8-2010

Arsenic in drinking water causes gene expression changes in the liver related to inflammation and metabolic dysfunction and accelerates atherosclerosis in ApoE-/-mice

Matthew R. Zajack
University of Louisville

ARSENIC IN DRINKING WATER CAUSES GENE EXPRESSION CHANGES IN THE LIVER RELATED TO INFLAMMATION AND METABOLIC DYSFUNCTION AND ACCELERATES ATHEROSCLEROSIS IN APOE-/- MICE

By

Matthew R. Zajack
B.S., University of Wisconsin–Stout, 2007

A Thesis
Submitted to the Faculty of the
Graduate School of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Master of Science

Department of Pharmacology and Toxicology
University of Louisville
Louisville, Kentucky

August 2010

ARSENIC IN DRINKING WATER CAUSES GENE EXPRESSION CHANGES IN THE LIVER RELATED TO INFLAMMATION AND METABOLIC DYSFUNCTION AND ACCELERATES ATHEROSCLEROSIS IN APOE-/- MICE.

By

Matthew R. Zajack
B.S. Mathematics and Computer Science, 2007

A Thesis Approved on

August 5, 2010

by the following Thesis Committee:

J. Christopher States, Ph.D.

Gavin E. Arteel, Ph.D.

Eric C. Rouchka, D.Sc.

ACKNOWLEDGEMENTS

First of all, I would like to thank the Department of Pharmacology and Toxicology for giving me the opportunity to study at an institute with such rich traditions of academic excellence. I would like to thank my committee Drs. Gavin Arteel and Eric Rouchka and , in particular, my advisor Dr. J. Christopher States, for providing feedback and insight into my thesis and defense. I would like to acknowledge my lab mates, Ntube "Olive" Ngalame, Clarisse Muenyi, and Erica Rogers for their insight and constructive criticisms. I would also like to thank a few people who helped with experiments on this study: Heather Miller for maintaining and exposing the animals and Drs. Yulan Piao and Minoru Ko's lab for performing the microarray hybridization.

ABSTRACT

ARSENIC IN DRINKING WATER CAUSES GENE EXPRESSION CHANGES IN THE LIVER RELATED TO INFLAMMATION AND METABOLIC DYSFUNCTION AND ACCELERATES ATHEROSCLEROSIS IN APOE-/- MICE.

Matthew R Zajack

August 5, 2010

Arsenic exposure in drinking-water is a significant worldwide health problem. It causes adverse human health effects, such as cancer, increases the risks for others such as cardiovascular disease, and accelerates atherosclerosis. In this study, we analyze arsenic-induced gene expression changes in the liver of ApoE-knockout mice given 49 ppm arsenic in drinking-water. We hypothesize chronic arsenic exposure accelerates atherosclerosis by disrupting liver homeostasis, causing aberrant gene expression changes. Networks revealed hubs on 3 stress-response MAP kinase pathways, ERK, JNK, and p38. Pathways revealed mitochondrial dysfunction and oxidative phosphorylation enrichment from the gene set. Transcription factor binding site analysis revealed specific transcription factors Foxd1, estrogen receptors, heat shock factor, Myc, and Pparg. Taken together, results suggest metabolic function disruption in the mitochondria related to oxidative stress and increased MAPK activity. Future

studies should focus on arsenic's effects on mitochondria dealing with metabolism and activation of ERK, JNK, and p38 MAPK pathways.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Arsenic exposure

Arsenic exposure through drinking-water is a significant worldwide health problem. Arsenic contaminates drinking water primarily through both natural mineral deposits, tapped by groundwater wells, and anthropogenic sources, through mining and industrial use[1-2].

The US Agency for Toxic Substances and Disease Registry (ATSDR)—a division of the CDC—has, for over 13 years, listed arsenic as the number one priority substance based on frequency, toxicity, and potential for human exposure at facilities located at National Priorities List sites. To put this in perspective, arsenic is ranked above lead (2), mercury (3), benzo(a)pyrene (9), and DDT (12) in 2007. The US Environmental Protection Agency (EPA) set the maximum contaminant level (MCL) for arsenic at 0.010 ppm (parts per million; 0.010 mg/L As) or 10 ppb (parts per billion; 10 µg/L As) in 2001, which is the same as the World Health Organization (WHO) provisional guideline value set in 1993. Prior to this regulation, the MCL was set at 0.050 ppm for both the EPA in 1976 and WHO in 1963 [3]. In addition, the EPA set the public health goal for arsenic

exposure at 0 ppm. There are many areas in the United States in which groundwater is naturally contaminated with arsenic (0.010 ppm or more) including, but not limited to, Michigan, Texas, and parts of the southwestern states of Arizona, Nevada, and California, although this estimate excludes many rural, private, and unregulated wells for which data are unavailable [4-5]. In a study by the EPA in 2000, of the 43,443 ground water systems queried, 2,302 had arsenic at concentrations > 10 ug/L (0.010 ppm), which is above the current EPA MCL [6]. Additionally, there are reports of arsenic drinking-water contamination from around the world, including Southeast Asia, China, Iran, and Mexico, and many others [7-8], with some of these areas containing concentrations as high as 1000 µg/L arsenic dissolved in groundwater. Long-term arsenic exposure increases the risk of many diseases, including hyperpigmentation, keratosis, skin cancer, and internal cancers [9]. Additionally, the EPA states health effects from long-term arsenic exposure include skin damage, problems with circulatory systems, and an increased risk of cancer [10].

Long-term exposure to arsenic in drinking-water has a plethora of human health effects including increased risks of skin, bladder, and lung cancer [1-2]. Arsenic is excreted in the urine, primarily through the kidneys, within a few days after ingestion. In a study by Chen et al conducted in northeastern Taiwan, arsenic exposure at <10 µg/L (0.010 ppm) arsenic for over 50 years from birth significantly increased the risk of urinary cancer with a relative risk of 4.12 [11]. It is also hypothesized that arsenic acts as a skin carcinogen by enhancing the effects of ultraviolet (UV) radiation [12]. This is especially likely considering the

aberrant effects that arsenic has in combination with polymorphisms in nucleotide excision repair genes such as *XPA* and *XPD* [13]. Arsenic may also enter the body through inhalation. A study at the National Cancer Institute by Lubin et al reveals a linear relationship between lifetime cumulative arsenic exposure and respiratory cancer mortality [14]. In fact, this study argues against a previously postulated exposure threshold for carcinogenesis that was set at 150 µg/L (0.150 ppm) arsenic [15] by showing cumulative lifetime arsenic intake in smelter workers in Tacoma, Washington unequivocally corresponded to < 30 µg/L (0.030 ppm) arsenic. However, cancer is not the only concern from long-term arsenic exposure, cardiovascular disease risk is also an important consideration.

Non-cancer studies of the reproductive, neurological, cardiovascular, respiratory, hematological, diabetic, and dermal effects of arsenic are of significant interest, especially in cost-to-benefit estimates for drinking water standards [2]. Early reports concerning the noncancer effects of inorganic arsenic focused primarily on the blackfoot disease endemic in Taiwan that has occurred since the early 20^{th} century and peaked in the last 1950's [16]. Bangladesh and West Bengal, in an effort to stave off bacterial contamination persistent in surface drinking water sources, groundwater wells were dug by the WHO for use by residents. After a period of years investigators started seeing gangrene of the extremities, similar to the blackfoot disease seen in Taiwan, among other diseases, and related it back to the artesian well water being consumed by residents [17]. Blackfoot disease advances from coldness or numbness of the extremities, to intermittent claudication, and finally to gangrene

and surgical or spontaneous amputation. Engel et al [16] went on to explain that manifestations such as blackfoot disease are only "the tip of the iceberg," and that it is likely there is unseen systemic disease—he goes on to support it with reports of mortality due to vascular causes such as atherosclerosis. Indeed, vascular effects should be taken into consideration with respect to arsenic exposure and will be discussed more fully below.

In addition to arsenic's direct effects on the vascular system, there are studies implicating an association of arsenic with cardiovascular disease. Arsenic induces atherogenesis by causing endothelial damage, endothelial dysfunction, enhanced inflammatory activity, increased coagulability and decreased fibrinolysis, smooth muscle cell proliferation, increased oxidative stress, impaired nitric oxide balance, and induction of apoptosis. Arsenic has also been found in epidemiological studies to be associated with hypertension and diabetes mellitus [18]. Low-to-moderate levels of arsenic (11 μg/L) in drinking water in Michigan revealed an association with many diseases by standardized mortality ratio (SMR); e.g. circulatory system (SMR, 1.11) cerebrovascular disease (SMR, 1.19), diabetes mellitus (SMR, 1.28), and kidney disease (SMR, 1.28) [5]. All of these could predispose, trigger, or aggravate atherogenesis or are a consequence to atherosclerosis and are discussed below in more detail.

Cardiovascular disease and atherosclerosis

Cardiovascular disease is the leading cause of death worldwide. Cardiovascular disease accounted for more than 34% of all deaths in the United States in 2006, and the estimated direct and indirect cost of cardiovascular disease is over $500 billion for 2010 [19]. Cardiovascular disease is characterized by the dysfunction of blood vessels and the heart. Diseases caused by cardiovascular disease are high blood pressure (hypertension), coronary heart disease, heart failure, renal failure, and stroke. Atherosclerosis, a disease of the large arteries, is the underlying cause of the majority of cardiovascular events and in westernized societies is also the underlying cause of about 50% of all deaths [20]. Atherosclerosis is a condition in which plaque—composed of cholesterol, fat, calcium, and other substances in the blood—builds up in the inner linings of arteries.

