DIFFERENTIAL GENE EXPRESSION IN *FUNDULUS HETEROCLITUS* DUE TO ARSENIC EXPOSURE

HORACIO GONZALEZ

Department of Biological Sciences

APPROVED

_____
Lisa Bain, Ph.D., Chair

_____
Joanne T. Ellzey, Ph.D.

_____
William S. Baldwin, Ph.D.

_____
Peter Golding, Ph.D.

_____
Pablo Arenaz, Ph.D.
Dean of the Graduate School

DIFFERENTIAL GENE EXPRESSION IN *FUNDULUS HETEROCLITUS* DUE TO ARSENIC EXPOSURE

By

HORACIO GONZALEZ, M.S.

Presented to the Faculty of the Graduate School

of The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Biological Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

May 2007

UMI Number: 3262914

UMI Microform 3262914

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGMENTS

I am very grateful to all the persons that contributed to my success during my doctoral program. I would like to say thanks to Dr. Lisa Bain, my advisor, for her constant support and encouragement and the opportunity to work with her for several years in an especially friendly environment. I also want to say thanks to the members of my committee, Dr. Joanne T. Ellzey, Dr. William S. Baldwin, and Dr. Peter Golding who were always a source of advice and direction during my years in the program.

Special thanks to Dr. Joanne T. Ellzey and Marian Viveros for assistance with the morphological analyses using the Zeiss Axioskop microscope and Pascal software in the Analytical Cytology Core Facility, Dr. Jorge Gardea, Chairman of the Chemistry department, for his cooperation in the chemical analysis of our project, Dr. Elizabeth Walsh, who let us use her incubation chambers for our fish, and Dr. Renato Aguilera and Armando Varela for facilitating the Axiocam MRm camera and Zeiss AxioVision 4.4 software to analyze our specimens.

I am also grateful to the staff and faculty of the ESE program, especially Cindy Conroy and Dr. John Walton, who were always a source of guidance. I wish to thanks all the members of the Biology Department and UTEP for providing me support through the Excellence Scholarship and the Dodson Dissertation Fellowship. I also extend my gratitude to all my co-workers and friends: Janice Peterson, Jonathan Roling, Laura Chapman, Catalina Ordonez, Jeffrey Sivils, Caroline Krestchmetcher, Michelle Garcia and many others who always promoted an interactive and cooperative environment.

ABSTRACT

Arsenic is present as a contaminant in drinking water supplies throughout the world, which may cause unhealthy exposures to people. Several epidemiological studies have linked arsenic ingestion in drinking water and adverse health effects such as cancer, respiratory illnesses and diabetes. As a result of this, the United States Environmental Protection Agency lowered the drinking water standards from 50ppb to 10ppb in 2001. There was a strong opposition from public utilities to lower the arsenic limits due to high implementation costs and uncertainty about its benefits. As a result, the Environmental Protection Agency and other research centers requested research to be conducted in order to understand the effects of arsenic at low doses under different perspectives, including the use of animal models. In this context, this dissertation project investigated the effects of arsenic in gene expression using the fish mummichog (*Fundulus heteroclitus*). Three different environmentally relevant doses were employed (230ppb, 575ppb and 1720ppb) and the effects of the exposure in offspring of the exposed fish were analyzed. There was an increase in the incidence of fish with morphological abnormalities such as curved tails in the 230ppb and 1720ppb doses. Several genes were differentially expressed as a result of parental exposure, including genes involved in muscle contraction such as myosins and tropomyosins, structural cytokeratin genes, signaling genes such as parvalbumin, genes involved in immune responses and oxidative stress and some expressed sequence tags. The differential expression of some of these genes may be related to the observed developmental abnormalities and enhances our knowledge about the mode of action of arsenic at the molecular level.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

INTRODUCTION

Purpose

Risk assessments require continuous refinement in order to better understand environmental disturbances. A possible tool to accomplish this is to incorporate gene expression profiles as an essential step in the assessments. Measurements in gene expression changes following pollutant exposure have increased the sensitivity of ecological assessments and have been used as effective biomarkers of toxicological insult.

