10, and 12 ESTs respectively)

Table 3: Distribution of putative transcripts into different GO categories. Only 3 hierarchical levels are shown.

Name	Count
molecular function	226
antioxidant activity	2
binding	105
catalytic activity	108
motor activity	3
signal transducer activity	2
structural molecule activity	37
transcription regulator activity	2
translation regulator activity	7
transporter activity	62
cellular component	113
ATPase complex	15
extracellular space	13
intercellular bridge	12
intracellular	78
intracellular organelle	67
membrane	31
membrane-bound organelle	33
non-membrane-bound organelle	42
organelle envelope	18
ribonucleoprotein	29
biological process	192
cellular physiological process	183
development	8
localization	68
metabolism	156
regulation of cellular process	7
regulation of physiological process	8
reproduction	3
response to abiotic stimulus	5
response to stress	8

showed an E-value between 0 and 2 against proteins that are predicted from genome sequencing without assigned function. Six of the most commonly sequenced contigs showed highly significant similarity to cytochrome c oxidase subunit I (E=1E-116), alpha-amylase (E=1E-111), cathepsin (E=1E-106), hemocyanin 2 (E=1E-103), vitellogenin (E=1E-91), and phosphoenolpyruvate carboxykinase (E=4E-91). Five of them came from *Litopenaeus vannamei*, *Macrobrachium rosenbergii*, and *Pandalus borealis*.

GOblet software (Hennig et al., 2003; Groth et al., 2004) was used to assign probable GO terms to all assembled sequences (661) by similarity searching against GO-annotated invertebrate protein database with an E-value cutoff of less than 1E-10. A total of 226, 113, and 192 sequences were assigned to the three main groups in GO: molecular function, cellular components, and biological process. Table 3 shows the more detailed assignment of sequences to the first and second hierarchical levels of the GO functional categories. According to molecular functions, the majority of expressed genes were involved in catalytic activity, binding, transporter activity, and structural molecule activity. As far as cellular components are concerned, 78, 67, and 29 putative proteins were listed as intracellular, intracellular organelle, and ribonucleoprotein complex, respectively. For biological processes 183, 156, and 68 genes were assigned to cellular physiological processes, metabolism, and localization, respectively.

Venn diagrams were used to identify the number of genes that were up- (Figure 8) or down-regulated (Figure 9) in chronic (1.5 mg/L and 2.5 mg/L) and cyclic DO exposures. Several differentially expressed ESTs usually seem to be unique to a particular exposure, however, there was some overlap. For up-regulated transcripts, cytochrome c oxidase subunit I is the only gene in all three DO treatments, and hence might be an indicator of hypoxia, be it severe, moderate or cyclic. Cytochrome c oxidase subunit III was up-regulated in both chronic (1.5 mg/L) and cyclic exposure, whereas translation elongation factor 2 (TEF-2) was up-regulated in both cyclic and chronic hypoxia (2.5 mg/L). Phosphoenolpyruvate carboxykinase and hemocyanin are the common genes between chronic

(2.5 mg/L) and cyclic exposures (Figure 8). For down-regulated putative transcripts, vitellogenin was identified among all three exposures. Cathepsin L, hemocyanin 2, heat shock protein 70, and crustapain were down-regulated in both chronic hypoxia (2.5 mg/L) and cyclic hypoxia (Figure 9).

Discussion

With increasing urban, industrial and agricultural development along the southeastern coast of the United States and Gulf of Mexico, estuarine organism such as grass shrimp are likely to be affected by anthropogenic and natural stressors. To explore effects of deteriorating water quality conditions we generated 6 EST libraries from grass shrimp exposed to conditions that represent broad categories of natural (chronic and cyclic hypoxia and molting) and anthropogenic (copper-metal and PAH-pyrene) stressors. This approach may as a first step provide an indication of how single stressors affect gene expression without considerations of interaction effects among different stressors. The ESTs from six libraries were coded by specific letters to indicate the laboratory exposure types, and forward or reverse directions. The identity of their origins wasn't lost after all ESTs were assembled together. This allows inferences regarding the differences in gene expression among exposures (Table 2). 96 out of 167 ESTs in the up-regulated cyclic DO library were 16S mitochondrial RNA sequences. The down-regulated cyclic DO library had 8 16S and 19 28S rRNA sequences. These sequences were excluded from the assembly and annotation process.

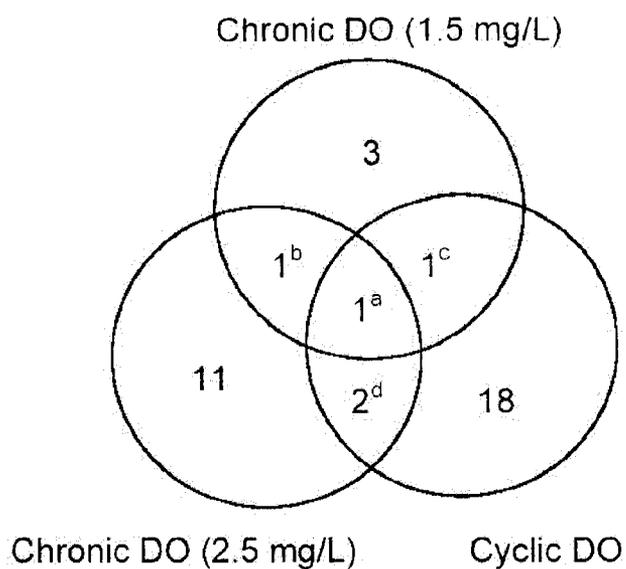


Figure 8: Venn diagram of up-regulated transcripts in three DO treatments. a=cytochrome c oxidase subunit I. b=translation elongation factor 2. c=cytochrome c oxidase subunit 3. d=hemocyanin and phosphoenolpyruvate carboxykinase.

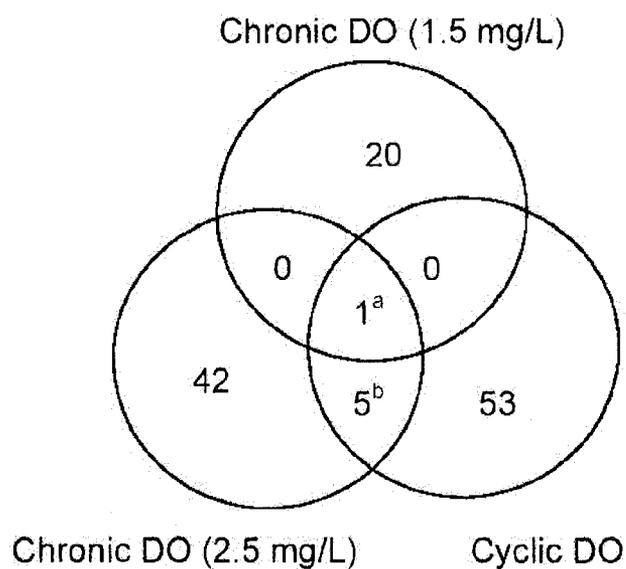


Figure 9: Venn diagram of down-regulated transcripts in three DO treatments. a=vitellogenin. b=cathepsin L; heat shock protein 70; hemocyanin 2 and crustapain.

Table 4: Summary of unique genes found only in particular libraries.

Clone Name	Accession	Putative Protein	E Value
Up-regulated chronic DO (1.5 mg/L)			
18A-E07.g	P05661	Myosin heavy chain	7.00E-36
18A-F04.g	P05547	Troponin I	8.00E-15
18A-G07.g	P13280	Glycogenin-1	3.00E-12
Down-regulated chronic DO (1.5 mg/L)			
18A-B06.g	P12387	Complement C3	2.00E-26
18A-B10.g	O57523	Apolipoprotein A-I	7.00E-38
18A-C05.g	P14448	Fibrinogen alpha chain	3.00E-41
18F-B12.g	P07195	LDH B chain	6.00E-93
18F-C08.g	P19021	Peptidyl-glycine alpha-amidating monoxygenase (PAM)	4.00E-97
18F-H09.g	Q9XYN1	Innexin inx2	1.00E-19
Up-regulated chronic DO (2.5 mg/L)			

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
16B-C03.g	Q15392	24-dehydrocholesterol reductase	2.00E-59
16B-C04.g	O46107	Lipase 1	2.00E-18
16B-D10.g	P84293	Hemocyanin subunit 2	2.00E-35
16B-F11.g	P53713	Integrin beta-1	1.00E-04
16B-G04.g	P12259	Coagulation factor V	7.00E-07
16B-G07.g	P50544	Very-long chain specific Acyl-CoA dehydrogenase (MVLCAD)	2.00E-05
16B-H06.g	P36417	G-box binding factor (GBF)	2.00E-05
16B-H12.g	Q9Y5Z4	Heme-binding protein 2	7.00E-08
Down-regulated chronic DO (2.5 mg/L)			
16A-A03.g	Q9U572	Hemolymph clottable protein	7.00E-38
16A-B07.g	P06708	Troponin C gamma	5.00E-67
16A-B10.g	Q9U943	Apolipoproteins	1.00E-07
16A-C01.g	Q5E9E4	Signal sequence receptor beta (SSR-beta)	9.00E-32
16A-C07.g	P54985	Peptidyl-prolyl cis-trans isomerase (PPIase)	4.00E-25

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
16A-D01.g	Q9MFN9	Cytochrome b	3.00E-37
16A-F05.g	Q9NJ98	Beta-1,3-glucan-binding protein	4.00E-15
16A-F07.g	P54369	Ornithine decarboxylase antizyme	2.00E-13
16A-G07.g	P12026	Acyl-CoA-binding protein (ACBP)	2.00E-23
16A-H01.g	Q94572	Tubulin alpha-3 chain	5.00E-74
16A-H11.g	Q811F4	Fibrillarin	5.00E-57
Up-regulated cyclic DO			
45A-A07.g	P38040	Guanine nucleotide-binding protein gamma subunit	8.00E-16
45A-B04.g	Q91WS0	Protein C10orf70 homolog	3.00E-22
45A-C07.g	P65427	S-adenosyl-methyltransferase mraW	2.00E-26
45A-C09.g	P30710	Epididymal secretory glutathione peroxidase	1.00E-29
45A-D03.g	P38135	Acyl-CoA synthetase	3.00E-26
Down-regulated cyclic DO			

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
35B-C01.g	Q20591	Vacuolar ATPase subunit H	7.00E-18
35B-C02.g	P42279	Trypsinogen Y precursor	2.00E-07
35B-E04.g	P08962	CD63 antigen	3.00E-09
35B-F07.g	Q5ZJH9	H/ACA ribonucleoprotein complex subunit 4	4.00E-69
45B-A03.g	Q95108	Thioredoxin	2.00E-30
45B-A07.g	Q07832	Serine/threonine-protein kinase PLK1	1.00E-86
45B-A09.g	Q5XIP1	Protein pelota	1.00E-31
45B-B03.g	Q99NF1	Beta-carotene dioxygenase 2	9.00E-37
45B-B04.g	P46432	Glutathione S-transferase 3	2.00E-16
45B-D05.g	Q9Y6N5	Sulfide:quinone oxidoreductase	2.00E-68
45B-D10.g	P35520	Cystathionine beta-synthase	6.00E-52
45Rev2-A01.g	Q9FKK7	Xylose isomerase	2.00E-25
45Rev2-A11.g	P35623	Serine hydroxymethyltransferase	1.00E-30
Up-regulated post molt			
			continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
37A-B05.g	P13284	Gamma-interferon-inducible protein IP-30	1.00E-04
37A-B09.g	Q918U0	Cytochrome c oxidase subunit 4 isoform 1 (COX IV-1)	1.00E-21
37A-D04.g	Q3U0B3	Dehydrogenase/reductase SDR family	2.00E-30
37A-E12.g	Q6P3H7	Retinoblastoma-binding protein 4 (RBBP-4)	4.00E-59
37A-G05.g	P41822	Ferritin heavy chain-like protein	9.00E-19
37A-H05.g	Q6P635	COP9 signalosome complex subunit 5	6.00E-20
37AR-A01.g	Q9Y3Q0	N-acetylated-alpha-linked acidic dipeptidase 2 (NAALADase II)	9.00E-11
37AR-A05.g	P10606	Cytochrome c oxidase polypeptide Vb	3.00E-08
37AR-B06.g	P09117	Fructose-bisphosphate aldolase C	1.00E-18
37AR-B09.g	P80971	Cytochrome c oxidase subunit 4 isoform 2 (COX IV-2)	1.00E-19
37AR-C05.g	P81583	Cuticle protein CP1499	1.00E-06
37AR-G08.g	Q9BZP6	Chitinase	6.00E-63
37AR-G09.g	Q91YT0	NADH-ubiquinone oxidoreductase 51 kDa subunit	3.00E-26
37AR-H03.g	Q8C0L6	Polyamine oxidase	7.00E-10
37C-B07.g	Q9QXE5	Thymus-specific serine protease	6.00E-13

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
37C-D08.g	P79401	Cytochrome P450 3A	1.00E-06
37C-G05.g	P12821	Angiotensin-converting enzyme	2.00E-12
37C-H09.g	P41340	Actin-3	3.00E-77
Down-regulated post molt			
37B-A04.g	Q8NI27	THO subunit 2 (Tho2)	4.00E-38
37B-A05.g	Q02645	Adducin-like protein	9.00E-58
37B-A10.g	P00940	Triosephosphate isomerase	2.00E-61
37B-B07.g	Q02379	NADH-ubiquinone oxidoreductase 15 kDa subunit	2.00E-06
37B-B08.g	Q91ZY8	Tripartite motif protein 9	3.00E-19
37B-B11.g	P07764	Fructose-bisphosphate aldolase	1.00E-09
37B-B12.g	Q9D3D0	Protein C20orf121 homolog	2.00E-06
37B-C01.g	Q04212	Hypothetical oxidoreductase	2.00E-04
37B-C05.g	P49762	Serine/threonine-protein kinase Doa	2.00E-66
37B-D06.g	P17644	Acetylcholine receptor protein	1.00E-07

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
37B-E03.g	Q25008	Tubulin alpha-1 chain	8.00E-90
37B-F10.g	P30041	Glutathione peroxidase	4.00E-50
37B-F12.g	O08638	Myosin-11	3.00E-04
37B-G06.g	P26221	Endoglucanase E-4	9.00E-46
37B-H03.g	P17861	X-box binding protein 1	4.00E-18
37B-H05.g	Q8JZX4	Splicing factor 45	5.00E-27
37D-C01.g	O02649	60kD heat shock protein	9.00E-83
37D-D03.g	O74976	Peroxisomal-coenzyme A synthetase	1.00E-25
37D-D04.g	P28799	Granulin protein	9.00E-06
37D-E08.g	Q28983	Zonadhesin	2.00E-07
37D-H02.g	P49948	Ferritin heavy chain 2	7.00E-34
37D-H12.g	P17663	Ferritin heavy chain 1	3.00E-45
Up-regulated copper			
39A-E10.g	P80029	Crustacyanin-C1 subunit	4.00E-15

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
39A-F05.g	Q9JJW3	Up-regulated during skeletal muscle growth protein 5	2.00E-07
39A-F08.g	Q08790	Lipoprotein b1c	9.00E-05
39A-G03.g	Q5FVJ8	Serpin-specific protease 8	9.00E-14
39A-G06.g	Q99LT0	Dpy-30-like protein	1.00E-06
39A-H02.g	O15371	Eukaryotic translation initiation factor 3 subunit 7	2.00E-17
39A-H03.g	Q4QXT9	Coagulation factor X	3.00E-06
39A-H11.g	P60370	Keratin-associated protein 10-5	2.00E-05
39C-C03.g	P49247	Ribose 5-phosphate isomerase A	2.00E-53
39C-D11.g	Q3T0E3	Coiled-coil domain-containing protein 56	3.00E-08
39C-F01.g	O00399	Dynactin subunit 6	1.00E-09
39C-F10.g	Q9U1N0	Hrp65 protein	4.00E-33
39C-G04.g	P00637	Fructose-1,6-bisphosphatase 1	5.00E-39
39C-G05.g	Q08473	Heterogeneous nuclear ribonucleoprotein 40	2.00E-20
Down-regulated copper			

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
39B-C06.g	P79403	Glucosidase II subunit alpha	4.00E-41
39B-C09.g	Q99814	Hypoxia-inducible factor 2 alpha (HIF-2 alpha)	2.00E-30
39B-E02.g	Q94522	Succinyl-CoA synthetase alpha chain (SCS-alpha)	7.00E-17
39B-E04.g	P56616	Ubiquitin-protein ligase	2.00E-46
39B-E09.g	Q8WXI8	C-type lectin	6.00E-05
39B-G11.g	P22985	Xanthine dehydrogenase	4.00E-37
39B-G12.g	Q17004	Serine protease SP24D	5.00E-13
39B-H11.g	Q5TGI0	Protein C6orf168	6.00E-06
39D-A07.g	P52175	Nucleoside diphosphate kinase NBR-B (NDK NBR-B)	1.00E-46
39D-B02.g	P25229	Actin-binding protein chain A (ABP-A)	2.00E-27
39D-B10.g	Q7ZYB4	Polyadenylate-binding protein (PABP)-interacting protein 1	9.00E-09
39D-C07.g	Q5TTG1	Vacuolar ATPase subunit A	5.00E-22
39D-D09.g	Q96Q77	Calcium and integrin-binding protein	1.00E-04
39D-E09.g	Q68HB4	Profilin	3.00E-20
39D-E10.g	Q9D0R2	Threonyl-tRNA synthetase	3.00E-72

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
39D-F04.g	O80433	Citrate synthase	4.00E-31
Up-regulated pyrene			
41A-A04.g	Q5RFM9	10-formyltetrahydrofolate dehydrogenase	2.00E-28
41A-B04.g	P54886	Pyrroline-5-carboxylase synthase	2.00E-10
41A-B05.g	Q24407	ATPase subunit F6	8.00E-25
41A-E05.g	P22464	Annexin-B9	3.00E-22
41A-E10.g	Q7QCK2	Translationally controlled tumor protein	1.00E-14
41A-G01.g	Q8WZ42	Titin	3.00E-10
41A-H08.g	Q8WQA4	Chloride intracellular channel 6-like protein	5.00E-17
41C-D01.g	P17276	Phenylalanine-4-hydroxylase (PAH)	4.00E-21
Down-regulated pyrene			
41B-A10.g	Q96QE2	Proton myo-inositol cotransporter	8.00E-33
41B-C08.g	P37235	Hippocalcin-like protein 1	1.00E-12

continued on next page

Table 4 – continued from previous page

Clone Name	Accession	Putative Protein	E Value
41B-D02.g	Q9QXW2	F-box/WD repeat protein 5	1.00E-12
41B-D10.g	O76757	Adenosylhomocysteinase	6.00E-42
41B-E03.g	P52826	Carnitine acetyltransferase	2.00E-44
41B-E05.g	O75643	U5 small nuclear ribonucleoprotein helicase	3.00E-29
41B-F05.g	Q6QBQ4	Phospholipid scramblase 3	6.00E-15
41B-G09.g	Q5E9K0	Proteasome subunit beta type 2	2.00E-09
41B-H07.g	P62898	Cytochrome c	8.00E-18
41B-H09.g	Q6A068	Cell division cycle 5-related protein	2.00E-13
41B-H12.g	Q8T635	Pupal cuticle protein 20	7.00E-10
41D-B12.g	P40620	HMG1/2-like protein	9.00E-07
41D-E10.g	P78344	Eukaryotic translation initiation factor 4 gamma 2	6.00E-12

Most of the crustacean EST sequences deposited in GenBank are limited to those obtained from commercially important shrimp species: *Litopenaeus setiferus* (1042), *Litopenaeus vannamei* (7506), *Marsupenaeus japonicus* (3156), *Fenneropenaeus chinensis* (10446), and *Penaeus monodon* (7800). A total of 8821 ESTs from a normalized cDNA library made for grass shrimp has been uploaded to the Marine Genomics Project web site at the University of South Carolina (<http://www.marinegenomics.org/index.php>). Hansen et al. (2007) described a SSH-based library of the copepod *Calanus finmarchicus* exposed to sublethal environmental stressors. A total of 887 clones were sequenced, and 189 unigenes were annotated based on similarity search against arthropod sequences in the NCBI protein database. Finally, sequences of 1503 and 1192 cDNA clones obtained from an early blastemal and late proecdysial limb bud cDNA library of the fiddler crab (*Celuca pugilator*) have been determined in the lab of David Durica from the University of Oklahoma (<http://www.genome.ou.edu/crab.html>).

In this study a collection of 1553 ESTs from grass shrimp was generated from six cDNA subtraction libraries. After clustering and assembly, the results represent 661 putative transcripts. Of these, 480 ESTs were singlets and 181 were contigs. An average of 7 ESTs per contig accounted for the larger fraction of ESTs, 85 out of 165 clusters contain 2 sequences, and 4 clusters were composed of more than 20 ESTs. The cDNA libraries constructed here reduced the number of highly abundant transcripts, and increased the efficiency of random sequencing which is essential for rare gene discovery.

Of the 661 putative transcripts 312 (47%) have at least one significant BLASTX hit against the NCBI non-redundant protein database (nr) with an E-value less than $1E-5$. Six of the most commonly sequenced contigs with highly significant similarity came from *Litopenaeus vannamei* (phosphoenolpyruvate carboxykinase and alpha-amylase), *Macrobrachium rosenbergii* (cytochrome c oxidase subunit I and vitellogenin), *Pacifastacus leniusculus* (hemocyanin 2), and *Pandalus borealis* (cathepsin L). A total of 238 genes were linked to GO terms using GOBlet. Large proportions of sequences have no

significant BLAST hits (53%) or GO terms (64%). This implies, not surprisingly, that the genome/transcriptome of grass shrimp as a non-model organism is poorly explored at present. On the other hand, identification of ESTs with no database matches may facilitate the discovery of proteins with novel functions or biochemical pathways relative to model organisms.

The use of gene expression profiles as bioindicators of exposure to environmental stressors has led to increased attention on house-keeping genes which are assumed to be expressed relatively consistently regardless of experimental conditions, and are therefore often used as internal standards for quantifying the relative expression levels of target genes. Examples of such genes among the sequenced ESTs were beta-actin, cyclophilin, elongation factor 1 alpha, and beta-tubulin. Cyclophilin was found to be down-regulated in chronic hypoxia exposure (2.5 mg/L DO), and tubulin was found to be down-regulated in chronic hypoxia exposure (2.5 mg/L DO), copper exposure and post molt. Actin expression was altered after molting, copper, and pyrene exposures. Since their expression levels varied under experimental conditions, relying on one or few housekeeping genes can result in skewed data in future gene expression studies using qPCR and microarrays.

Few genes were found in 1 particular library and not in the other 5. These genes might therefore be specific for the particular stressor, although it is still possible that, after more extensive sequencing of additional clones, these genes might be found in the other libraries as well (Table 4). Two muscle proteins (myosin S1 heavy chain and troponin I) appear to be specifically up-regulated in response to chronic hypoxia (1.5 mg/L DO) exposure, whereas the heterotrimeric guanine nucleotide-binding protein (G protein) gamma subunit and glutathione peroxidase are up-regulated in response to cyclic exposure. It is rather striking that several genes involved in sulfur redox and (homo)cysteine metabolism (thioredoxin, sulfide:quinone oxidoreductase, glutathione-S-transferase, cystathionine beta-synthase) are all down-regulated in response to cyclic hypoxia. Not surprisingly, the molting process is accompanied by changes in expression of many genes, none of which

are found in the hypoxia/copper/pyrene libraries.

More detailed GOblet analysis of GO annotation from the individual SSH cDNA libraries showed almost 33% of sequences putatively coded for enzymes that are responsible for catalytic activity. The majority of those (50%) represent proteolytic enzymes (hydrolases: cathepsin C and L) which are abundant in the down-regulated chronic (2.5 mg/L) and cyclic hypoxia libraries, and the post molt libraries in both directions. Another 32% of the ESTs are mapped to the GO “binding” term, which includes nucleotide binding, protein binding, and ion binding. These genes are more prevalent in the down-regulated chronic hypoxia library (1.5 mg/L), relative to the up-regulated library. Approximately 25% of the ESTs represent genes involved in transporter activity, including lipid, oxygen and ion transport. These genes are found in all libraries with the exception of the cyclic hypoxia up-regulated SSH library. However, 11 transporter proteins, including 6 involved in lipid transport, were identified in the down-regulated cyclic library. Finally, 23 ESTs were mapped to cytochrome c oxidase activity (GO Level 5).

KEGG has been widely used for pathway mapping (Kanehisa et al., 2006), and enzyme commission (EC) numbers were used to identify which sequences related to a specific pathway. A total of 208 out of 1553 sequences matched enzymes with an EC number. 32, 21, 16, and 12 sequences had a match in KEGG map for oxidative phosphorylation, antigen processing and presentation, pyruvate metabolism, and glycolysis/gluconeogenesis, respectively. Genes putatively involved in oxidative phosphorylation, such as cytochrome c oxidase subunit I and III, were exclusively found in up-regulated cyclic and chronic (1.5 mg/L) libraries. Eight ESTs that mapped to the TCA cycle were present in chronic (DO 2.5 mg/L, up-regulated), cyclic (up-regulated), and post molt and copper (down-regulated) libraries. Phosphoenolpyruvate carboxykinase, which is part of the PPAR (peroxisome proliferator-activated receptor) signaling pathway, and plays a crucial role in gluconeogenesis, was found only in up-regulated chronic (DO 2.5 mg/L) and cyclic hypoxia libraries, and in the down-regulated post molt library. 21 sequences were

matched to antigen processing and presentation, such as cathepsin L, cysteine proteinase, and disulfide-isomerase. 19 of these ESTs were present in down-regulated cyclic DO and up-regulated post molt libraries, while two sequences were identified in down-regulated copper library.

SSH PCR-based libraries described here only qualitatively represent transcripts that may be affected by the exposures, while microarrays can quantitatively indicate how much transcript is affected by the exposures. The 661 putative sequences from cDNA libraries, along with several mRNA sequences identified and characterized in our laboratory before this project were applied to design a cDNA microarray. PCR-amplified cDNA arrays have been printed, and hybridization experiments will analyze changes in gene expression in response to chronic and cyclic hypoxic exposures. Together with the sequence information and annotation given here, the results may increase our understanding of adaptation to hypoxic conditions at the molecular level, and may identify gene expression profiles that can be used to assess hypoxia exposure in the field.

CHAPTER IV

CHRONIC HYPOXIA

Abstract

DNA microarrays have become an important tool to measure the global gene expression changes and genetic pathways involved in response to environmental stressors and toxicants. In this study a cDNA microarray was designed and constructed from six libraries of expressed sequence tags generated by suppression subtractive hybridization from the grass shrimp, *Palaemonetes pugio*, exposed to a variety of natural and anthropogenic stressors: chronic hypoxia, cyclic hypoxia, molt, copper, and pyrene. The microarrays were used to examine differentially expressed genes in hypoxic vs. normoxic groups at 6 (H6), 12 (H12), 24 (H24), 48 (H48), 120 (H120), and 240 (H240) hours exposure to chronic hypoxia. The initial response to hypoxia is an up-regulation of 29 genes. Only 6 hours later, a dramatic down-regulation of 47 genes was observed. After 24 hours there is another reversal with 19 genes being up-regulated and none down-regulated. 34 and 22 genes are up-regulated after 2 and 5 days, respectively. 24 genes are down-regulated and 6 up-regulated by day 10. Cluster analysis confirmed two response patterns, one composed of an up-regulated dominated cluster, including H6, H24, and H120, the other composed of a down-regulated dominated cluster, including H12, H48, and H240. Venn diagrams of differentially expressed genes showed there is no gene up- or down-regulated common to all six groups. Hemocyanin transcription is up-regulated after 24, 48, and 120 hours, but down-regulated after 12 hours. Some genes appear unique for specific time points. Phosphoenolpyruvate carboxykinase is up-regulated in the H120 and H240 groups. Cytochrome c oxidase subunit I and C-type lectin are uniquely up-regulated in H12, whereas vitellogenin and tracheless are uniquely down-regulated in H48. org.Dm.eg.db and GOstats packages from R were used to assign GO terms to significantly expressed genes. A total of 291, 129, and 219 genes were assigned to biological

process, cellular components, and molecular function, respectively. The most abundant groups of genes were associated with transport, metabolic process, defense response, and proteolysis. Pathways were analyzed using *Drosophila* metabolic pathways in KEGG database. Oxidative phosphorylation/Citrate cycle and Ribosome were the most abundant categories for chronic DO exposure. Of 19 selected genes that showed differential expression on the microarrays, 17 showed similar up- or down-regulated patterns in both microarray and qPCR.

The custom cDNA microarray is a valid and useful tool to investigate the changes in gene expression of grass shrimp during chronic hypoxia exposure, and show the gene expression profiles affected by hypoxia. Some genes, such as hemocyanin genes, ATP synthase, phosphoenolpyruvate carboxykinase, vitellogenin, cytochrome c oxidase subunit I, Lysosomal thiol reductase, and C-type lectin, may be used as molecular indicators at certain time points of chronic hypoxia treatment in grass shrimp. However, changes of significant genes are too dynamic to serve as biomarkers of hypoxia stress in grass shrimp.

Keywords - *Palaemonetes pugio*; grass shrimp; crustacean; microarray; hypoxia; gene expression; annotation.

Introduction

The recent sequencing of whole genomes of model species has accelerated the development of various transcriptomic profiling techniques, including microarray-based gene expression profiling, that allows scientists to reveal the expression of all the genes simultaneously in an organism. The application of these techniques to toxicology, toxicogenomics, enables biologists to measure the global gene expression changes and genetic pathways involved in stress response (Currie et al., 2005). Toxicogenomics is the evolving science which includes genomic-scale mRNA expression (transcriptomics), cell and tissue-wide protein expression (proteomics), metabolite profiling (metabonomics),

and bioinformatics (Marchant, 2002). Combined with toxicogenomics, DNA microarrays have become most popular and important method to allow biologists to monitor the activities of thousands of genes simultaneously, and characterize biological functions of genomes in response to environmental stressors and toxicants (Ju et al., 2007b). Therefore, microarrays can possibly provide more sensitive, immediate, comprehensive markers of toxicity than traditional toxicological methods, such as morphological changes, carcinogenicity, and reproductive toxicity (Marchant, 2002).

The majority of toxicogenomic research has been focused on the gene-environment interactions of model organisms, such as human (Waters et al., 2003; Kronick, 2004), mouse (Carter et al., 2005), zebrafish (van der Ven et al., 2005), and nematode (Reichert and Menzel, 2005). This limitation has been recognized and more non-model species, especially ecologically relevant ones, have been sequenced recently. Most of the crustacean sequences deposited in GenBank are from commercially important organisms (Chapter 3).

One alternative method to work around this limited genomic information is sequencing cDNA libraries and designing cDNA microarrays by spotting genes identified by various techniques (Chapter 3). Expressed sequence tag (EST) libraries have been successfully utilized in partial sequencing of various aquatic invertebrates, such as water flea *Daphnia magna* (Watanabe et al., 2005), blue crab *Callinectes sapidus* (Coblentz et al., 2006), and copepod *Calanus finmarchicus* (Hansen et al., 2007). ESTs contain enough sequence information to design and construct DNA microarrays for determining gene expression patterns. EST-microarray approach is limited to only a few aquatic species, including *Fundulus* (Oleksiak et al., 2001), channel catfish (*Ictalurus punctatus*) (Ju et al., 2002), European flounder (*Platichthys flesus*) (Williams et al., 2003), fathead minnow (*Pimephales promelas*) (Miracle et al., 2003), common carp (*Cyprinus carpio*) (Gracey et al., 2004), salmon (Rise et al., 2004), rainbow trout (*Oncorhynchus mykiss*) (Krasnov et al., 2005), zebrafish (*Danio rerio*) (van der Ven et al., 2005), goldfish (*Carassius aur-*

tus) (Martyniuk et al., 2006), and Japanese medaka (*Oryzias latipes*) (Kimura et al., 2004; Ju et al., 2007a).