Atherosclerosis is a progressive, multifactorial disease that is characterized by the accumulation of lipids, inflammation, and scar tissue and other fibrous elements in the arterial walls—the process below is summarized from *Harrison's Principles of Internal Medicine* [21]. An abundance of circulating cholesterol promotes accumulation of low-density lipoprotein (LDL) particles on the outside and invading into the endothelial layer of the arterial wall, forming fatty streaks. Once LDL particles are deposited below the endothelial layer, away from circulating blood, they favor oxidative modification. These oxidized LDL (oxLDL) particles may trigger a local inflammatory response, activating inflammatory mediators such as leukocytes and macrophages. Macrophages will

engulf oxLDL and become cholesterol-rich foam cells. Accumulation of foam cells forms a plaque and progresses the atherosclerotic lesion, releasing more inflammatory markers. Under stable conditions, smooth muscle cells will eventually proliferate into the macrophage-rich lesion area creating a fibrous lesion and narrowing the artery. Alternatively, under certain conditions this plaque may remain unstable, not forming a fibrous lesion, and rupture; a rupture triggers blood clotting that may block the entire artery. Atherosclerosis is a chronic process that occurs in multiple "steps" and is able to go into quiescence for long periods of time to resume when conditions again favor lesion formation.

Inflammation plays an established fundamental role in mediating all stages of atherosclerosis, from initiation through progression, and finally, thrombotic complications of atherosclerosis [22]. Atherosclerosis is characterized by recruitment of white blood cells, such as monocytes and leukocytes, to early atherosclerotic lesions [20]. Normally, however, the endothelium does not support binding of white blood cells. According to Libby et al [22], an event or trigger, such as hypertension, smoking, insulin resistance, or an atherogenic diet (a diet high in cholesterol and fats) is needed to stimulate endothelial cells to begin to express pro-inflammatory molecules, including adhesion molecules that will bind various leukocytes; once blood-derived inflammatory cells participate with endothelial cells, they may perpetuate a local inflammatory response [22]. In this way, the inflammatory signal to initiate atherosclerosis need not be found locally.

Distant sources of inflammation may contribute to atherosclerosis [23]. The idea that a seemingly unrelated condition may contribute to complications of another disease is not novel. For example, the hypothesis that an infection may be associated with atherosclerosis, in fact, was suggested over 100 years ago by Hektoen [24]. One such example is from the teeth in the case of gingivitis by Ford et al [25]. Ford injected 6 week-old apoE-/- mice each either *P. gingivalis* or *C. pneumoniae*, killed the mice at 18 or 34 weeks of age, and assessed atherosclerotic lesion area. They showed all immunized mice developed lesions greater than those of controls ($p<0.05$), but that those immunized with *P. gingivalis* developed larger atherosclerotic lesions than those with *C. pneumoniae* ($p<0.043$).

Systemic inflammation through liver damage also contributes to atherogenesis. In a case-control study by Brea et al, nonalcoholic fatty liver disease (NAFLD) is associated with atherosclerosis [26]. They assessed 93 patients with fatty liver and 40 controls for frequency and magnitude of cardiovascular risk factors and measured carotid atherosclerosis by intima-media thickness (IMT). They found patients with NAFLD had significantly more abnormalities related to metabolic syndrome (MetS), including visceral adiposity, hypertension, abnormal glucose metabolism, insulin resistance, hypertriglyceridemia, and low HDL cholesterol ($p<0.005$). They suggest that NAFLD is atherogenic beyond its association with MetS. They propose mechanisms related to enhanced oxidative stress, reactive oxygen species

causing increased circulating inflammatory milieu, and circulating C-reactive protein (CRP) contributing to the inflammatory status in NAFLD.

Systemic inflammation contributing to atherosclerosis is widely accepted and now being linked with many other diseases. Berg et al reviews evidence supporting the hypothesis that adipose tissue plays a role in development of systemic inflammatory state that contributes to CVD [27]. The review touches on subjects such as innate immunity mediated solely through fat bodies activated by NF-kB in metazoan species, toll-like receptors in mammals activating signal transduction pathways activating mediators such as interleukin-6 (IL-6) and TNF-alpha, and response to infections by IL-6, IL-11, and interferon-gamma (IFN-gamma). Berg et al explains that obesity contributes to systemic inflammation by causing increased expression of TNF-alpha, IL-6, PAI-1, CRP, fibrinogen, and increases macrophage infiltration, and that the liver contributes to this condition.

Obesity also contributes to CVD via atherosclerosis, hypercholesterolemia, insulin resistance and type 2 diabetes, and metabolic syndrome [27-28]. Libby et al review a study that looks at clinical applications for inflammation in atherosclerosis to limit cardiovascular events [29]. The JUPITER trial is a placebo-controlled study to evaluate if people with some inflammation (indicated by CRP) but below median levels of low-density lipoprotein (LDL) would benefit from statin therapy. Endpoints of myocardial infarction (MI), stroke, or cardiovascular death fell by 47% in rosuvastatin group compared to placebo group ($p<0.00001$) with little cost in terms of unwanted effects. The study provided further evidence that cardiovascular events can be reduced by LDL

lowering and an LDL-independent anti-inflammatory effect. The concept of inflammation in atherosclerosis is no longer only theory and laboratory investigation, it shows a promising role as a tool in the clinic to aid prevention and management of cardiovascular disease [29]. Hotamisligil, in a review, addresses metabolic and inflammatory responses as the core of metabolic disorders such as obesity and CVD [28]. He discusses several topics related to the mechanistic core of metabolic and inflammatory responses: Tumor Necrosis Factor alpha (TNF-alpha), a proinflammatory cytokine, is overexpressed in adipose tissues and acts to inhibit insulin action. Peroxisome-proliferator activated receptor (PPAR) and liver X receptor (LXR) transcription factors inhibit the inflammatory response in macrophages and adipocytes. c-Jun amino-terminal kinase (JNK) mediates insulin resistance, plays a critical role in liver tissues, and promotes inflammatory gene expression through activator protein-1 (AP-1) and NF-kB. Obesity and type 2 diabetes are critically involved in many processes that interact with metabolic dysfunction and inflammation[28]. Among many transcription factors involved in stress-response, NF-kB is very sensitive to arsenic-induced oxidative stress [30]. Normally, up-regulation of IKK leads to degradation of IkB and in turn translocation of NF-kB to the nucleus where it is activated. Ghosh et al have shown that arsenic in cardiomyocytes increased ROS production and increased apoptotic cell death [30]. When they administered a JNK and p38 MAPK inhibitor, NF-kB and IKK phosphorylation was attenuated, suggesting that arsenic increases phosphorylation of p38 MAPK and JNK MAPK that leads to NF-kB by IKK activation.

Apolipoprotein E-knockout mouse is a model for atherosclerosis

The apolipoprotein E-knockout (ApoE-/-) mouse [31] is a powerful investigatory model for atherosclerosis, especially considering that normal mice are highly resistant to atherosclerosis. ApoE-/- mice are used extensively to study the development of atherosclerosis.

Apolipoprotein E (apoE) is one of several evolutionarily conserved lipoprotein genes. Its primary function is to mediate the receptor-mediated removal of lipoproteins from the blood. ApoE is the primary ligand for the low-density lipoprotein receptor (LDLr) which mediates the removal of lipoprotein remnants from circulation. ApoE-deficiency results in decreased removal of LDL cholesterol, i.e. increasing "bad" cholesterol. This results in an increase in overall cholesterol levels from 75 mg/dL in the normal (background strain) to 500 mg/dL in ApoE-/- mice. High levels of LDL and/or low levels of HDL cholesterol are risk factors for atherosclerosis and CVD. Plump et al showed the atherosclerotic lesions in ApoE-/- mice also accurately represent the extent and severity of human atherosclerosis [31]. The ApoE-/- mouse develops massive fibroproliferative atherosclerosis [31]. The increased lipid levels and disease progression mimic human atherogenesis making the ApoE-/- mouse a desirable model to study various effects on the cardiovascular system.

Simeonova et al evaluated atherosclerotic plaque formation in arsenic exposed ApoE-/- mice, and C57BL/6 (background strain), and inflammatory signaling in human aortic endothelial cell (HAECs) culture [32]. Mice were given

20 or 100 μg/mL sodium arsenite in drinking water for up to 24 weeks with either regular chow or high fat diet. Body weight and serum cholesterol levels did not significantly differ between arsenic and control animals fed a regular diet, whereas the high-fat diet significantly increased only cholesterol levels; moreover, 20 ug/mL sodium arsenite did not increase serum cholesterol or weight but 100 ug/mL sodium arsenite had a slight but significant increase in cholesterol. C57BL/6 mice did not develop atherosclerotic lesions with any exposure or diet; however, mice with 20 or 100 μg/mL sodium arsenite on high fat diet induced a 1.6 and 2.3 fold increase in lesion size, respectively. Inflammatory genes IL-8, NF-kB, and AP-1 were induced in HAECs in response to arsenic treatment. They also suggest the role of oxidative stress needs further evaluation as a mechanism for arsenic-related vascular effects. Similar results were seen by Bunderson et al, who treated mice with a lower arsenic concentration (10 ppm) [33]. They also saw a significant increase in plaque size between arsenic and control, and saw evidence of oxidative stress and inflammation. These data suggest the ApoE-/- mouse is a good model for studying the effects of arsenic exposure on atherosclerosis.

Mechanisms of arsenic-induced atherosclerosis

We have already discussed that atherosclerosis is a multifactorial disease that is characterized by fat accumulation, inflammation, and thrombotic events [21]. We have also discussed how inflammation is central to every step of atherogenesis, including initiation, progression, and ultimately the thrombotic events [22]. Arsenic is a known carcinogen, causing skin, bladder, and lung

cancer at concentrations < 150 µg/L arsenic [1-2, 12-15]. Arsenic also has the noncancer effects of vascular disease, represented as blackfoot disease [17] and diabetes mellitus[5]. A couple reviews have been published that review epidemiological evidence implicating arsenic exposure in cardiovascular disease and atherosclerosis [34-35]. Briefly, these two reviews look not only at the epidemiological evidence, but also at *in vitro* cell culture studies and *in vivo* ApoE-/- mouse studies. Simeonova et al reviews epidemiological studies on arsenic exposure from 1989-2004 and examines animal models of atherosclerosis and arsenic. States et al summarizes several epidemiological studies in Bangladesh and Taiwan, as well as provides an updated review of arsenic exposure in the ApoE-/- mouse. Now we look at several studies implicating arsenic exposure in accelerated atherosclerosis, chronic inflammation, and cardiovascular disease risk factors such as diabetes mellitus.