During the last decade, risk assessments due to arsenic exposure have received special attention from governments, research centers and the scientific community. This is due to the presence of arsenic in the environment, as a natural element and as an industrial contaminant. Arsenic occurs in underground water as a result of leaching from minerals during storage for extended periods of time. The use of this arsenic-contaminated water for drinking purposes in several parts of the world has been linked to adverse human health effects such as cancer, increases in abortions, hypertension, and heart disease. Extensive epidemiological risk assessments have been conducted in order to relate arsenic exposure levels and length of exposure to negative health outcomes. But, whether current standards are actually protective of human health is unknown.

The use of gene expression changes as a result of arsenic exposure could provide a very sensitive endpoint to understand the molecular effects of arsenic that may be the basis of disease. At the same time, this novel research approach will provide important information to understand the doses that are ecologically relevant as well as the long term effects of exposure. In this context, this dissertation research investigated the effects of arsenic exposure in gene expression using a fish species termed mummichog (*Fundulus heteroclitus*) as a model. We used microarray

technology and several environmentally relevant arsenic concentrations to detect profiles of differential gene expression and to correlate these changes to abnormal reproductive and developmental outcomes.

Arsenic as an environmental contaminant

Arsenic is a metalloid present in the environment as a natural element and as a pollutant (Mandal and Suzuki, 2002; Smedley and Kinniburgh, 2002). Its name has been historically related to poison and has been considered one of the main chemical substances of environmental concern (ATSDR, 1999; NRC, 1999). One of the reasons for this is its presence in natural water bodies that are used for agricultural and drinking purposes throughout the world (NRC, 2001). As a chemical element, arsenic belongs to the group Va of the periodic chart. In aqueous environments it exists as oxidation states (+V) arsenate, (+III) arsenite, arsenous acid, arsenic acid and their salts. Industrial arsenic contamination of water and soil is originated mainly from smelting of arsenic bearing materials (Cu, Ni, Pb and Zn ores), with copper being the most important and accounting for approximately 80% of these emissions. Additionally, the burning of fossil fuels, especially coal, the disposal of industrial waste, and its use in chemical products such as pesticides, paintings, glass manufacturing, medical compounds, and warfare materials all contribute to the release of arsenic in the environment (Bissen and Frimmel, 2003). For these reasons, human populations that are considered with the highest risk of arsenic exposure are those living in the proximities of smelters or who consume arsenic-contaminated drinking water (Hindmarsh, 2000).

The consumption of drinking water from arsenic contaminated sources was increased during the last decades of the twentieth century, especially in developing countries, due to

population growth pressures and the lack of adequate technological systems to reduce the levels of arsenic in the water. Figure 0.1 shows a worldwide distribution of the sites where underground arsenic-contaminated water occurs. Notice the extension of arsenic contaminated aquifers distributed widely across the earth, as well as the arsenic polluted sites due to mining activities.

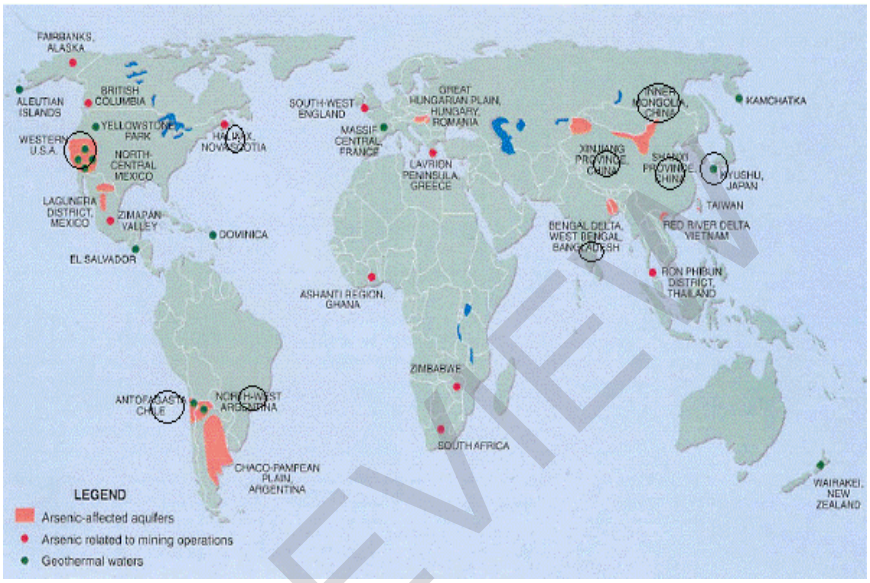

Figure 0.1. Arsenic in groundwater around the world.
(http://phys4.harvard.edu/~wilson/arsenic/arsenic_project_introduction.html)