Grass shrimp, *Palaemonetes pugio*, is an ecologically important crustacean, which can be found from Maine to Texas. Little is known about the toxicogenomic information of grass shrimp exposed to environmental stressors. Nine grass shrimp sequences from previous studies in our laboratory were submitted to GenBank (Brouwer et al., 2007; Li and Brouwer, 2007). Additional 42 expressed sequence tags (ESTs) from SSH libraries prepared from grass shrimp exposed to three different xenobiotics have been deposited in GenBank (Griffitt et al., 2006). A total of 8821 ESTs from a normalized cDNA library made for grass shrimp have been uploaded to the Marine Genomics Project web site at the University of South Carolina. Griffitt et al. (2007) employed serial analysis of gene expression (SAGE) to study gene expression profiles of adult male grass shrimp exposed to three environmental stressors, fipronil, endosulfan, and cadmium.

Here six libraries of ESTs were constructed by SSH from the grass shrimp exposed to environmental stress: moderate (DO 2.5 mg/L) and severe (1.5 mg/L) hypoxia, cyclic hypoxia (1.5→7 mg/L), contaminant-induced stress (pyrene and copper), and biological stress (molt) (Chapter 3). A total of 661 annotated transcripts were selected based on certain criteria described in Chapter 3.

The purpose of the present study was to demonstrate the utility of purpose-designed DNA microarray as a useful tool to monitor gene expression changes in hepatopancreas of grass shrimp during the chronic hypoxia exposure, and identify potential biomarkers and validate them using qPCR.

Materials and Methods

Laboratory Exposures

Collection and maintenance of grass shrimp prior to exposure experiments were conducted as described in Chapter 3.

Exposures were conducted in an intermittent flow-through system described by Manning et al. (1999). Normoxic (DO 7.5 mg/L) and chronic hypoxic (DO 1.5 mg/L) conditions within the treatment aquaria were established and maintained as described before (Brouwer et al., 2004, 2005, 2007; Brown-Peterson et al., 2005, Chapter 3). In all experiments, oxygen was monitored continuously in one hypoxic flow-through aquarium, and DO, temperature and salinity were measured in all flow-through aquaria once or twice daily using a YSI Model 600XLM data sonde. All exposures were conducted in triplicate for both controls and treatments. The thorax/hepatopancreas of 20 shrimp per treatment was removed after 0, 6, 12, 24, 48, 120, and 240 hours of exposure and stored in 1 mL RNALater (Ambion Inc. Austin, TX, USA) at -20°C.

Isolation and Quantification of Total RNA

Total RNA was isolated from grass shrimp hepatopancreas using Stat-60 (Tel-Test, Friendswood, TX, USA) according to the manufacturer's protocol. Three or four hepatopancreas tissues were pooled from each treatment, and then total RNA was extracted. These pools were used for microarray analysis as well as quantitative real-time PCR (qPCR). After precipitation, RNA was DNase-treated and stored in RNA Storage Solution (Ambion). RNA was quantified using a NanoDrop Spectrophotometer (ND-1000, NanoDrop Technologies, Wilmington, DE, USA), and quality was assessed on a 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). All samples used had ratios of 260/280 and 260/230 greater than 1.8. RNA was non-degraded as confirmed using the Bioanalyzer.

SSH Libraries and cDNA Clones

Suppression subtractive hybridization (SSH) was performed as described in Chapter 3 to generate six libraries of DNA fragments from grass shrimp. A total of 661 unique cDNA fragments, including 480 singlets and 181 contigs, were PCR-amplified in a 100 μ l

reaction containing 10 μ M nested PCR primer 1 and 2R (Table 5), 10 μ M each deoxynucleotide triphosphate, 10 \times ThermoPol Reaction Buffer, and 2.5 units Taq DNA Polymerase (New England Biolabs, Ipswich, MA, USA). The PCR reaction conditions were 1 cycle of 94°C for 1 min; 35 cycles of 94°C for 30 sec, 68°C for 30 sec, and 72°C for 1 min; 1 cycle of 72°C for 5 min, and then hold at 4°C.

Several potentially hypoxia-responsive genes, identified and characterized in our laboratory from previous studies with grass shrimp, were also amplified. These genes include beta-actin (AY935989), cadmium metallothionein (CdMT, AY935987), mitochondrial superoxide dismutase (mSOD, AY935986), cytosolic manganese superoxide dismutase (cSOD, AY211084), tracheless, and hypoxia-inducible factor 1 alpha (HIF-1 α , AY655698) (Brouwer et al., 2007; Li and Brouwer, 2007).

For beta-actin, CdMT, mSOD, and cSOD, single-stranded cDNA was generated from total RNA by reverse transcription (RT) with Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) and oligo dT primers (5'-TTT TTT TTT TTT TTV N-3'). For tracheless and hypoxia-inducible factor 1 alpha, templates were the glycerol stocks of previously preserved clones. The forward and reverse primers for PCR amplification were designed using Primer3, and the summary of primers is listed in Table 5 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). PCR conditions were 94°C for 2 min for 1 cycle, 35 cycles of 94°C for 30 sec, 55°C for 30 sec, and 68°C for 30 sec, followed by 68°C for 7 min, and then hold at 4°C. The PCR reactions using AccuPrime™ SuperMix I (Invitrogen) gave only a single product. Bands were visualized by ethidium bromide on a 1% agarose gel, purified with a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA), and T/A cloned using pGEM®-T Easy Vector System (Promega, Madison, WI, USA), and sequenced using a Beckman CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA, USA). To confirm identity of PCR products sequence analyses and homology searches were performed using the online BLAST suite of programs (NCBI).

After completion of the PCR, the PCR products were purified using a Millipore Montage Plasmid Miniprep Kit (Billerica, MA, USA). Aliquots of products were run on a 1.5% agarose gel containing ethidium bromide to confirm the correct insert size of amplicon as well as to ascertain minimal primer dimer content. Concentrations were measured using a NanoDrop Spectrophotometer (ND-1000). 1500ng of DNA was speed vacuum dried using an Eppendorf Vacufuge (Westbury, NY, USA), and resuspended in 10 μ L 50% DMSO to get a final concentration of 150ng/ μ l for each clone. Samples were stored in 96-well polypropylene plates to be spotted on microarrays.

Table 5: Primers used for grass shrimp PCR. 11, 47, 52, 59, 124, 173, and 197 are clone names for trachealess and hypoxia-inducible factor 1 alpha.

Primer Name	Nucleotide Sequences
Primer 1	5'-TCG AGC GGC CGC CCG GGC AGG T-3'
Primer 2R	5'-AGC GTG GTC GCG GCC GAG GT-3'
ActinF	5'-GCC TCC TCC TCT TCC CTA GA-3'
ActinR	5'-GTG TTG GCG TAC AGG TCC TT-3'
CdMTF	5'-GAA ACT GAC TGC TCC AAG G-3'
CdMTR	5'-ATG ACT TAC AAA CGC GCA CA-3'
mSODF	5'-GAC TTC GGA ACC ATC AAC AAA-3'
mSODR2	5'-CCA ACC AGC CCC AGC C-3'
cSODF	5'-CAG CTT ATG TTG CCG GTA T-3'
cSODR3	5'-ACA AAT GTG AGG TTC CAG-3'
11F	5'-TCG CAT TCC TCA AGA CCA-3'
11R	5'-ATG AGT TCC GAG TCC AG-3'
47F	5'-GAG GTC GTG AAC AAC AAG CA-3'
47R	5'-GAC TTC TCT TTC CGC TTT-3'
52F	5'-CAG ACG GAA GCA TCT TAG-3'
52R	5'-CCG GAA TGG TGA TCT CG-3'
59F	5'-TTC AAC CTC CTA CCC CAG-3'
59R	5'-GAT TGG GAT CTG GGT AA-3'
124F	5'-TCT CCC CAA AGA AGT CC-3'
124R	5'-CGA TAT TGG CAG GAG CAT TT-3'
173F	5'-GGC TGC CAC TTC AAG AGT TC-3'
173R	5'-CGC AAT TTG TCA ACA TCC TG-3'
197F	5'-AGG CTA CAC ATG GAT GCA A-3'
197R	5'-TAG GGG CGT TGA TAT CTG-3'

Array Printing

For microarray analysis, each array was printed with various sets of controls provided by the Lucidea Universal ScoreCard system (GE Healthcare, Piscataway, NJ, USA) which include ten calibration, eight ratio, and two negative controls designed to validate and normalize experimental data in microarrays and to compare data across experiments using pre-determined fold changes. The negative controls were used to assess the degree of non-specific hybridization and to estimate background. Any spot on the array whose signal was not significantly stronger than that of the negative controls was removed from further analysis. An MSP, multiple sample pool, was also constructed from the grass shrimp clone library by combining cDNA samples together in equal quantities to make a heterogeneous pool. As an unbiased control, MSP won't differentially change and usually have the maximum signals. The sample and control DNAs were prepared by the same methods and printed in an identical fashion.

The PCR products and control DNAs were robotically spotted in duplicate arrays on UltraGAPS (Corning, Corning, NY, USA) coated glass slides using the Bio-Rad ChipWriter system (Bio-Rad, Hercules, CA, USA) equipped with 4 Stealth SMP3 pins (Telechem, Sunnyvale, CA, USA). Each DNA clone was printed side by side in quadruplicate. After printing, slides were kept overnight at room temperature to dry, and were ultraviolet cross-linked at 300 mJoules using the UV Stratalinker 1800 (Stratagene, La Jolla, CA, USA). Slides were stored in desiccator and used within six months of fabrication.

Prior to large-scale printing, a few of slides were printed following the method described above to check the quality and consistency of the microarrays using SYBR 555 nucleic acid stain (Invitrogen).

Preparation of Labeled cDNA for Hybridization

For a single reaction, 15 μ g of DNase-treated (DNA-free, Ambion) RNA was mixed with 2 μ L of 50 μ M Oligo(dT)₁₈ primer (5'-TTT TTT TTT TTT TTT TTT-3', IDT) and spiked with 0.5 μ L of Lucidea Universal ScoreCard control mRNA spike mix, which complemented the Lucidea DNA spotting samples. The mRNA spike mix contained ten individual controls at pre-determined concentrations that span 4.5 orders of magnitude for both the Cy3 and Cy5 channels, and eight ratio controls with target ratios of 1:3, 3:1, 1:10, and 10:1 at both high and low expression levels. The reaction mixture was incubated at 75°C for 7 minutes and then kept at room temperature. The remaining components of the reverse transcription reaction were added as follows: 2 μ L of 10 \times RT Buffer, 1 μ L of RNase Inhibitor (40U/ μ L), 1 μ L of dNTP mix without dTTP (containing 10 mM each of dATP, dCTP, and dGTP), 1 μ L of dTTP and aminoallyl dUTP mix (containing 10 mM each of dTTP and aminoallyl-dUTP, and Tris, pH 8.0), and 2 μ L of ArrayScript Reverse Transcriptase (Ambion). Aminoallyl cDNA was synthesized at 42°C for 2 hours. The template RNA was removed by heating at 65°C for 15 min with the addition of 4 μ L of 1M NaOH, and the reaction was neutralized by adding 10 μ L of 1M HEPES (pH 7.0). cDNA was purified by ethanol precipitation overnight at -20°C.

The precipitated cDNA was resuspended in 4.5 μ L Coupling Buffer (0.1M sodium bicarbonate, pH 9.0, stored at -20°C for up to 3 months), and then mixed with 3 μ L of 40 nmol of either Cy3 or Cy5 dye (GE Healthcare) in DMSO. The dye coupling reaction was incubated at room temperature for 60 min in the dark and briefly mixed every 15 min. The coupling reaction was terminated by the addition of 6 μ L of 4M hydroxylamine and the reaction was allowed to proceed for 15 min at room temperature in the dark. The aminoallyl labeled cDNA was purified using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's protocol. The cDNA was concentrated by ethanol precipitation overnight at -20°C. The purified cDNA was resuspended in nuclease-free water and stored in the dark at -20°C. Exposure of the labeled cDNA to ambient light during

this and all subsequent steps was kept at a minimum to avoid photobleaching. Concentrations of labeled cDNA were measured using the NanoDrop Spectrophotometer (ND-1000). FOI, frequency of incorporation, was calculated as the number of cyanine-labeled nucleotides incorporated per 1000 nucleotides of cDNA to assess the sensitivity of the assay.

Array Hybridization and Washing

Prehybridization was done immediately preceding the application of the target cDNA onto the arrays. The preparation of the hybridization solutions was completed during the time arrays were prehybridized. Each time prehybridization buffer containing 50% formamide (Amresco, Solon, Ohio, USA), $5 \times$ SSC, 0.1% SDS, and 0.1 mg/mL BSA was freshly made from stock and pre-warmed to 42°C. Arrays were immersed in prehybridization buffer and incubated at 42°C for 45 to 60 minutes. Prehybridized arrays were transferred to $0.1 \times$ SSC and incubated at room temperature (22 to 25°C) for 5 minutes (twice). Next, arrays were incubated in DEPC-treated water at room temperature for 30 seconds, and then dried by centrifuging at 1,000 rpm for 3 minutes at room temperature and blowing compressed gas over the array. Arrays were kept in a dust-free environment while completing the preparation of the hybridization solution. Slides were used immediately following prehybridization to ensure optimal hybridization efficiency.

A loop design with direct and indirect comparisons was used for chronic exposure (Table 6). Each individual sample was hybridized to each of two different samples in two different dye orientations, and some direct sample-to-sample comparisons were used for comparing important samples.

The Cy3 and Cy5 labeled cDNAs were mixed in equal amounts measured as the number of pmoles/mixture (about 25 pmoles). The mixture was dried in the dark by using a Vacufuge (Eppendorf) for 1-2 hr, dissolved in 6 μ l of nuclease-free water, denatured by heating at 95°C for 2 min, and then cooled on ice for 30 sec. The cDNA was mixed

Table 6: Loop design for comparison of normoxic (N) and chronic hypoxic (H) samples (0, 5, and 10 days; 6, 12, 24, and 48 hours). There were 3 pooled samples for each time point (n=3), for a total of 39 arrays.

Array	Cy3	Cy5	Array	Cy3	Cy5
1	N0	N5	1	H6	N6
2	N5	H5	2	N6	N12
3	H5	H10	3	N12	H12
4	H10	N10	4	H12	H24
5	N10	N0	5	H24	N24
			6	N24	N48
			7	N48	H48
			8	H48	H6

with 1.5 μ l of oligonucleotide dA₈₀ (1 μ g/ μ l, GE Healthcare), incubated at 75°C for 45 min, and then added to pre-warmed (42°C) 7.5 μ l of Microarray Hybridization Buffer (GE Healthcare) and 15 μ l of 100% (v/v) formamide (Amresco).

Hybridization was carried out in a-Hyb Hybridization Station (Miltényi Biotec, Auburn, CA, USA) equipped with four sealed chambers to hold four slides. After sealing, the hybridization chambers were individually programmed for the selected protocols. Hybridization buffer containing about 25 pmoles of the Cy3 and Cy5 labeled cDNAs was pipetted into sample reservoir, and slides were hybridized at 42°C for 16 hours in a humid hybridization chamber. After hybridization, the slides were rinsed in 2 \times SSC and 0.1% SDS at 42°C for 5 minutes (twice). The slides were washed at room temperature in 0.1 \times SSC with 0.1% SDS for 10 minutes, followed by five washes at room temperature in 0.1 \times SSC for 1 minute, rinsed in 0.01 \times SSC for 10 seconds at room temperature, and dried by centrifuging at 1000 \times g for 3 minutes at room temperature and blowing compressed gas over the array. Slides were protected from light during hybridization and subsequent steps until ready to scan.

Image Analysis

Microarray slides were scanned using a Bio-Rad VersArray ChipReader (Bio-Rad). The excitation setting was at 635 nm (Cy5) and 532 nm (Cy3) and the images were scanned at 10 μ m pixel resolution using a simultaneous dual-laser scanning system and captured in 16 bit TIFF (Tagged Image File Format) format.

The tab-delimited plate files containing the individual plate information about the samples in each well were generated using the BioRad VersArray ChipWriter. These files were converted to a standard GAL (GenePix Array List) file which includes the position and identity for each sample (Block, Column, Row, ID, and Name), the printing pattern, and the size of the spots. Initial Spot finding and quantification was carried out with GenePix Pro 6.1 microarray image analysis software (Molecular Devices, Sunnyvale, CA, USA) using an automatic method in which each spot was checked for proper alignment within the grid. Finally, spot quantification was performed to obtain the expression level of each gene on the array and the results were saved as GPR (GenePix Results) files. Foreground intensities were background corrected using the morphological opening method in GenePix Pro 6.1. Ratio images were displayed at standard 24-bit composite RGB overlay images using the GenePix square root transformation method.

Data and Statistical Analyses

The data analysis was carried out using Acuity 4.0 microarray informatics system (Molecular Devices). Spots flagged as Bad, Not Found, or Absent were excluded from the analysis. For each spot, signal-to-noise ratio (SNR) was calculated using the following formula: $(F635Mean - B635Mean)/B635SD$, where F635Mean is the mean of all the feature pixel intensities at 635nm, B635Mean is the mean of all the background pixel intensities at 635nm, and B635SD is the standard deviation of the background pixel intensities at 635nm. A SNR threshold of 3 is established as a measure of detectability for each channel, and any spots which were less than 3 were marked as bad and excluded from

downstream analysis. Several quality control conditions, such as spots with only a small percentage of saturated pixels, spots with relatively uniform intensity and background, intensities larger than the negative controls on the array, were established to select spots for further analysis. The data were then imported into R (R Development Core Team, 2008) and normalized using weighted LOWESS in LIMMA (version 2.7) (Oshlack et al., 2007) and controls before calculating the gene expression log₂ ratios.

To make full use of the within-array spots, the residual maximum likelihood model (REML) from LIMMA (Smyth et al., 2005) was used to estimate the spatial correlation between the adjacent four spots. This approach calculated a separate linear model to the expression data of each gene and a common correlation for all the genes for between within-array replicates.

Given the normalized data, the significance of differential expression between time points was determined using the empirical Bayes (Smyth, 2004) method to shrink the gene-wise sample variances towards a common value. Resulting statistics were corrected for multiple hypothesis testing using Benjamini and Hochberg false discovery rates (Benjamini and Hochberg, 1995) and significant changes in gene expression were identified at $p < 0.05$. Significance test was performed with hypoxia vs. normoxia samples at the same time point.

Hierarchical clustering was performed on the most differentially expressed genes ($p < 0.05$) using two-way hierarchical cluster analysis in R. The uncertainty in the hierarchical clustering analysis was assessed using the pvclust package. The significant clusters were calculated using multiscale bootstrap resampling ($\alpha=0.95$).

To identify the biological process and functions changed by exposures, the differentially expressed genes were classified into three Gene Ontology (GO) categories (Ashburner et al., 2000) using GOstats package from R, and annotated by genome-wide *Drosophila melanogaster* annotation package available from org.Dm.eg.db. This organism-specific package provided detailed information and mappings among different identifiers

that they are directly associated with.

Pathway analysis was carried out according to KEGG mapping (Kanehisa et al., 2006). Custom scripts were used to create pathway categories of *Drosophila* from http://www.genome.jp/kegg/KGML/KGML_v0.6/dme/ and map *Drosophila* identifiers to KEGG pathways.

Several genes from the arrays were selected for gene quantification using quantitative real-time PCR to check for consistency in gene expression patterns observed with the microarrays.

Quantitative Real-Time PCR (qPCR) and Statistical Analysis

Quantitative real-time PCR was performed on cDNA generated from 1 μ g of RNA obtained from the same pooled RNA samples used in the microarray study. cDNA was generated by reverse transcription using Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen) and random hexamers (50ng/ μ L). Primers (Table 7) were designed to amplify a short fragment (100-150 bp) using Beacon Designer (PREMIER Biosoft, Palo Alto, CA, USA), and synthesized by IDT (Coralville, IA, USA). Real-time PCR was performed using the iCycler iQ Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). 50 μ l qPCR reactions using Bio-Rad iQ SYBR Green Supermix (Bio-Rad) containing 10 μ M forward and reverse primers, and cDNA template. To generate a standard curve and determine PCR efficiency, cDNA was pooled and 1:1, 1:5, 1:25, 1:125, 1:625, and 1:3125 dilutions were made. Individual samples were diluted (1:25) and amplification performed in triplicate in a 96 well PCR plate (Bio-Rad). Relative quantities of each transcript were determined using grass shrimp 18s rRNA as the internal standard. Gene expression data for each time point are expressed as fold change relative to the mean of the same-day controls. The following PCR program was used for gene amplification: 95°C for 2 min; 95°C for 15 sec and 58°C for 15 sec (50 cycles); 72°C for 20 sec, followed by a melt curve. Melt curve analysis was performed to ensure single products by using a temperature step

gradient from 55°C to 80°C in 0.5°C increments with fluorescence measured after a 10 sec incubation at each temperature. The threshold cycle values (C_t) for each sample was calculated using the iCycler iQ Real Time Detection software (Bio-Rad). For each standard curve, the correlation coefficient ranged from 0.995 to 0.999 indicating a high degree of confidence for the measurement of the copy number of molecules in the samples. The relative starting concentration was calculated using a standard curve for each primer set and then normalized to the internal control 18S rRNA (Bio-Rad).

The sequence analysis was carried out using EMBOSS (Rice et al., 2000), and statistical analysis was performed using R (R Development Core Team, 2008). The data are shown as mean \pm S.E.M. For qPCR, statistical differences in data sets between the control and treated groups were determined using a one sample Student's t-test (n=3 independent samples) against the baseline value of 1 (equal gene expression in both samples). Data was log-transformed for homogeneity of variance where necessary. Significance level was set at $p < 0.05$.

Results

Water Quality Parameters and Grass Shrimp

The average DO concentration during chronic hypoxia was 1.59 ± 0.24 (mg/L). The average length and weight (without eggs) of control grass shrimp and shrimp exposed to chronic hypoxia were 36.30 ± 2.39 (mm) and 378.00 (mg), and 36.85 ± 2.50 (mm) and 380.35 (mg), respectively. At the end of the exposure hepatopancreas tissues were dissected from grass shrimp, weighed, and stored in 1 mL RNALater at -20°C. The mean hepatopancreas weight of control shrimp and shrimp exposed to chronic hypoxia was 8.82 and 8.41 (mg), respectively.

Table 7: Summary of qPCR primers.

Gene	Clone	Forward	Reverse
18s RNA		5'-GCA ATT CGC CGT CGT TAT TCC-3'	5'-ACT CCT GGT GGT GCC CTT C-3'
vitellogenin	16A-E12.g	5'-CTG GTG TCT TTA CTT CTA TAA TCC-3'	5'-CGT TCA TCG GCT GTC AAG-3'
cathepsin L	16A-H05.g	5'-CAG CCG TTG TAA TGG AGG AG-3'	5'-CGA AGC GGC AGG TAT TGT C-3'
phosphoenolpyruvate carboxykinase	16B-C06.g	5'-TCC AAA GTT GTA GCC AAA GAA G-3'	5'-GAG ACA CGG AGT AAT GGT AGG-3'
hemocyanin 2	16B-F08.g	5'-AAG TAG TTG TTG CTG GAG GAG-3'	5'-AGA CAG GAG AGT ATT TGA TGA AAC-3'
cytochrome c oxidase subunit I	18A-H01.g	5'-ATG ATT GGA GGA TTT GGA AAC TG-3'	5'-CTG GAG AGT AGA AGT GTG AGA G-3'
chloride intracellular channel 6-like protein	41A-H08.g	5'-TTG CGA AGG TGG GTC ATT AG-3'	5'-CAG GAT AAG GAT GTG GCT CAG-3'
hemocyanin	41C-G09.g	5'-CGT TCT TAA TGT CAA TGC TGT ATG-3'	5'-TCT TGT TAA TGC TGT GGA TGA TAC-3'
Ferredoxin	45A-B01.g	5'-ACT CGT GGG ACA GAT ATT ATG C-3'	5'-GAG GTT GGT CAG TCG TGT AG-3'
kinase	45B-A07.g	5'-ATC TCT GTG AAT GAC CTG TCT G-3'	5'-ATA ATA CTG GAA CTC TGC CGT AG-3'
crustapain	45B-A11.g	5'-GTG TTC TCC AGT CAA CTT CTT C-3'	5'-CCT CGG AAT CAA CGC CTA C-3'
dioxygenase	45B-B03.g	5'-TGG TTG AGA CAG TGT GAA GAG-3'	5'-AAT CTT GCC AGG TCC ATT CC-3'
chitinase 1	45B-B06.g	5'-ACA GAA GGA GTA GAT GAG ATG AG-3'	5'-TTC GGT TGT AGC AGT GAG TG-3'

Array Construction

The maximum volume of total RNA allowed in the reverse transcription is 11 μL . 15 μg of total RNA was needed in a single reaction, so the concentrations of samples must be 1.5 $\mu\text{g}/\mu\text{l}$ or higher. Ethanol precipitated RNA samples were measured using a NanoDrop Spectrophotometer (ND-1000). The RNA concentration was $2.37 \pm 0.34 \mu\text{g}/\mu\text{l}$. The ratios of 260/280 and 260/230 were 2.13 ± 0.01 and 2.27 ± 0.03 , respectively. No RNA degradation was observed as evidenced by running diluted samples on RNA Nano Chips (Agilent Bioanalyzer 2100). The same RNA samples were used for microarray and qPCR. After synthesis and purification, the absorbance of labeled cDNA was measured at 260, 550, and 650 nm, and used to calculate the frequency of incorporation (FOI, the number of cyanine-labeled nucleotides incorporated per 1000 nucleotides of cDNA). The average FOI for Cy3 and Cy5 was 22.93 and 16.91, respectively.

The array was constructed from cDNA fragments generated from grass shrimp exposed to environmental stress: moderate (DO 2.5 mg/L) and severe (1.5 mg/L) hypoxia, cyclic hypoxia (1.5 \rightarrow 7 mg/L), contaminant-induced stress (pyrene and copper), and biological stress (molt). The average length of the cDNA was 476 base pairs (bp), and these sequences were annotated by BLAST searches against the public protein databases. The genes chosen for the array were broadly categorized into functional groups including binding, catalytic activity, metabolism, transporter activity, and structural molecule activity (See Chapter 3).

Self Hybridization

The same pooled RNA samples were reversed transcribed separately in the presence of aminoallyl-UTP. Each cDNA was independently labeled with Cy3 and Cy5. 25 pmoles of cyanine labeled cDNAs were mixed and hybridized to three replicate arrays. The self-self hybridization was performed to calculate the ratio between raw intensities of red and green channels among different slides (Figure 10). Ratios of background

subtracted red (Figure 10A) or green (Figure 10B) mean intensities from different slides are roughly equal to one (slope=0.98 and $R^2=0.87$ for red channels, and slope=1.06 and $R^2=0.88$ for green channels). Large portion of spots are located at the low part of intensities, whereas the spots were spread out at the high end of intensities. The intensities from different slides generally have the same intensity range for each channel. If considering the batch effects of slides, printing, and dye labeling, the method of microarray data collection in this study does accurately maintain the original microarray (intensity) information, and high reproducibility of intensities across slides were consistent during the experiment. Further analysis with normalized data showed no differentially expressed genes from self-self hybridization.

Control Hybridization

Additional hybridizations were set up to quantitatively assess the correlation between different amounts of the labeled sample and signal intensity. Two pooled RNA samples were reverse-transcribed separately in the presence of aminoallyl-UTP, and each cDNA was labeled with Cy3 and Cy5, respectively. Different amount of labeled samples, 0.5 pmole, 2.5 pmole, 5 pmole, and 25 pmole, were mixed and hybridized on microarrays. For 0.5 pmole dyes, the maximum intensities were less than 600 and 100 for red and green channels, respectively. More meaningful data were generated starting from 5 pmoles. Figure 11 shows individual selected spots with high raw intensities at 25pmole and their corresponding intensities at other concentrations in the red channel (Figure 11A) and green channel (Figure 11B). No individual intensity was saturated at the above concentrations. Since signal intensities measured from arrays hybridized with low pmole samples appear to be unreliable, hybridizations were performed with 25 pmole Cy3/Cy5-labeled samples which provided a better data quality and a clear visual picture in this study.

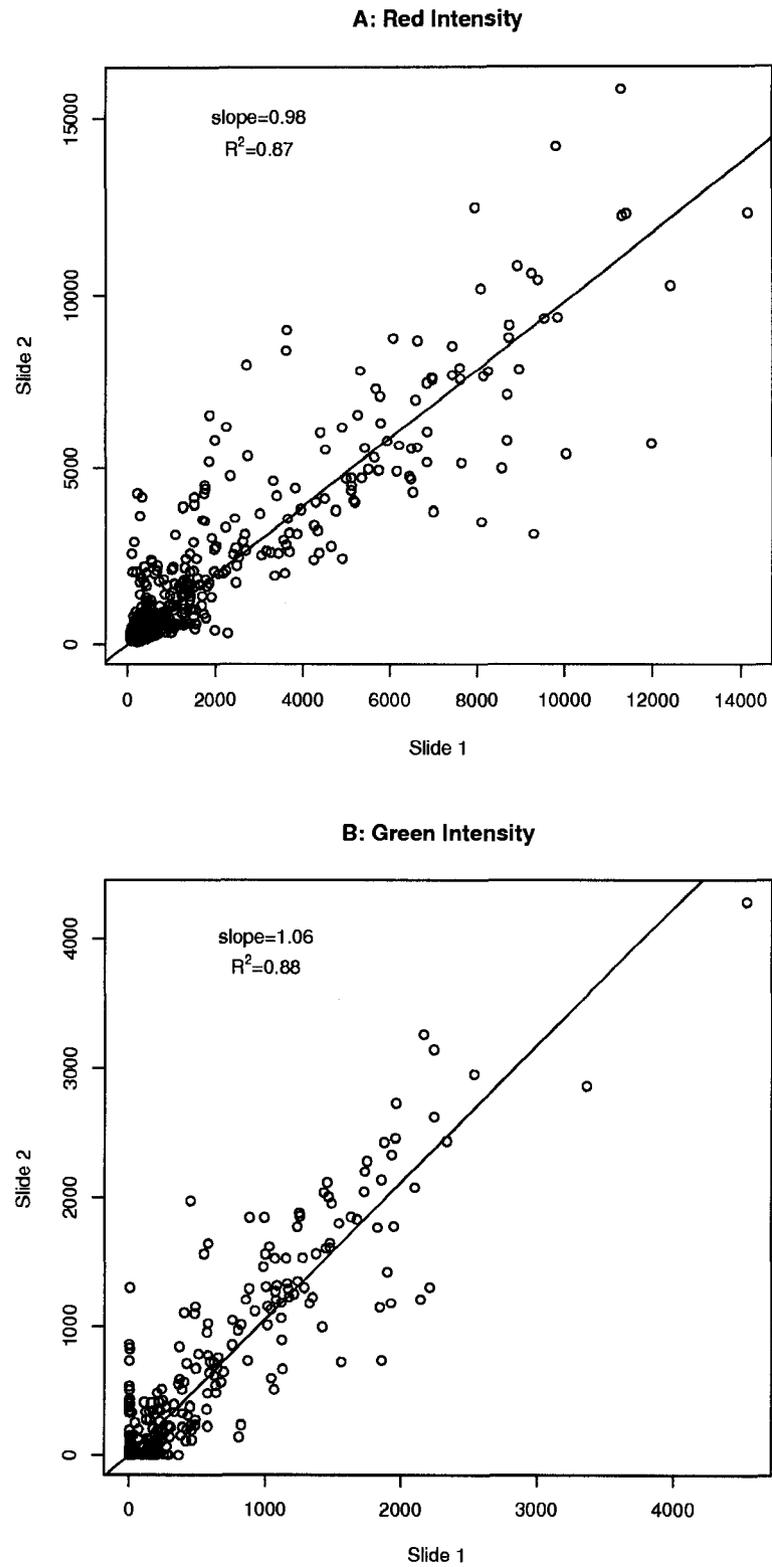


Figure 10: Ratios of background subtracted red (A) or green (B) mean intensities from different slides for self hybridization.