A study by Wu et al showed that a high fat diet exacerbates arsenic-induced liver fibrosis in mice [36]. This study exposed Kunming white mice to water containing 200 ppm arsenic as sodium arsenite for 10 months and collected blood and livers. They found that arsenite significantly increased serum aspartate aminotransferase (AST)—a clinical sign of liver injury—levels 5 fold, and when combined with a high fat diet increased by 10 fold over unexposed controls. Mild fibrosis (5 fold increase over unexposed controls) occurred in the arsenite alone group; combined with a high fat diet, however, fibrotic lesions were more severe (more than 10 fold increase over unexposed controls). RT-PCR was used to reveal an increase in expression of genes related to hepatic

inflammatory response, namely TNF-alpha and IL-6 by 6 and 7.5 fold, respectively, which was drastically increased in a high fat diet. TGF-beta was also enhanced for a high-fat diet with arsenic exposure (13 fold) and played a role in arsenic hepatofibrogenesis. A high fat diet enchanced arsenic-induced liver fibrosis in this study and caused an array of related aberrant gene expression changes [36]. Arteel et al demonstrated similar results using four to 6 week-old C57BL/6J mice were exposed to 49 ppm arsenic as sodium arsenite in drinking water for seven months [37]. They showed ALT and AST levels were significantly increased 2 and 3 fold, respectively in response to lipopolysaccharide (LPS) challenge. Arsenic pre-exposure in this study significantly increased liver damage due to LPS. This, combined with a review in the Waalkes lab, show that the liver is a major target organ of arsenic toxicity and carcinogenesis [38]. Many of the aberrant genes that were differentially expressed in these experiments are related to inflammatory signals and that cause restructuring of the liver.

A study by Srivastava et al shows arsenic exacerbates atherosclerotic lesion formation and inflammation in ApoE-/- mice [39]. Three-week-old mice were exposed to arsenic in drinking water at 1, 4.9, and 49 ppm arsenic concentrations for 13 through 33 weeks. 10 week-old control mice showed few lesions in aortic arch, while arsenic-exposed mice showed >5 fold larger lesions ($p<0.01$). Exposure to arsenic increased plaque formation in the aortic valve and aortic arch in a dose-dependent manner at 1 and 4.9 ppm arsenic in mice

exposed for 13 weeks; macrophages also accumulated in lesions in a dose-dependent manner.

Various signal transduction pathways are arsenic-induced, and reviews can be found here [40-41]. The three major classes of MAPKs are extracellular signal-regulated kinases (ERK)s, c-jun N-terminal kinases (JNK), and p38. ERKs are growth factor-responsive and induce cell differentiation, proliferation, and transformation signaling. JNK and p38 primarily mediate cytokines and stress-related responses such as cell growth arrest and apoptosis. These pathways affect genes that perform a variety of cellular responses that have mechanisms leading to cancer-causing effects. Nuclear Factor-kappa B (NF-kB) is a transcription factor that mediates cellular processes such as cell-to-cell interaction, intracellular communication, cell recruitment or transmigration, amplification of pathogenic signals, and initiation or acceleration of carcinogenesis. MAP kinases play a role in arsenic-incuded NF-kB activation but are complicated and have been reviewed by Qian et al [40].

There are also reports that *in utero* (developing fetus in mother) arsenic exposure induces the early onset of atherosclerosis in ApoE-/- mice [42]. Pregnant mice were exposed to 85 µg/L sodium arsenite (49 ppm arsenic) in drinking water for 2 weeks, and the aortic trees of male offspring were analyzed at 10 and 16 weeks after birth. A few small lesions were seen in unexposed control mice. Atherosclerotic lesions in the aortic valve were seen in 5 of 10 arsenic-exposed mice, and as a group had 2 fold greater lesion size than controls (0.5-1.2% in controls vs. 0.6-3.8% in arsenic-exposed; $p<0.05$). The

authors speculate arsenic has lasting effects on susceptibility to atherosclerosis in adulthood related to endothelial dysfunction [42]. In addition, a study by Fry et al monitored gene expression profiles in newborns whose mothers were arsenic exposed during pregnancy and has proposed 11 transcripts that are potential biomarkers of prenatal arsenic exposure [43]. CXL1, DUSP1, EGR-1, IER2, JUNB, MIRN21, OSM, PTGS2, RNF149, SFRS5, and SOC3 all show a dose response to prenatal arsenic exposure. Stress response and cell cycle regulation are the associated molecular functions of this gene set. Furthermore, this study looks at other genes and found that gene networks centered on NF-kB and IL1-beta, an inflammation response gene, transcription factors STAT1 and HIF-1alpha, and proinflammatory cytokine TNF-alpha [43].

Kozul et al found chronic arsenic exposure in drinking water alters immune response genes expression in mouse lung [44]. Six-week-old C57BL/6J mice were exposed to 0.010 ppm arsenic as sodium arsenite in chow and 0.010 or 0.100 ppm arsenic as sodium arsenite in drinking water for 5 weeks and RNA from lung samples were collected for microarray analysis. They saw alterations in immune-related gene expression using Ingenuity Pathway Analysis (IPA) in 0.010 and 0.100 ppm arsenic exposure. Pathways including TNF-alpha (which was upregulated) and IL-1/TLR signaling were also altered. This study shows that even low-to-moderate levels of arsenic are causing aberrant gene expression using *in vivo* models.

Arsenic is metabolized in the liver by humans and rodents, and the liver is a known target of arsenic toxicity causing abnormalities such as fibrosis and

hepatocarcinoma The Waalkes lab has published articles elucidating mechanisms of arsenic-induced liver injury [36, 45] including a literature review [38]. Among potential mechanisms for arsenic hepatocarcinogenesis are oxidative damage and stress, altered DNA repair, protein interaction with zinc finger proteins, hypo- and hypermethylation, and aberrant estrogen linked gene expression. As mentioned earlier, Arteel et al have studies implicating arsenic exposure to liver injury [37]. They demonstrated that even subhepatotoxic levels of arsenic are able to enhance inflammatory liver damage caused by lipopolysaccharide (LPS) and elicited a hyper-response. These studies suggest arsenic may be a risk modifying factor for other diseases.

We have proposed a body of evidence (above) linking arsenic exposure with cardiovascular complications. Cardiovascular disease is prevalent throughout the world and in the United States. Even a small increase in risk attributed to arsenic exposure would translate into a large number of excess deaths and would therefore be of great significance to public health. Providing insight into the biological mechanisms by which arsenic is able to exacerbate atherosclerosis may reveal targets for preventative measures or therapies. Additionally, learning more about the effects of a known water contaminant would aid policy-making decisions based on risk-cost-benefit models.

We hypothesize that chronic arsenic exposure accelerates atherosclerosis by disrupting liver homeostasis. Homeostasis in the liver is disrupted in a way that makes the liver prone to aberrant gene expression changes that result metabolic dysregulation and inflammation. The inflammation in the liver is able

to impact the rest of the body, and thereby atherosclerotic lesions, through systemic circulation because of the chronic nature of the immune response. The following study examines the transcriptome of the livers of ApoE-/- mice exposed to 49 ppm arsenic in drinking-water as sodium arsenite ($NaAsO_4$) from 3 weeks (PND21) to 10 weeks (PND70) of age.

CHAPTER 2

METHODS

Animal housing and treatment protocols

Male ApoE-/- mice (B6.129P2-ApoE^{tm1Unc}/J, Jackson Laboratory, Bar Harbor, ME) were used in this study. Mice were housed in pathogen-free conditions in the University of Louisville vivarium under temperature control in a 12 h light/12 h dark cycle. All mice were maintained on a normal chow diet and tap water prior to and during exposure as described previously [42]. Tap water contains arsenic below the detection limit of 2 ppb, according to the Louisville Water Company. Arsenic concentration in the normal chow was not measured in the present study; however, arsenic up to 390 ppb has been reported in non-purified diet [46]. This study and protocols therein were approved by the University of Louisville Animal Care and Use Committee.

Three week old mice (PND21) were weaned and maintained on tap water (control) or tap water supplemented with Sodium Arsenite, $NaAsO_2$, (85 mg/L; 49 ppm As; 49 mg/L As) for 7 weeks by Heather Miller. Mice were euthanized immediately at the end of the exposure period, when they were 10 weeks old

(PND70). Livers and other organs were collected and snap frozen in liquid nitrogen and subsequently stored at -80 C until analysis.

RNA collection and preparation

Livers of ApoE-/- mice with or without exposure to 49 ppm arsenic in drinking-water were collected. RNA was extracted, and sample quality was checked by A260/A280 ratio. Yulan Piao in Minoru Ko's lab at the National Institute on Aging (NIA) generated fluorescently-labeled double-stranded cDNA using a T7 reverse transcriptase. A reference total mouse standard RNA (Cy5) was obtained from Stratagene. Labeled samples at the NIA were applied to the NIA Mouse 44K Microarray v2.1 (Whole Genome 60-mer Oligo) GEO accession GPL2552, as previously described [47]. Each of the control (Cy5) and experimental samples (Cy3) were used on one array, for a total of 6 arrays.

Microarray analysis

Microarray output data was processed by the National Institute on Aging (NIA) and provided in a tab-delimited text file. Data files have been uploaded to NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/), and the GEO accession number is pending approval. The data file contained numerical Cy3 and Cy5 signal values per chip for each of the probes, which were represented as proprietary Agilent Feature ID's (also called Probeset ID's or Probe ID's in some systems). Feature ID's were combined with Agilent annotation data. The annotation data allowed propriety Feature ID's to be

identified by open standards, where available, in the form of RefSeq accession, Genbank accession, Entrez gene symbol, among others.