The current arsenic limit for drinking water in the United States is 10 parts per billion (10ppb) and is subject to revision. Consumption of contaminated reservoirs has resulted in exposure levels that surpass this safety standard for arsenic in several parts of the world, as summarized in Table 0.1. Since millions of individuals are exposed to arsenic, there is a pressing need for research to determine what concentrations of arsenic in drinking water are safe. As seen in the table, one of the most extensive cases of arsenic exposure through drinking water occurs in

Bangladesh. The situation has been reported as the largest mass poisoning of a population in history (Atkins *et al.*, 2007).

Table 0.1. **Population exposed to arsenic due to underground water use (Nordstrom, 2002)**

Country	Potential population under exposure	Estimated arsenic concentration (µg/L)
Bangladesh	30,000,000	< 1 to 2,500
India	6,000,000	10 t0 3,200
Vietnam	> 1,000,000	1 to 3,050
Thailand	15 000	1 to > 5,000
Taiwan	100,000 to 200,000	10 to 1,820
Mongolia	1,000 to 600,000	<1 to 2,400
Argentina	2,000,000	<1 to 9,900
Chile	400,000	100 to 1,000
Mexico	400,000	8 to 620
Hungary-Romania	400,000	<2 to 176
Spain	>50,000	<1 to 100
Ghana	<100,000	<1 to 175

Human health effects due to arsenic-polluted water

Risk assessment research indicates that long term exposure to arsenic in drinking water can result in adverse health effects such as cancer, hypertension, diabetes mellitus, skin lesions, cardiovascular disease and neuro-cognitive function damage (NRC, 2001). Still, there is great uncertainty in the scientific community about the exact modes of action of arsenic in the body and the concentrations that must be achieved in drinking water in order to minimize damage from exposure.

Several epidemiological studies have been conducted in order to find any correlation of arsenic in drinking water and increased risk of developing cancer, as well as the dose-response model for these effects. Through these studies, chronic arsenic exposure through drinking water

has been linked to liver, bladder, lung and skin cancers (NRC, 2001). For example, one of the earliest epidemiological studies that linked arsenic exposure through drinking water and cancer was performed in Taiwan in 1989. A group of villages in the southwestern part of the country were placed in three groups according to the average level of arsenic in their drinking water (<300µg/L, 300-590µg/L, and ≥600µg/L) and their relationship to elevated cancer risks was evaluated. This study showed increased mortality rates for lung cancers in males by 49.2, 100.7 and 104.1-fold, respectively. For women, these rates were 25.6, 57.0 and 111.3. Similarly, increases for kidney cancer in women were 3.4, 19.4 and 58.0-fold and in men were 8.4, 18.9 and 58.0-fold (Wu *et al.*, 1989). Other important studies that linked arsenic exposure in drinking water with increased cancer cases included a case-control study in northern Chile (Ferreccio *et al*., 2000) and a standardized mortality ratio test study in Argentina (Hopenhayn-Rich *et al*., 1998). These and other epidemiological studies have shown a strong dose-response relationship for arsenic-polluted water and lung, skin, bladder and kidney cancers (NRC, 1999; IARC, 2004).

One of the major limitations of these epidemiological risk assessments is the great level of variation factors among the populations being considered. In addition to differences in arsenic levels and in exposure times, individual factors such as diet, weight, personal habits such as smoking condition or exercise habits, exposure to other chemicals and many others change can be considerable depending on the geographic area and the socioeconomic status of the population. For these reasons, the epidemiological risk assessments can be enhanced by incorporating new technological tools that allow more sensitive endpoint detection of exposure and the relationship of these endpoints to disease. Together, these types of results will better support the establishment of safety levels and the environmental legislation that enforces discharge levels.

EPA arsenic drinking water limits

The United States Environmental Protection Agency lowered the arsenic limit for drinking water from 50 to 10 ppb. This regulation became effective in 2006. There were strong debates among scientists, technicians and legislators concerning the actual need of the new limit. The former 50ppb limit was established in 1942. There were multiple recommendations to lower the standard over the next decades, but it was not until 2001, in one of the last acts by the Clinton administration, that the new limit of 10ppb was enacted. The new administration delayed the establishment of the standard based on the uncertainty of the scientific evidence and the cost of the technology required to reach the new limit (Smith *et al*., 2002).