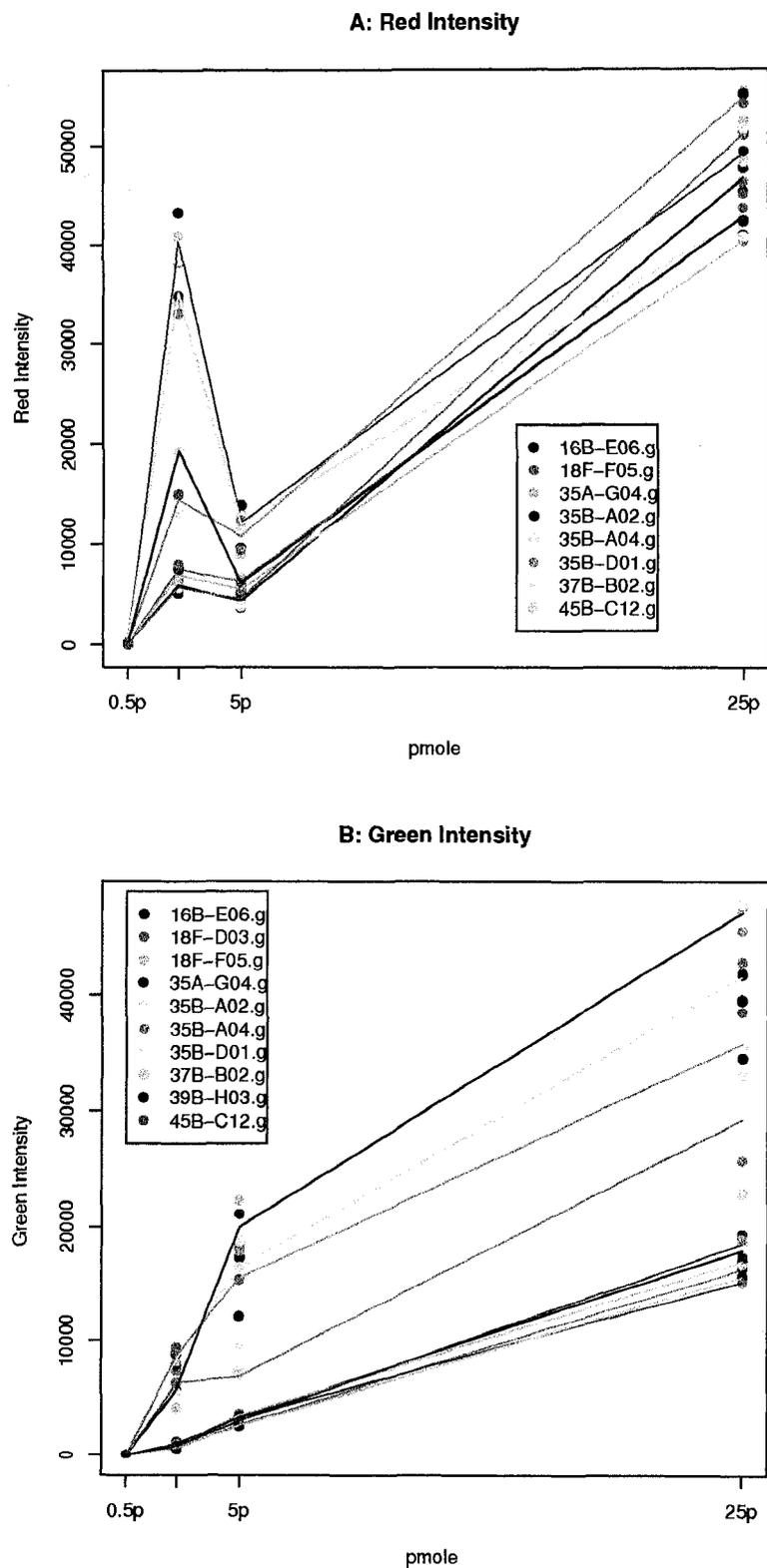


Figure 11: Individual intensities of red (A) and green (B) channel at different hybridization concentrations.

Chronic Hypoxia Exposures

This study determined differentially expressed genes at a p value of less than 0.05 that occurred in hypoxic vs. normoxic groups at the same time point (Table 8). All sequences on the arrays were annotated (see Chapter 3) by similarity search against the non-redundant protein database (nr) and Swiss-Prot database using BLAST algorithm (Altschul et al., 1997). Combined with the annotation tables, Table 9 shows the differentially expressed genes with BLAST E value less than 1E-5 in hypoxic vs. normoxic groups at the same time point.

Observed changes in gene expression revealed a rather dynamic pattern (Table 8). The initial response to hypoxia is an up-regulation of gene expression. A total of 29 genes were up-regulated after 6 hours of hypoxic exposure, whereas none were down-regulated. Only 6 hours later, a dramatic down-regulation of 47 genes, including several hemocyanin genes, was observed, whereas only 2 genes were up-regulated. After 24 hours there is another reversal with 19 genes, including several hemocyanin genes, being up-regulated and none down-regulated. 34 and 22 genes are up-regulated after 2 and 5 days, respectively. After 10 days exposed to chronic hypoxia treatment, 24 genes are down-regulated and 6 up-regulated.

Table 8: Differentially expressed genes during chronic hypoxia exposures at the same time points.

Chronic Hypoxia vs. Normoxia						
Time (Hours)	6	12	24	48	120	240
Up-regulated	29	2	19	34	22	6
Down-regulated	0	47	0	17	1	24

Cluster analysis of the 100 most differentially expressed genes ($p < 0.05$) with BLAST E value less than 1E-5 confirmed the patterns described above. The heat map shown in Figure 12 represents the chronic hypoxia-regulated genes, clustered by the cor-

relation coefficient according to their similarity in expression pattern by gene and by treatments (time exposed to chronic hypoxia). The response to chronic hypoxia appears to have two patterns, one composed of an up-regulated dominated cluster, including H6, H24, and H120, the other composed of a down-regulated dominated cluster, including H12, H48, and H240. Within the two “overall” clusters H24 was most similar to H120 and H48 to H240. Multiscale bootstrap resampling shows that none of the genes clusters significantly ($\alpha=0.95$).

Venn diagrams were used to display the differentially expressed genes that were up- or down-regulated during chronic hypoxic exposure (Figure 13). There is no gene up- or down-regulated common to all six groups. However, there are some genes whose expression is commonly changed in two or three groups. Lysosomal thiol reductase [*Amblyomma americanum*] and hemocyanin [*Litopenaeus vannamei*] are the common down-regulated genes in H12 and H240. A total of 12 genes, including vitellogenin [*Macrobrachium rosenbergii*] and serine protease SP24D precursor [*Drosophila pseudoobscura*] are shared between H12 and H48. Vitellogenin [*Macrobrachium rosenbergii*] and trachea-less [*Tribolium castaneum*] are unique genes in H48 (Figure 13A).

In contrast to 12 hours hypoxia exposure, hemocyanin transcription is up-regulated after 24, 48, and 120 hours. Specifically, hemocyanin 2 [*Pacifastacus leniusculus*] is up-regulated in the H24, H48, and H120 groups, whereas two other hemocyanin genes, closely related to *Litopenaeus vannamei* and *Penaeus monodon* hemocyanin, are up-regulated in H24 and H48 (Figure 13B). In addition, phosphoenolpyruvate carboxykinase [*Litopenaeus vannamei*] is up-regulated after long-term hypoxia exposure (H120 and H240 groups). Cytochrome c oxidase subunit I [*Metapenaeopsis barbata*] and C-type lectin [*Litopenaeus vannamei*] are unique genes in H12 and H240, respectively (Figure 13B).

Table 9: Differentially expressed genes in hypoxia vs. normoxia at the same time points.

ID	logFC	P Value	Definition	E Value
Chronic DO: Hypoxia vs. Normoxia 6 hrs				
up-regulated				
37D-D07.g	1.64	0.031	vitellogenin [Cherax quadricarinatus]	8.00E-12
35B-F07.g	1.64	0.016	H/ACA ribonucleoprotein complex subunit 4 [Danio rerio]	3.00E-69
18A-H01.g	0.53	0.015	cytochrome c oxidase subunit I [Macrobrachium rosenbergii]	1.00E-116
18A-G02.g	0.43	0.020	ferritin subunit (glycosylated) [Nilaparvata lugens]	8.00E-04
Chronic DO: Hypoxia vs. Normoxia 12 hrs				
up-regulated				
39B-D07.g	0.73	0.040	cytochrome oxidase subunit I [Metapenaeopsis barbata]	2.00E-13
down-regulated				
Actin Clone 215	-2.99	0.000	actin 2 [Heterorhabditis bacteriophora]	1.00E-31

continued on next page

Table 9 – continued from previous page

ID	logFC	P Value	Definition	E Value
37B-A01.g	-2.47	0.000	hemocyanin alpha-subunit [Homarus americanus]	1.00E-69
37C-H07.g	-1.76	0.006	ATP synthase beta chain [Caenorhabditis briggsae]	3.00E-47
37D-H04.g	-1.67	0.044	vitellogenin [Macrobrachium rosenbergii]	1.00E-79
39B-G12.g	-1.41	0.009	serine protease SP24D precursor [Drosophila pseudoobscura]	1.00E-16
37D-E12.g	-1.39	0.009	phosphoenolpyruvate carboxykinase [Nephrops norvegicus]	3.00E-26
45Rev2-A12.g	-1.34	0.032	vitellogenin [Macrobrachium rosenbergii]	4.00E-65
16B-D10.g	-1.17	0.000	hemocyanin [Palaeomonetes pugio]	2.00E-46
16A-D04.g	-1.11	0.000	hemocyanin [Litopenaeus vannamei]	3.00E-65
41C-F10.g	-0.98	0.004	hemocyanin [Penaeus monodon]	7.00E-76
37C-B02.g	-0.96	0.018	hemocyanin 2 [Pacifastacus leniusculus]	1.00E-69
37D-G08.g	-0.92	0.040	hemocyanin [Penaeus monodon]	2.00E-66
37A-H06.g	-0.92	0.001	hemocyanin [Litopenaeus vannamei]	3.00E-74
16A-A02.g	-0.88	0.001	hemocyanin [Penaeus monodon]	2.00E-66
16A-H05.g	-0.80	0.000	cathepsin L [Pandalus borealis]	1.00E-106
37A-C08.g	-0.80	0.044	gamma-interferon inducible lysosomal thiol reductase [Amblyomma americanum]	0
16B-A11.g	-0.79	0.023	hemocyanin [Penaeus monodon]	2.00E-75

continued on next page

Table 9 – continued from previous page

ID	logFC	P Value	Definition	E Value
41C-B09.g	-0.77	0.017	hemocyanin [Litopenaeus vannamei]	1.00E-72
41C-G09.g	-0.75	0.043	hemocyanin [Litopenaeus vannamei]	5.00E-69
37B-G08.g	-0.63	0.017	hemocyanin [Litopenaeus vannamei]	7.00E-72
18A-H01.g	-0.56	0.005	cytochrome c oxidase subunit I [Macrobrachium rosenbergii]	1.00E-116
18F-F01.g	-0.55	0.031	hemocyanin [Litopenaeus vannamei]	2.00E-33
37B-C02.g	-0.51	0.040	cathepsin L [Pandalus borealis]	2.00E-33
41A-H08.g	-0.51	0.043	chloride intracellular channel 6-like protein [Aedes aegypti]	2.00E-71
37B-G01.g	-0.50	0.043	glucoamylase GLU1 (Glucan 1,4-alpha-glucosidase)	3.00E-05
41A-G07.g	-0.50	0.033	cathepsin L [Pandalus borealis]	1.00E-106
Chronic DO: Hypoxia vs. Normoxia 24 hrs				
up-regulated				
41C-A02.g	1.06	0.002	hemocyanin [Penaeus monodon]	2.00E-53
16A-D04.g	0.92	0.001	hemocyanin [Litopenaeus vannamei]	3.00E-65
41C-B09.g	0.91	0.006	hemocyanin [Litopenaeus vannamei]	1.00E-72

continued on next page

Table 9 – continued from previous page

ID	logFC	P Value	Definition	E Value
41C-G09.g	0.90	0.016	hemocyanin [Litopenaeus vannamei]	5.00E-69
16A-A02.g	0.88	0.002	hemocyanin [Penaeus monodon]	2.00E-66
41C-F10.g	0.83	0.047	hemocyanin [Penaeus monodon]	7.00E-76
41C-E06.g	0.78	0.047	hemocyanin [Penaeus monodon]	8.00E-76
16B-F08.g	0.74	0.011	hemocyanin 2 [Pacifastacus leniusculus]	2.00E-46
18F-F01.g	0.63	0.017	hemocyanin [Litopenaeus vannamei]	2.00E-33
39B-F04.g	0.56	0.047	C-type lectin [Litopenaeus vannamei]	1.00E-04
Chronic DO: Hypoxia vs. Normoxia 48 hrs				
up-regulated				
16B-F08.g	1.30	0.000	hemocyanin 2 [Pacifastacus leniusculus]	2.00E-46
45B-A11.g	1.25	0.000	crustapain [Pandalus borealis]	9.00E-50
45B-B03.g	1.21	0.000	beta-carotene dioxygenase 2 [Apis mellifera]	1.00E-37
45B-B06.g	1.18	0.000	chitinase 1 [Penaeus monodon]	3.00E-45
37D-H12.g	1.10	0.001	ferritin [Litopenaeus vannamei]	7.00E-47

continued on next page

Table 9 – continued from previous page

ID	logFC	P Value	Definition	E Value
37B-G08.g	0.95	0.000	hemocyanin [Litopenaeus vannamei]	7.00E-72
41C-F10.g	0.88	0.012	hemocyanin [Penaeus monodon]	7.00E-76
16B-A11.g	0.88	0.007	hemocyanin [Penaeus monodon]	2.00E-75
16A-D04.g	0.86	0.001	hemocyanin [Litopenaeus vannamei]	3.00E-65
16A-H05.g	0.63	0.002	cathepsin L [Pandalus borealis]	1.00E-106
45B-D04.g	0.56	0.007	vitellogenin [Macrobrachium rosenbergii]	1.00E-18
37B-C02.g	0.55	0.018	cathepsin L [Pandalus borealis]	2.00E-33
45A-A07.g	0.49	0.009	guanine nucleotide-binding protein gamma subunit [Sitobion avenae]	3.00E-17
16B-H12.g	0.39	0.043	heme-binding protein 2 [Balanus amphitrite]	3.00E-12
18A-G02.g	0.39	0.027	ferritin subunit (glycosylated) [Nilaparvata lugens]	8.00E-04
down-regulated				
TRH Clone 174	-1.83	0.018	tracheless [Tribolium castaneum]	1.00E-114
45Rev2-A10.g	-1.60	0.045	vitellogenin [Macrobrachium rosenbergii]	1.00E-79
39B-G12.g	-1.22	0.031	serine protease SP24D precursor [Drosophila pseudoobscura]	1.00E-16
16A-E12.g	-0.55	0.022	vitellogenin [Cherax quadricarinatus]	6.00E-04

continued on next page

Table 9 – continued from previous page

ID	logFC	P Value	Definition	E Value
Chronic DO: Hypoxia vs. Normoxia 120 hrs				
up-regulated				
45Rev2-A12.g	2.43	0.000	vitellogenin [Macrobrachium rosenbergii]	4.00E-65
35B-F07.g	1.63	0.009	H/ACA ribonucleoprotein complex subunit 4 [Danio rerio]	3.00E-69
45A-A11.g	1.59	0.002	hemocyanin 2 [Pacifastacus leniusculus]	1.00E-103
37D-E12.g	1.39	0.009	phosphoenolpyruvate carboxykinase [Nephrops norvegicus]	3.00E-26
16B-C06.g	0.83	0.013	phosphoenolpyruvate carboxykinase [Litopenaeus vannamei]	4.00E-91
16B-F08.g	0.66	0.017	hemocyanin 2 [Pacifastacus leniusculus]	2.00E-46
Chronic DO: Hypoxia vs. Normoxia 240 hrs				
up-regulated				
39D-G11.g	1.42	0.011	C-type lectin [Litopenaeus vannamei]	3.00E-18
16B-C06.g	0.79	0.019	phosphoenolpyruvate carboxykinase [Litopenaeus vannamei]	4.00E-91

continued on next page

Table 9 – continued from previous page

ID	logFC	P Value	Definition	E Value
down-regulated				
37D-G06.g	-1.53	0.001	asparaginyl endopeptidase [Xenopus laevis]	2.00E-39
41C-G09.g	-1.46	0.000	hemocyanin [Litopenaeus vannamei]	5.00E-69
41C-E06.g	-1.33	0.000	hemocyanin [Penaeus monodon]	8.00E-76
41A-F07.g	-1.16	0.004	ATP synthase coupling factor 6 [Anopheles gambiae]	1.00E-29
37A-C08.g	-1.07	0.001	gamma-interferon inducible lysosomal thiol reductase [Amblyomma americanum]	0
37A-G01.g	-1.03	0.002	chitinase [Litopenaeus vannamei]	7.00E-22
37B-C01.g	-0.91	0.017	oxidoreductase [Anopheles gambiae]	2.00E-11
45B-B07.g	-0.87	0.000	chitinase [Marsupenaeus japonicus]	5.00E-46
37B-D06.g	-0.84	0.037	acetylcholine receptor protein [Tribolium castaneum]	9.00E-07
45B-A09.g	-0.82	0.000	protein pelota [Tribolium castaneum]	2.00E-32
45B-A07.g	-0.79	0.000	polo-like kinase 1 [Drosophila melanogaster] (PLK1)	3.00E-86

Gene Ontology and KEGG

The differentially expressed genes identified at all time points during chronic DO exposure were annotated by sequence similarity comparison against the genomic *Drosophila melanogaster* RefSeq databases with BLAST algorithm (BLASTX and BLASTN) (Altschul et al., 1997). The matched RefSeq identifiers were mapped to the corresponding Entrez Gene identifiers and GO terms assigned in org.Dm.eg.db package. GOstats package from R was employed to assign probable GO terms to all annotated genes. A total of 291, 129, and 219 genes were assigned to the three main groups in GO: biological process, cellular components, and molecular function. Table 10 lists the detailed assignment of genes to GO functional categories. Only GO terms that have at least 2 genes assigned to them are shown.

Table 10: Distribution of differentially expressed genes into different GO categories. Only GO terms with at least two genes assigned to are shown.

GO Terms	GO IDs	Counts
Biological Process		
transport	GO:0006810	20
metabolic process	GO:0008152	15
defense response	GO:0006952	14
proteolysis	GO:0006508	9
mitotic spindle organization and biogenesis	GO:0007052	6
protein catabolic process	GO:0030163	5
salivary gland cell autophagic cell death	GO:0035071	5
autophagic cell death	GO:0048102	5
translation	GO:0006412	4
lipid transport	GO:0006869	4

continued on next page

Table 10 – continued from previous page

GO Terms	GO IDs	Counts
mitotic spindle elongation	GO:0000022	3
cytokinesis	GO:0000910	3
mitochondrial electron transport, cytochrome c to oxygen	GO:0006123	3
regulation of transcription, DNA-dependent	GO:0006355	3
phagocytosis, engulfment	GO:0006911	3
mitosis	GO:0007067	3
aerobic respiration	GO:0009060	3
ATP synthesis coupled proton transport	GO:0015986	3
proton transport	GO:0015992	3
sleep	GO:0030431	3
ribosome biogenesis and assembly	GO:0042254	3
meiotic spindle organization and biogenesis	GO:0000212	2
nuclear mRNA splicing, via spliceosome	GO:0000398	2
pseudouridine synthesis	GO:0001522	2
chitin metabolic process	GO:0006030	2
chitin catabolic process	GO:0006032	2
gluconeogenesis	GO:0006094	2
regulation of transcription from RNA polymerase II promoter	GO:0006357	2
rRNA processing	GO:0006364	2
protein amino acid phosphorylation	GO:0006468	2
iron ion transport	GO:0006826	2
cellular iron ion homeostasis	GO:0006879	2
cytoskeleton organization and biogenesis	GO:0007010	2
actin filament organization	GO:0007015	2
male meiosis	GO:0007140	2
cell adhesion	GO:0007155	2

continued on next page

Table 10 – continued from previous page

GO Terms	GO IDs	Counts
germ cell development	GO:0007281	2
germarium-derived egg chamber formation	GO:0007293	2
germarium-derived oocyte fate determination	GO:0007294	2
axon guidance	GO:0007411	2
open tracheal system development	GO:0007424	2
visual perception	GO:0007601	2
visual behavior	GO:0007632	2
protein localization	GO:0008104	2
tRNA pseudouridine synthesis	GO:0031119	2
olfactory behavior	GO:0042048	2
phototaxis	GO:0042331	2
apical protein localization	GO:0045176	2
male courtship behavior, veined wing generated song production	GO:0045433	2
neuron development	GO:0048666	2
dendrite morphogenesis	GO:0048813	2
asymmetric neuroblast division	GO:0055059	2
Cellular Component		
extracellular region	GO:0005576	10
nucleus	GO:0005634	9
lipid particle	GO:0005811	9
mitochondrion	GO:0005739	7
larval serum protein complex	GO:0005616	5
cytoplasm	GO:0005737	5
lysosome	GO:0005764	5

continued on next page

Table 10 – continued from previous page

GO Terms	GO IDs	Counts
membrane	GO:0016020	5
integral to membrane	GO:0016021	5
mitochondrial respiratory chain complex IV	GO:0005751	4
extracellular space	GO:0005615	3
intracellular	GO:0005622	3
nucleolus	GO:0005730	3
mitochondrial inner membrane	GO:0005743	3
plasma membrane	GO:0005886	3
integral to plasma membrane	GO:0005887	3
cytosolic large ribosomal subunit	GO:0022625	3
kinetochore	GO:0000776	2
spliceosome	GO:0005681	2
mitochondrial proton-transporting ATP synthase, central stalk	GO:0005756	2
ferritin complex	GO:0008043	2
Molecular Function		
oxygen transporter activity	GO:0005344	33
protein binding	GO:0005515	7
zinc ion binding	GO:0008270	6
oxidoreductase activity	GO:0016491	6
nucleotide binding	GO:0000166	5
cathepsin L activity	GO:0004217	5
nutrient reservoir activity	GO:0045735	5
nucleic acid binding	GO:0003676	4
transcription factor activity	GO:0003700	4

continued on next page

Table 10 – continued from previous page

GO Terms	GO IDs	Counts
RNA binding	GO:0003723	4
mRNA binding	GO:0003729	4
cytochrome-c oxidase activity	GO:0004129	4
lipid transporter activity	GO:0005319	4
iron ion binding	GO:0005506	4
calcium ion binding	GO:0005509	4
ATP binding	GO:0005524	4
chitin binding	GO:0008061	4
heme binding	GO:0020037	4
structural constituent of ribosome	GO:0003735	3
actin binding	GO:0003779	3
binding	GO:0005488	3
lipid binding	GO:0008289	3
hydrogen-exporting ATPase activity, phosphorylative mechanism	GO:0008553	3
cation binding	GO:0043169	3
hydrogen ion transporting ATP synthase activity, rotational mechanism	GO:0046933	3
hydrogen ion transporting ATPase activity, rotational mechanism	GO:0046961	3
RNA polymerase II transcription factor activity	GO:0003702	2
mRNA 3'-UTR binding	GO:0003730	2
catalytic activity	GO:0003824	2
endonuclease activity	GO:0004519	2
chitinase activity	GO:0004568	2
phosphoenolpyruvate carboxykinase (GTP) activity	GO:0004613	2
protein kinase activity	GO:0004672	2
protein serine/threonine kinase activity	GO:0004674	2
pseudouridylate synthase activity	GO:0004730	2

continued on next page

Table 10 – continued from previous page

GO Terms	GO IDs	Counts
GTP binding	GO:0005525	2
poly-pyrimidine tract binding	GO:0008187	2
ferrous iron binding	GO:0008198	2
ferric iron binding	GO:0008199	2
electron carrier activity	GO:0009055	2
tRNA-pseudouridine synthase activity	GO:0016439	2
protein homodimerization activity	GO:0042803	2
sequence-specific DNA binding	GO:0043565	2

Chronic DO regulated the expression of genes associated with a broad range of biological processes. The most abundant groups of genes were associated with transport, metabolic process, defense response, and proteolysis. Some genes were involved in lipid transport, aerobic respiration, iron ion transport and homeostasis, and mitochondrial electron transport (cytochrome c to oxygen). Chronic DO also altered the expression of genes in a variety of cellular component locations, with extracellular region, nucleus, lipid particle, and mitochondrion genes being the top four groups. Additional genes were listed as mitochondrial respiratory chain complex IV, mitochondrial inner membrane, and ferritin complex. According to molecular functions, 33 genes were assigned to oxygen transport activity. Additional genes were mapped to GO binding terms, such as nucleotide binding, protein binding, DNA binding, ATP binding, and various ion bindings.

The Entrez Gene identifiers assigned in org.Dm.eg.db package were used to map the corresponding computed gene (CG) accession numbers in FlyBase. Custom scripts were used to retrieve pathway ID and descriptions associated with CG numbers. The differentially expressed genes were mapped to KEGG metabolic and regulatory pathways according to the gene distribution in *Drosophila* pathway database. Oxidative phosphorylation/Citrate cycle and Ribosome were the most abundant categories for chronic DO

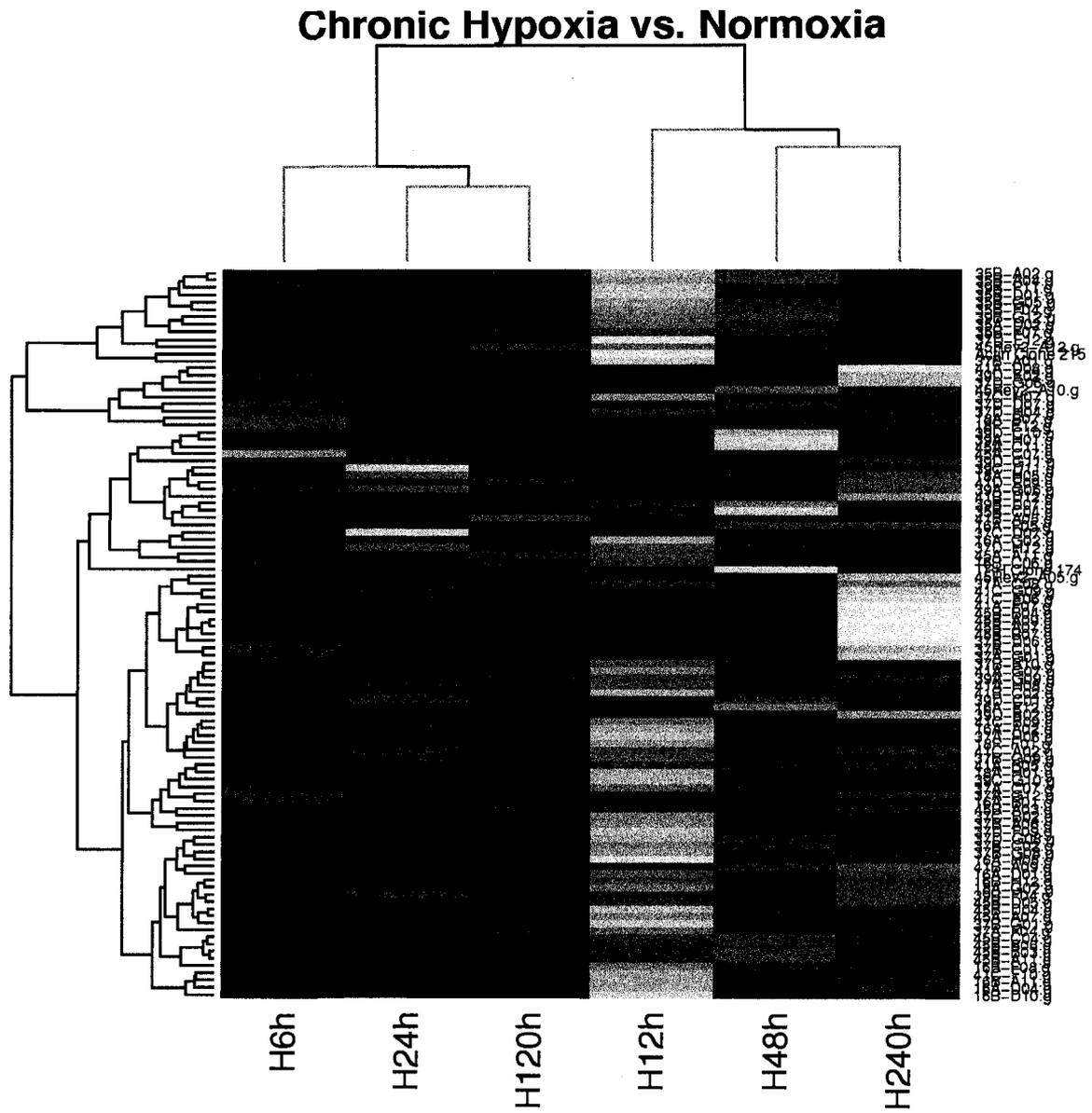


Figure 12: Hierarchical clustering using the differentially expressed genes ($p < 0.05$) in hypoxia vs. normoxia at the same time points.

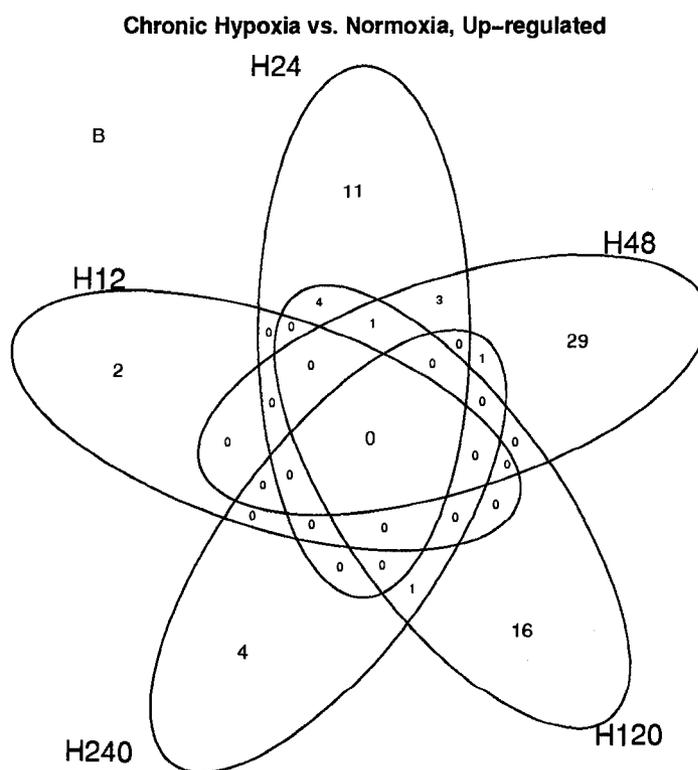
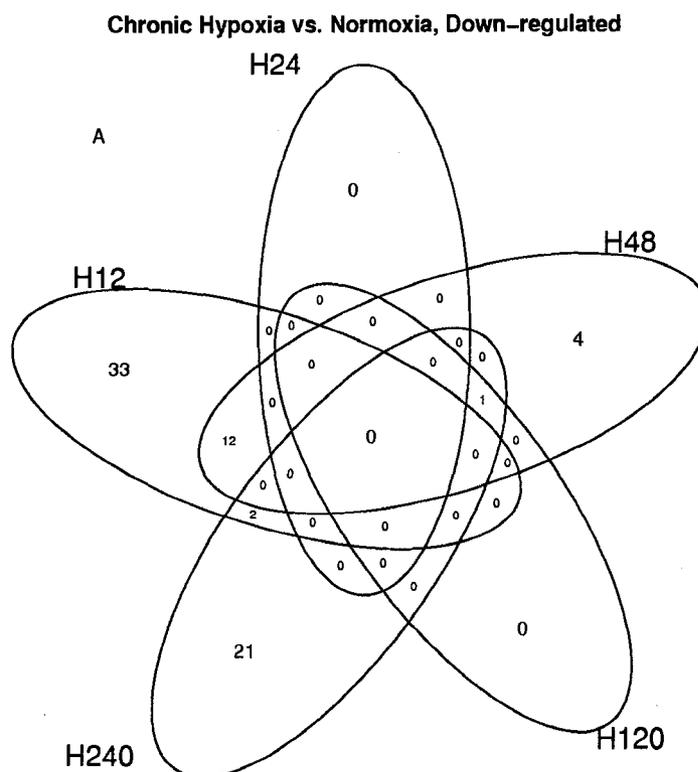


Figure 13: Venn diagram of chronic hypoxia vs. normoxia at the same time points, down-regulated (A) and up-regulated genes (B) genes.

exposure (Figure 14).

qPCR

Gene expression levels of selected genes that were significantly up- or down-regulated ($p < 0.05$) after 12 and 48 hours of chronic hypoxia exposure according to microarray analysis were also measured using qPCR. Grass shrimp 18S rRNA was used as internal standard and gene expression data for each time point are shown in Table 11. Of the 19 genes examined, the change in expression of 17 genes corresponded in direction (up or down) to the change observed on the microarrays, the expression of 1 was unaltered whereas another was down on microarrays but up according to PCR. However none of the changes measured by qPCR was statistically significant ($p=0.05$).