The signal values were processed by both GeneSpringGX 10 (Agilent Technologies) and Partek Genomics Suite (Partek Inc.) for comparison. The process was similar in both programs: We used each program's import feature on our signal values file, and linked it to the Agilent annotation data. Next we defined each column of signals as either Cy5 (standard RNA) or Cy3 (experimental sample). Then we formed groups for different experimental parameters; e.g. treatment conditions (arsenic-exposed and unexposed control) and age when euthanized (PND70). Finally, we needed to define interpretations for the data—that is, compare and contrast arsenic-exposed vs. unexposed control groups. Each program provided various, slightly different statistical methods to interpret the data. In GeneSpring we used the unpaired TTest of equal variance, and in Partek we used the ANOVA to obtain *p*-values—both GeneSpring and Partek averaged signal value over replicates. Programs were compared and contrasted (explained in results). Fold-change ratios were calculated identically in each program by dividing the replicate average Cy3 (experimental samples) by the replicate average Cy5 (standard RNA). Fold-change ratios were log2-transformed into fold-change values that are reported as either up or down in arsenic-exposed vs. control unexposed.

MicroRNA microarray analysis

A microRNA microarray analysis was performed by Exiqon using miRCURY™ LNA Array Version 8.1. Results were provided in a spreadsheet in Microsoft Excel format with signal values for each probe with Hy3 as arsenic-exposed and Hy5 as RNA standard from Ambion. Probes were labeled with miRNA names that match miRBase ID's. Filtering criteria was based on based on students *t*-test used in Microsoft Excel and defined as all microRNA's where $p \leq 0.05$. Those that passed these criteria were subsequently used in further analysis.

Next, we found mRNA targets to the microRNA. We used two different targets databases, TargetScan (http://www.targetscan.org/) and microCosm (http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/). TargetScan provides a browser interface for searching each microRNA individually, outputting matching mRNA in a text file with columns for Entrez gene symbol, number of conserved sites, and representative microRNA. MicroCosm Targets provides a tab-delimited text file containing all microRNA-mRNA pairs in a specified species. Using this, we filtered out genes that were targets of our microRNA, obtaining a text file with columns corresponding to Entrez gene symbols and binding characteristics. We filtered these lists down by finding microRNA-mRNA pairs that were in common between the two lists.

Functional analysis

Ingenuity Pathway Analysis (IPA; Ingenuity Systems) allows us to gain insight into the biological significance to our large gene set. IPA takes advantage of several different data sources including, but not limited to, Argonaute for microRNA-mRNA interactions, BIND for protein-protein interactions, and Gene Ontology. We uploaded the entire gene set for each GeneSpring and Partek, including all annotations and specifically columns for p-value and fold-change. We used the Genbank Accession number as the identifier to allow IPA to map additional annotation information. We performed a New Core Analysis with each set of data where $p \leq 0.01$ with any fold-change (FC); the settings were defaulted (all databases) with species set to both human and mouse.

DAVID Bioinformatics Resource (Database for Annotation, Visualization and Integrated Discovery; http://david.abcc.ncifcrf.gov/) is an application that enhances data analysis in functional annotation. Similar to IPA, DAVID allows us to extract biological meaning from our large gene set. DAVID uses several different publicly available tools to create function annotation including, but not limited to, GoMiner, GSEA, Onto-express. It will then display only the most significantly enriched terms. We used DAVID according to protocols published in *Nature Protocols* [48] and *Genome Biology* [49]. Both our GeneSpring and Partek annotated gene sets data were filtered for genes with $p \leq 0.01$ with any fold-change (FC). DAVID only allows the identifiers—it does not use *p*-value or FC information, so we uploaded the filtered Genbank Accession numbers into DAVID. It automatically detected mouse as our host and set the background

gene list to the whole mouse genome. The Functional Annotation view was used to evaluate function based on all databases available in DAVID.

Transcription factor binding sites analysis

We also utilized computational promoter analysis to find significantly enriched transcription factor binding sites (TFBS) using Expander (EXpression Analyzer and DisplayER; http://acgt.cs.tau.ac.il/expander/). Expander is a full-featured suite for genomics analysis. It allows uploading of a raw signal file that takes the data through processing, normalization, mapping, filtering, and then may perform a plethora of enrichment-finding procedures—the enrichment-finding is similar to IPA and DAVID. We, however, are exclusively interested in Expander's promoter analysis feature. We followed a protocol published in *Nature Protocols* as a reference [50]. Expander requires an organism specific data file to provide gene information, including the transcription factor fingerprint files. The information in the data file is obtained from publicly available databases such as NCBI, Ensembl, and others. We used the mouse organism file.

Promoter analysis utilizes the PRIMA algorithm, which works by finding transcription factors whose binding sites are enriched in a given set of promoters. To stay within the restraints of the program, we were restricted to looking at a region 2000 bp (base pairs) upstream and 200 bp downstream of the transcription start site (TSS) in our genes of interest. We filtered GeneSpring genes where $p \leq 0.05$ and FC$\geq \pm 1.5$—genes that had a "good chance" of being "highly differentially regulated" in our gene set. Expander requires genes to be

identified by Entrez gene ID's. To convert them, we uploaded the Entrez gene name into DAVID and used the batch name conversion tool to obtain the gene ID's for mouse. We are then able to import the gene set without any expression values into Expander. Expander returns TFBS names, along with number of genes enriched with it and a *p*-value against the total mouse genome as the background. To relate TFBS to the transcription factor (TF) that binds it, we used MAPPER (http://snpper.chip.org/mapper/mapper-main) to resolve the TFBS numerical IDs into most-likely binding TFs.

CHAPTER 3

RESULTS

Gene expression profiles, using RNA derived from livers of apoE-/- mice exposed and unexposed to arsenic, were run on the whole-genome mouse NIA 44K 2-color oligo-DNA microarrays. When we began this study, we had to choose among many resources for analyzing our microarray data, including the ability to choose between two different genomics suite applications, GeneSpringGX 10 and Partek Genomics Suite. Both applications provided comparable features for our needs, so we compared and contrasted how these applications differed in the processing of our data. Signal files containing raw Cy5 and Cy3 readings were imported to GeneSpring and Partek as described in the Materials and Methods. We received the signal files from NIA (materials and methods), and they had already normalized the data and no further normalization was necessary. Figure 1 is an MvA plot that shows the difference versus the average of probe measurements between two samples and assesses the relation between the Cy5 and Cy3 channels of each hybridization. Each small square represents an individual probe. The original plot (left) and GeneSpring normalized plot (right) did not differ, indicating the signal file was already normalized, as we expected.

Each of the expression files contained signal values for 41,961 oligonucleotide probes. Using NIA annotation file data, probes were mapped to 28,256 (21,484 unique) Entrez gene symbols and 28,225 GenBank (20,686 unique) accessions. Although there were more gene symbols mapped to probes, when we uploaded the dataset to Ingenuity Pathway Analysis (IPA) we received over 4000 fewer IPA mapped IDs with the gene symbols set as the ID column than with GenBank accessions. Less IPA mapped IDs means that fewer of our probes, and subsequently fewer of our genes are available for analysis. Since the GenBank accessions provided more annotation information in IPA than gene symbols, we subsequently used GenBank accessions for all our functional analyses.

Significantly expressed genes were identified by TTest in GeneSpring and by ANOVA in Partek. Fold-change was calculated with log2 transformed Cy3/Cy5 ratios. For functional analysis we focused primarily on genes passing the criteria filter $p \leq 0.01$ with any fold-change. We were primarily concerned with populating our functional analysis programs with all of the genes that were significantly regulated, regardless of the amount of change. These sets we loaded into both IPA and DAVID for comparative analysis.

Comparing and contrasting GeneSpring and Partek

Principal component analysis (PCA) allows a broad overview of how the data sets relate to each other and is one way to express the variability among the sets. In Figure 2, we provided a 3D PCA image for both GeneSpring (top) and

Partek (bottom). The figure contains information for this postnatal study (exposures: none-pnd70 and pnd21-pnd70) and for a prenatal study (exposures: none-pnd1, gd8-pnd1, and gd8-pnd70) that was performed at the same time. Both programs indicate similar groupings among the data sets (similar colors group together and similar shapes group together). The biggest variability among the data sets was the age at which the mice were euthanized (called age at sacrifice), on the x-axis. In GeneSpring the pnd1 group is represented by squares in the foreground, and in Partek the pnd1 group is isolated on the left side of the graph. Additionally, including all of the components, there are groupings among replicate sets; however, the replicates corresponding to no arsenic exposure (none-pnd1 and none-pnd70) have the greatest spread and therefore greatest variability.

Both GeneSpring and Partek contain functional analysis abilities similar to IPA and DAVID; however, these methods were not further explored. GeneSpring requires different results interpretations to have specific annotation data that was not available to us in our Agilent annotation data file; for example, GO Analysis in GeneSpring requires Entrez gene ID's in the form of their Homo sapiens numerical ID's, which were unavailable because we were using a mouse genome microarray chip. Partek works similarly; however, we had the additional drawback of not having direct access to the program at our leisure. The best course of action was to use the readily available IPA and DAVID applications, which specialize in functional analysis of gene sets.

We initially compared the lists generated from GeneSpring and Partek with filtering criteria where $p \leq 0.01$ and any fold-change (FC) by directly comparing number of probes that were in common between the two programs. We chose any FC because both programs computed it similarly, so there was no difference. After filtering, the lists have 1,461 and 1,346 probes in GeneSpring and Partek, respectively. To compare these lists, we intersected the two filtered lists based on the Agilent Feature ID's column. This way, the specific probes needed to be represented in both lists, eliminating ambiguity accessions or gene symbols would have created. The intersection produced 599 probes between GeneSpring and Partek with $p \leq 0.01$ and any FC, which is less than half of either list. This shows that the statistical methods each program uses are not equivalent, as far as a significance cutoff is concerned; however, this does not give any insight as to whether this relates to a biological difference in the two sets. To determine if the two sets differ in function, we will need to use our functional analysis programs IPA and DAVID.