In order to comply with the new standard, the water treatment utilities around the country needed to develop state-of-the-art technology in their purification systems. The Environmental Protection Agency evaluated the following technologies as the best in order to reach the new standards: ion exchange, activated alumina, reverse osmosis, modified coagulation/filtration, modified lime softening, electrodialysis reversal and oxidation/filtration (EPA, 2001). These technologies are expensive and hard to implement. Most public utilities argued that the former arsenic limits in drinking water were safe enough in order to prevent any damage to people. In spite of this, using a report from the National Research Council as the main scientific evidence, the Environmental Protection Agency finally lowered the drinking water standard to 10ppb in October of 2001 (Smith *et al*., 2002). This limit has also been recommended by the European Community and by the World Health Organization (Kouras *et al*., 2007).

The new drinking-water levels became effective in 2006 and will be subject to revision every six years in order to assess the need for further reduction in arsenic levels (Hughes *et al*., 2007). One of the major points of uncertainty regarding the limits is the lack of knowledge about

the potential for arsenic poisoning at very low exposure levels. In addition, the epidemiological risk assessments have multiple external factors that make extrapolation and generalizations difficult. One possible scientific alternative to support the legislation process is to investigate the effects of arsenic at the molecular level, since this type of research allows dose manipulation at very low exposure levels and is extremely sensitive.

Molecular effects due to arsenic

Experiments in animals and in cell cultures have shown that arsenic has many biochemical and cytotoxic effects. Among these negative affects are increased oxidative damage to DNA (Liu *et al*., 2001), altered DNA methylation and gene expression (Chen *et al*., 2001), changes in intracellular levels of proteins (Hamadeh *et al*., 1999), enzymatic inhibition (Petrick *et al*., 2000), increases in protein-DNA cross links (Ramirez *et al*., 2000), induction of apoptosis (Rousselot *et al*., 1999), and altered regulation of DNA-repair genes and other stress-response pathways (NRC, 1999). Whether these biochemical effects occur in humans due to arsenic in drinking water remains to be determined.

Kaltreider, (Kaltreider *et al*., 1999) showed that low doses of sodium arsenite altered nuclear-binding levels of several transcription factors in rat hepatoma cells. In other experiment, Parrish found an increase AP-1 DNA-binding activity and increasing c-*fos*, c-*jun*, and c-*myc* gene expression in rabbit renal cortical slices (Parrish *et al*., 1999). Liu, (Liu *et al*., 2001) reported that mice exposed to arsenic exhibited significant changes in expression of genes related to stress, DNA damage and xenobiotic biotransformations. Other experiments established that sodium arsenate produced significant dose-responsive hypermethylation within a fragment of the promoter p53 (Mass and Liangjun, 1997). Kuo, (Kuo *et al*., 1997) found p53 overexpression in

human populations presumably exposed to arsenic. Studies with human cells in culture determined that arsenic alters DNA methylation at 0.002-0.2 μM (Zhong and Mass, 2001). In addition, Li (Li *et al*., 2002) reported increased levels of the oxidative stress-inducible protein peroxiredoxin I in mice that were treated with sodium arsenate.

Arsenic has been reported to act as an estrogenic element in several experiments. For example, Liu, (Liu *et al*., 2007) investigated mice fetal gene expression after in-utero exposure (85ppm) and reported the differential expression of several estrogen-controlled genes such as over-expression of the X-inactive-specific transcript, anterior gradient-2, trefoil factor-1, CRP-ductin, ghrelin, small praline-rich protein-2A, cytokeratin 1-19 and Cyp 2a4 genes. This study also showed the differential expression of steroid metabolism genes such as increased expression of 17β-hydroxysteroid dehydrogenase-7 gene (HSD17β7) that is involved in the production of estradiol and down regulation of HSD17β5, important in testosterone production. These authors also reported the down regulation of methylation related genes: methionine adenosyltransferase-1a, betaine-homocysteine methyltransferase and thioether S-methyltransferase.