Discussion

The first aim of this study was to determine whether HIF 1 α expression can be used as biomarker to monitor the oxygen stress in aquatic organism. However, four HIF 1 α clones don't show differentially expression during chronic hypoxia exposures. Similarly, early studies using custom cDNA macroarray also showed there was no significant difference among the expression levels of HIF 1 α under normoxic, moderate (2.5 ppm DO), and severe (1.5 ppm DO) chronic hypoxia (Li and Brouwer, 2007, Chapter 3).

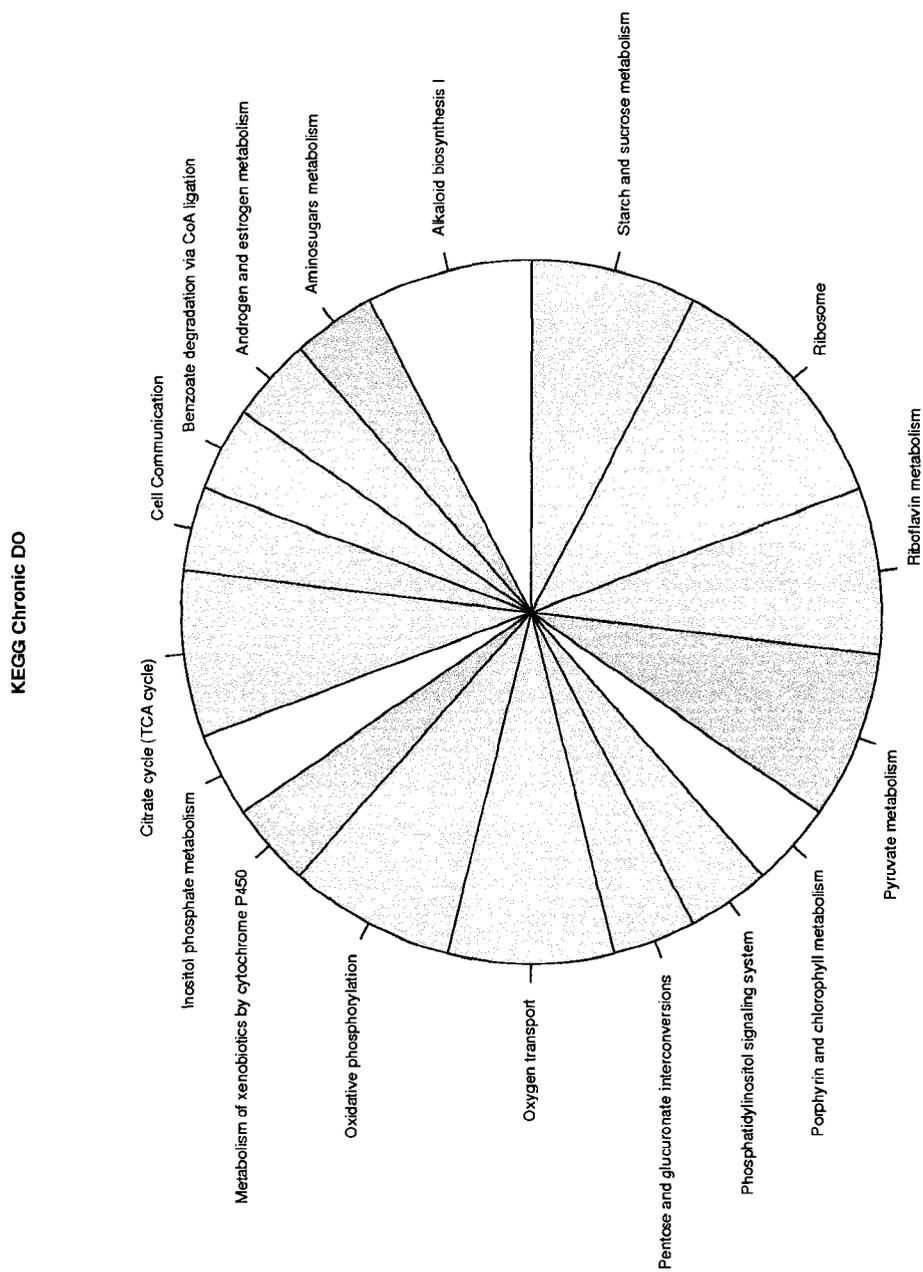


Figure 14: Pie chart of KEGG pathways.

Table 11: Summary of qPCR.

Name	ID	H12 vs. N12		H12 vs. N12		H48 vs. N48		H48 vs. N48	
		Fold Change qPCR	Fold Change Array						
vitellogenin	16A-E12					0.444		0.683	
cathepsin L	16A-H05	0.768		0.573		1.066		1.549	
phosphoenolpyruvate carboxykinase	16B-C06					1.548		1.409	
hemocyanin 2	16B-F08	0.563		0.751		2.022		2.466	
cytochrome c oxidase subunit I	18A-H01	0.454		0.680					
chloride intracellular channel 6-like protein	41A-H08	0.456		0.705					
hemocyanin	41C-G09	0.383		0.595					
Ferredoxin	45A-B01					1.238		1.311	
kinase	45B-A07					2.406		1.328	
crustapain	45B-A11	0.486		0.871		1.635		2.380	
dioxygenase	45B-B03	0.558		0.863		1.434		2.305	
chitinase 1	45B-B06	0.396		0.848		1.211		2.270	

Beta-actin and trachealess characterized from previous studies are the only genes that are significantly down-regulated at H12 and H48, respectively. BLASTX search against non-redundant protein database (nr) revealed 74% identity with protein trachealess from *Tribolium castaneum* (red flour beetle, XP_967112). The fragment of grass shrimp trachealess on the microarray encoded 291 amino acids without start and stop codons. It shows a high level of conservation with other trachealess proteins in PAS domains (Li and Brouwer, 2007). The PAS-A (64-116 AA) shows 86% identity with PAS-A from *Drosophila melanogaster* trachealess (AAA96754), and PAS-B (235-291 AA) domain is most similar to that of *Tribolium castaneum* (78%, XP_967112).

Like HIF 1 α , trachealess protein is also a member of bHLH-PAS family (Isaac and Andrew, 1996). PAS domains are recognized as the signaling domains widely distributed in proteins from archaea, eubacteria, fungi, plants, insects, and vertebrates. PAS proteins are always located intracellularly, and can monitor the external and internal environments simultaneously, including changes in oxygen, light, redox potential, other stimuli, and overall energy level of a cell (Glagolev, 1980; Baryshev et al., 1981; Taylor, 1983). However, the mechanism of PAS regulation has not been determined. Oxygen is both a terminal acceptor for oxidative phosphorylation with high ATP yield, and a toxic agent that forms harmful reactive free radicals when partially reduced. It will be interesting to determine whether the down-regulated trachealess gene at 48 hours is caused by PAS domains, which may serve as an early warning system for any reduction in cellular energy levels by detecting changes in the electron transport system.

It is rather remarkable that grass shrimp appear to respond so differently over time to chronic hypoxia. Different softwares were used to analyze microarray data, and comparison of chronic hypoxia induced transcription profiles across the six different time points by unsupervised hierarchical clustering showed similar patterns as illustrated in Figure 12; one up-regulated dominated cluster, including H6, H24, and H120, and one down-regulated dominated cluster, including H12, H48, and H240. Additional analysis

indicated that the response to chronic hypoxia was pronounced and transient at the above experimental time points; grass shrimp don't respond gradually to chronic exposure. Additional cluster analyses performed on three replicate samples at each time point reveal the similarity in transcription profiling among the replicate shrimp at each time point, illustrating the high reproducibility of gene response among individual sample exposed to chronic hypoxia. Interestingly, Brouwer et al. (2007) also observed that the up-regulation of both mitochondrial and Fe-metabolism genes at day 7 was completely reversed by day 14.

Since large portions of potential transcripts generated from subtractive libraries don't have significant hits against different databases, more differentially expressed sequence tags show up in Venn diagrams than annotated tables. The maximum number of regulated genes occurs at 12 and 48 hours with 47 being down-regulated at 12 hours and 34 up-regulated at 48 hours. Both H12 and H48 are in the down-regulated cluster with 12 genes shared between them. Hemocyanin genes, ATP synthase beta chain [*Caenorhabditis briggsae*], phosphoenolpyruvate carboxykinase [*Nephrops norvegicus*], vitellogenin [*Macrobrachium rosenbergii*], cytochrome c oxidase subunit I [*Macrobrachium rosenbergii*], and glucoamylase GLU1 (Glucan 1,4-alpha-glucosidase) are the unique down-regulated genes in H12. Generally, more genes are down-regulated than up-regulated at the above experimental time points, and few genes were altered after 120 hours. Thus, hemocyanin genes, ATP synthase, phosphoenolpyruvate carboxykinase, vitellogenin, cytochrome c oxidase subunit I, Lysosomal thiol reductase, and C-type lectin may be used as molecular indicators at certain time points of chronic hypoxia treatment in grass shrimp (Brouwer et al., 2007). However, changes of significant genes are too dynamic to serve as biomarkers of hypoxia stress in grass shrimp.

All putative transcripts on microarrays were selected by the E values of sequence similarity search against protein databases, however, few clones also show a high nucleotide similarity with 18S and 28S subunit ribosomal RNA genes related to *Palaemon-*

etes vulgaris, *Palaemonetes paludosus*, and *Palaemon serenus*. Since these clones were derived from SSH libraries, prepared from mRNA purified from total RNA by 2 binding steps on oligo(dT) columns, this suggest that grass shrimp 28S and 18S rRNA may have poly-A tails just as mRNA. In this study, the RNA used for microarray was reverse-transcribed with Oligo(dT)₁₈ primer, while the same RNA was amplified in the presence of random hexamers for qPCR. For microarray, 18S and 28S rRNA appears to be up-regulated after 6, 24, and 120 hours, and down-regulated after 12 and 24 hours. None of them show up after 10 days. Similar up- and down-regulated 16S mitochondrial rRNA was observed after 7- and 14-day exposure to chronic hypoxia (Brouwer et al., 2007). For qPCR, grass shrimp 18S rRNA was constantly expressed as evidenced by highly reproducible constant C_t value. Eukaryotic mRNA is transcribed and then polyadenylated at the 3' end by poly-A polymerase upon termination of transcription of the primary mRNA transcript. The exact role of poly-A tails is still unclear but it seems to play a key role in translation initiation, stability, and nuclear export (Proudfoot and O'Sullivan, 2002; Slomovic et al., 2006). Recent studies demonstrated that polyadenylated 28S rRNA undergoes extensive post-transcriptional processing, and the degree of rRNA polyadenylation can vary between different strains and life stages in *Leishmania* (Decuypere et al., 2005). The purpose and mechanism of polyadenylation in crustacean is not clear. We tentatively conclude that the degree of polyadenylation of 18S and 28S rRNA is oxygen dependent, and that the increased levels of 18S and 28S rRNA on the arrays are due to increased polyadenylation, but not to increased 18/28S rRNA gene transcription.

The information on individual genes obtained from microarray experiments needs to be translated into knowledge of the biological processes and molecular pathways affected. The Gene Ontology (GO) Consortium (Ashburner et al., 2000) has developed a controlled vocabulary that describes the biological processes, molecular functions, and cellular components associated with a particular gene product, and so acts as a repository of the known functional biological information on each gene. Several software packages

are created to visualize metabolic and signaling pathways, such as Gene Map Annotator and Pathway Profiler (GenMAPP) (Dahlquist et al., 2002; Doniger et al., 2003; Salomonis et al., 2007), PANTHER of Applied Biosystems, and Ingenuity Pathways Analysis of Ingenuity. Recently released org.Dm.eg.db package can determine which GO terms or biological pathways of genome-wide *Drosophila* are associated with differentially expressed genes from a microarray experiment, hereinafter referred to as GO and pathway mapping. It provides another valuable way to gain an understanding of the molecular processes affected by hypoxia exposures.

Detailed GO analysis of individual significant genes up-regulated at 6 hours showed the genes are associated with biological processes, such as mitochondrial electron transport, cytochrome c to oxygen, aerobic respiration, iron ion transport, and cellular iron ion homeostasis. At 12 hours, the largest groups of down-regulated biological processes are transport and proton transport, defense response, metabolic process and chitin metabolic process, protein amino acid phosphorylation, ATP synthesis, gluconeogenesis, proteolysis, and autophagic cell death. Protein amino acid phosphorylation and proteolysis are the common processes in down-regulated H12 and H48. This rather simplistic GO analysis merely quantifies the numbers of clones associated with a particular annotation, and it reflects the distribution of annotations related with genes expressed in response to chronic hypoxia.

Additional pathway mapping can facilitate the interpretation of significant gene data derived from complex biological processes and systems, especially characterize differentially expressed genes related to environmental toxicants or stressors (Heinloth et al., 2004; Moggs et al., 2004; Currie et al., 2005). 28% of significantly expressed genes have a match in KEGG pathway targeted by chronic hypoxia. The most abundant categories were associated with ribosome, citrate cycle (TCA cycle), oxidative phosphorylation, and pyruvate metabolism. As discussed in Chapter 3, genes putatively involved in ribosome were exclusively found in cyclic DO libraries in both directions (Chapter 3). Phospho-

enolpyruvate carboxykinase mapped to TCA and pyruvate metabolism pathways was also presented in up-regulated chronic (DO 2.5 mg/L) and down-regulated post molt libraries (Chapter 3).

According to GO-based molecular function, tyrosinase is a copper-dependent phenol oxidase widely occurring in plants and animals, which can oxidize the phenols such as tyrosine and catechol using oxygen (Sussman, 1949). Hydrogens from catechol can combine with oxygen to form water. Tyrosinases are also key components of the primary immune response in arthropods (Johansson and Soderhall, 1996; Soderhall and Cerenius, 1998), and the highly reactive quinones produced by tyrosinase serve to sclerotize the protein matrix of the arthropod cuticle after molting (Sugumaran, 1998; Decker et al., 2007). Since tyrosinase and hemocyanin belong to the same protein superfamily of type 3 copper proteins, all significantly expressed hemocyanin genes of grass shrimp exposed to hypoxia were found in tyrosine metabolism category. Jaenicke and Decker (2003) described the purification of tyrosinases from two crustacean species, *Palinurus elephas* (European spiny lobster) and *Astacus leptodactylus* (freshwater crayfish). The tyrosinase hexamers appear to be similar to the hemocyanins, based on electron microscopy. Because of the structural similarities of tyrosinase and hemocyanin on the level of tertiary and quaternary structure, the tyrosinase proteins appear to have been the ideal predecessors from which to develop the oxygen-carrier protein hemocyanin.

In Chapter 3, a partial cDNA sequence of hemocyanin was identified and characterized by suppression subtractive hybridization (SSH). Hemocyanin is a multi-subunit protein complex, which is conserved in arthropods and mollusks. An arthropod hemocyanin complex is composed of hexamers formed by similar or identical subunits. 38 hemocyanin clones annotated using SwissProt database were amplified and printed on microarray. 17 of these hemocyanin genes are significantly expressed during chronic hypoxia. Two hemocyanin 2 genes [*Pacifastacus leniusculus*] which are up-regulated at H24, H48, and H120 come from up-regulated chronic and cyclic libraries. Generally

chronic hypoxia can force crustaceans to increase the ventilatory flow of water to accelerate the diffusion of oxygen into the blood for the first five days (deFur and Pease, 1988). For long term adaptation, active shrimp favor the most primitive subunits of hemocyanin, such as hemocyanin 2, to increase oxygen affinity (Mangum, 1997). During chronic exposure, hemocyanin doesn't change at 6 hours, however 12 hemocyanin genes are down-regulated at 12 hours, including the only down-regulated hemocyanin from the severe chronic hypoxia SSH library (1.2 ppm DO). Upon longer exposure hemocyanin gene remains up-regulated at 24, 48, and 120 hours. The range of logFC values is 0.63-0.92, 0.86-0.95, and 0.66-1.59 at the above hours, respectively. After 10 days, hemocyanin is down-regulated again.

Quantification of gene expression levels by qPCR confirmed the microarray patterns and ensured the microarray analysis provides a generally accurate picture of gene expression response to chronic hypoxia. Of all the genes validated by qPCR, the expression levels correlated well between qPCR and microarray, however no of the changes measured by qPCR was found to be significant ($p < 0.05$).

This research represents the first grass shrimp cDNA microarray constructed to detect global gene expression changes from normoxic and chronic hypoxic exposed grass shrimp. Hypoxia affected a wide range of cellular processes, and microarray data analysis identified the significantly expressed genes at different time points. GO-based and pathway-based mapping of hypoxia-responsive genes to biological pathways and processes represents a key step in microarray data mining to illustrate why and how genes respond to hypoxia. The utility of shrimp microarray was confirmed in this study. In the future, more genetic information and studies focusing on selected genes and pathways found here will provide further molecular understanding regarding the genetic response to hypoxia.

CHAPTER V

CYCLIC HYPOXIA

Abstract

Laboratory exposures were performed to examine the genomic responses of grass shrimp, *Palaemonetes pugio*, exposed to cyclic hypoxia. Differentially expressed genes were determined in hypoxic vs. normoxic groups after 1, 2, 5 and 10 days exposure to cyclic hypoxia. Sampling on each day was conducted at two different time series, one in the morning (representing low DO, CA) and one in the afternoon (representing high DO, CP). There are distinct differences between the number and identity of specific genes that are significantly down- or up-regulated in shrimp collected at the low DO and high DO points of the cyclic DO cycle. However, cluster analysis showed that the overall response patterns of high (CP) and low DO (CA) exposures were in the same cluster at 24 hrs, 48 hrs, and 120 hrs. In contrast, the response patterns at different time points were in different clusters. After 10 days of exposure to cyclic DO the high DO samples show a dramatic gene up-regulation and do not cluster with any of the other treatment groups. There is no gene shared by any of the eight exposure groups. For genes differentially expressed in samples collected in the morning, 9 of 11 down-regulated genes in day 1 corresponded to hemocyanin. Vitellogenin, cathepsin L, cytochrome c oxidase subunit III, and fatty acid binding protein 10 are the signature down-regulated genes at day 10. For cyclic (low) DO exposure, a total of 127, 44, and 101 genes were assigned to the three main groups in GO: biological process, cellular components, and molecular function. For biological processes, 18, 12, and 11 genes were associated with transport, defense response, and metabolic process. The most abundant group of genes was associated with oxygen transport activity. For cyclic (high) DO exposure, a total of 276, 122, and 229 genes were assigned to the three main groups in GO. Cyclic (high) DO regulated the expression of genes associated with a more broad range of functional categories. The most abundant groups of genes

were associated with transport, defense response, and metabolic process. For molecular functions, 33 and 8 genes were assigned to oxygen transport activity and ATP binding, respectively. The differentially expressed genes were mapped to KEGG metabolic and regulatory pathways according to the gene distribution in *Drosophila* pathway database. Cyclic (high) DO affected a broad range of pathways compared to cyclic (low) DO.

Keywords - *Palaemonetes pugio*; grass shrimp; crustacean; microarray; hypoxia; gene expression; annotation.

Introduction

Hypoxia refers to a state of oxygen deficiency, which is observed frequently in estuarine waters of Gulf of Mexico. Over the last few decades, there have been increases in the frequency, duration, and spatial extent of hypoxic events, which are regarded as one of the major factors responsible for declines in habitat quality and harvestable resources in estuarine ecosystems (Rabalais et al., 1999, Chapter 1). Hypoxia can profoundly affect aquatic ecosystem and have a variety of impacts on associated species. The responses of estuarine fishes and crustaceans to hypoxia can lead to behavioral, physiological, and cellular and molecular changes depending on the duration and severity of hypoxia (Chapter 1).

Chronic and cyclic (intermittent) hypoxia occurs naturally in shallow estuarine ecosystem. Adaptations to chronic hypoxia include avoidance or escape for some mobile species (Wannamaker and Rice, 2000; Wu et al., 2002; Bell and Eggleston, 2005; Craig et al., 2005), respiratory regulations via physiological mechanisms (Johnson et al., 1984; Hagerman, 1986; deFur and Pease, 1988; Mangum and Rainer, 1988; Mangum, 1997), and molecular responses of differentially expressed genes in fishes (Gracey et al., 2001; Ton et al., 2002, 2003; van der Meer et al., 2005) and invertebrates (Brouwer et al., 2005, 2007; Brown-Peterson et al., 2005; David et al., 2005).

In addition to chronic hypoxia, oxygen concentrations may vary throughout the

day. The daily cyclical pattern of hypoxia and normoxia caused by biotic and abiotic factors occurs naturally, but are exacerbated by eutrophication, in microtidal estuaries of Gulf of Mexico, especially in the summer. Thus, estuarine organisms face increases in amplitude and frequency of hypoxia/normoxia cycles (Gupta et al., 1996). Cyclic hypoxia can cause severe organ damage in mammalian species through the generation of reactive oxygen species (ROS), however, little is known about the effects of cyclic hypoxia on aquatic species. Compared to severe hypoxia or anoxia that cause mortality instantly, sublethal levels of cyclic hypoxia in aquatic ecosystem are commonly less severe, longer-lasting, and more widespread.

A comparative study was conducted by Coiro et al. (2000) to evaluate the effects of diurnal, semidiurnal, and constant hypoxia on the growth of first stage larval marsh grass shrimp, *Palaemonetes vulgaris*. Compared to normoxia, any hypoxia can cause growth impairment. Moreover, there is a significant difference in growth impairment between chronic and cyclic exposures, and cyclic exposure results in less growth impairment than chronic exposure. Stierhoff et al. (2006) measured the growth and feeding rates of two estuary-dependent juvenile flounders exposed to sublethal hypoxia over a range of temperatures. Generally growth rates of both fishes were reduced as DO decreased, and also as temperature increased. Growth was significantly reduced by 90% or 100% at 2.0 mg/L DO. Cyclic hypoxia also caused significant growth limitation. Growths of both fishes were significantly reduced (35-60%) in cyclic hypoxia (2.0-11.0 mg/L DO). Tyler and Targett (2007) reported the ecological impacts of short-term cyclic hypoxia (<2 to 20 mg/L DO) on juvenile weakfish, *Cynoscion regalis*. The distribution and abundance of weakfish demonstrate fish can frequently abandon the preferred habitats whenever DO was <2 mg/L, and return within 2 hours of DO exceeding 2 mg/L. More recently, Brown-Peterson et al. (2008) described a 77-day laboratory experiment demonstrating the effects of cyclic hypoxia on gene expression and reproduction in grass shrimp using a custom cDNA macroarray.

Grass shrimp, *Palaemonetes pugio*, is a hypoxia-tolerant species that is abundant in estuaries along the Gulf of Mexico, and it is an excellent model for investigating the molecular responses to hypoxia (Chapter 1). In the present study we used grass shrimp to examine the effects of cyclic hypoxia on gene expression in laboratory exposures using DNA microarray. We hypothesize grass shrimp exposed to cyclic hypoxia will show different and more complicated changes in gene expression than those exposed to chronic hypoxia. The differences in gene expression profiles between chronic vs. cyclic hypoxia may provide potential biomarkers which can be used to assess and monitor the impacts of cyclic and chronic hypoxia on estuarine resident organisms. Further, the results of laboratory cyclic hypoxia will be compared with those of field cyclic hypoxia, as described in Chapter 6.

Materials and Methods

If not described below, please see Chapter 4.

Laboratory Exposures

Collection and maintenance of grass shrimp prior to exposure experiments were conducted as described in Chapter 3.

Exposures were conducted in an intermittent flow-through system described by Manning et al. (1999). Normoxic (DO 7.5 mg/L) and cyclic hypoxic (DO 1.5-8 mg/L) conditions within the treatment aquaria were established and maintained as described before (Brouwer et al., 2004, 2005, 2007; Brown-Peterson et al., 2005, Chapter 3). In all experiments, oxygen was monitored continuously in one hypoxic flow-through aquarium, and DO, temperature and salinity were measured in all flow-through aquaria once or twice daily using a YSI Model 600XLM data sonde. All exposures were conducted in triplicate for both controls and treatments. The thorax/hepatopancreas of 20 shrimp per treatment was removed after 0, 1, 2, 5, and 10 days of exposure and stored in 1 mL RNALater (Am-

bion Inc. Austin, TX, USA) at -20°C. Sampling on each day was conducted in two different fixed time points, one in the morning (representing low DO) and one in the afternoon (representing high DO).

Experimental Design

A loop design with direct and indirect comparisons was used for cyclic DO exposures (Table 12). Each individual sample was hybridized to each of two different samples in two different dye orientations, and some direct sample-to-sample comparisons were used for comparing important samples.

Table 12: Loop design for comparison of normoxic (N) and cyclic hypoxic (H) samples (0, 1, 2, 5 and 10 days). There were 3 pooled samples for each time point (n=3). This design was used twice. Once for samples collected at low DO and once for samples collected at high DO, for a total of 54 arrays.

Array	Cy3	Cy5
1	N0	N1
2	N1	H1
3	H1	H2
4	H2	N2
5	N2	N5
6	N5	H5
7	H5	H10
8	H10	N10
9	N10	N0

Results

Water Quality Parameters and Grass Shrimp

The range of DO concentration during cyclic hypoxia was 1.0-8.6 (mg/L) (Figure 15). Shrimp were exposed on average to DO <2 mg/L and DO 2-3 mg/L for 11 and 2 hours for every 24 hours cycle. The average length and weight (without eggs) of control grass shrimp and grass shrimp exposed to cyclic hypoxia over the entire exposure

period were 35.28 ± 2.09 (mm) and 354.59 (mg), and 35.20 ± 2.09 (mm) and 346.42 (mg), respectively. At the end of the exposure hepatopancreas tissues were dissected from grass shrimp, weighed, and stored in 1 mL RNALater at -20°C . The mean hepatopancreas weight of control shrimp and shrimp exposed to cyclic hypoxia was 9.15 and 8.88 (mg), respectively.

RNA Extraction and Labeling

RNA was extracted and concentrations were determined using a NanoDrop Spectrophotometer (ND-1000, see Chapter 4). The RNA concentration was 2.47 ± 0.33 $\mu\text{g}/\mu\text{l}$. The ratios of 260/280 and 260/230 were 2.11 ± 0.01 and 2.25 ± 0.07 , respectively. No RNA degradation was shown by running diluted samples on RNA Nano Chips (Agilent Bioanalyzer 2100). After labeling and purification, the absorbance of labeled cDNA was measured at 260, 550, and 650 nm, and the absorbance readings were used to calculate the frequency of incorporation (FOI). The average FOI for Cy3 and Cy5 was 24.38 and 17.82 per 1000 nucleotides, respectively.

Gene Expression During Cyclic Hypoxia Exposures

This study determined differentially expressed genes at a p value of less than 0.05 that occurred in hypoxic vs. normoxic groups at the same time point (Table 14). Combined with the annotation tables (see Chapter 3) generated by sequence similarity search against the protein databases using BLAST algorithm (Altschul et al., 1997), Table 13 shows the differentially expressed genes with BLAST E value less than $1\text{E}-5$ in hypoxic vs. normoxic groups at the same time point. The genes identified here were derived from two different time series, one in the morning (representing low DO, indicated as CA) and one in the afternoon (representing high DO, indicated as CP).

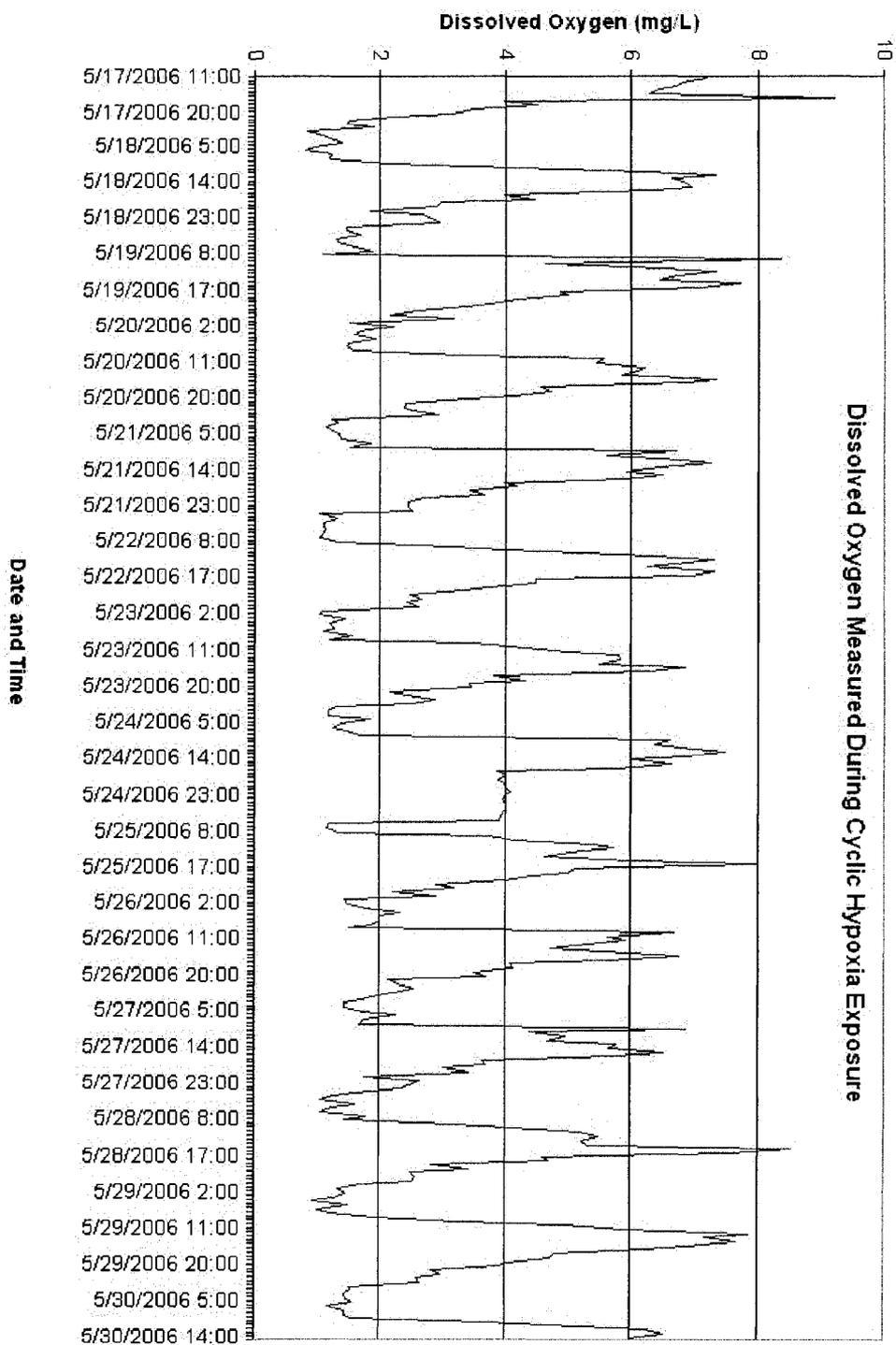


Figure 15: Dissolved oxygen measured during cyclic DO exposure.

Table 13: Differentially expressed genes in cyclic hypoxia vs. normoxia at the same time points.