Ingenuity Pathway Analysis (IPA) functional analysis

We uploaded the GeneSpring and Partek gene set lists into IPA. GenBank accessions were marked as the ID field that IPA uses to map genes to additional annotation data, such as gene function and protein location in the cell. IPA also accepts the data corresponding to the *p*-value and fold-change (FC) of each gene. IPA was able to map 27,174 (non-unique) probes on each data set, which is nearly all of the probes available with GenBank accessions. We ran a new core analysis on each dataset with criteria set to $p \leq 0.01$ and any FC. With

these settings, IPA reported having 775 molecules available for functions/pathways and 374 networks eligible for GeneSpring and 690 molecules for functions/pathways and 324 network eligible for Partek. After running the core analysis on both datasets, we used the comparison analysis feature to compare information from both datasets at the same time. The top function in common between the two was cardiovascular disease. Top canonical pathways are oxidative phosphorylation and mitochondrial dysfunction. Taking a look at the networks in the comparison revealed numerous overlaps among genes.

The comparison analysis provides an overlapping networks view where we are able to view what networks are connected within and between analyses and how many molecules are in each overlap. Out of 50 networks between both analyses, only 4 were orphans (not connected to any other network). Many of the networks contained overlaps that accounted for a third of the network composition. For example, Figure 3 shows the top 2 networks from each analysis merged. Both networks have 10 molecule overlaps and are highlighted in the figure—the top network contains 2 networks composed of 32 molecules and 28 molecules for GeneSpring and Partek analyses, respectively, and the bottom 2 networks are composed of 27 molecules each. The top networks overlap on and contain network hubs on Creb, ERK1/2, GSTM3, GSTM4, MBD2, PPARG, among others, and the bottom networks overlap and contain hubs on INF Beta, Interferon alpha, Jnk, MAP2K6, NFkB, MAPK14 (P38 MAPK), among others. Top functions in the top network are cellular development, growth, proliferation, and movement, connective tissue development and function,

embryonic development, and cell cycle. Top functions in the bottom network are cellular movement, hematological system development, function, and disease, immune cell trafficking, cell death, and post-translational modification. The top network functions and top pathways are listed in Table 1 and Table 2, respectively. Finally, the "Top Tox functions" (toxicity phenotypes and clinical pathologies) for both datasets are mitochondrial dysfunction, oxidative stress, and LPS/IL-1 mediated inhibition of RXR function. These results indicate that both lists have common biological functions, despite having less than half the significantly regulated probes in common.

DAVID functional analysis

We corroborated results obtained in IPA by using DAVID Bioinformatics Resource. DAVID does not accept p-value or FC. In that respect, DAVID exclusively performs enrichment analysis—it only finds functions and/or networks that are significantly overrepresented by our gene list over a background gene list. Each of our gene sets was filtered in Microsoft Excel where $p \leq 0.01$ with any FC and output in text-only format containing only GenBank accessions. DAVID was able to find annotation data for 1026 and 949 GeneSpring and Partek probes, respectively. We set total mouse genome as the background gene list. The next step is to run the functional annotation clustering analysis and view the combined view. This list is composed of various clusters, arranged by function and score, containing categories that are database names and terms represented by short descriptions of the function/pathway that is enriched,

followed by various statistical measurements including number of molecules found in the term, percent of term filled, *p*-value, among others.

Table 3 lists representative terms per cluster for each GeneSpring (A) and Partek (B). Comparing the tables from each dataset, both lists contain similar terms that are significantly enriched including mitochondrion, translation, and ribosome function enrichment. In contrast, the Partek list contains much lower enrichment scores and lower enrichment percentages, although the values are still statistically significant ($p \leq 0.01$). The top functions in the DAVID analysis also matches what was seen in the IPA analysis. For example, the mitochondrion term in DAVID matches with the mitochondrial dysfunction and oxidative phosphorylation with oxidative stress. Additionally, several of the genes that appear in corresponding pathways from IPA and DAVID overlap Table 4, showing that the enrichment is due to the same genes and same functional pathway being enriched by both applications. This overlap lends evidence for taking a deeper look at these corroborated pathways and how they are regulated. Additionally, it implicates the mitochondrion in and provides a list of genes for mitochondrial dysfunction and oxidative phosphorylation.

MicroRNA analysis

MicroRNA analysis revealed a different challenge in differential expression analysis. The Exiqon microRNA microarray contained only 496 probes, of which only 344 had readable signals, which is substantially fewer probes than the cDNA microarray. We performed students t-test in Microsoft Excel on the

readable signals and were left with 8 probes where $p \leq 0.05$ and any FC. However, of these only 4 microRNA probes had annotation data to continue analysis, which were mmu-miR-206, mmu-miR-18, mmu-miR-140, and mmu-miR-337. We were then interested in what regulatory networks these microRNAs were most likely to regulate. To do this, we found all of the gene targets for each microRNA, but we found there are many different databases that provide prediction of microRNA targets and compared two popular methods, miRBase (http://www.mirbase.org/) and TargetScan (http://www.targetscan.org/).

There are fundamental differences in how each microRNA target prediction algorithm works. As a result, the predicted genes vary widely between miRBase and TargetScan Table 5. TargetScan produced substantially fewer predicted target genes in general—761 genes with TargetScan vs. 3912 genes with miRBase total. Additionally, the overlap or intersection between the two gene lists was low, resulting in 175 total genes that were common between the two methods for the 4 microRNAs. The mmu-miR-337 did not have any predicted genes in either database at the time of analysis.

Using the gene set of the intersection between miRBase and TargetScan, and we ran it through the same paces as our original transcript gene sets—we used IPA and DAVID to lend insight into the biological workings of the genes predicted to be targeted by our significantly expressed microRNA. IPA analysis revealed networks in Figure 4 with top functions tissue morphology, cellular growth and proliferation, and gene expression for network 1 (top) and cancer, cell morphology, and DNA replication, recombination, and repair for network 2

(bottom). Notable genes in these networks include TNF, which is a hub in networks, TP53, HIF1A, and many other DNA binding transcription factors, and IL10 in network 1. DAVID analysis reveals several different functional groups significantly enriched with the microRNA targets gene set. Additionally, top canonical pathways are listed as RAR activation, aryl hydrocarbon receptor signaling, role of BRCA1 in DNA damage response, estrogen receptor signaling, and hypoxia signaling in cardiovascular system. Listed in Table 6 are functions significantly enriched by this gene set. Listed in the table are many functions related to transcriptional regulatory networks, including the term transcriptional regulation and DNA binding. If the predicted genes targeted by the microRNAs that were significantly involved in our system were involved in regulatory networks such as transcriptional regulation, we may see these results corroborated with a transcription factor binding site enrichment analysis.

We also wanted to find out how the microRNA predicted targets compared to our transcript genes. So, we intersected the miRBase and TargetScan intersection genes (175 genes) with all genes where $p \leq 0.01$ and any FC (1461 genes). This intersection resulted in only 8 unique genes that are similar between the two lists, and they are Mylk, Mfn1, Pgrmc1, Commd2, E2f5, Usp33, Dhx15, and Pdap1. These are not enough to populate enriched pathways or to construct networks, so we had to use a relaxed filtering method. We then looked at microRNA predicted target genes from miRBase (3912 genes) and intersected those with our filtered genes where $p \leq 0.01$ and any FC (1461 genes). This intersection resulted in 158 unique genes. We ran this set through IPA to find if

any networks were populated and found results similar to what we saw in our transcript gene set Figure 5. The top 2 networks report functions for cell death, hematological system development and function, and inflammatory disease (top) and cellular development, cell death, and cell signaling (bottom). The networks had hubs at TNF, JNK, and p38 MAPK (top) and TNF, IL1B, STAT1, STAT3, and SOCS3 (bottom).

Transcription Factor Binding Sites (TFBS) analysis

Initially, we ran our TFBS enrichment analysis with all of the genes from our IPA/DAVID analysis set (where $p \leq 0.01$ with any FC). However, this resulted in hundreds of TFBS that were only marginally enriched (data not shown). We decided to be more selective with the dataset we input and decided to check only transcripts that were likely to be significantly involved with a large change against control, so we filtered the genes where $p \leq 0.05$ and $FC \geq |1.5|$ which resulted in 94 genes passing the criteria. In Table 7 are listed the most significantly enriched transcription factor binding sites, along with the transcription factor that binds the site and the fold-change of the TF, if it was available in our original gene data. Only a few of the transcription factors that were resolved from the TFBS were differentially expressed, indicated by our transcript dataset—Bhlhe40, Prdm1, Pparg, Zcrb, and Gtf2a1 were changed by at least 1.1 fold up or down. Additionally, Figure 6 is a hierarchical cluster map showing each gene and the TFBS it contains. With this, we are able to look at specific genes in relation to transcription factor regulation and identify sets of genes with potential common

regulators. Thus, sets of potentially coordinately regulated genes can be readily identified.

CHAPTER 4

DISCUSSION

Arsenic in drinking-water is a worldwide health problem, and millions of people are exposed to the detrimental effects of chronic arsenic exposure with drinking water levels exceeding the US EPA and WHO guidelines [2]. Arsenic exposure in human populations is associated with long-term health consequences [51]. We would like to address the molecular mechanisms pertinent to arsenic-induced cardiovascular disease. In this study, we characterize transcription-based molecular events occurring in the liver that may impact and contribute to the acceleration of atherosclerosis in the vascular system.

The current study shows chronic arsenic exposure in drinking-water induced gene expression changes in the liver of apolipoproteinE-knockout (ApoE-/-) mice related to functions dealing with metabolism and inflammation response. This study is an extension of one previously done in our lab where we observed chronic arsenic exposure in drinking water exacerbates atherosclerosis in ApoE-/- mice [39], and was discussed in a review [35]. Using an inbred mouse model provides us a way to eliminate many variables associated with population-

based arsenic exposure studies—it also reduces confounding factors that may occur in genome-wide expression studies.