As explained before, arsenic exposure has been related to carcinogenesis based on epidemiological data. Some experiments have investigated possible carcinogenic effects due to arsenic exposure. For example, an experiment investigated the effects of chronic exposure of rat liver epithelial TRL1215 cells to arsenic. Arsenic induced malignant transformation in a concentration-dependent manner. At carcinogenic concentrations (500nM) arsenic exposure increased the expression of α-fetoprotein (AFP), Wilm's tumor protein-1 (WT-1), c-jun, c-myc, H-ras, c-met and hepatocyte growth factor, heme oxygenase-1, superoxide dismutase-1, glutathione-S-transferase-π and metallothionein-1 (MT) by 3 to 12-fold, while expressions of

insulin-like growth factor II (IGF-II) and fibroblast growth factor receptor (FGFR1) were diminished significantly (Liu *et al*., 2006).

Other scientists investigated early hepatic events associated with oncogenic transformation in fetal liver samples of male mice exposed *in-utero* to arsenic (85ppm). Global methylation of hepatic DNA was not altered by the exposure, however there was an important reduction in methylation of GC-rich regions. The methylation status of these regions is an important epigenetic mechanism that controls gene expression. The authors were able to relate the changes in methylation to increased expression of genes coding for glutathione production and abnormal expression patterns of genes involved in insulin growth factor signaling pathways and cytochrome P450 enzymes (Xie *et al*., 2007).

Arsenic effects on reproduction

Arsenic has also been found to alter reproduction, which may be through an estrogenic mode of action. Treatment of rats with sodium arsenite in water at 0.4ppm/rat/day, an environmentally relevant water concentration, resulted in a significant reduction in plasma levels of LH, FSH and estrogen, along with a decrease in steroid hydroxylase activities. This treatment decreased ovarian, uterine and vaginal weights, while it increased the diestrous phase of the estrous cycle (Chattopadhyay *et al*., 1999). In Bangladesh, a human epidemiological study showed a reduction to 89.1% of live births in women who had drunk arsenic-contaminated well water compared to women that drank uncontaminated water (Ahmad *et al*., 2001). The study also showed an increase in stillbirths in exposed women by 2.6-fold and an increase in spontaneous abortions by 3.3-fold. Similar increases were found by Pinney (Pinney and Lemasters, 1996) in female employees exposed to arsenic. Hopenhayn-Rich, (Hopenhayn-Rich

et al., 2000) reported an elevation of late fetal, neonatal, and post neonatal mortality rates in Antofagasta, Chile in populations that have been chronically exposed to arsenic. Other studies associated spontaneous abortions with arsenic exposure and considered this exposure as a primary risk to the developing fetus (Golub *et al*., 1998). In addition, Wlodarczyk, (Wlodarczyk *et al*., 2001) induced exencephaly in mice after *in utero* sodium arsenate exposure.

Exposure to arsenic during development has been linked to carcinogenesis. As mentioned above, Liu (Liu *et al*., 2007) detected differential expression of estrogen and steroid related genes in male fetal livers using microarray analysis. These changes occurred during early stages of embryonic development and were linked to hepatocarcinogenesis later in the adult mice.

Toxicogenomics in risk assessment

The process of conducting epidemiological and environmental risk assessments can be greatly enhanced by the incorporation of gene expression studies. Microarrays are ideal tools in gene expression studies because they allow the simultaneous analysis of expression of multiple genes under several experimental conditions. In addition, microarrays allow the detection of changes in gene expression at low level of exposure and produce extremely sensitive results. Toxicogenomics combines genomic and bioinformatic tools to obtain profiles of gene expression in organisms resulting from chemical exposure. These profiles may be used to predict the potential toxicity of new chemicals or to elucidate mechanisms of toxicity of known toxicants (Gant and Zhang, 2005).

Toxicogenomic analysis will be combined in the future with more sophisticated techniques such as proteomics or metabolomics. Proteomics is the branch of toxicology that describes the expression of proteins in cells or organisms under different experimental

conditions. Similarly, metabolomics investigates the level of metabolic responses in an organism due to toxic or physiological challenge (Nicholson *et al.*, 1999). But there is strong parallelism in the effects that can be measured using these three different approaches. For this reason, the integration of toxicogenomics, proteomics and metabolomics will provide the most effective technological tool in environmental risk assessments in a new integrated area of knowledge called ecotoxicogenomics (Snape *et al.*, 2004).