ID	logFC	P Value	Definition	E Value
Cyclic (low) DO: Hypoxia vs. Normoxia 24 hrs				
up-regulated				
41A-D08.g	1.76	0.049	beta actin [Cherax quadricarinatus]	5.00E-22
39D-G11.g	1.51	0.011	C-type lectin [Litopenaeus vannamei]	3.00E-18
down-regulated				
37D-G08.g	-0.95	0.049	hemocyanin [Penaeus monodon]	2.00E-66
41C-F10.g	-0.84	0.037	hemocyanin [Penaeus monodon]	7.00E-76
41C-G09.g	-0.80	0.042	hemocyanin [Litopenaeus vannamei]	5.00E-69
16B-D10.g	-0.79	0.049	hemocyanin [Palaemonetes pugio]	2.00E-46
16A-A02.g	-0.77	0.009	hemocyanin [Penaeus monodon]	2.00E-66
16B-A11.g	-0.77	0.048	hemocyanin [Penaeus monodon]	2.00E-75
41C-B09.g	-0.73	0.042	hemocyanin [Litopenaeus vannamei]	1.00E-72

continued on next page

Table 13 – continued from previous page

ID	logFC	P Value	Definition	E Value
16A-D04.g	-0.70	0.022	hemocyanin [Litopenaeus vannamei]	3.00E-65
18F-F01.g	-0.66	0.008	hemocyanin [Litopenaeus vannamei]	2.00E-33
Cyclic (low) DO: Hypoxia vs. Normoxia 48 hrs				
up-regulated				
16B-F08.g	0.79	0.008	hemocyanin 2 [Pacifastacus leniusculus]	2.00E-46
45A-A07.g	0.47	0.035	guanine nucleotide-binding protein gamma subunit [Sitobion avenae]	3.00E-17
45B-A11.g	0.44	0.045	crustapain [Pandalus borealis]	9.00E-50
Cyclic (low) DO: Hypoxia vs. Normoxia 120 hrs				
Cyclic (low) DO: Hypoxia vs. Normoxia 240 hrs				
down-regulated				
18F-E12.g	-3.20	0.003	vitellogenin [Macrobrachium rosenbergii]	1.00E-91
39D-C10.g	-1.43	0.025	cytochrome c oxidase subunit III [Farfantepenaeus notialis]	3.00E-32

Table 13 – continued from previous page

ID	logFC	P Value	Definition	E Value
41B-D12.g	-1.30	0.019	fatty acid binding protein 10 [Litopenaeus vannamei]	1.00E-24
37A-G12.g	-1.23	0.026	cathepsin L [Pandalus borealis]	4.00E-39
Cyclic (high) DO: Hypoxia vs. Normoxia 24 hrs				
up-regulated				
45Rev2-A12.g	2.88	0.000	vitellogenin [Macrobrachium rosenbergii]	4.00E-65
16A-G02.g	1.88	0.010	hemocyanin 2 [Pacifastacus leniusculus]	1.00E-103
Cyclic (high) DO: Hypoxia vs. Normoxia 48 hrs				
Cyclic (high) DO: Hypoxia vs. Normoxia 120 hrs				
down-regulated				
37B-G01.g	-0.79	0.000	glucoamylase GLU1 (Glucan 1,4-alpha-glucosidase)	3.00E-05
45B-B03.g	-0.67	0.000	beta-carotene dioxygenase 2 [Apis mellifera]	1.00E-37
45B-B07.g	-0.63	0.002	chitinase [Marsupenaeus japonicus]	5.00E-46

continued on next page

Table 13 – continued from previous page

ID	logFC	P Value	Definition	E Value
45B-A07.g	-0.62	0.002	polo-like kinase 1 [Drosophila melanogaster] (PLK1)	3.00E-86
45B-A09.g	-0.60	0.002	protein pelota [Tribolium castaneum]	2.00E-32
18A-G02.g	-0.57	0.000	ferritin subunit (glycosylated) [Nilaparvata lugens]	8.00E-04
16B-H12.g	-0.57	0.001	heme-binding protein 2 [Balanus amphitrite]	3.00E-12
45B-B06.g	-0.54	0.003	chitinase 1 [Penaeus monodon]	3.00E-45
45A-A07.g	-0.53	0.004	guanine nucleotide-binding protein gamma subunit [Sitobion avenae]	3.00E-17
45B-A11.g	-0.51	0.004	crustapain [Pandalus borealis]	9.00E-50
Cyclic (high) DO: Hypoxia vs. Normoxia 240 hrs				
up-regulated				
35B-C04.g	2.45	0.000	PMAV [Penaeus monodon]	4.00E-04
39D-C01.g	1.94	0.000	PMAV [Penaeus monodon]	5.00E-23
39A-H01.g	1.86	0.000	PMAV [Penaeus monodon]	5.00E-23
39D-G11.g	1.69	0.001	C-type lectin [Litopenaeus vannamei]	3.00E-18
41C-A02.g	1.48	0.000	hemocyanin [Penaeus monodon]	2.00E-53

Table 13 – continued from previous page

ID	logFC	P Value	Definition	E Value
35B-F07.g	1.36	0.032	H/ACA ribonucleoprotein complex subunit 4 [Danio rerio]	3.00E-69
41C-E06.g	1.36	0.000	hemocyanin [Panaeus monodon]	8.00E-76
37A-F04.g	1.35	0.000	hemocyanin subunit 1 [Gammarus roeseli]	5.00E-33
45A-C11.g	1.27	0.028	cathepsin C [Marsupenaeus japonicus]	3.00E-30
41C-F10.g	1.26	0.000	hemocyanin [Panaeus monodon]	7.00E-76
18F-F01.g	1.26	0.000	hemocyanin [Litopenaeus vannamei]	2.00E-33
16B-A11.g	1.26	0.000	hemocyanin [Panaeus monodon]	2.00E-75
37A-H06.g	1.20	0.000	hemocyanin [Litopenaeus vannamei]	3.00E-74
16A-A02.g	1.20	0.000	hemocyanin [Panaeus monodon]	2.00E-66
18A-H05.g	1.18	0.014	cytochrome c oxidase subunit III [Farfantepenaeus notialis]	3.00E-25
37B-C01.g	1.17	0.001	oxidoreductase [Anopheles gambiae]	2.00E-11
16A-D04.g	1.14	0.000	hemocyanin [Litopenaeus vannamei]	3.00E-65
16A-E12.g	1.13	0.000	vitellogenin [Cherax quadricarinatus]	6.00E-04
41A-H08.g	1.11	0.000	chloride intracellular channel 6-like protein [Aedes aegypti]	2.00E-71
37D-H12.g	1.10	0.001	ferritin [Litopenaeus vannamei]	7.00E-47

continued on next page

Table 13 – continued from previous page

ID	logFC	P Value	Definition	E Value
37A-C08.g	1.10	0.001	gamma-interferon inducible lysosomal thiol reductase [Amblyomma americanum]	0
39A-G09.g	1.09	0.000	cathepsin L [Pandalus borealis]	1.00E-105
41A-G07.g	1.08	0.000	cathepsin L [Pandalus borealis]	1.00E-106
16B-D10.g	1.06	0.001	hemocyanin [Palaemonetes pugio]	2.00E-46
16B-F08.g	1.05	0.000	hemocyanin 2 [Pacifastacus leniusculus]	2.00E-46
41A-F07.g	1.04	0.013	ATP synthase coupling factor 6 [Anopheles gambiae]	1.00E-29
16A-H05.g	1.04	0.000	cathepsin L [Pandalus borealis]	1.00E-106
37B-B10.g	1.03	0.000	cathepsin L [Pandalus borealis]	1.00E-105
41C-B09.g	1.01	0.000	hemocyanin [Litopenaeus vannamei]	1.00E-72
37B-C02.g	1.01	0.000	cathepsin L [Pandalus borealis]	2.00E-33
41C-G09.g	0.95	0.003	hemocyanin [Litopenaeus vannamei]	5.00E-69
37C-B02.g	0.86	0.027	hemocyanin 2 [Pacifastacus leniusculus]	1.00E-69
39C-G10.g	0.82	0.009	cytochrome c oxidase subunit I [Macrobrachium rosenbergii]	1.00E-116
18A-H01.g	0.81	0.000	cytochrome c oxidase subunit I [Macrobrachium rosenbergii]	1.00E-116
16A-F05.g	0.77	0.044	beta-1,3-glucan binding protein [Penaeus monodon]	6.00E-46

Table 13 – continued from previous page

ID	logFC	P Value	Definition	E Value
41A-B05.g	0.76	0.008	ATP synthase coupling factor 6 [Anopheles gambiae]	1.00E-29
16B-C06.g	0.76	0.021	phosphoenolpyruvate carboxykinase [Litopenaeus vannamei]	4.00E-91
39B-D07.g	0.73	0.025	cytochrome oxidase subunit I [Metapenaeopsis barbata]	2.00E-13
37B-G08.g	0.65	0.008	hemocyanin [Litopenaeus vannamei]	7.00E-72
45B-D05.g	0.53	0.024	sulfide:quinone oxidoreductase [Anopheles gambiae]	1.00E-64
18A-G02.g	0.45	0.005	ferritin subunit (glycosylated) [Nilaparvata lugens]	8.00E-04
45B-A11.g	0.42	0.020	crustapain [Pandalus borealis]	9.00E-50
16B-H12.g	0.40	0.028	heme-binding protein 2 [Balanus amphitrite]	3.00E-12
down-regulated				
18A-B07.g	-2.11	0.029	vitellogenin [Macrobrachium rosenbergii]	1.00E-91

Genes listed with BLAST E-value < 1E-5 and protein names from swissprot or nr databases.

There are distinct differences between the number and identity of genes that are significantly down- or up-regulated in shrimp collected at the low DO (morning) and high DO (afternoon) points of the cyclic DO cycle (Table 14). For the H24 and H48 samples 19 and 6 genes are up-regulated in the morning and only 4 and 0 in the afternoon. 11 and 10 genes are down-regulated in the H24 and H48 “morning” samples and none are down-regulated in the afternoon. After 5 days of cyclic DO the major response is down-regulation of 40 genes in the “afternoon” samples, followed by a dramatic up-regulation of 85 genes in the “afternoon” samples after 10 days.

Table 14: Differentially expressed genes during cyclic hypoxia exposures at the same time points ($p < 0.05$).

Cyclic Low Hypoxia (CA) vs. Normoxia				
Time (Hours)	24	48	120	240
Up-regulated	19	6	0	0
Down-regulated	11	10	1	7
Cyclic High Hypoxia (CP) vs. Normoxia				
Up-regulated	4	0	0	85
Down-regulated	0	0	40	1

Even through there are differences in expression of specific genes between CA and CP, cluster analysis of the 100 most differentially expressed genes ($p < 0.05$) in each exposure group with BLASTE value less than $1E-5$ shows overall gene response patterns to be similar (Figure 16). The heat map represents the cyclic hypoxia-regulated genes, clustered by the correlation coefficient according to their similarity in expression pattern by gene and by treatments (time exposed to cyclic hypoxia). For hypoxia vs. normoxia at 24 hrs, 48 hrs, and 120 hrs, the response patterns of high (CP) and low DO (CA) exposures were in the same cluster, but not significantly, whereas the response patterns at different time points were in different clusters. This indicates that the overall gene expression patterns of cyclic low and high DO on the same day are similar to each other, but the gene expression patterns on different days are dissimilar. After 10 days of exposure to

cyclic DO the low DO samples cluster with the 5 day samples, whereas the 10 day high DO samples show a dramatic gene up-regulation and do not cluster with any of the other treatment groups. There are no genes that cluster significantly using multiscale bootstrap resampling ($\alpha=0.95$).

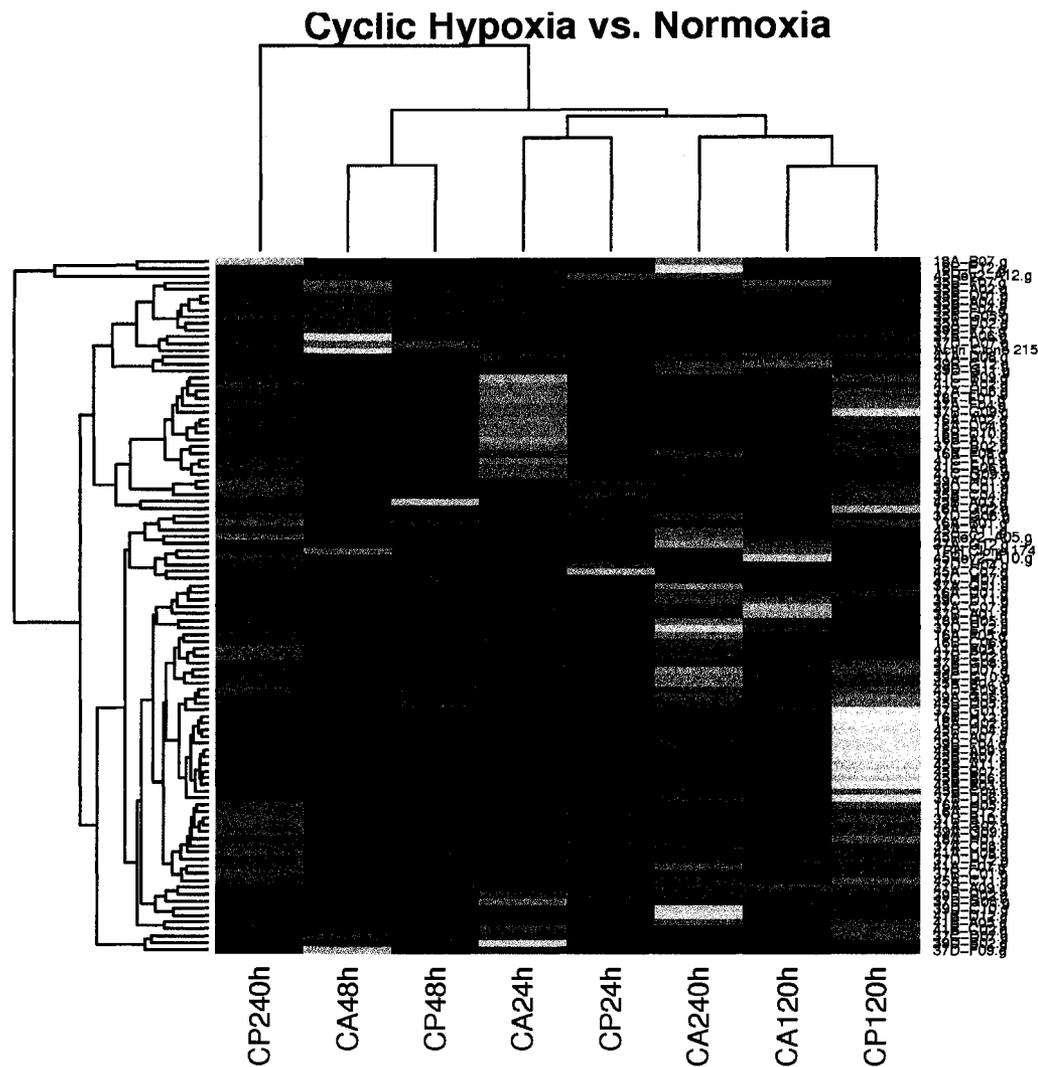


Figure 16: Hierarchical clustering using the differentially expressed genes ($p < 0.05$) in hypoxia vs. normoxia at the same time points.

Gene Expression Changes in Cyclic Low DO Hypoxia Exposures (CA)

Venn diagrams were used to display the differentially expressed genes that were up- or down-regulated during cyclic DO exposures (Figure 17). There is no gene shared

by any of the eight exposure groups. For genes differentially expressed in samples collected in the morning, 9 of 11 down-regulated genes in H24 corresponded to hemocyanin. Hemocyanin 2 [*Pacifastacus leniusculus*], guanine nucleotide-binding protein gamma subunit [*Sitobion avenae*], and crustapain [*Pandalus borealis*] are up-regulated in H48. Vitellogenin [*Macrobrachium rosenbergii*], cathepsin L [*Pandalus borealis*], cytochrome c oxidase subunit III [*Farfantepenaeus notialis*], and fatty acid binding protein 10 [*Litopenaeus vannamei*] are the signature down-regulated genes in H240.

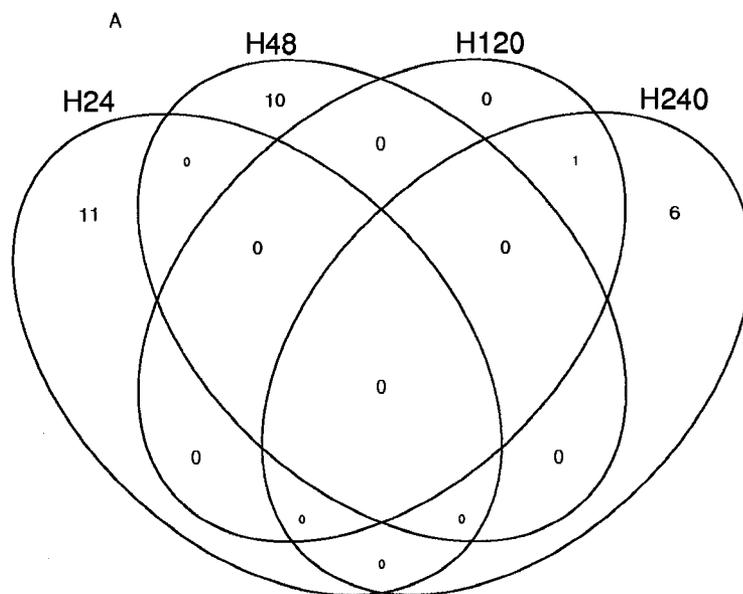
Gene Expression Changes in Cyclic High DO Hypoxia Exposures (CP)

For differentially expressed genes found in samples collected in the afternoon, hemocyanin 2 [*Pacifastacus leniusculus*] and vitellogenin [*Macrobrachium rosenbergii*] are among the up-regulated genes found in H24. Beta-carotene dioxygenase 2 [*Apis mellifera*], heme-binding protein 2 [*Balanus amphitrite*], and crustapain [*Pandalus borealis*] are among the down-regulated genes in H120. Vitellogenin [*Macrobrachium rosenbergii*] is the only down-regulated gene found in H240. A pronounced down-regulation of 40 genes as observed at H120, and 120 hours later there is a striking reversal with 85 genes being up-regulated (Figure 17). Combined with annotation table with BLAST E value less than 1E-5 (Table 13), there are 10 and 43 differentially expressed genes shown in H120 and H240, respectively. Of 43 genes up-regulated in H240, 15, 6, 4, 3, and 2 annotated genes are hemocyanine, cathepsin C or L, cytochrome c oxidase subunit I or III, PmAV [*Penaeus monodon*], and ATP synthase coupling factor 6 [*Anopheles gambiae*], respectively.

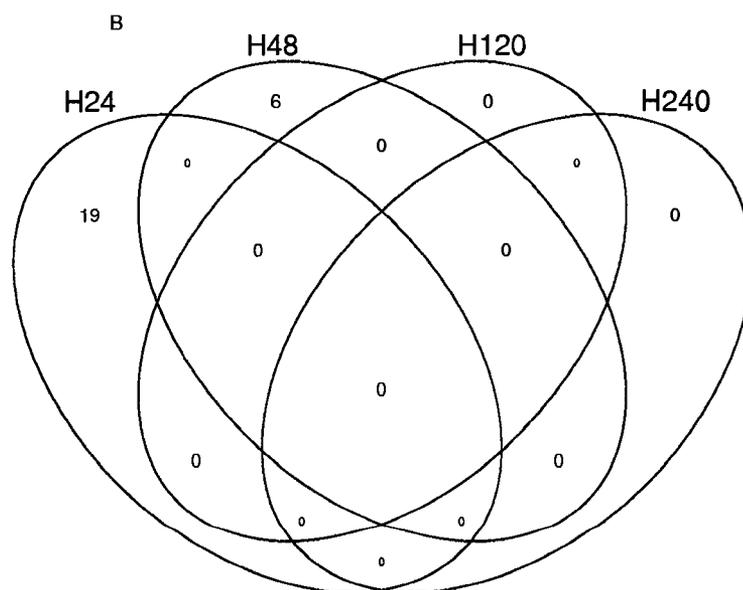
Gene Ontology and KEGG

The differentially expressed genes identified during cyclic DO exposures were annotated by sequence similarity comparison against the genomic *Drosophila melanogaster* RefSeq database with BLAST algorithm (BLASTX and BLASTN) (Altschul et al., 1997). The matched RefSeq identifiers were mapped to the corresponding Entrez Gene identi-

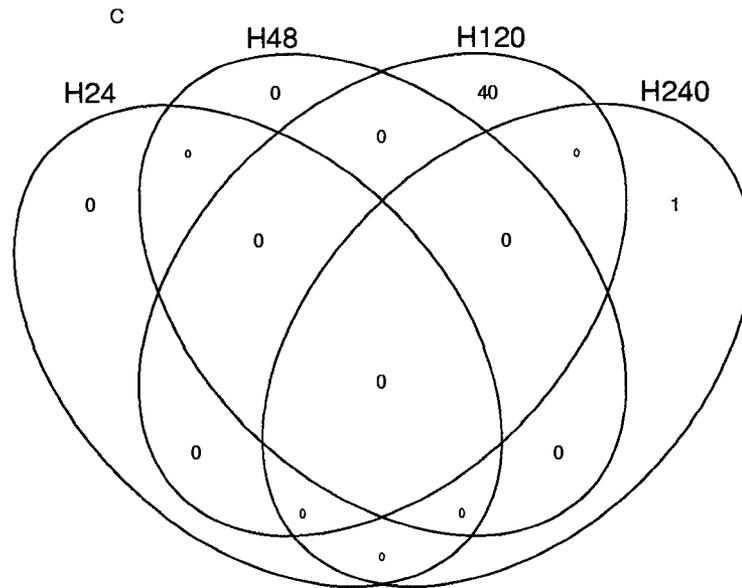
Cyclic (Low) Hypoxia vs. Normoxia, Down-regulated



Cyclic (Low) Hypoxia vs. Normoxia, Up-regulated



Cyclic (High) Hypoxia vs. Normoxia, Down-regulated



Cyclic (High) Hypoxia vs. Normoxia, Up-regulated

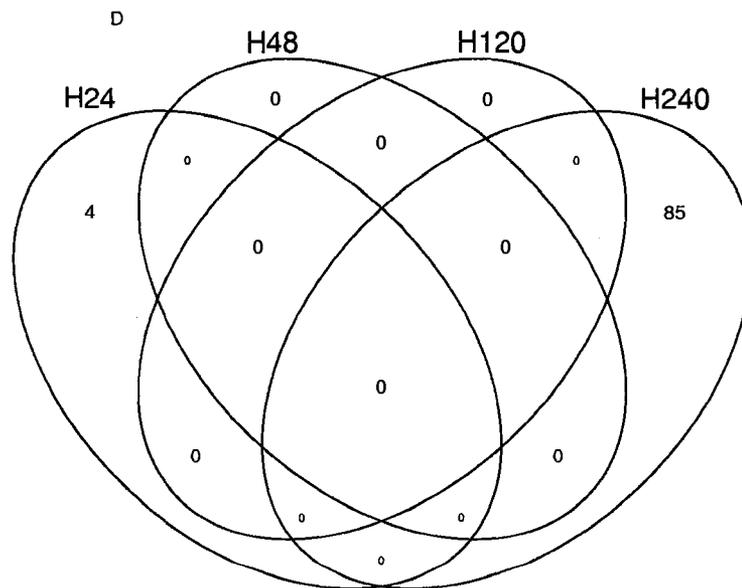


Figure 17: Distribution of differentially expressed genes during cyclic (low) exposure into different GO categories. Only GO terms with at least two genes assigned to are shown.

fiers and GO terms assigned in org.Dm.eg.db package. GOSTats package from R was employed to assign probable GO terms to all annotated genes. Table 15 and Table 16 show the detailed assignment of genes to GO functional categories of cyclic low and high DO exposures, respectively. Only GO terms that have at least 2 genes assigned to them are shown.

For cyclic (low) DO exposure (Table 15), a total of 127, 44, and 101 genes were assigned to the three main groups in GO: biological process, cellular components, and molecular function. For biological processes, 18, 12, and 11 genes were associated with transport, defense response, and metabolic process. According to cellular components, the largest group of genes encoded proteins that are located in the extracellular region. Cyclic (low) DO regulated the expression of genes in various molecular functions. The most abundant group of genes was associated with oxygen transport activity, with additional genes being listed in binding terms, such as DNA binding, ATP binding, protein binding, nucleic acid binding, and various ion binding.

For cyclic (high) DO exposure (Table 16), a total of 276, 122, and 229 genes were assigned to the three main groups in GO: biological process, cellular components, and molecular function. Cyclic (high) DO regulated the expression of genes associated with a more broad range of functional categories. The most abundant groups of genes were associated with transport (19), defense response (16), and metabolic process (16). Additional genes encode proteins involved in proteolysis, protein amino acid phosphorylation, aerobic respiration, iron ion transport and homeostasis, and mitochondrial electron transport (cytochrome c to oxygen). As far as cellular components are concerned, 11 and 8 genes were listed as encoding proteins located in mitochondrion and extracellular region. Additional genes were associated with mitochondrial respiratory chain complex IV, mitochondrial inner membrane, mitochondrial matrix, and ferritin complex. For molecular functions, 33 and 8 genes were assigned to oxygen transport activity and ATP binding, respectively. Various genes were involved in oxidoreductase activity, cytochrome c oxidase

activity, heme binding, cathepsin L/B/K activities, and various ion bindings.

The Entrez Gene identifiers assigned in org.Dm.eg.db package were used to map the corresponding computed gene (CG) accession numbers in FlyBase. Custom scripts were used to retrieve pathway ID and descriptions associated with CG numbers. The differentially expressed genes were mapped to KEGG metabolic and regulatory pathways according to the gene distribution in *Drosophila* pathway database. Cyclic (high) DO regulated a broad range of pathways compared to cyclic (low) DO. Alkaloid biosynthesis I, Riboflavin metabolism, Oxygen transport, and Ribosome were the most abundant categories for cyclic DO exposures (Figure 18).

Table 15: Distribution of differentially expressed genes during cyclic (low) exposure into different GO categories. Only GO terms with at least two genes assigned to are shown.

GO Terms	GO IDs	Counts
Biological Process		
transport	GO:0006810	18
defense response	GO:0006952	12
metabolic process	GO:0008152	11
protein amino acid phosphorylation	GO:0006468	3
proteolysis	GO:0006508	3
mitotic spindle organization and biogenesis	GO:0007052	3
pseudouridine synthesis	GO:0001522	2
DNA metabolic process	GO:0006259	2
rRNA processing	GO:0006364	2
continued on next page		

Table 15 – continued from previous page

GO Terms	GO IDs	Counts
lipid transport	GO:0006869	2
sister chromatid cohesion	GO:0007062	2
mitotic chromosome condensation	GO:0007076	2
cell adhesion	GO:0007155	2
germ cell development	GO:0007281	2
open tracheal system development	GO:0007424	2
protein catabolic process	GO:0030163	2
tRNA pseudouridine synthesis	GO:0031119	2
salivary gland cell autophagic cell death	GO:0035071	2
ribosome biogenesis and assembly	GO:0042254	2
autophagic cell death	GO:0048102	2
Cellular Component		
extracellular region	GO:0005576	5
lipid particle	GO:0005811	5
larval serum protein complex	GO:0005616	4
nucleus	GO:0005634	4
extracellular space	GO:0005615	3
mitochondrion	GO:0005739	3
condensin complex	GO:0000796	2
intracellular	GO:0005622	2
nucleolus	GO:0005730	2
lysosome	GO:0005764	2
continued on next page		

Table 15 – continued from previous page

GO Terms	GO IDs	Counts
cohesin complex	GO:0008278	2
membrane	GO:0016020	2
Molecular Function		
oxygen transporter activity	GO:0005344	26
ATP binding	GO:0005524	6
protein binding	GO:0005515	4
nutrient reservoir activity	GO:0045735	4
RNA binding	GO:0003723	3
transporter activity	GO:0005215	3
zinc ion binding	GO:0008270	3
nucleic acid binding	GO:0003676	2
DNA binding	GO:0003677	2
cathepsin L activity	GO:0004217	2
protein serine/threonine kinase activity	GO:0004674	2
pseudouridylate synthase activity	GO:0004730	2
lipid transporter activity	GO:0005319	2
lipid binding	GO:0008289	2
tRNA-pseudouridine synthase activity	GO:0016439	2
ATPase activity	GO:0016887	2

Table 16: Distribution of differentially expressed genes during cyclic (high) exposure into different GO categories. Only GO terms with at least two genes assigned to are shown.

GO Terms	GO IDs	Counts
Biological Process		
transport	GO:0006810	19
defense response	GO:0006952	16
metabolic process	GO:0008152	16
proteolysis	GO:0006508	7
mitotic spindle organization and biogenesis	GO:0007052	6
translation	GO:0006412	5
protein catabolic process	GO:0030163	5
salivary gland cell autophagic cell death	GO:0035071	5
autophagic cell death	GO:0048102	5
protein amino acid phosphorylation	GO:0006468	4
mitotic spindle elongation	GO:0000022	3
nuclear mRNA splicing, via spliceosome	GO:0000398	3
carbohydrate metabolic process	GO:0005975	3
chitin catabolic process	GO:0006032	3
mitochondrial electron transport, cytochrome c to oxygen	GO:0006123	3
actin filament organization	GO:0007015	3
aerobic respiration	GO:0009060	3
proton transport	GO:0015992	3
sleep	GO:0030431	3
cytokinesis	GO:0000910	2
chitin metabolic process	GO:0006030	2
regulation of transcription, DNA-dependent	GO:0006355	2

continued on next page

Table 16 – continued from previous page

GO Terms	GO IDs	Counts
iron ion transport	GO:0006826	2
lipid transport	GO:0006869	2
cellular iron ion homeostasis	GO:0006879	2
phagocytosis, engulfment	GO:0006911	2
cytoskeleton organization and biogenesis	GO:0007010	2
mitosis	GO:0007067	2
male meiosis	GO:0007140	2
cell adhesion	GO:0007155	2
germ cell development	GO:0007281	2
germarium-derived egg chamber formation	GO:0007293	2
axon guidance	GO:0007411	2
visual perception	GO:0007601	2
visual behavior	GO:0007632	2
determination of adult life span	GO:0008340	2
ATP synthesis coupled proton transport	GO:0015986	2
chondroitin sulfate biosynthetic process	GO:0030206	2
olfactory behavior	GO:0042048	2
ribosome biogenesis and assembly	GO:0042254	2
phototaxis	GO:0042331	2
male courtship behavior, veined wing generated song production	GO:0045433	2
Cellular Component		
mitochondrion	GO:0005739	11
extracellular region	GO:0005576	8
nucleus	GO:0005634	8

continued on next page

Table 16 – continued from previous page

GO Terms	GO IDs	Counts
cytoplasm	GO:0005737	8
lipid particle	GO:0005811	8
mitochondrial respiratory chain complex IV	GO:0005751	5
lysosome	GO:0005764	5
membrane	GO:0016020	5
integral to membrane	GO:0016021	4
larval serum protein complex	GO:0005616	3
mitochondrial inner membrane	GO:0005743	3
mitochondrial matrix	GO:0005759	3
integral to plasma membrane	GO:0005887	3
cytosolic large ribosomal subunit	GO:0022625	3
kinetochore	GO:0000776	2
extracellular space	GO:0005615	2
intracellular	GO:0005622	2
spliceosome	GO:0005681	2
nucleolus	GO:0005730	2
mitochondrial proton-transporting ATP synthase, central stalk	GO:0005756	2
plasma membrane	GO:0005886	2
ferritin complex	GO:0008043	2
Molecular Function		
oxygen transporter activity	GO:0005344	33
ATP binding	GO:0005524	8
protein binding	GO:0005515	7
zinc ion binding	GO:0008270	6
continued on next page		

Table 16 – continued from previous page

GO Terms	GO IDs	Counts
oxidoreductase activity	GO:0016491	6
nucleotide binding	GO:0000166	5
nucleic acid binding	GO:0003676	5
cytochrome-c oxidase activity	GO:0004129	5
cathepsin L activity	GO:0004217	5
calcium ion binding	GO:0005509	5
chitin binding	GO:0008061	5
cation binding	GO:0043169	5
mRNA binding	GO:0003729	4
protein serine/threonine kinase activity	GO:0004674	4
binding	GO:0005488	4
iron ion binding	GO:0005506	4
heme binding	GO:0020037	4
RNA binding	GO:0003723	3
structural constituent of ribosome	GO:0003735	3
actin binding	GO:0003779	3
catalytic activity	GO:0003824	3
chitinase activity	GO:0004568	3
hydrogen-exporting ATPase activity, phosphorylative mechanism	GO:0008553	3
nutrient reservoir activity	GO:0045735	3
DNA binding	GO:0003677	2
transcription factor activity	GO:0003700	2
mRNA 3'-UTR binding	GO:0003730	2
GTPase activity	GO:0003924	2
endonuclease activity	GO:0004519	2
alpha-amylase activity	GO:0004556	2

continued on next page

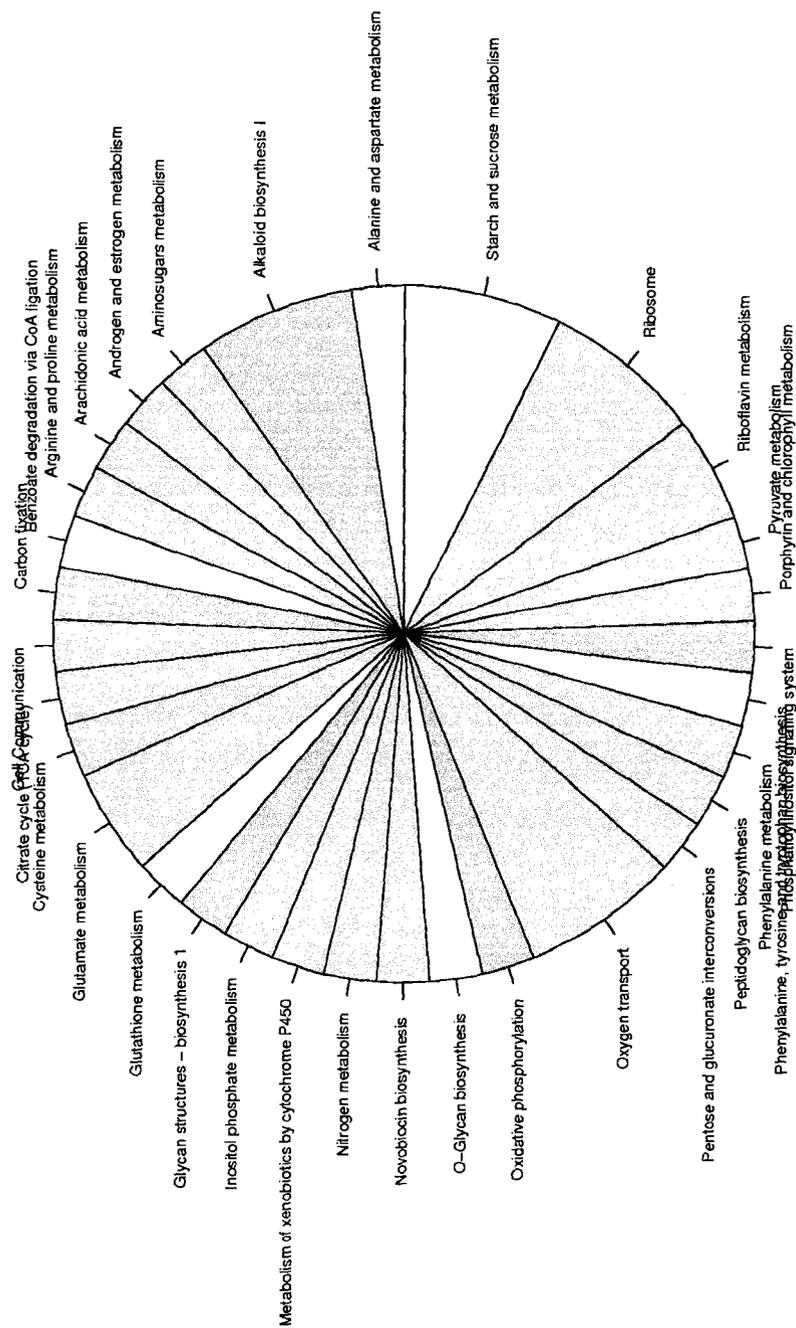
Table 16 – continued from previous page

GO Terms	GO IDs	Counts
protein kinase activity	GO:0004672	2
lipid transporter activity	GO:0005319	2
GTP binding	GO:0005525	2
poly-pyrimidine tract binding	GO:0008187	2
ferrous iron binding	GO:0008198	2
ferric iron binding	GO:0008199	2
lipid binding	GO:0008289	2
glucuronosyltransferase activity	GO:0015020	2
hydrogen ion transporting ATP synthase activity, rotational mechanism	GO:0046933	2
hydrogen ion transporting ATPase activity, rotational mechanism	GO:0046961	2

Discussion

Several potentially hypoxia-responsive genes, including CdMT, mSOD, cSOD, HIF, and tracheless, were cloned and sequenced from previous studies using gene-by-gene method, and amplified and printed on microarrays. None of these genes are statistically significantly differentially expressed according to cDNA microarray during cyclic hypoxia exposures. HIF 1 α expression, which also doesn't change during chronic hypoxia (Chapter 4), can thus not be used as biomarker of chronic and cyclic exposures. However, HIF 1 α expression levels do change considerably during hypoxic exposures, which suggests the changes may be biologically significant. For example, HIF 1 α expression levels increase in cyclic (high) DO, with a 2.83- ($p=0.053$) and 2.15-fold ($p=0.25$) upregulation observed after 5- and 10-day, respectively. Additionally, HIF 1 α is consistently down-regulated at all time points during chronic hypoxia ($p > 0.05$).