In this study, we evaluated the effects of arsenic exposure on gene expression in livers of ApoE-/- mice. GO categories in DAVID found to be enriched with significantly differentially expressed genes between arsenic-exposed and unexposed groups included broad terms such as mitochondrion, translation, and ribosome function. Additionally, Ingenuity Pathway Analysis (IPA) produced networks with functions for cellular development, cellular growth, cellular proliferation, and cellular movement, cell cycle, hematological system development, hematological system function, and hematological system disease, immune cell trafficking, and post-translational modification that contained hubs on Creb, ERK1/2, GSTM3 and GSTM4, MBD2, and PPARG for one network and INF Beta, Interferon alpha, Jnk, MAP2K6, NFkB, and MAPK14 (P38 MAPK) on another network. These genes play an important role in mitochondrial function and inflammation regulation. Additionally, we included microRNA data into this study. There is much that is not known about the functions of microRNA, especially those found in the mouse. However, we were able to use predictive techniques to shed some light on the processes occurring here. In general, microRNA pointed to broad functions such as biological regulation, transcription regulatory activity, and blood vessel morphogenesis that support their supposed functions. Gene alterations seen in the present study are largely consistent with previous analyses in gene expression on arsenic in drinking water [33, 36, 43-45, 52]. Below is a critical discussion of the information contained in the results.

Atherosclerosis develops in "steps," when the conditions are present in a way to promote atherogenesis. We have previously discussed diseases such as diabetes [20], obesity [19, 27, 53], nonalcoholic fatty liver disease (NAFLD) [26], and chronic inflammation [28], among others that contribute to the risk atherosclerosis and cardiovascular disease. We propose that chronic arsenic exposure predisposes the mice to atherosclerosis when they are exposed to a second atherogenic insult, such as high fat diet or infection, if not directly.

The first and most significantly enriched functional Gene Ontology (GO) category in our gene set is the mitochondrion term. The mitochondrion (mitochondria, plural) is an organelle responsible for aerobic tissue respiration and is the primary site for energy production in eukaryotic cells via the electron transport chain. Oxidation of glucose takes place to produce energy in the form of adenosine triphosphate (ATP) in a process called oxidative phosphorylation. It is likely not a coincidence we also observed the oxidative phosphorylation GO term to be significantly enriched. Mitochondrial oxidative phosphorylation (OXPHOS) dysfunction leads to oxidative stress and is associated with atherogenesis and cardiac failure [54]. Although these processes are occurring in the liver, it is likely that arsenic reaching other tissues, such as the heart or vascular tissue, will show similar disruptions in oxidative phosphorylation. Additionally, the mitochondrion is responsible for a plethora of other processes, including signaling, cellular differentiation, cell death, and cell cycle control.

Our data also indicate that mitochondrial dysfunction and oxidative stress are occurring. It is not surprising that these two processes pop up at the same

time given that mitochondria are the primary consumers of oxygen, containing many redox enzymes capable of generating reactive oxygen species (ROS) [55]. ROS are chemically-reactive, highly reactive molecules containing oxygen that also contains an unpaired valence shell electron. ROS is a natural byproduct of respiration (metabolism of oxygen) but also can be dramatically increased during times of environmental stress, e.g. ultraviolet (UV) light, and ionizing radiation. Accumulation of ROS results in oxidative stress. Since the cell is exposed regularly to ROS, there are defense systems in the form of antioxidants. Enzymes such as superoxide dismutase (SOD) and glutathione S-transferase (GST) and small molecules such as ascorbic acid, glutathione, and sulfur containing molecules will prevent ROS damage by scavenging free radicals, leaving little net ROS production. If mitochondrial damage occurs with a decrease in antioxidant defense capacity, there will be net ROS production. If enough ROS accumulates to trigger oxidative stress it will result in additional damage to the mitochondria. Once this occurs, a vicious cycle can ensue, causing even more damage to the mitochondria and further reduce antioxidant capacity. In this study, many genes of interest related to antioxidant defense (SOD2 and various GSTs) are up-regulated, suggesting the livers of arsenic-exposed mice are experiencing increased ROS levels. This was expected, because arsenic is known to produce oxidative stress as a mechanism of hepatotoxicity, and possibly, carcinogenesis [1-2].

Above we spoke briefly on various arsenic-induced signal transduction pathways such as mitogen-activated protein kinase (MAPK) pathways that

regulate the expression of a variety of genes that mediate cell apoptosis, differentiation, proliferation, and transformation, and Nuclear Factor-kappa B (NF-kB) that mediates cellular processes such as cell-to-cell interaction, intracellular communication, cell recruitment or transmigration, amplification of pathogenic signals, and initiation or acceleration of carcinogenesis. The three major classes of MAPKs are enriched in our IPA networks, including extracellular signal-regulated kinases (ERK)s, c-jun N-terminal kinases (JNK), and p38. Oxidative stress is able to activate p38, which leads to cardiac dysfunctions, via generation of malondialdehyde [56]. Gosh et al have shown that arsenic not only increased phosphorylation of MAP kinases p38 and JNK, but they, in turn, activated NF-kB via the IKK pathway [30]. We have also seen enrichment of the ERK1/2 network with arsenic exposure. These relate directly with the enriched network functions dealing with cellular development, cellular growth, and cellular movement, cell death, immune cell trafficking, among others.

Peroxisome-proliferator activated receptor gamma (PPARG) is a transcription factor that plays a critical role with lipids in inflammation and insulin resistance [28]. PPARG can be activated through Creb and ERK1/2 and we see this association enriched in our networks. Most notably of PPARG actions, ligands to all PPAR family members suppress production of proinflammatory cytokines through suppression of NF-kB. It could be that these enriched networks are compensatory mechanisms to inflammation or damage in the liver. In that sense, then, arsenic is working upstream of these signals. Indeed, in the IPA network, PPARG appears to be downstream of the MAP kinase, ERK1/2.

Additionally, PPARG is regulated by Myc, which is seen to have enrichment in the transcription factor binding site analysis. A constant, if slight, elevation of PPARG may have wide-reaching effects on the liver because of PPARG transcription factor activity. Indeed, in our dataset we saw that the PPARG binding site was significantly enriched in significantly differentially expressed ($p \leq 0.05$ and $FC > |1.5|$) genes. PPARG is able to interact with some of our most highly-regulated genes, so it may be exerting substantial effects in our arsenic-exposed livers.

In addition to the molecular mechanisms above, it is beneficial not to forget the forest for the trees. There are many other risk factors associated with arsenic-induced atherosclerosis that may not be explained explicitly by our liver transcriptome analysis, but are supported by some of the pathways that are enriched. For example, we have seen arsenic affects normal mitochondrial functions. Mitochondrial dysfunction has been implicated in a number of diseases including neurodegenerative disorders [55], cardiovascular disease, and diabetes [5, 28]. Lane et al have also stated, "Mitochondrial deficiency can theoretically give rise to any symptom, in any organ, at any age" [57]. Indeed, it makes sense considering the important functions the mitochondria provide to eukaryotic cells. Hotamisligil reviewed literature on inflammation and metabolic disorders and showed these two can be triggered by excess of lipids and glucose and that the inflammatory response can trigger further metabolic dysfunction, creating a vicious cycle. Arsenic could be slightly stimulating both of these pathways, as we have seen. In that case, it would only require a small, chronic

amount before the system began to propagate the aberrant activity on itself. This is corroborated with the fact that ROS will activate JNK, discussed above, causing stress to the endoplasmic reticulum, critical in initiation of inflammation and insulin action in obesity and diabetes. Hotamisligil discusses therapeutic targets to manipulate metabolic and immune systems by targeting not single molecules but networks such as those governed by JNK or even to target organelles, specifically the mitochondrion. It seems likely that such a multifactorial disease process brought on by arsenic as we have seen here would have to be treated at a network level, rather than at single genes.

We worked to minimize limitations and confounding factors in this study by using a well-established mouse model [35] and treatment protocol [39]. However, transcript levels can vary greatly from one animal to another [46]. Normalization cannot fix all of these problems, and for that reason we used data that met significance criteria ($p \leq 0.01$ or 0.05) to be sure there was consistency across replicates and only real changes were taken into consideration. Additionally, there have been reports that microarrays dilute levels of gene expression changes when compared to gene-specific quantitative RT-PCR methods, although directionality seems to remain [44]. Tools such as IPA and DAVID are crucial in experiments that contain whole-transcriptome information. They allow researchers to use large gene sets to get insight into modeling, analyzing, and understanding complex biological systems that would be impossible by looking at single genes. The microRNA provided a unique challenge because the data we could collect for analysis was purely based on

computational predictions [58]. There are many databases based on different algorithms that provide predicted microRNA targets, however the two very popular options we chose only overlapped on a few genes. Once the genes are selected, what to do with them creates another problem because the large sets of seemingly unrelated genes have to have a meaning attached. The TFBS was also limited in that it does not report all transcription factors that are significantly enriched in a gene set of interest. TFBS looks at the nucleotide sequence present in the region 2000 base-pairs upstream and 200 base-pairs downstream of the transcription start site. Transcription factors that contribute to a gene may be tens of thousands of nucleotides up- or downstream of the transcription start site. However, the TFBS analysis does allow us to see another level of biological activity that a transcriptome analysis is not able to provide. That is, the transcription factors in our list mostly do not have any fold-changes associated. This is not surprising, since many transcription factors are activated through post-translational modification, such as phosphorylation (as in the case of targets of MAP kinases such as ERK, JNK, and p38). Using this technique, we were able to extend the amount of information we are able to obtain beyond that purely related to the transcript measurements.