Toxicogenomic analysis must be performed in an objective biological context. In order to do this, better bioinformatics methods must be designed to link gene expression changes with any relevant transcriptional, biochemical or biological process. One example of this has been the use of Gene Ontology (GO) pathway analysis (Currie *et al.*, 2005). However, toxicogenomics has many intrinsic challenges such as the development of standard protocols and the normalization of data to minimize systematic and experimental variation. In addition, the vast amount of information obtained in toxicogenomic experiments presents a formidable task for traditional statistical analyses, since statistical significance must be calculated from datasets that contain thousands of variables but few replicates (Maggioli *et al.*, 2006). In spite of this, this technology is maturing rapidly and will become an essential step in risk assessments.

Research needs for arsenic in drinking water

As explained above, the United States Environmental Protection Agency lowered the arsenic drinking water standards from 50 to 10 ppb in 2001. A strong factor that motivated legislators to implement this strict limit was scientific evidence that demonstrated that arsenic was correlated to increasing risk of carcinogenesis in a dose-responsive manner. The link of arsenic exposure to carcinogenesis has been supported by several additional experiments using

animal models. But the limits of arsenic in drinking water will be subject to periodic revisions. There is still a need to conduct experiments in order to understand the molecular effects of arsenic in a dose-responsive manner. Many areas of research will work together to accomplish this.

Clewell reported efforts done by the EPA and different research centers in order to explore more sensible areas of research regarding the controversial arsenic drinking water standards. These efforts can be summarized as the study of physiologically based pharmacokinetic (PBPK) models of kinetics of arsenic and its metabolites in mice and humans; the development of mathematical models to investigate multiple pathways of cell transformation; the evaluation of literature regarding dose-response effects of arsenic using genomic technologies; and the investigation of dose-response patterns for changes in gene expression in urinary bladder following arsenic exposure (Clewell *et al*., 2007).

Even as the specific carcinogenic mode of arsenic action at the molecular level is not clearly understood, there is evidence that arsenic may interfere with proteins that are important to control the cell cycle (Snow *et al*., 2001; Snow *et al*., 2003) or with epigenetic mechanisms of gene expression (Xie *et al*., 2007). Some experiments indicate that these effects may be transient and adaptive at low submolar concentrations while toxic at higher levels (Snow *et al*., 2005). Evidence from *in vitro* and physiologically based models indicate that the arsenic concentrations in the cell at which the transition between adaptive and toxic effect occurs have been associated with arsenic exposure levels on the range of the former arsenic-drinking water standard (Clewell *et al*., 2007).

In 2006, a workshop was sponsored by the United States Environmental Protection Agency entitled “Research and Risk Assessment for Arsenic”. In this workshop, scientists

identified the main research points that need to be performed in order to optimize the risk assessment studies related to arsenic in drinking water (Hughes *et al*., 2007). Table 0.2 summarizes some of these points.

Table 0.2. **Main research points related to arsenic in drinking water (Hughes *et al*., 2007)**

- Identify source contribution of diet and drinking water to inorganic arsenic exposure.
- Investigate arsenical species distribution in food.
- Develop analytical methods to quantify trivalent and pentavalent arsenicals in tissue.
- Establish biomarkers of arsenic exposure and their link to human health.
- Investigate the mechanisms of arsenic-induced carcinogenesis.
- Develop exposure-dose-response models for arsenic exposure.
- Identify individual factors that influence arsenic metabolism such as ethnicity or polymorphisms and their relationship to adverse human health outcomes.
- Use animal models to investigate mechanisms of arsenic toxicity and carcinogenicity.
- Detect susceptible human populations to the negative effects of arsenic exposure.

Given the above rationale, this dissertation project investigated the effects of arsenic at the molecular level using a species of fish called mummichog (*Fundulus heteroclitus*). Radiolabeled microarrays were used to detect differential gene expression under different arsenic doses, which were correlated to negative reproductive outcomes. Fish provide excellent models in toxicogenomic studies due to their reproduction patterns and relatively low cost. Mummichogs

reproduce in a synchronous pattern based upon the lunar cycle, and the eggs are translucent and easy to monitor (Lotrich, 1975). It is possible to get large numbers of offspring in a short period of time, making this fish more convenient to investigate relative to other animal models. Mummichogs were exposed to different arsenic concentrations, and the differential gene expression patterns were analyzed in the livers of exposed fish and in the hatchlings of the first generation.