KEGG Cyclic High DO



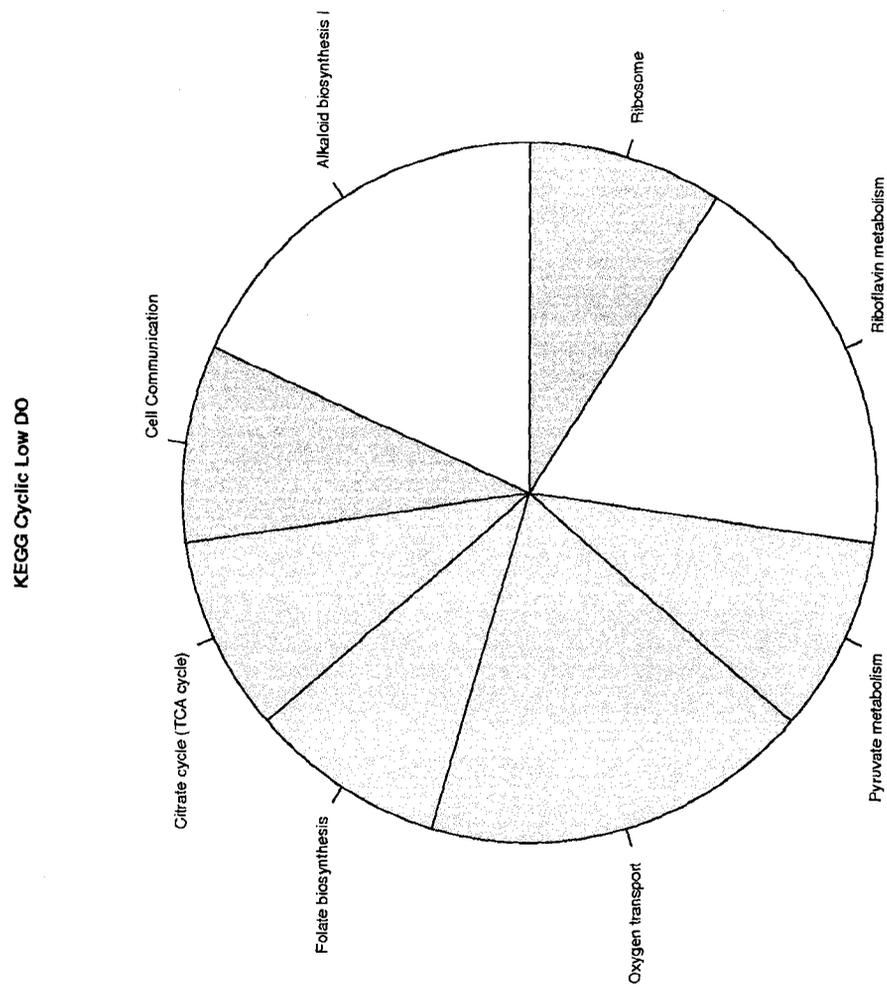


Figure 18: Pie chart of KEGG pathways.

Since hemocyanin is the oxygen carrier which is abundantly expressed in grass shrimp, arrays include 38 hemocyanin clones. For the first 24 hours, all of the down-regulated genes in low DO are hemocyanin genes. However hemocyanin genes are up-regulated after 10 days in high DO.

In general, cycles of hypoxia and reoxygenation can lead to increased production of reactive oxygen species and tissue injury, and down-regulated cellular antioxidant defense systems (Li and Jackson, 2002), which include cSOD, mSOD, glutathione (GSH) peroxidase and GSH reductase. In cyclic (high) DO exposure, mSOD is 1.74-fold ($p=0.3$) up-regulated after 24 and 48 hours, and cSOD is 2.3-fold up-regulated after 24 hour exposure. For chronic hypoxia, mSOD is 2.6- and 1.6-fold up-regulated at 120 and 240 hours ($p=0.23$), respectively (Chapter 4). In previous studies of grass shrimp exposed to hypoxia, cSOD was 19-fold down-regulated after 14-day exposure to chronic hypoxia (Brouwer et al., 2007, 2.5 mg/L) using more sensitive P^{33} cDNA labeling, and mSOD was 60-fold up-regulated after only 3-day exposure to cyclic hypoxia (Brown-Peterson et al., 2008). However, similar significant expression level changes were not observed at any time points during chronic and cyclic exposures in this study, suggesting Cy3/Cy5 labeling may not be sensitive enough to detect these changes.

Glutathione (GSH) peroxidase is a selenium-dependent enzyme that catalyzes breakdown of H_2O_2 and various peroxides and can be found in cytoplasm and mitochondria. In this study, glutathione peroxidase was identified in the up-regulated in the cyclic library (Chapter 3). However, no statistically significant upregulation was revealed using microarrays. Nevertheless, glutathione peroxidase is consistently up-regulated at 24, 48, and 120 hours, with a peak 2.11-fold change at 48 hours ($p=0.19$, CA), which may be biological relevant. In contrast to cyclic hypoxia, glutathione peroxidase is down-regulated for the initial 48 hours, then up-regulated at 5- and 10-day chronic exposures, with a peak 1.9-fold change ($p=0.19$) at day 5 (Chapter 4).

Several genes involved in sulfur redox and (homo)cysteine metabolism (thiore-

doxin, sulfide:quinone oxidoreductase, glutathione-S-transferase, cystathionine beta-synthase) are found in down-regulated libraries in response to cyclic hypoxia (Chapter 3). Glutathione-S-transferase is barely changed during the first 48 hours, then down-regulated at day 5 (1.56-fold, $p=0.15$) and 10 (1.4-fold change, $p=0.06$) in cyclic (high) hypoxia. According to microarray data, sulfide:quinone oxidoreductase is slightly up-regulated in both cyclic hypoxia at all time points, and a similar up-regulation also occurs in response to chronic hypoxia at all time points ($p=0.14$). Thioredoxin is about 1.6-fold up-regulated in cyclic (low) hypoxia ($p=0.28$).

Metallothionein belongs to a family of small, cysteine-rich, and heat stable proteins involved in the cellular regulation of essential metals (Zn, Cu, Se, etc.), and in detoxification of heavy metals (Cd, Hg, Ag, etc.). Studies relate metallothionein proteins with diverse physiological functions including protection against oxidative stress (Suzuki and Cherian, 2000, Chapter 1). English and Storey (2003) identified a 100 amino acid protein belonging to the metallothionein family using differential cDNA libraries constructed from the foot muscle of marine snails *Littorina littorea*. Metallothionein was up-regulated in both foot muscle and hepatopancreas in response to anoxia stress. After 24 h recovery from anoxia stress, transcript levels were reduced but remained elevated in hepatopancreas from anoxia-treated snails. CdMT obtained from previous study (Brouwer et al., 2007) was included in microarray. In this study, expression levels of metallothionein in cyclic (high) hypoxia are 1.4-fold up-regulated. In chronic DO, it is 1.74-fold down-regulated for the first 48 hours. Since metallothionein can be induced by multiple environmental stressors including hypoxia and heavy metals, and since its expression levels change minimally in response to cyclic and chronic hypoxia, metallothionein isn't a candidate biomarker to monitor biological impacts of long or short time oxygen changes in aquatic ecosystem.

Like other invertebrates, crustaceans lack specific immunity, and their innate immune system must rely on non-self-recognition molecules to ensure efficient defense re-

sponses against infectious pathogens, and environmental toxicants and stressors (Rowley and Powell, 2007). So crustacea are regarded as an appropriate species for studying the innate immune system, which plays an important role in host defense system.

Lectins were first discovered in plants, and now are found throughout nature. These sugar-binding proteins typically bind to a carbohydrate molecule which is a part of a glycoprotein or glycolipid, and also agglutinate certain animal cells and/or precipitate glycoconjugates. Lectins serve many different biological functions, including regulation of cell adhesion during infection, glycoprotein synthesis, and the control of protein levels in the blood. Lectins from the hemolymph of invertebrates, including crustaceans, have also been regarded as molecules involved in immune recognition. Several lectin clones were identified from down-regulated copper and pyrene exposures, and up-regulated cyclic library (Chapter 3). In this study, lectin is up-regulated in response to cyclic (low) DO at day 1 (2.85-fold, $p=0.01$) and 2 (2.03-fold, $p=0.19$), to cyclic (high) DO at day 10 (3.23-fold, $p < 0.01$), and to chronic DO at day 10 (2.68-fold, $p < 0.01$).

PmAV, a novel gene involved in virus resistance of shrimp *Penaeus monodon*, is highly expressed in the hepatopancreas and up-regulated on day 2 in response to viral infection (Luo et al., 2003, 2007). PmAV is closely related to C-type lectin. Earlier studies suggest PmAV, obtained from a chronic hypoxia SSH library, is also a hypoxia-responsive gene, which is down-regulated in grass shrimp after 14-day chronic exposure (Brouwer et al., 2007, 1.5 mg/L). Here we find expression of this PmAV is barely changed during cyclic hypoxia, with the only exception of day 10 in cyclic (high) DO (1.8-fold up-regulation, $p < 0.05$). About 4-fold upregulation is observed for several additional PmAV clones which were originally identified in the down-regulated cyclic DO SSH library, and in the up- and down-regulated copper SSH library (Chapter 3). In this study, a 2.1- and 1.8-fold up-regulation of PmAV occurs at 12 hours and day 10 in chronic hypoxia ($p=0.098$), respectively (Chapter 4). It is interesting that both lectin and PmAV genes which are supposed to be immunity-related in crustaceans, are also significantly

up-regulated in response to cyclic (high) hypoxia ($p < 0.01$).

Phosphoenolpyruvate (PEP) is an important metabolite which has the highest energy phosphate bond found in living organisms, and is involved in glycolysis and gluconeogenesis. In gluconeogenesis, phosphoenolpyruvate carboxykinase (PEPCK) converts oxaloacetate into phosphoenolpyruvate and carbon dioxide. This reaction is a rate-limiting step in gluconeogenesis. However, PEPCK levels alone were found to be not highly correlated with gluconeogenesis in the mouse liver (Burgess et al., 2007). Therefore, the role of PEPCK in gluconeogenesis may be more complex and more factors may be involved than was previously believed.

The response of PEPCK to chronic and cyclic hypoxia exposure is rather complex. In this study, PEPCK is 1.5-fold up-regulated ($p=0.3$) after 2 days exposure to cyclic (low) DO, followed by 2.0-fold down-regulation on day 10 ($p=0.26$). PEPCK appears 1.8-fold up-regulated on day 10 ($p=0.02$) during cyclic (high) hypoxia. Similar up-regulation of PEPCK was also observed after 7 day cyclic exposure of grass shrimp (Brown-Peterson et al., 2008). In response to chronic hypoxia, a 2.05-fold up-regulation of PEPCK is observed for the initial 6 hours ($p=0.13$). For the next 6 hours, PEPCK expression is completely reversed to significant downregulation (2.6-fold, $p=0.01$), then followed by up-regulation at day 5 (2.5-fold, $p=0.01$) and day 10 (1.9-fold, $p=0.2$).

Cluster analysis indicated the response patterns of high (CP) and low DO (CA) exposures were in the same cluster at 24 hrs, 48 hrs, and 120 hrs, however the correlation isn't statistically significant ($p < 0.05$). At the above 3 time points, more up- and down-regulated genes are found in low DO than high DO (Figure 16 and Table 14). For low DO, few genes are significantly expressed after 5 days, while for high DO few genes are differentially expressed during the first 5 days.

In this study a poor correlation was observed between results from microarray and qPCR data. Generally microarray results underestimate the actual gene expression changes, and qPCR is more accurate and sensitive than microarray. However, data gen-

erated by microarray and qPCR often result in disagreement (Morey et al., 2006). Which of the two methods is more accurate is debatable (Allison et al., 2006). It suggests that certain microarray platforms don't need any further validation because they perform very accurate (Epstein et al., 2002). Several factors can contribute to the lack of correlation between array and qPCR, such as up- vs. down-regulation, spot intensity, fold change, cycle threshold (C_t), array averaging, and tissue type and preparation (Morey et al., 2006). Sometimes it is reverse transcription PCR, not microarray data, results in the disagreement between microarray and qPCR. In this study, the RNA used for microarray was reverse-transcribed with Oligo(dT)₁₈ primer, while the same RNA was amplified in the presence of random hexamers for qPCR (Chapter 4). Random hexamer gives full length transcripts on average, while Oligo(dT) generates first strand cDNA from the 3' end of the transcripts. Better validation may be achieved if more primers designed from 3' end of the transcripts.

The results reported in this study provide a preliminary basis for a better understanding of gene expression changes of grass shrimp in response to cyclic hypoxia conditions. It illustrates a general picture of the molecular response to hypoxia exposures under laboratory conditions, and of the regulation pathways affected by hypoxia exposures. The observation that gene expression patterns are not only dependent on the duration of exposure but also on the time of day makes interpretation of the data generated by this study a very challenging undertaking.

CHAPTER VI

FIELD EXPOSURES

Abstract

Grass shrimp, *Palaemonetes pugio*, offer many advantages for ecological and toxicological research. A cDNA microarray was utilized to investigate the changes in gene expression in grass shrimp collected from two areas in Weeks Bay (Mobile, Alabama). One is a traditionally normoxic location (WBM), and the other is a traditionally cyclic hypoxic location (WC). There were no significant differences in egg counts between grass shrimp at WBM and WC, although the WBM shrimp had more eggs than WC shrimp. Shrimp from WC had a significant higher number of parasites than those from WBM. Six genes were significantly down-regulated in WC shrimp relative to WBM. All of those genes were also found in the SSH down-regulated cyclic DO library. A putative vitellogenin was the most significantly up-regulated gene in WC shrimp. Hemocyanin was found in both up- and down-regulated genes. It appears that grass shrimp hemocyanin has at least two distinct subunits. The up-regulated hemocyanin subunit showed high sequence similarity to the WSSV-inducible hemocyanin gene of *Marsupenaeus japonicus* and may function in pathogen defense. Hypoxia conditions in field locations were less severe than those in laboratory exposures, which may account for the observation that few significant genes were identified in field studies.

Keywords - *Palaemonetes pugio*; grass shrimp; crustacean; microarray; hypoxia; gene expression; annotation.

Introduction

In addition to chronic hypoxia in stratified deeper water, less severe episodic and often cyclic hypoxic conditions can occur during the summer time in shallow waters in

coastal estuaries. Smith and Able (2003) recorded DO concentrations in marsh salt pools to characterize the responses of fishes to rapid changes of DO. The greatest DO range of 0-20 mg/L occurred in mid-July, and from mid- to late August, DO remained < 1.0 mg/L for a long period. The effect of cyclic hypoxia is limited to a few aquatic species, such as marsh grass shrimp *Palaemonetes vulgaris* (Coiro et al., 2000), flounders (Stierhoff et al., 2006), weakfish *Cynoscion regalis* (Tyler and Targett, 2007), and grass shrimp (Brown-Peterson et al., 2008).

Many crustacean species are frequently exposed to chronic and cyclic hypoxia in their natural habitats. To evaluate the sublethal effects of cyclic hypoxia exposure on aquatic organisms in the laboratory, research should ideally mimic cyclic conditions that occur in the field. However, laboratory cyclic hypoxia with its rigid 24 hours DO cycle can at best only approach the actual field exposure. The significantly expressed genes identified during laboratory exposures must therefore be validated using field exposure to determine if potential biomarkers identified in laboratory exposures can be used in field studies as well.

In this study, DNA microarrays were used to measure differentially expressed genes in grass shrimp collected from normoxic and cyclic hypoxic field sites. Since grass shrimp are well adapted to hypoxia, and most of the shrimp caught for field study came from the surface, there may be fewer genes changed in shrimp from cyclic hypoxic field sites compared to laboratory cyclic hypoxia.

Materials and Methods

If not described below, please see Chapter 4.

Field Exposures

Grass shrimp were collected on September 11, 2006 from two areas (Figure 19) in Weeks Bay (Mobile, Alabama, USA). One is a traditionally normoxic location, Weeks

Bay Mouth (WBM, 30°22'66"N, 87°50'25"W), and the other is a traditionally cyclic hypoxic location, Weeks Creek (WC, 30°22'25"N, 87°49'95"W). Twenty grass shrimp were collected by hand-held dip net from each location and the thorax/hepatopancreas was removed and immediately fixed in the field in 1 mL RNALater for subsequent RNA extraction and microarray analysis. DO, temperature and salinity were measured continuously for 7 days (September 6-13) in both areas using a YSI Model 600XLM data sonde prior to collection of animals.

Shrimp were collected again on September 13, 2006 from the same locations in Weeks Bay. Twenty grass shrimp were collected by hand-held dip net from each location and placed into buckets with aerated water and transported to the laboratory for egg counts and dissection, following standard protocols.

Experimental Design

A simple reference design with dye swap was employed for field studies. A series of replicates were hybridized with the grass shrimp collected from normoxic (WBM) and hypoxic (WC) locations. Each sample was hybridized equally with Cy3 or Cy5 dye.

Results

Water Quality Parameters and Grass Shrimp

The average length of grass shrimp collected on September 11, 2006 in WBM and WC were 28.6 ± 3.30 (mm) and 25.6 ± 2.96 (mm), respectively. Shrimp collected at WBM were longer than those at WC ($p < 0.01$). The measured DO and salinity for 5 days prior to field collection was 6.34 ± 1.84 (mg/L) and 15.85 ± 2.15 (‰) for WBM (Figure 20), and 1.05-8.87 (mg/L) and 16.85 ± 2.15 (‰) for WC (Figure 21), respectively. During the collection (10-11 AM), the temperature, pH, DO, and salinity were 27.77 (°C), 7.46 (pH), 7.76 (mg/L DO), and 13.36 (‰) for WBM, 26.32 (°C), 6.95 (pH), 3.32 (mg/L DO), and 12.89 (‰) for WC, respectively. The time periods during which measured DO

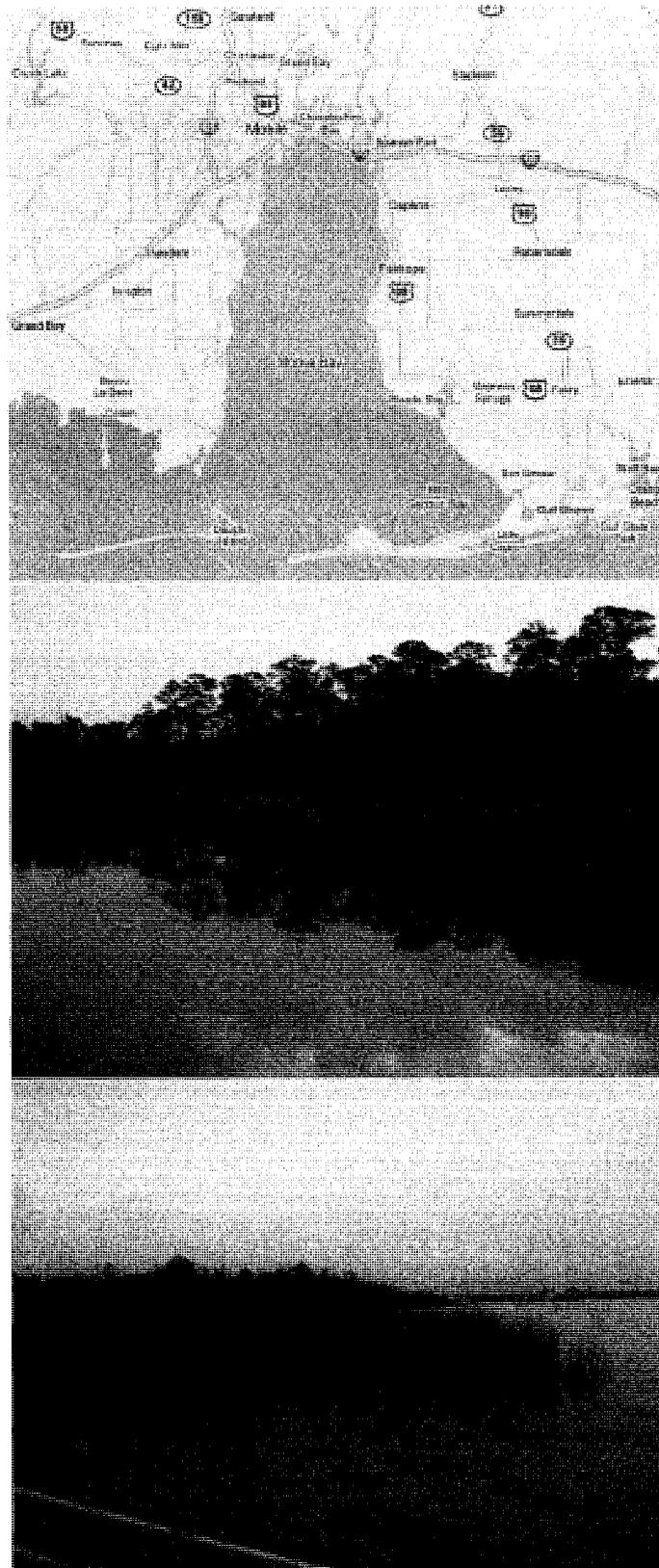


Figure 19: Map of collection sites. Top: Weeks Bay located near Mobile Bay's eastern shore. Center: Weeks Creek, cyclic hypoxic site. Bottom: Weeks Bay Mouth, normoxic location.

was in the ranges of < 2 (mg/L DO) and between 2-3 (mg/L DO) over the 5 days recording time were 0 and 255 min for WBM, and 285 and 1005 min for WC, respectively.

The average length and weight of grass shrimp collected on September 13, 2006 were 28.5 ± 3.63 (mm) and 202.51 (mg) for WBM, and 25.9 ± 4.95 (mm) and 157.45 (mg) for WC, respectively. Shrimp collected at WBM were slightly longer and heavier than those at WC ($p=0.04$). During the collection (8-9 AM), the temperature, pH, DO, and salinity recorded as 27.70 ($^{\circ}\text{C}$), 7.85 (pH), 5.21 (mg/L DO), and 16.33 (‰) for WBM, 27.98 ($^{\circ}\text{C}$), 7.17 (pH), 2.38 (mg/L DO), and 20.38 (‰) for WC, respectively. Total number of eggs and parasites were 2691 (134.55 ± 10.95) and 42 (2.1 ± 0.8) for WBM, and 2381 (119.05 ± 14.00) and 138 (6.9 ± 1.9) for WC, respectively.

RNA Extraction

RNA concentrations were determined using a NanoDrop Spectrophotometer (ND-1000, see Chapter 4). The RNA concentration was 2.04 ± 0.77 $\mu\text{g}/\mu\text{l}$. The ratios of 260/280 and 260/230 were 2.06 ± 0.04 and 2.22 ± 0.04 , respectively. No RNA degradation was shown by running diluted samples on RNA Nano Chips (Agilent Bioanalyzer 2100).

Microarray Hybridization and Analysis

Dye swaps were performed for replicate arrays using the RNA samples from WBM and WC. A model was established using the spatial correlation estimated between within-array duplicates (Smyth et al., 2005, Chapter 4). The fold changes and significant tests reported in Table 17 were corrected by any probe-specific dye effects. The significant genes were selected at $p = 0.05$. It is rather striking that most significantly down-regulated genes in WC grass shrimp were also in the SSH down-regulated cyclic DO library. A putative vitellogenin was the most significantly up-regulated gene in WC shrimp.

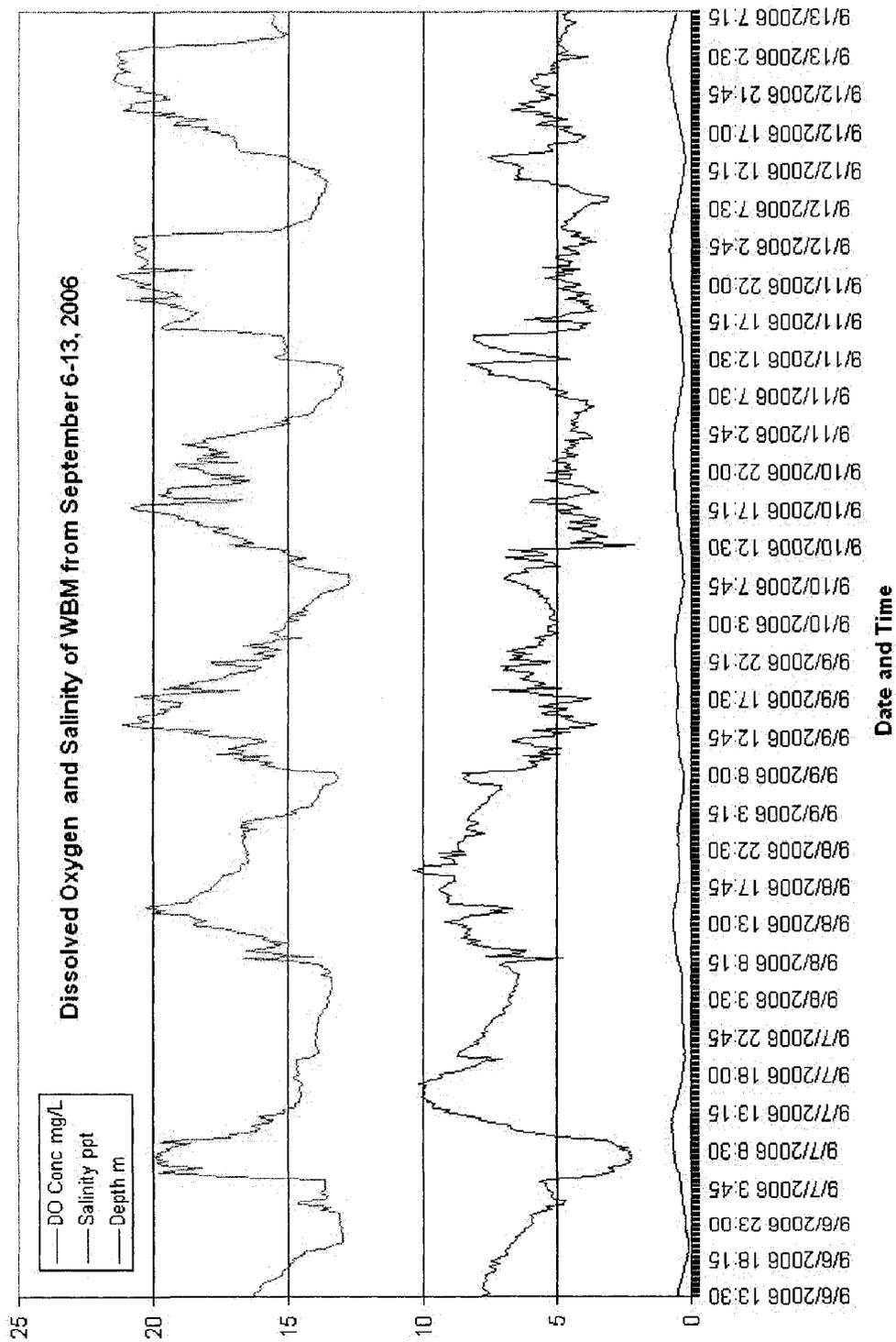


Figure 20: Dissolved Oxygen and Salinity at WBM from September 6 to 13, 2006.