This study may have implications on human health, as well. Although concentrations of arsenic in drinking water was much higher than current EPA and WHO standards (0.010 ppm), our lab has performed dose-response studies of arsenic in drinking water to atherosclerosis development and found it to hold true for exposures as low as 1 ppm to 4.9 ppm, suggesting the effect may scale

down to lower concentrations. In fact, it has already been proposed that there is no threshold for effect for noncancer diseases in arsenic exposure [1-2]. This may be because even small levels of arsenic causing small changes in gene expression over a long period of time may result in significant structural and functional changes in the liver. Arsenic may also be acting to "prime" the liver for an exaggerated reaction to a second atherosclerotic event, such as a high fat diet or an infection.

Future directions

The results in this study give a broad-view of changes taking place on the transcript-level. It is important to keep in mind that transcript levels do not have direct effects on cellular actions, protein drive cellular activity. This should be taken into consideration when proposing future directions, as these results only hint at mechanisms taking place in a cell.

The next steps for this study should be to verify inferences that have been made. The mitochondrion is implicated in dysfunction as well as dysregulation of oxidative phosophorylation. Experiments focusing on these elements would clarify whether or not these are occurring. Reactive oxygen species are likely culprits of problems to the mitochondrion and should also be checked. We mentioned the effects from arsenic that we have seen may only be small changes and may actually require another stimulus in order to trigger noticeable effects. In this way, experiments could be done for mock infections, possibly using lipopolysaccharides (LPS) and then metrics corresponding to mitochondrial

dysfunction should be analyzed. In any case, direct reactions of the mitochondrion organelle should be a good place to start before concentrating on more specific areas, such as signaling pathways.

Networks composed of significantly expressed genes contained hubs on MAP kinase signaling complexes such as ERK, JNK, and p38. These are stress-response genes and are able to regulate gene expression, cell differentiation, and apoptosis, and their action may be causing aberrant gene expression in the liver. These genes are activated through phosphorylation pathways, which would explain why we did not see their transcript level significantly differentially expressed. Since we are not able to see direct changes, western blots may be performed to check levels of phosphorylated proteins to assess activation. Additionally, we have seen increases in gene expression in genes such as PPARG, which are regulated by these MAP kinases. Transcript levels should be confirmed by protein levels in order to validate that the changes seen directly impact enzymes driving cellular actions.

Conclusion and summary

In summary, we have characterized gene expression changes occurring in the liver of ApoE-/- mice with arsenic exposure in drinking water. These results support the hypothesis that chronic arsenic exposure is able to disrupt liver homeostasis and cause aberrant gene changes. Among the changes, metabolic functions, especially those in dealing with the electron transport chain, appear to be disrupted, and the evidence points towards an induction of oxidative stress

that acts as both cause and effect in the dysregulation. This study suggests a few mechanisms that are disrupting the liver in a way to promote atherosclerosis in the vascular system and suggests approaches and mechanisms for further research.

To date, there are still millions of people consuming water that exceeds 50 µg/L arsenic and possibly twice as many consuming 10 µg/L arsenic in many parts of the world [51]. The association between chronic arsenic exposure and atherosclerosis has far-reaching implications in research and public health. Because atherosclerosis is a major public health concern, the consequences of a causal association with arsenic would affect large numbers of people worldwide.

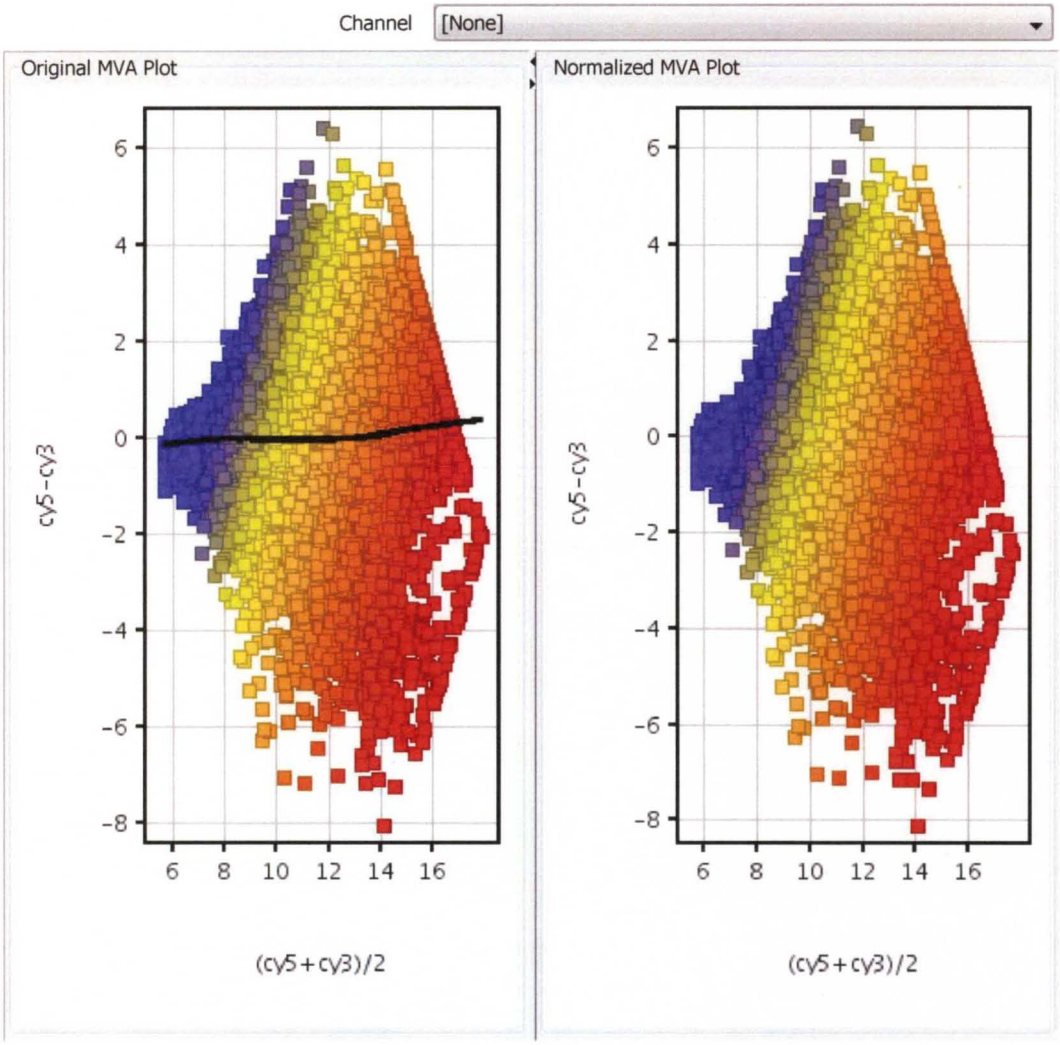


Figure 1: MvA plot on gene expression. Normalized MA plot of postnatal arsenic exposure shows no change after normalization, indicating the signal value data was previously normalized. (Left) is the original MvA plot. (Right) is the GeneSpring-normalized MvA plot. This plot shows the difference versus the average of probe measurements between two samples and assesses the relation between the Cy5 and Cy3 channels of each hybridization. Each small square represents an individual probe.