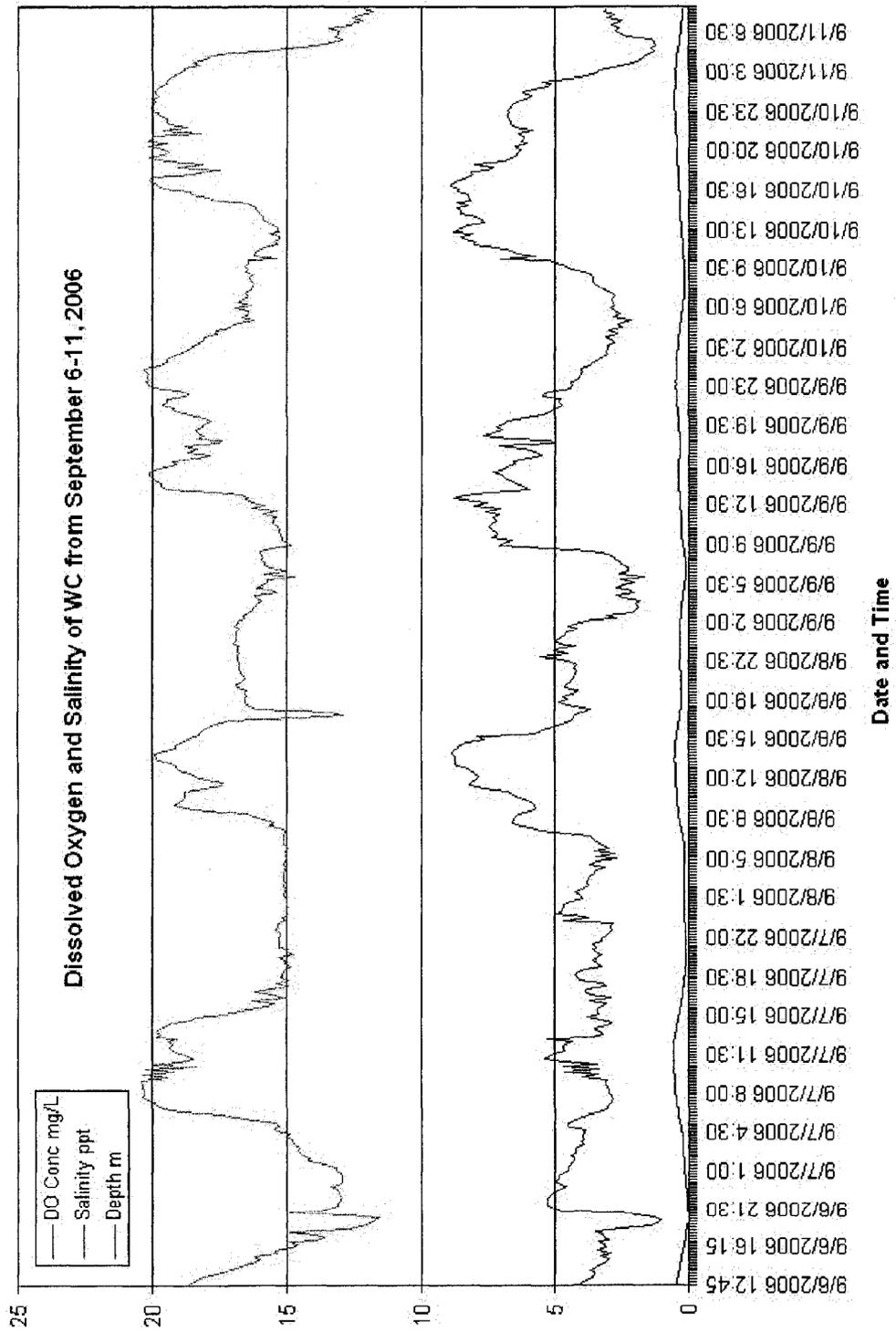


Figure 21: Dissolved Oxygen and Salinity at WC from September 6 to 11, 2006.

Table 17: List of differentially expressed genes from field samples.

Libraries	ID	logFC	P Value	Definition	E Value
Down-regulated					
Cyclic DO	45B-A11.g	-2.97	0.009	crustapain [Pandalus borealis]	1.00E-50
Cyclic DO	45B-B03.g	-2.96	0.009	beta-carotene dioxygenase 2 [Apis mellifera]	1.00E-37
Cyclic DO	45B-A07.g	-2.29	0.044	polo-like kinase 1 [Drosophila melanogaster] (PLK1)	1.00E-86
Cyclic DO	45B-A09.g	-1.88	0.023	protein pelota [Tribolium castaneum]	2.00E-32
Cyclic DO	45B-B06.g	-1.74	0.032	chitinase 1 [Penaeus monodon]	3.00E-45
Cyclic DO	45B-A01.g	-1.58	0.044	trypsinogen [Rana pirica]	0
Post molt	37A-F04.g	-1.16	0.068	hemocyanin subunit 1 [Gammarus roesei]	5.00E-33
Up-regulated					
Cyclic DO	35B-A12.g	1.65	0.025	vitellogenin [Ornithorhynchus anatinus]	3.00E-05
Copper	39D-D09.g	1.13	0.011	calcium and integrin binding protein CIB [Bombyx mori]	4.00E-06
Copper	39C-A03.g	0.93	0.019	putative ferrichrome iron receptor precursor [Plectonema boryanum]	0.04
Cyclic DO	35B-C02.g	0.88	0.025	trypsinogen Y precursor [Gadus morhua]	2.00E-07
Pyrene	41B-D08.g	0.82	0.055	beta-glucocerebrosidase [Drosophila melanogaster]	2.00E-38
Post molt	37D-D07.g	0.80	0.055	vitellogenin [Cherax quadricarinatus]	8.00E-12
Pyrene	41C-B09.g	0.76	0.068	hemocyanin [Litopenaeus vannamei]	1.00E-72

Discussion

The role of oxygen in development, physiology, and pathology is an area of long-standing biological interest. Most of our understanding about the role of oxygen in mechanisms and gene expression patterns comes from studies on mammals or well-characterized model organisms, such as human (Gao et al., 2002), *Caenorhabditis elegans* (Jiang et al., 2001), *Drosophila melanogaster* (Bacon et al., 1998; Lavista-Llanos et al., 2002), and *Danio rerio* (Ton et al., 2003). In contrast, the effect of changing oxygen tension in aquatic environment has been little studied, especially for crustaceans. Oxygen enters the water body by photosynthesis of aquatic biota or by the physical transfer across the air-water interface. The maximum oxygen concentration in estuarine water (~35 ppt salinity) is about 7 mg/L at 20°C and 1 atmosphere pressure. The critical oxygen concentration at which grass shrimp can't uptake sufficient amount of oxygen from water to sustain aerobic metabolism, so they must turn to anaerobic metabolism, is 30-35 torr, or 1.8 mg/L (Cochran and Burnett, 1996b).

Oxygen concentrations in estuarine waters often change in a diel pattern and can drop periodically below 1.8 mg/L. In previous laboratory studies we have measured the effects of DO cycles on gene expression. However, these artificial cycles can only approach but not reproduce DO cycles as they occur in the field. For example, in this field study, DO, temperature, and salinity were recorded continuously using a multi-parameter YSI sonde, every 15 min for 5 days. For WBM, the minimal DO is 2.08 mg/L. During this period, 2.6% of measured DO is in the range of 2-3 (mg/L DO). For WC, the minimal DO is 1.05 mg/L. 2.2%, 2.0%, and 14.7% of DO concentrations measured in WC are in the ranges of < 1.5, 1.5-2, and 2-3 (mg/L DO), respectively. For cyclic laboratory exposures in Chapter 5, 21.2%, 11.1%, and 14.9% of DO are in the ranges of < 1.5, 1.5-2, and 2-3 (mg/L DO), respectively. Clearly, grass shrimp exposed to cyclic DO in the laboratory experienced more severe hypoxia than those collected from cyclic DO field site.

In addition, the grass shrimp from normoxic and hypoxic locations were usually

collected from the water surface, whereas DO was measured at the bottom using a YSI sonde. Finally, grass shrimp can climb out of water during periods of oxygen deficiency (Welsh, 1975; Anderson, 1985). In view of all this, we would expect fewer genes to be changed in grass shrimp from our field cyclic hypoxia site relative to laboratory cyclic hypoxia.

For grass shrimp collected on September 13, 2006, there were no significant differences in egg counts between grass shrimp at WBM and WC, although the WBM shrimp had more eggs than WC shrimp (134.55 ± 10.95 vs. 119.05 ± 14.00 eggs). A similar result was also observed in laboratory cyclic exposure. There is no significant difference in the number of eggs of shrimp in hypoxic and normoxic conditions. However, it takes shrimp in cyclic hypoxia longer to produce the second brood (Brown-Peterson et al., 2008). Furthermore, shrimp from WC had a significant higher number of parasites than those from WBM (6.9 ± 1.0 vs. 2.1 ± 0.8 parasites per shrimp, $p = 0.015$). This suggests shrimp located in the cyclic DO area may be more stressed and susceptible to parasite infestation than shrimp from the normoxic site.

One family of genes that shows significant changes in crustaceans in response to hypoxia is the hemocyanin gene family (Brown and Terwilliger, 1999; Terwilliger et al., 1999, 2006; Brown-Peterson et al., 2008). Hemocyanins are multisubunit oxygen-transporting proteins in arthropods and mollusks. The hemocyanin subunits of arthropods can self-assemble into hexamers or multi-hexamer complexes. In addition to its role as an oxygen carrier, hemocyanin has been found to function in the innate immune response. Phenoloxidase, a related copper protein, catalyzes the hydroxylation of monophenols such as tyrosine to diphenols and further oxidizes them to highly reactive quinones (Johansson and Soderhall, 1996; Soderhall and Cerenius, 1998). These quinones are compounds with antimicrobial and antifungal properties that are critical in the arthropod immune response. Terwilliger et al. (2006) cloned and sequenced the complete cDNA of six hemocyanin subunits, two cryptocyanins, and one prophenoloxidase of *Cancer magister*. Alignment

of all nine amino acid sequences shows high similarity and high degree of conservation in this family. Hemocyanin transports oxygen used by phenoloxidase, and even functions as a phenoloxidase under certain conditions (Zlateva et al., 1996; Decker and Rimke, 1998; Decker and Tuzcek, 2000). Moreover, hemocyanin subunits have recently been found to contribute to the immune function. Crayfish hemocyanin without domain I becomes an active phenoloxidase (Lee et al., 2004), and some antimicrobial peptides are derived from the C-terminal hemocyanin domain (Decker and Jaenicke, 2004).

Finally, a specific hemocyanin subunit is up-regulated in shrimp infected with WSSV virus (Lei et al., 2008). In this study, one hemocyanin gene (37A-F04, most similar to *Gammarus* hemocyanin subunit 1) is down-regulated in shrimp exposed to cyclic hypoxia, while a second distinct hemocyanin gene (41C-B09, most similar to *Litopenaeus* hemocyanin) is up-regulated. It appears that grass shrimp hemocyanin has at least two distinct subunits. Considering grass shrimp from the cyclic DO location are heavily infected with parasites, we speculate the up-regulated hemocyanin subunit, which shows high sequence identity (75%, $E=4E-77$) with WSSV-inducible hemocyanin gene of *Marsupenaeus japonicus*, may function in pathogen defense.

In conclusion, only a few genes are differentially expressed in grass shrimp exposed to cyclic hypoxia in the field study relative to those collected from a normoxia reference site. Compared with laboratory chronic and cyclic exposures, gene expression of grass shrimp was not as much affected by cyclic hypoxia probably because field DO conditions at our study site were less severe than those used in our laboratory studies.

CHAPTER VII

SUMMARY

Occurrence and severity of hypoxia is increasing in coastal and estuarine environments, and recovery of impacted habitats and living resources is slow. Detection of early effects of hypoxia is needed for timely remedial action to be taken. The overall objectives of this research was to develop molecular indicators of dissolved oxygen stress to assess the biological impact of hypoxia in coastal estuaries and validate their use through a combination of laboratory and field studies. In order to assess if the molecular indicators can be used widely along coastal ecosystems, the hypoxia responsive biomarkers must be identified from an oxygen-sensitive and hypoxia-tolerant species that has a wide range of geographical distribution. Grass shrimp, *Palaemonetes pugio*, is one of the most intensively studied crustaceans in eco-toxicology, and dissolved oxygen regulates its distribution and abundance along the shores of the Atlantic and Gulf of Mexico. It has been recognized as an excellent model for the search of molecular biomarkers of oxygen stress in estuarine systems. Therefore, this study determined which genes are significantly up- or down-regulated in grass shrimp exposed to chronic and cyclic hypoxia in the laboratory, and evaluated if these hypoxia-responsive genes can be used as indicators of DO stress in the aquatic environment. To achieve these goals, this study addressed the following specific research objectives.

Objective 1: Clone and sequence grass shrimp HIF.

All organisms possess mechanisms to maintain oxygen homeostasis, which are essential for survival. In aquatic ecosystem, hypoxia is a state of oxygen deficiency in which the concentration of dissolved oxygen in the water column decreases from normoxic 7 mg/L to below 2 mg/L (NSTC, 2003). The hypoxia inducible factor (HIF), conserved during evolution from nematodes to flies to vertebrates, is a key transcription fac-

tor that controls a variety of cellular and systemic homeostatic responses to hypoxic stress (Semenza, 1998). The existence of HIF-1 α has not been characterized in crustaceans, which experience wide fluctuations of oxygen tensions in their aquatic environment.

HIF-1 α homolog in grass shrimp was cloned and sequenced using RT-PCR and RACE thus demonstrating the existence of HIF-1 α gene in crustaceans. The 3822 bp full-length HIF-1 α cDNA encodes a protein sequence of 1057 amino acids that shows a 46% similarity to *Tribolium castaneum* HIF-1 α protein. Grass shrimp HIF-1 α protein also shows a high level of conservation with other HIF-1 α proteins in the bHLH domain, two PAS domains, an ODD domain with two proline hydroxylation motifs (LTHLAP and MRAPFIP), and a C-TAD with an asparagine hydroxylation motif (EVNAP). However, grass shrimp HIF-1 α protein lacks N-TAD, and has a unique 230 amino acid sequence that isn't found in any vertebrate HIF proteins. Similar conserved domains and motifs suggest that grass shrimp HIF-1 α protein is regulated by similar molecular mechanism as other HIF-1 α proteins. Successful cloning of grass shrimp HIF-1 α gene is only the first step to fully understanding the response mechanisms in crustaceans exposed to hypoxia. More studies are needed to determine whether HIF expression at the mRNA (see Objective 3) and/or protein levels can be used as a potential molecular indicator of hypoxic stress in laboratory and field studies.

Objective 2: Perform sequence and bioinformatic analysis of ESTs from SSH libraries.

Genomics and related approaches have increasingly enhanced our understanding of the mechanisms that underlie toxic effects of chemicals on living tissues of various organisms and may help to identify gene expression profiles that may serve as biomarkers of exposure (Calzolari et al., 2007). Six libraries of expressed sequence tags (ESTs) were constructed by suppression subtractive hybridization (SSH) from the grass shrimp exposed to environmental (moderate, severe, and cyclic hypoxia, copper, and pyrene) and biological (molt) stress. An in-house pipeline of cleaning, clustering, and assembling was

built for sequence analysis, and the ESTs were annotated by similarity searches against different public databases. GO and pathway analysis of the resulting potential transcripts revealed that stressor specific genes were present in different libraries. Several genes involved in sulfur redox and (homo)cysteine metabolism were all down-regulated in response to cyclic hypoxia. Up-regulation of cytochrome c oxidase subunit I and down-regulation of vitellogenin was a common response to chronic (1.5 mg/L and 2.5 mg/L) and cyclic DO exposures. The molting process was accompanied by changes in expression of many genes not found in the hypoxia/copper/pyrene libraries. The cDNA clones and sequence information can be used for future functional analysis and microarray design to monitor the environmental stressors using grass shrimp in coastal ecosystems.

Objective 3: Determine if HIF can be used as an indicator of chronic and cyclic hypoxia exposure by analyzing HIF expression levels using microarrays.

Potential putative genes selected from SSH libraries and HIF-1 α were used to construct a cDNA microarray to measure gene expression changes and genetic pathways involved in response to hypoxia. Four HIF 1 α clones didn't show differentially expression during chronic hypoxia exposures. Similarly, none of the HIF clones showed significant changes in expression during cyclic hypoxia exposures and field studies. Thus HIF can't be used as biomarker of chronic and cyclic exposures in both laboratory and field studies.

Objective 4: Determine if expression of hypoxia responsive genes can be used as an indicator of chronic and cyclic hypoxia exposure in the laboratory and field using microarrays.

The microarrays were used to examine differentially expressed genes in hypoxic vs. normoxic groups at 6 (H6), 12 (H12), 24 (H24), 48 (H48), 120 (H120), and 240 (H240) hours exposure to chronic hypoxia. Cluster analysis showed two response patterns, composed of an up- (including H6, H24, and H120) and down-regulated (including H12, H48,

and H240) dominated cluster. Venn diagrams of differentially expressed genes showed there is no gene up- or down-regulated common to all six groups. Changes of significant genes are too dynamic to serve as biomarkers of hypoxia stress in grass shrimp. However, some genes appear unique for specific time points. Most selected differentially expressed genes on the microarrays also showed similar up- or down-regulated patterns in qPCR.

Differentially expressed genes were determined in hypoxic vs. normoxic groups after 1, 2, 5 and 10 days exposure to cyclic hypoxia. Sampling on each day was conducted at two different time series, one in the morning (representing low DO, CA) and one in the afternoon (representing high DO, CP). There are distinct differences between the number and identity of specific genes that are significantly down- or up-regulated in shrimp collected at the low DO and high DO points of the cyclic DO cycle. However, cluster analysis showed that the overall response patterns of high (CP) and low DO (CA) exposures were in the same cluster at 24 hrs, 48 hrs, and 120 hrs. In contrast, the response patterns at different time points were in different clusters. There is no gene shared by any of the eight exposure groups. None of significant differentially expressed genes can serve as biomarkers of cyclic hypoxia stress in grass shrimp. Cyclic (high) DO regulated the expression of genes associated with a broader range of functional categories and pathways compared to cyclic (low) DO.

During cyclic hypoxia, microarray results had poor agreement with qPCR data. One possible explanation is that different modes of reverse transcription PCR results in the disagreement between microarray and qPCR. RNA used for microarray was reverse-transcribed with Oligo(dT)₁₈ primer, while the same RNA was amplified in the presence of random hexamers for qPCR. Random hexamer gives full length transcripts on average, and Oligo(dT) generates first strand cDNA from 3' end of the transcripts. Better validation may be achieved if primers designed from 3' end of the transcripts are used for qPCR.

In addition to genes obtained from the SSH libraries, several potentially hypoxia-

responsive genes, including CdMT, mSOD, cSOD, and trachealess, were cloned using a gene-by-gene method, and showed significant differences in expression at certain time points under chronic and cyclic hypoxia exposures. However, none of the genes are significantly differentially expressed at all time points according to the cDNA microarray.

Only a few genes are differentially expressed in grass shrimp exposed to cyclic hypoxia in the field study relative to those collected from a normoxic reference site. Compared with laboratory chronic and cyclic exposures, gene expression of grass shrimp was not as much affected by cyclic hypoxia probably because field DO conditions at our study site were less severe than those used in our laboratory studies. Hemocyanin was found in both up- and down-regulated genes. It appears that grass shrimp hemocyanin has at least two distinct subunits. Since hemocyanin has been shown to play a role in immune defense we suggest that up-regulation of hemocyanin may have been due to the presence of high levels of parasite infestation at the cyclic DO field site.

Studies presented here demonstrated that while a HIF-1 α homolog was successfully cloned from the grass shrimp, grass shrimp HIF is constitutively expressed and not induced by chronic and cyclic hypoxia exposures in both laboratory and field studies. While several potentially hypoxia-responsive genes were cloned and sequenced, none of genes are consistently and significantly differentially expressed according to cDNA microarray during chronic and cyclic hypoxia exposures. Some differentially expressed genes were identified at certain time points during laboratory and field exposures, however, the lack of consistency limits their use as sensitive biomarkers of hypoxia stress in aquatic ecosystem.

This study presents the first grass shrimp cDNA microarray constructed from a limited number of cDNA clones to detect gene expression changes in response to hypoxia. The utility of shrimp microarray was confirmed in this study. GO-based and pathway-based mapping of hypoxia-responsive genes to biological pathways and processes represents a key step in microarray data mining to illustrate why and how genes response to

hypoxia. Microarray provides a general picture of the molecular response to hypoxia exposures under laboratory and field conditions, and of the regulation pathways affected by hypoxia exposures. Data analysis revealed that the response to chronic hypoxia was pronounced and transient at the experimental time points, and grass shrimp don't respond gradually to chronic exposure. Gene expression changes of grass shrimp in response to cyclic hypoxia conditions were even more dynamic at different time points. The expression levels are not only dependent on the duration of exposure but also on the time of day. The dynamic nature of the response to hypoxia precludes a biologically meaningful interpretation of the limited data generated by this study.

REFERENCES

- Allison, D. B., Cui, X., Page, G. P., Sabripour, M., 2006. Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews Genetics* 7 (1), 55–65. 151
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J., 1997. Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Research* 25 (17), 3389–3402. 46, 91, 100, 122, 133
- American Public Health Association, 1975. Standard methods for the examination of water and waste water. 14th Edition. American Public Health Association. 14, 38
- Anderson, G., 1985. Species profiles: Life histories and environmental requirements of coastal fishes and invertebrates (gulf of mexico), grass shrimp. U.S. Fish and Wildlife Service Biological Report 82 (1), 11–35. 13, 32, 38, 161
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. *Nature Genetics* 25 (1), 25–29. 46, 83, 113
- Bacon, N. C. M., Wappner, P., O'Rourke, J. F., Bartlett, S. M., Shilo, B., Pugh, C. W., Ratcliffe, P. J., 1998. Regulation of the drosophila bhlh-pas protein sima by hypoxia: Functional evidence for homology with mammalian hif-1a. *Biochemical and Biophysical Research Communications* 249 (3), 811–816. 160
- Baryshev, V. A., Glagolev, A. N., Skulachev, V. P., 1981. Sensing of uH^+ in phototaxis of halobacterium halobium. *Nature* 292 (5821), 338–340. 111
- Bell, G. W., Eggleston, D. B., 2005. Species-specific avoidance responses by blue crabs and fish to chronic and episodic hypoxia. *Marine Biology* 146 (4), 761–770. 118

- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* 57 (1), 289–300. 83
- Berra, E., Benizri, E., Ginouves, A., Volmat, V., Roux, D., Pouyssegur, J., 2003. Hif prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of hif-1 α in normoxia. *The EMBO Journal* 22 (16), 4082–4090. 8, 9, 11, 19, 20
- Berra, E., Roux, D., Richard, D. E., Pouyssegur, J., 2001. Hypoxia-inducible factor-1 α (hif-1 α) escapes o₂-driven proteasomal degradation irrespective of its sub-cellular localization: nucleus or cytoplasm. *EMBO Reports* 2 (7), 615–620. 8, 19
- Bilton, R., Mazure, N., Trottier, E., Hattab, M., Dery, M.-A., Richard, D. E., Pouyssegur, J., Brahimi-Horn, M. C., 2005. Arrest-defective-1 protein, an acetyltransferase, does not alter stability of hypoxia-inducible factor (hif)-1 α and is not induced by hypoxia or hif. *Journal of Biological Chemistry* 280 (35), 31132–31140. 10
- Bracken, C. P., Whitelaw, M. L., Peet, D. J., 2003. The hypoxia-inducible factors: key transcriptional regulators of hypoxic responses. *Cellular and Molecular Life Sciences* 60 (7), 1376–1393. 7, 8, 19
- Brouwer, M., Brown-Peterson, N., Larkin, P., Manning, S., Denslow, N., Rose, K., 2005. Molecular and organismal indicators of chronic and intermittent hypoxia in marine crustacea. *Estuarine Indicators*, 261–276. 25, 26, 41, 75, 118, 120
- Brouwer, M., Brown-Peterson, N. J., Larkin, P., Patel, V., Denslow, N., Manning, S., Brouwer, T. H., 2007. Molecular and whole animal responses of grass shrimp, *palaeomonetes pugio*, exposed to chronic hypoxia. *Journal of Experimental Marine Biology and Ecology* 341 (1), 16–31. 15, 25, 26, 39, 40, 41, 74, 75, 76, 112, 113, 118, 120, 147, 148, 149

- Brouwer, M., Larkin, P., Brown-Peterson, N., King, C., Manning, S., Denslow, N., 2004. Effects of hypoxia on gene and protein expression in the blue crab, *Callinectes sapidus*. *Marine Environmental Research* 58 (2-5), 787–792. 25, 41, 75, 120
- Brown, A. C., Terwilliger, N. B., 1999. Developmental changes in oxygen uptake in cancer magister in response to changes in salinity and temperature. *Journal of Experimental Marine Biology and Ecology* 241 (2), 179–192. 161
- Brown-Peterson, N. J., Larkin, P., Denslow, N., King, C., Manning, S., Brouwer, M., 2005. Molecular indicators of hypoxia in the blue crab *Callinectes sapidus*. *Marine Ecology Progress Series* 286, 203–215. 25, 41, 75, 118, 120
- Brown-Peterson, N. J., Manning, C. S., Patel, V., Denslow, N. D., Brouwer, M., 2008. Effects of cyclic hypoxia on gene expression and reproduction in a grass shrimp, *Palaeomonetes pugio*. *Biological Bulletin* 214 (1), 6–16. 119, 147, 150, 153, 161
- Bruick, R. K., 2003. Oxygen sensing in the hypoxic response pathway: regulation of the hypoxia-inducible transcription factor. *Genes and Development* 17 (21), 2614–2623. 10, 12, 20, 21, 33, 34
- Bruick, R. K., McKnight, S. L., 2002. Oxygen sensing gets a second wind. *Science* 295 (5556), 807–808. 10, 11, 12, 33, 34
- Burgess, S. C., He, T., Yan, Z., Lindner, J., Sherry, A. D., Malloy, C. R., Browning, J. D., Magnuson, M. A., 2007. Cytosolic phosphoenolpyruvate carboxykinase does not solely control the rate of hepatic gluconeogenesis in the intact mouse liver. *Cell Metabolism* 5 (4), 313–320. 150
- Burton, D. T., Fisher, D. J., 1990. Acute toxicity of cadmium, copper, zinc, ammonia, 3,3'-dichlorobenzidine, 2,6-dichloro-4-nitroaniline, methylene chloride, and 2,4,6-trichlorophenol to juvenile grass shrimp and killifish. *Bulletin of Environmental Contamination and Toxicology* 44 (5), 776–783. 14, 38

- Calzolari, L., Ansorge, W., Calabrese, E., Denslow, N., Part, P., Lettieri, T., 2007. Transcriptomics and proteomics. applications to ecotoxicology. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 2 (3), 245–249. 39, 164
- Carter, M. G., Sharov, A. A., VanBuren, V., Dudekula, D. B., Carmack, C. E., Nelson, C., Ko, M. S., 2005. Transcript copy number estimation using a mouse whole-genome oligonucleotide microarray. *Genome Biology* 6 (7), R61. 73
- Coblentz, F. E., Towle, D. W., Shafer, T. H., 2006. Expressed sequence tags from normalized cDNA libraries prepared from gill and hypodermal tissues of the blue crab, *Callinectes sapidus*. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 1 (2), 200–208. 39, 50, 73
- Cochran, R. E., Burnett, L. E., 1996a. Respiratory responses of the salt marsh animals, *Fundulus heteroclitus*, *Leiostomus xanthurus*, and *Palaemonetes pugio* to environmental hypoxia and hypercapnia and to the organophosphate pesticide, azinphosmethyl. *Journal of Experimental Marine Biology and Ecology* 195 (1), 125–144. 14
- Cochran, R. E., Burnett, L. E., 1996b. Respiratory responses of the salt marsh animals, *Fundulus heteroclitus*, *Leiostomus xanthurus*, and *Palaemonetes pugio* to environmental hypoxia and hypercapnia and to the organophosphate pesticide, azinphosmethyl. *Journal of Experimental Marine Biology and Ecology* 195 (1), 125–144. 160
- Coiro, L. L., Poucher, S. L., Miller, D. C., 2000. Hypoxic effects on growth of *Palaemonetes vulgaris* larvae and other species: Using constant exposure data to estimate cyclic exposure response. *Journal of Experimental Marine Biology and Ecology* 247 (2), 243–255. 119, 153
- Craig, J. K., Crowder, L. B., 2005. Hypoxia-induced habitat shifts and energetic consequences in Atlantic croaker and brown shrimp on the Gulf of Mexico shelf. *Marine Ecology Progress Series* 294, 79–94. 4

- Craig, J. K., Crowder, L. B., Henwood, T. A., 2005. Spatial distribution of brown shrimp (*farfantepenaeus aztecus*) on the northwestern gulf of mexico shelf: effects of abundance and hypoxia. *Canadian Journal of Fisheries and Aquatic Sciences* 62 (6), 1295–1308. 4, 118
- Currie, R. A., Bombail, V., Oliver, J. D., Moore, D. J., Lim, F. L., Gwilliam, V., Kimber, I., Chipman, K., Moggs, J. G., Orphanides, G., 2005. Gene ontology mapping as an unbiased method for identifying molecular pathways and processes affected by toxicant exposure: Application to acute effects caused by the rodent non-genotoxic carcinogen diethylhexylphthalate. *Toxicological Sciences* 86 (2), 453–469. 72, 114
- Czyzyk-Krzeska, M. F., 1997. Molecular aspects of oxygen sensing in physiological adaptation to hypoxia. *Respiration Physiology* 110 (2-3), 99–111. 6
- Dahlquist, K. D., Salomonis, N., Vranizan, K., Lawlor, S. C., Conklin, B. R., 2002. Genmapp, a new tool for viewing and analyzing microarray data on biological pathways. *Nature Genetics* 31 (1), 19–20. 114
- David, E., Tanguy, A., Pichavant, K., Moraga, D., 2005. Response of the pacific oyster *crassostrea gigas* to hypoxia exposure under experimental conditions. *FEBS Journal* 272 (21), 5635–5652. 118
- Dawson, T. L., Gores, G. J., Nieminen, A. L., Herman, B., Lemasters, J. J., 1993. Mitochondria as a source of reactive oxygen species during reductive stress in rat hepatocytes. *American Journal of Physiology - Cell Physiology* 264 (4), C961–C967. 6, 7
- Decker, H., Jaenicke, E., 2004. Recent findings on phenoloxidase activity and antimicrobial activity of hemocyanins. *Developmental and Comparative Immunology* 28 (7-8), 673–687. 162
- Decker, H., Rimke, T., 1998. Tarantula hemocyanin shows phenoloxidase activity. *Journal of Biological Chemistry* 273 (40), 25889–25892. 162

- Decker, H., Schweikardt, T., Nillius, D., Salzbrunn, U., Jaenicke, E., Tuzcek, F., 2007. Similar enzyme activation and catalysis in hemocyanins and tyrosinases. *Gene* 398 (1-2), 183–191. 115
- Decker, H., Tuzcek, F., 2000. Tyrosinase/catecholoxidase activity of hemocyanins: structural basis and molecular mechanism. *Trends in Biochemical Sciences* 25 (8), 392–397. 162
- Decuypere, S., Vandesompele, J., Yardley, V., Doncker, S. D., Laurent, T., Rijal, S., Llanos-Cuentas, A., Chappuis, F., Arevalo, J., Dujardin, J.-C., 2005. Differential polyadenylation of ribosomal rna during post-transcriptional processing in leishmania. *Parasitology* 131 (3), 321–329. 113
- deFur, P. L., Pease, A. L., 1988. Metabolic and respiratory compensation during long term hypoxia in blue crab, *Callinectes sapidus*. *Understanding the Estuary: Advances in Chesapeake Bay Research* 128, 608–616. 5, 116, 118
- Diaz, R., Solow, A., 1999. Ecological and economic consequences of hypoxia: Topic 2 report for the integrated assessment on hypoxia in the gulf of mexico. NOAA Coastal Ocean Program Decision Analysis Series 16, 45 pp. 4
- Doniger, S. W., Salomonis, N., Dahlquist, K. D., Vranizan, K., Lawlor, S. C., Conklin, B. R., 2003. Mappfinder: using gene ontology and genmapp to create a global gene-expression profile from microarray data. *Genome Biology* 4 (1), R7. 114
- English, T. E., Storey, K. B., 2003. Freezing and anoxia stresses induce expression of metallothionein in the foot muscle and hepatopancreas of the marine gastropod *Littorina littorea*. *The Journal of Experimental Biology* 206 (14), 2517–2524. 148
- EPA, 2003. Hypoxia and wetland restoration. 2, 3

- Epstein, J. R., Biran, I., Walt, D. R., 2002. Fluorescence-based nucleic acid detection and microarrays. *Analytica Chimica Acta* 469 (1), 3–36. 151
- Erez, N., Stambolsky, P., Shats, I., Milyavsky, M., Kachko, T., Rotter, V., 2004. Hypoxia-dependent regulation of *phd1*: cloning and characterization of the human *phd1/egln2* gene promoter. *FEBS Letters* 567 (2-3), 311–315. 10, 20
- Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39 (4), 783–791. 26
- Felsenstein, J., 1989. Phylip - phylogeny inference package. *Cladistics* 5, 164–166. 26
- Finley, D., Scott, G., Daugomah, J., Layman, S., Reed, L., Sanders, M., Sivertsen, S., Strozier, E., 1998. Ecotoxicological assessment of urban and agricultural nonpoint source runoff effects on the grass shrimp, *palaemonetes pugio*. *Ecotoxicology and Risk Assessment for Wetlands*, pp. 243–274. 14, 22, 38
- Gao, N., Jiang, B.-H., Leonard, S. S., Corum, L., Zhang, Z., Roberts, J. R., Antonini, J., Zheng, J. Z., Flynn, D. C., Castranova, V., Shi, X., 2002. p38 signaling-mediated hypoxia-inducible factor 1alpha and vascular endothelial growth factor induction by *cr(vi)* in du145 human prostate carcinoma cells. *Journal of Biological Chemistry* 277 (47), 45041–45048. 160
- GEO, 2003. Global environment outlook year book. United Nations Environment Programme. 2
- GEO, 2004. Global environment outlook year book. United Nations Environment Programme. 2
- Glagolev, A. N., 1980. Reception of the energy level in bacterial taxis. *Journal of Theoretical Biology* 82 (2), 171–185. 111

- Gracey, A. Y., Fraser, E. J., Li, W., Fang, Y., Taylor, R. R., Rogers, J., Brass, A., Cossins, A. R., 2004. Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate. *PNAS* 101 (48), 10970–10975. 73
- Gracey, A. Y., Troll, J. V., Somero, G. N., 2001. Hypoxia-induced gene expression profiling in the euryoxic fish *gillichthys mirabilis*. *PNAS* 98 (4), 1993–1998. 118
- Griffitt, R. J., Chandler, G. T., Greig, T. W., Quattro, J. M., 2006. Cathepsin b and glutathione peroxidase show differing transcriptional responses in the grass shrimp, *palaemonetes pugio* following exposure to three xenobiotics. *Environmental Science and Technology* 40 (11), 3640–3645. 39, 74
- Griffitt, R. J., Greig, T. W., Chandler, G. T., Quattro, J. M., 2007. Serial analysis of gene expression reveals identifiable patterns in transcriptome profiles of *palaemonetes pugio* exposed to three common environmental stressors. *Environmental Toxicology and Chemistry* 26 (11), 2413–2419. 13, 38, 74
- Groth, D., Lehrach, H., Hennig, S., 2004. Goblet: a platform for gene ontology annotation of anonymous sequence data. *Nucleic Acids Research* 32 (Web Server Issue), W313–W317. 46, 53
- Gupta, B. K. S., Turner, R. E., Rabalais, N. N., 1996. Seasonal oxygen depletion in continental-shelf waters of louisiana; historical record of benthic foraminifers. *Geology* 24 (3), 227–230. 6, 119
- Hagerman, L., 1986. Haemocyanin concentration in the shrimp *crangon crangon* (l.) after exposure to moderate hypoxia. *Comparative Biochemistry and Physiology Part A: Physiology* 85 (4), 721–724. 5, 118
- Hansen, B. H., Altin, D., Nordtug, T., Olsen, A. J., 2007. Suppression subtractive hybridization library prepared from the copepod *calanus finmarchicus* exposed to a sub-

- lethal mixture of environmental stressors. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics* 2 (3), 250–256. 39, 67, 73
- Hansson, L. O., Friedler, A., Freund, S., Rudiger, S., Fersht, A. R., 2002. Two sequence motifs from hif-1 α bind to the dna-binding site of p53. *PNAS* 99 (16), 10305–10309. 9, 11, 20
- Harper, S. L., Reiber, C., 1999. Influence of hypoxia on cardiac functions in the grass shrimp (*palaemonetes pugio holthuis*). *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology* 124 (4), 569–573. 14, 38
- Heidbreder, M., Frohlich, F., Jöhren, O., Dendorfer, A., Aadri, F., Dominiak, P., 2003. Hypoxia rapidly activates hif-3 α mrna expression. *FASEB Journal* 17 (11), 1541–1543. 35
- Heinloth, A. N., Irwin, R. D., Boorman, G. A., Nettesheim, P., Fannin, R. D., Sieber, S. O., Snell, M. L., Tucker, C. J., Li, L., Travlos, G. S., Vansant, G., Blackshear, P. E., Tennant, R. W., Cunningham, M. L., Paules, R. S., 2004. Gene expression profiling of rat livers reveals indicators of potential adverse effects. *Toxicological Sciences* 80 (1), 193–202. 114
- Henikoff, S., Henikoff, J. G., Alford, W. J., Pietrokovski, S., 1995. Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene* 163 (2), GC17–GC26. 22
- Hennig, S., Groth, D., Lehrach, H., 2003. Automated gene ontology annotation for anonymous sequence data. *Nucleic Acids Research* 31 (13), 3712–3715. 46, 53
- Hu, C.-J., Sataur, A., Wang, L., Chen, H., Simon, M. C., 2007. The n-terminal transactivation domain confers target gene specificity of hypoxia-inducible factors hif-1 and hif-2 α . *Molecular Biology of the Cell* 18 (11), 4528–4542. 12

- Huang, L. E., Gu, J., Schau, M., Bunn, H. F., 1998. Regulation of hypoxia-inducible factor 1 is mediated by an o₂-dependent degradation domain via the ubiquitin-proteasome pathway. *PNAS* 95 (14), 7987–7992. 9, 10, 20, 33
- Huang, X., Madan, A., 1999. Cap3: A dna sequence assembly program. *Genome Research* 9 (9), 868–877. 45
- Isaac, D. D., Andrew, D. J., 1996. Tubulogenesis in drosophila: a requirement for the tracheless gene product. *Genes and Development* 10 (1), 103–117. 32, 111
- Isaacs, J. S., Jung, Y.-J., Mimnaugh, E. G., Martinez, A., Cuttitta, F., Neckers, L. M., 2002. Hsp90 regulates a von hippel lindau-independent hypoxia-inducible factor-1alpha-degradative pathway. *Journal of Biological Chemistry* 277 (33), 29936–29944. 9, 20
- Jaenicke, E., Decker, H., 2003. Tyrosinases from crustaceans form hexamers. *Biochemical Journal* 371 (2), 515–523. 115
- Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., Gibson, T. J., 1998. Multiple sequence alignment with clustal x. *Trends in Biochemical Sciences* 23 (10), 403–405. 26
- Jeong, J.-W., Bae, M.-K., Ahn, M.-Y., Kim, S.-H., Sohn, T.-K., Bae, M.-H., Yoo, M.-A., Song, E. J., Lee, K.-J., Kim, K.-W., 2002. Regulation and destabilization of hif-1alpha by ard1-mediated acetylation. *Cell* 111 (5), 709–720. 10
- Jiang, B. H., Semenza, G. L., Bauer, C., Marti, H. H., 1996. Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of o₂ tension. *American Journal of Physiology - Cell Physiology* 271 (4), C1172–C1180. 33
- Jiang, B.-H., Zheng, J. Z., Leung, S. W., Roe, R., Semenza, G. L., 1997. Transactiva-

- tion and inhibitory domains of hypoxia-inducible factor 1alpha. *Journal of Biological Chemistry* 272 (31), 19253–19260. 11, 34
- Jiang, H., Guo, R., Powell-Coffman, J. A., 2001. The *caenorhabditis elegans* hif-1 gene encodes a bhlh-pas protein that is required for adaptation to hypoxia. *PNAS* 98 (14), 7916–7921. 160
- Johansson, M., Soderhall, K., 1996. The prophenol oxidase activating system and associated proteins in invertebrates. *Prog. Mol. Subcell. Biol.* 15 (1), 46–66. 115, 161
- Johnson, B. A., Bonaventura, C., Bonaventura, J., 1984. Allosteric modulation of *callinectes sapidus* hemocyanin by binding of l-lactate. *Biochemistry* 23 (5), 872–878. 5, 118
- Ju, Z., Dunham, R. A., Liu, Z., 2002. Differentially expressed genes in the brain of channel catfish (*ictalurus punctatus*) in response to cold acclimation. *Molecular Genetics and Genomics* 268 (1), 87–95. 73
- Ju, Z., Wells, M. C., Heater, S. J., Walter, R. B., 2007a. Multiple tissue gene expression analyses in japanese medaka (*oryzias latipes*) exposed to hypoxia. *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology* 145 (1), 134–144. 74
- Ju, Z., Wells, M. C., Walter, R. B., 2007b. Dna microarray technology in toxicogenomics of aquatic models: Methods and applications. *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology* 145 (1), 5–14. 73
- Justic, D., Rabalais, N. N., Turnerb, R. E., Dortch, Q., 1995. Changes in nutrient structure of river-dominated coastal waters: stoichiometric nutrient balance and its consequences. *Estuarine, Coastal and Shelf Science* 40 (3), 339–356. 4
- Kallio, P. J., Okamoto, K., O'Brien, S., Carrero, P., Makino, Y., Tanaka, H., Poellinger, L., 1998. Signal transduction in hypoxic cells: inducible nuclear translocation and re-

- cruitment of the cbp/p300 coactivator by the hypoxia-inducible factor-1alpha. *EMBO Journal* 17 (22), 6573–6586. 12, 34
- Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., Hirakawa, M., 2006. From genomics to chemical genomics: new developments in kegg. *Nucleic Acids Research* 34 (Database Issue), D354–D357. 47, 69, 84
- Kehrer, J. P., 1993. Free radicals as mediators of tissue injury and disease. *Critical Reviews in Toxicology* 23 (1), 21–48. 7
- Key, P. B., Fulton, M. H., 1993. Lethal and sublethal effects of chlorpyrifos exposure on adult and larval stages of the grass shrimp, *palaemonetes pugio*. *Journal of Environmental Science and Health - Part B Pesticides, Food Contaminants, and Agricultural Wastes* 28 (5), 621–640. 14, 38
- Key, P. B., Fulton, M. H., Scott, G. I., Layman, S. L., Wirth, E. F., 1998. Lethal and sublethal effects of malathion on three life stages of the grass shrimp, *palaemonetes pugio*. *Aquatic Toxicology* 40 (4), 311–322. 14
- Kimura, M., 1983. *The neutral theory of molecular evolution*. Cambridge University Press. 26
- Kimura, T., Jindo, T., Narita, T., Naruse, K., Kobayashi, D., Shin-I, T., Kitagawa, T., Sakaguchi, T., Mitani, H., Shima, A., Kohara, Y., Takeda, H., 2004. Large-scale isolation of ests from medaka embryos and its application to medaka developmental genetics. *Mechanisms of Development* 121 (7-8), 915–932. 74
- Krasnov, A., Koskinen, H., Pehkonen, P., Rexroad, C. E., Afanasyev, S., Molsa, H., 2005. Gene expression in the brain and kidney of rainbow trout in response to handling stress. *BMC Genomics* 6 (1), 3. 73

- Kronick, M. N., 2004. Creation of the whole human genome microarray. *Expert Review of Proteomics* 1 (1), 19–28. 73
- Kung, A. L., Wang, S., Klco, J. M., Kaelin, W. G., Livingston, D. M., 2000. Suppression of tumor growth through disruption of hypoxia-inducible transcription. *Nature Medicine* 6 (12), 1335–1340. 12, 34
- Lando, D., Peet, D. J., Gorman, J. J., Whelan, D. A., Whitelaw, M. L., Bruick, R. K., 2002a. Fih-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. *Genes and Development* 16 (12), 1466–1471. 10, 11, 12, 21, 35
- Lando, D., Peet, D. J., Whelan, D. A., Gorman, J. J., Whitelaw, M. L., 2002b. Asparagine hydroxylation of the hif transactivation domain: A hypoxic switch. *Science* 295 (5556), 858–861. 10, 11, 12, 21, 35
- Larkin, P., Folmar, L. C., Hemmer, M. J., Poston, A. J., Denslow, N. D., 2003. Expression profiling of estrogenic compounds using a sheepshead minnow cDNA macroarray. *Environmental Health Perspectives* 111 (6), 839–846. 26, 27
- Lavista-Llanos, S., Centanin, L., Irisarri, M., Russo, D. M., Gleadle, J. M., Bocca, S. N., Muzzopappa, M., Ratcliffe, P. J., Wappner, P., 2002. Control of the hypoxic response in *Drosophila melanogaster* by the basic helix-loop-helix PAS protein similar. *Molecular and Cellular Biology* 22 (19), 6842–6853. 10, 33, 160
- Law, S., 2002. Expression studies of two hypoxia-inducible factor genes from the grass carp, *Ctenopharyngodon idellus*. *International Congress on the Biology of Fish*. 21
- Law, S. H., Wu, R. S., Ng, P. K., Yu, R. M., Kong, R. Y., 2006. Cloning and expression analysis of two distinct hif- α isoforms - gchif-1 α and gchif-4 α - from the hypoxia-tolerant grass carp, *Ctenopharyngodon idellus*. *BMC Molecular Biology* 7, 15. 9, 21, 35

- Lee, R., Kim, G. B., Maruya, K. A., Steinert, S. A., Oshima, Y., 2000. Dna strand breaks (comet assay) and embryo development effects in grass shrimp (*palaemonetes pugio*) embryos after exposure to genotoxicants. *Marine Environmental Research* 50 (1-5), 553–557. 38
- Lee, R. F., Maruya, K. A., Francendese, L., 1998. Final report: Dna strand damage, embryotoxicity and toxaphene concentrations in grass shrimp (*palaemonetes pugio*) from stations in terry/dupree creek, brunswick georgia, tyetra tech em inc., norcross. 14, 22, 38
- Lee, S. Y., Lee, B. L., Soderhall, K., 2004. Processing of crayfish hemocyanin subunits into phenoloxidase. *Biochemical and Biophysical Research Communications* 322 (2), 490–496. 162
- Lei, K., Li, F., Zhang, M., Yang, H., Luo, T., Xu, X., 2008. Difference between hemocyanin subunits from shrimp *penaeus japonicus* in anti-wssv defense. *Developmental and Comparative Immunology* 32 (7), 808–813. 162
- Li, C., Jackson, R. M., 2002. Reactive species mechanisms of cellular hypoxia-reoxygenation injury. *American Journal of Physiology - Cell Physiology* 282 (2), C227–C241. 147
- Li, T., Brouwer, M., 2007. Hypoxia inducible factor, *ghif*, of the grass shrimp *palaemonetes pugio*: Molecular characterization and response to hypoxia. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 147 (1), 11–19. 15, 39, 40, 74, 76, 108, 111
- Liu, Q., Zhang, X., Yin, W., Li, C., Huang, Y., Qiu, P., Su, X., Hu, S., Yan, G., 2006. A catalog for transcripts in the venom gland of the *agkistrodon acutus*: Identification of the toxins potentially involved in coagulopathy. *Biochemical and Biophysical Research Communications* 341 (2), 522–531. 46

- Luo, J. C., Shibuya, M., 2001. A variant of nuclear localization signal of bipartite-type is required for the nuclear translocation of hypoxia inducible factors (1 α , 2 α and 3 α). *Oncogene* 20 (12), 1435–1444. 8, 13, 34
- Luo, T., Li, F., Lei, K., Xu, X., 2007. Genomic organization, promoter characterization and expression profiles of an antiviral gene pmav from the shrimp *penaeus monodon*. *Molecular Immunology* 44 (7), 1516–1523. 149
- Luo, T., Zhang, X., Shao, Z., Xu, X., 2003. Pmav, a novel gene involved in virus resistance of shrimp *penaeus monodon*. *FEBS Letters* 551 (1), 53–57. 149
- Makino, Y., Cao, R., Svensson, K., Bertilsson, G., Asman, M., Tanaka, H., Cao, Y., Berkenstam, A., Poellinger, L., 2001. Inhibitory pas domain protein is a negative regulator of hypoxia-inducible gene expression. *Nature* 414 (6863), 550–554. 8, 20
- Mangum, C. P., 1997. Adaptation of the oxygen transport system to hypoxia in the blue crab, *callinectes sapidus*. *American Zoologist* 37 (6), 604–611. 5, 6, 116, 118
- Mangum, C. P., Rainer, J. S., 1988. The relationship between subunit composition and o₂ binding of blue crab hemocyanin. *Biological Bulletin* 174 (1), 77–82. 5, 118
- Manning, C. S., Schesny, A. L., Hawkins, W. E., Barnes, D. H., Barnes, C. S., Walker, W. W., 1999. Exposure methodologies and systems for long-term chemical carcinogenicity studies with small fish species. *Toxicology Mechanisms and Methods* 9 (3), 201–217. 25, 41, 75, 120
- Marchant, G. E., 2002. Toxicogenomics and toxic torts. *Trends in Biotechnology* 20 (8), 329–332. 73
- Martyniuk, C. J., Xiong, H., Crump, K., Chiu, S., Sardana, R., Nadler, A., Gerrie, E. R., Xia, X., Trudeau, V. L., 2006. Gene expression profiling in the neuroendocrine brain

- of male goldfish (*carassius auratus*) exposed to 17 α -ethinylestradiol. *Physiological Genomics* 27 (3), 328–336. 74
- Masson, N., Ratcliffe, P. J., 2003. Hif prolyl and asparaginyl hydroxylases in the biological response to intracellular o₂ levels. *Journal of Cell Science* 116 (15), 3041–3049. 12, 35
- Masson, N., Willam, C., Maxwell, P. H., Pugh, C. W., Ratcliffe, P. J., 2001. Independent function of two destruction domains in hypoxia-inducible factor- α chains activated by prolyl hydroxylation. *EMBO Journal* 20 (18), 5197–5206. 9, 10, 33
- Matsunami, K., Kokubo, H., Ohno, K., Xu, P.-X., Ueno, K., Suzuki, Y., 1999. Embryonic silk gland development in bombyx: molecular cloning and expression of the bombyx tracheless gene. *Development Genes and Evolution* 209 (9), 507–514. 28
- McKenney, C. L., Weber, D. E., Celestial, D. M., MacGregor, M. A., 1998. Altered growth and metabolism of an estuarine shrimp (*palaemonetes pugio*) during and after metamorphosis onto fenvalerate-laden. *Archives of Environmental Contamination and Toxicology* 35 (3), 464–471. 38
- Melillo, G., Musso, T., Sica, A., Taylor, L. S., Cox, G. W., Varesio, L., 1995. A hypoxia-responsive element mediates a novel pathway of activation of the inducible nitric oxide synthase promoter. *Journal of Experimental Medicine* 182 (6), 1683–1693. 32
- Metzen, E., Zhou, J., Jelkmann, W., Fandrey, J., Brune, B., 2003. Nitric oxide impairs normoxic degradation of hif-1 α by inhibition of prolyl hydroxylases. *Molecular Biology of the Cell* 14 (8), 3470–3481. 32
- Miracle, A. L., Toth, G. P., Lattier, D. L., 2003. The path from molecular indicators of exposure to describing dynamic biological systems in an aquatic organism: Microarrays and the fathead minnow. *Ecotoxicology* 12 (6), 457–462. 73

- Mississippi River Gulf of Mexico Watershed Nutrient Task Force, 2001. Action plan for reducing, mitigating, and controlling hypoxia in the northern gulf of Mexico. The Mississippi River Gulf of Mexico Task Force. 2
- Moggs, J. G., Tinwell, H., Spurway, T., Chang, H.-S., Pate, I., Lim, F. L., Moore, D. J., Soames, A., Stuckey, R., Currie, R., Zhu, T., Kimber, I., Ashby, J., Orphanides, G., 2004. Phenotypic anchoring of gene expression changes during estrogen-induced uterine growth. *Environmental Health Perspectives* 112 (16), 1589–1606. 114
- Morey, J. S., Ryan, J. C., Dolah, F. M. V., 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time pcr. *Biological Procedures Online* 8 (1), 175–193. 151
- Nambu, J. R., Chen, W., Hu, S., Crews, S. T., 1996. The drosophila melanogaster similar bhlh-pas gene encodes a protein related to human hypoxia-inducible factor 1alpha and drosophila single-minded. *Gene* 172 (2), 249–254. 10, 33
- NCAT, 1999. The gulf of Mexico dead zone. National Center for Appropriate Technology. 1, 3
- NCAT, 2000. The gulf of Mexico dead zone: Impacts on fisheries. National Center for Appropriate Technology. 2
- NOAA, 2006. Commercial landings data.
URL <http://www.st.nmfs.gov/st1/commercial> 3
- NSTC, 2003. An assessment of coastal hypoxia and eutrophication in U.S. waters. National Science and Technology Council. 2, 6, 163
- Oberdorster, E., Brouwer, M., Hoexum-Brouwer, T., Manning, S., McLachlan, J. A., 2000. Long-term pyrene exposure of grass shrimp, *Palaemonetes pugio*, affects molting

- and reproduction of exposed males and offspring of exposed females. *Environmental Health Perspectives* 108 (7), 641–646. 14, 38, 42
- Oleksiak, M. F., Kolell, K. J., Crawford, D. L., 2001. Utility of natural populations for microarray analyses: Isolation of genes necessary for functional genomic studies. *Marine Biotechnology* 3 (Supplement 1), S203–S211. 73
- Oshlack, A., Emslie, D., Corcoran, L. M., Smyth, G. K., 2007. Normalization of boutique two-color microarrays with a high proportion of differentially expressed probes. *Genome Biology* 8 (1), R2. 83
- Paracel, 2002. PTA: Paracel Transcript Assembler User Manual. Paracel Inc. 45
- Powell, W. H., Hahn, M. E., 2002. Identification and functional characterization of hypoxia-inducible factor 2 from the estuarine teleost, *fundulus heteroclitus*: Interaction of hif-2a with two arnt2 splice variants. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 294 (1), 17–29. 21
- Proudfoot, N., O’Sullivan, J., 2002. Polyadenylation: A tail of two complexes. *Current Biology* 12 (24), R855–R857. 113
- Pugh, C. W., O’Rourke, J. F., Nagao, M., Gleadle, J. M., Ratcliffe, P. J., 1997. Activation of hypoxia-inducible factor-1; definition of regulatory domains within the alpha subunit. *Journal of Biological Chemistry* 272 (17), 11205–11214. 11, 12, 34
- R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 47, 83, 85
- Rabalais, N. N., Turner, R. E., Jr., W. J. W., 2002. Gulf of Mexico hypoxia, a.k.a. the dead zone. *Annual Review of Ecology and Systematics* 33, 235–263. 2, 21
- Rabalais, N. N., Turner, R. E., Justic, D., Dortch, Q., Jr., W. J. W., 1999. Characterization

- of hypoxia: Topic 1 report for the integrated assessment of hypoxia in the gulf of mexico. NOAA Coastal Ocean Program Decision Analysis Series 15, 167 pp. 2, 118
- Rabalais, N. N., Turner, R. E., Justic, D., Dortch, Q., Jr., W. J. W., Gupta, B. K. S., 1996. Nutrient changes in the mississippi river and system responses on the adjacent continental shelf. *Estuaries* 19 (2), 386–407. 3
- Reichert, K., Menzel, R., 2005. Expression profiling of five different xenobiotics using a *caenorhabditis elegans* whole genome microarray. *Chemosphere* 61 (2), 229–237. 73
- Rice, P., Longden, I., Bleasby, A., 2000. Emboss: The european molecular biology open software suite. *Trends in Genetics* 16 (6), 276–277. 47, 85
- Rise, M. L., von Schalburg, K. R., Brown, G. D., Mawer, M. A., Devlin, R. H., Kuipers, N., Busby, M., Beetz-Sargent, M., Alberto, R., Gibbs, A. R., Hunt, P., Shukin, R., Zeznik, J. A., Nelson, C., Jones, S. R., Smailus, D. E., Jones, S. J., Schein, J. E., Marra, M. A., Butterfield, Y. S., Stott, J. M., Ng, S. H., Davidson, W. S., Koop, B. F., 2004. Development and application of a salmonid est database and cdna microarray: Data mining and interspecific hybridization characteristics. *Genome Research* 14 (3), 478–490. 73
- Rose, T. M., Schultz, E. R., Henikoff, J. G., Pietrokovski, S., McCallum, C. M., Henikoff, S., 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research* 26 (7), 1628–1635. 22
- Rowley, A. F., Powell, A., 2007. Invertebrate immune systemspecific, quasi-specific, or nonspecific? *The Journal of Immunology* 179 (11), 7209–7214. 149
- Ryan, H. E., Lo, J., Johnson, R. S., 1998. Hif-1alpha is required for solid tumor formation and embryonic vascularization. *EMBO Journal* 17 (11), 3005–3015. 8, 19

- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4 (4), 406–425. 26
- Salomonis, N., Hanspers, K., Zambon, A. C., Vranizan, K., Lawlor, S. C., Dahlquist, K. D., Doniger, S. W., Stuart, J., Conklin, B. R., Pico, A. R., 2007. Genmapp 2: new features and resources for pathway analysis. *BMC Bioinformatics* 8 (1), 217. 114
- Semenza, G. L., 1998. Hypoxia-inducible factor 1: master regulator of o₂ homeostasis. *Current Opinion in Genetics and Development* 8 (5), 588–594. 7, 10, 19, 33, 164
- Semenza, G. L., Wang, G. L., 1992. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Molecular and Cellular Biology* 12 (12), 5447–5454. 8
- Slomovic, S., Laufer, D., Geiger, D., Schuster, G., 2006. Polyadenylation of ribosomal rna in human cells. *Nucleic Acids Research* 34 (10), 2966–2975. 113
- Smith, K. J., Able, K. W., 2003. Dissolved oxygen dynamics in salt marsh pools and its potential impacts on fish assemblages. *Marine Ecology Progress Series* 258, 223–232. 153
- Smyth, G. K., 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* 3 (1), Article 3. 83
- Smyth, G. K., Michaud, J., Scott, H. S., 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 21 (9), 2067–2075. 83, 156
- Soderhall, K., Cerenius, L., 1998. Role of the prophenoloxidase-activating system in invertebrate immunity. *Current Opinion in Immunology* 10 (1), 23–28. 115, 161

- Soitamo, A. J., Rabergh, C. M. I., Gassmann, M., Sistonen, L., Nikinmaa, M., 2001. Characterization of a hypoxia-inducible factor (hif-1) from rainbow trout. accumulation of protein occurs at normal venous oxygen tension. *Journal of Biological Chemistry* 276 (23), 19699–19705. 21, 34
- Stajich, J. E., Block, D., Boulez, K., Brenner, S. E., Chervitz, S. A., Dagdigian, C., Fullen, G., Gilbert, J. G., Korf, I., Lapp, H., Lehvaslaiho, H., Matsalla, C., Mungall, C. J., Osborne, B. I., Pocock, M. R., Schattner, P., Senger, M., Stein, L. D., Stupka, E., Wilkinson, M. D., Birney, E., 2002. The bioperl toolkit: Perl modules for the life sciences. *Genome Research* 12 (10), 1611–1618. 47
- Stierhoff, K. L., Targett, T. E., Miller, K., 2006. Ecophysiological responses of juvenile summer and winter flounder to hypoxia: experimental and modeling analyses of effects on estuarine nursery quality. *Marine Ecology Progress Series* 325 (1), 255–266. 119, 153
- Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., 2002. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *PNAS* 99 (26), 16899–16903. 21
- Sugumaran, M., 1998. Unified mechanism for sclerotization of insect cuticle. *Advances in Insect Physiology* 27 (1), 229–334. 115
- Sussman, A. S., 1949. The functions of tyrosinase in insects. *The Quarterly Review of Biology* 24 (4), 328–341. 115
- Suzuki, Y., Cherian, M. G., 2000. Effect of ethanol on brain metallothionein in transgenic mice. *Alcoholism: Clinical and Experimental Research* 24 (3), 315–321. 148
- Taylor, B. L., 1983. Role of proton motive force in sensory transduction in bacteria. *Annual Review of Microbiology* 37 (1), 551–573. 111

- Terwilliger, N. B., Dangott, L., Ryan, M., 1999. Cryptocyanin, a crustacean molting protein: Evolutionary link with arthropod hemocyanins and insect hexamerins. *PNAS* 96 (5), 2013–2018. 161
- Terwilliger, N. B., Ryan, M., Phillips, M. R., 2006. Crustacean hemocyanin gene family and microarray studies of expression change during eco-physiological stress. *Integrative and Comparative Biology* 46 (6), 991–999. 161
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., Higgins, D. G., 1997. The clustal x windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25 (24), 4876–4882. 26
- Thompson, J. D., Higgins, D. G., Gibson, T. J., 1994. Clustal w: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22 (22), 4673–4680. 26
- Tian, H., McKnight, S. L., Russell, D. W., 1997. Endothelial pas domain protein 1 (epas1), a transcription factor selectively expressed in endothelial cells. *Genes and Development* 11 (1), 72–82. 8
- Ton, C., Stamatiou, D., Dzau, V. J., Liew, C.-C., 2002. Construction of a zebrafish cDNA microarray: gene expression profiling of the zebrafish during development. *Biochemical and Biophysical Research Communications* 296 (5), 1134–1142. 118
- Ton, C., Stamatiou, D., Liew, C.-C., 2003. Gene expression profile of zebrafish exposed to hypoxia during development. *Physiological Genomics* 13 (2), 97–106. 118, 160
- Turner, R. E., Rabalais, N. N., 1994. Coastal eutrophication near the mississippi river delta. *Nature* 368 (6472), 619–621. 4

- Tyler, R. M., Targett, T. E., 2007. Juvenile weakfish *Cynoscion regalis* distribution in relation to diel-cycling dissolved oxygen in an estuarine tributary. *Marine Ecology Progress Series* 333 (1), 257–267. 119, 153
- van der Meer, D. L. M., van den Thillart, G. E. E. J. M., Witte, F., de Bakker, M. A. G., Besser, J., Richardson, M. K., Spaink, H. P., Leito, J. T. D., Bagowski, C. P., 2005. Gene expression profiling of the long-term adaptive response to hypoxia in the gills of adult zebrafish. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* 289 (5), R1512–R1519. 118
- van der Ven, K., Wit, M. D., Keil, D., Moens, L., Leemput, K. V., Naudts, B., Coen, W. D., 2005. Development and application of a brain-specific cDNA microarray for effect evaluation of neuro-active pharmaceuticals in zebrafish (*Danio rerio*). *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 141 (4), 408–417. 73
- Wannamaker, C. M., Rice, J. A., 2000. Effects of hypoxia on movements and behavior of selected estuarine organisms from the southeastern United States. *Journal of Experimental Marine Biology and Ecology* 249 (2), 145–163. 118
- Watanabe, H., Tatarazako, N., Oda, S., Nishide, H., Uchiyama, I., Morita, M., Iguchi, T., 2005. Analysis of expressed sequence tags of the water flea *Daphnia magna*. *Genome* 48 (4), 606–609. 39, 73
- Water Marks, 2004. Hypoxia: The Gulf of Mexico's summertime foe. *Water Marks* 26, 3–5. 1, 2
- Waters, M. D., Olden, K., Tennant, R. W., 2003. Toxicogenomic approach for assessing toxicant-related disease. *Mutation Research/Reviews in Mutation Research* 544 (2-3), 415–424. 73

- Welsh, B. L., 1975. The role of grass shrimp, *palaemonetes pugio*, in a tidal marsh ecosystem. *Ecology* 56 (3), 513–530. 13, 14, 22, 161
- Williams, T. D., Gensberg, K., Minchin, S. D., Chipman, J. K., 2003. A dna expression array to detect toxic stress response in european flounder (*platichthys flesus*). *Aquatic Toxicology* 65 (2), 141–157. 73
- Wingrove, J. A., O'Farrell, P. H., 1999. Nitric oxide contributes to behavioral, cellular, and developmental responses to low oxygen in *drosophila*. *Cell* 98 (1), 105–114. 32
- Wu, R. S. S., Lam, P. K. S., Wan, K. L., 2002. Tolerance to, and avoidance of, hypoxia by the penaeid shrimp (*metapenaeus ensis*). *Environmental Pollution* 118 (3), 351–355. 118
- Yu, B. P., 1994. Cellular defenses against damage from reactive oxygen species. *Physiological Reviews* 74 (1), 139–162. 7
- Zlateva, T., Muro, P. D., Salvato, B., Beltramini, M., 1996. The o-diphenol oxidase activity of arthropod hemocyanin. *FEBS Letters* 384 (3), 251–254. 162