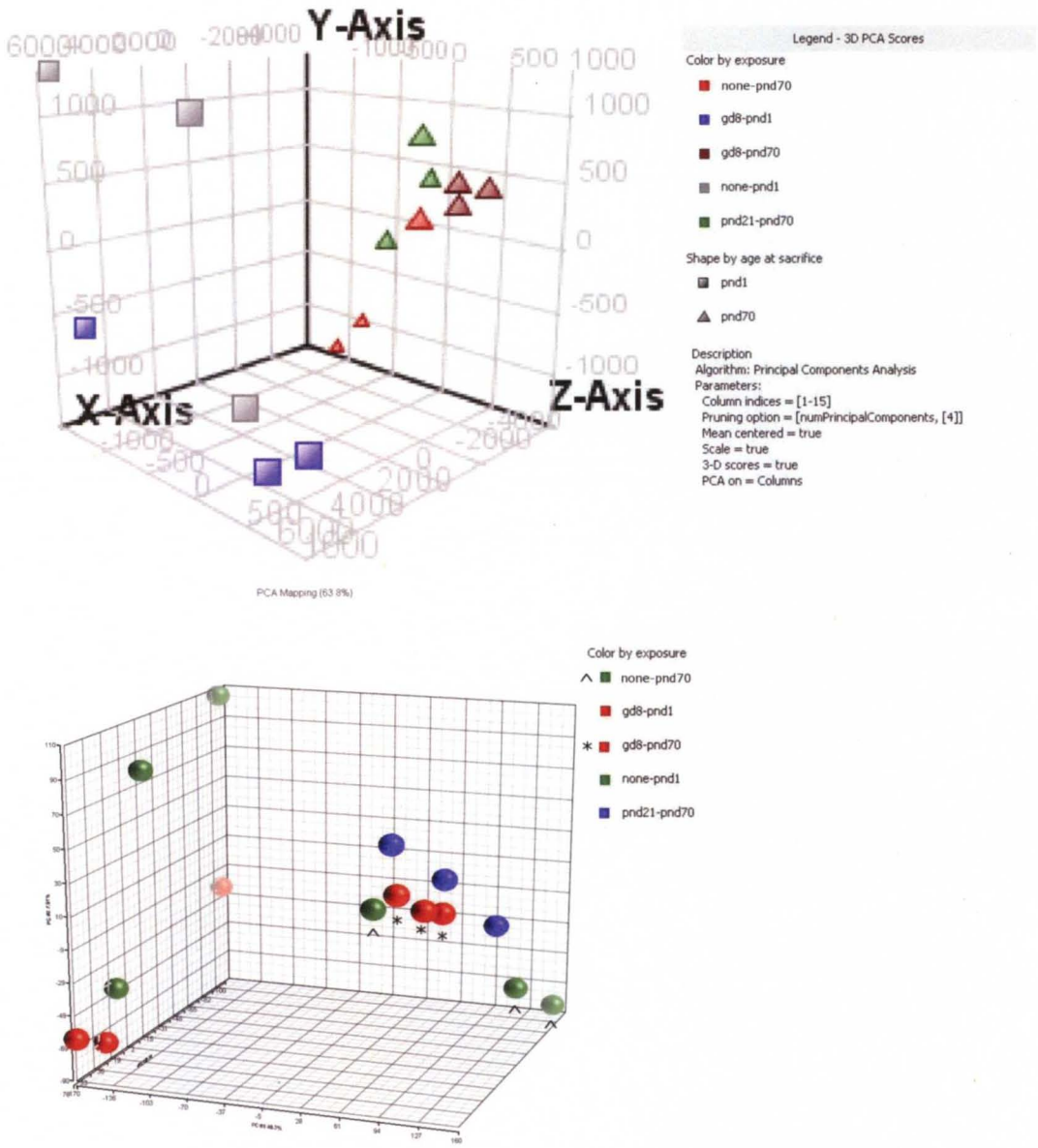
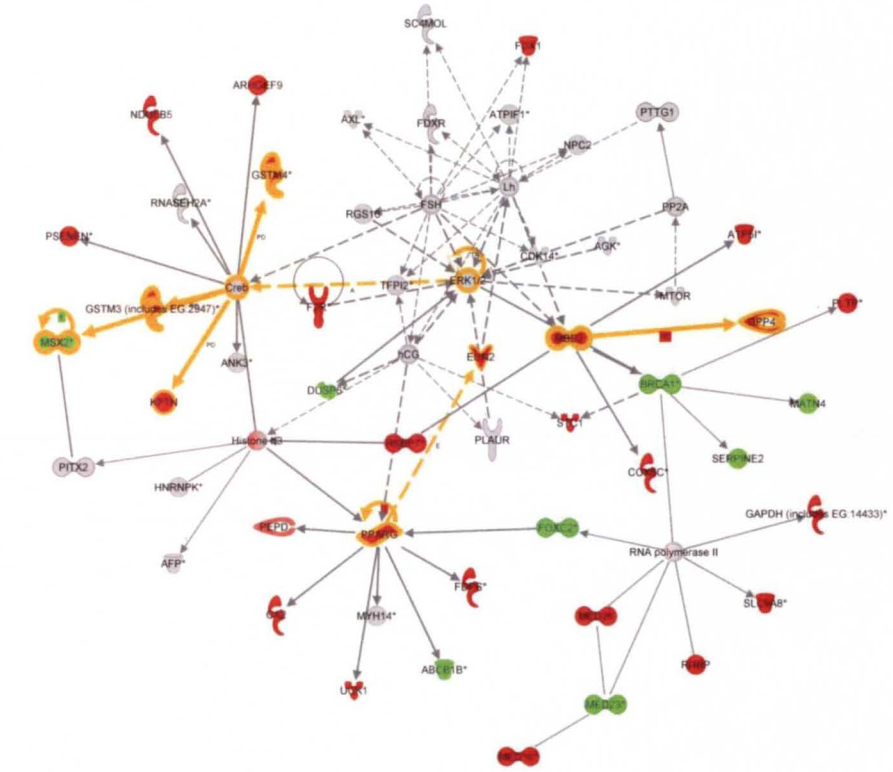


Figure 2: PCA analysis on gene expression. Principal Components Analyses (PCA) comparison for GeneSpring (top) and Partek (bottom) indicates similar grouping patterns among conditions. These are 3D PCA image for both GeneSpring (top) and Partek (bottom). Depicted above are two studies, a postnatal study (exposures: none-pnd70 and pnd21-pnd70) and a prenatal study

(exposures: none-pnd1, gd8-pnd1, and gd8-pnd70) that were performed at the same time. Similar colors indicate replicate samples (3 per color).

Path Designer Networks g1,p1 Merged 4



‹Path Designer Networks g2,p2 Merged 5

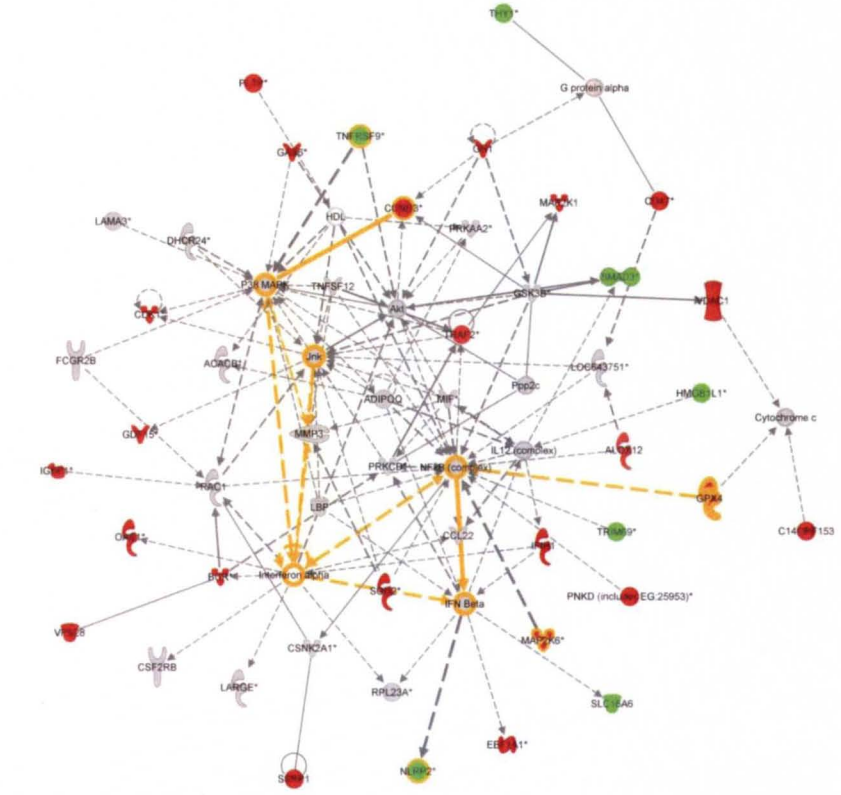


Figure 3: IPA networks on gene expression. Merged IPA networks comparing GeneSpring and Partek data indicate similarities between networks. (Top) The top functions of this network are cellular development, growth, proliferation, and movement, connective tissue development and function, and cell cycle. (Bottom) The top functions of this network are hematological system development, function, and disease, cellular movement, cellular movement, immune cell trafficking, cell death, and post-translational modification. Gene symbols are located in the center of each box; green shapes indicate down-regulated genes, and red shapes indicate up-regulated genes, compared with control. Gray shapes indicate genes that IPA inserted to obtain the most connections within the network, but are not regulated significantly regulated by the experiment. Highlighted shapes indicate those that were common between the merged GeneSpring and Partek networks in the comparison.

ID Analysis	Molecules in Network	Score	Focus Molec.	Top Functions
1-G	ABCB1B, ARHGEF9, ATP5I, BRCA1, CA2, COX6C, Creb, DPP4, DUSP6, EDN2, ERK1/2, F2R, FDPS, FOXC2, GAPDH (includes EG:14433), GSTM4, GSTM3 (includes EG:2947), KPTN, MATN4, MBD2, MED16, MED23, MED26, MSX2, NDUFB5, PEPD, PLTP, PPARG, PSENEN, RBBP7, RDBP, RNA polymerase II, SERPINE2, SLC9A8, UCK1	37	32	Cellular Development, Connective Tissue Development and Function, Embryonic Development
1-P	AFP, AGK, ANK3, ATPIF1, AXL, CDK14, Creb, DPP4, EDN2, ERK1/2, FDX1, FDXR, FSH, GSTM4, GSTM3 (includes EG:2947), hCG, Histone h3, HNRNPK, KPTN, Lh, MBD2, MSX2, MTOR, MYH14, NPC2, PITX2, PLAUR, PP2A, PPARG, PTTG1, RGS16, RNASEH2A, SC4MOL, STC1, TFPI2	31	28	Cellular Growth and Proliferation, Cellular Movement, Cell Cycle
2-P	ACACB, ADIPOQ, Akt, CCL22, CCND3, CSF2RB, CSNK2A1, DHCR24, FCGR2B, GPX4, GSK3B, HDL, IFN Beta, IGBP1, Interferon alpha, Jnk, LAMA3, LARGE, LBP, LOC643751, MAP2K6, MIF, MMP3, NFkB (complex), NLRP2, P38 MAPK, PLTP, Ppp2c, PRKAA2, PRKCD, RAC1, RPL23A, SSRP1, TNFRSF9, TNFSF12	29	27	Cellular Movement, Hematological System Development and Function, Immune Cell Trafficking
2-G	ALOX12, BCR, C14ORF153, CCND3, CD47, CDK1, Cytochrome c, EEF1A1, G protein alpha, GAS6, GDF15, GH1, GPX4, HMGB1L1, IFIH1, IFN Beta, IL12 (complex), Interferon alpha, Jnk, MAP2K1, MAP2K6, NFkB (complex), NLRP2, OAS1, P38 MAPK, PNKD (includes EG:25953), SLC16A6, SMAD3, SOD2, THY1, TNFRSF9, TRAF2, TRIM69, VDAC1, VPS28	27	27	Cell Death, Hematological Disease, Post-Translational Modification
3-G	ALPL, ATP9A, ATPIF1, BAG1, BMYC, DHPS, E2f, EXOSC8, FDX1, FSH, hCG, Histone h3, HRAS, Lh, MAP2K1, MBD2, NOL3, NPC2, PGRMC1, PRKAR1A, RAB14, RPA3, SGK1, SKA2, SMAD6, SMC2, SNRPC, SSRP1, STC1, SUPT16H, TK1, TMEM126A, TRIB1, TYMS, UXT	21	23	Embryonic Development, Tissue Development, Cell Death
3-P	AMPK, AP2M1, C14ORF166, CYP2A6, CYP7A1, DLAT, EIF3F, GFER, GLI1, HNF4A, IRS2, ME1, NCOA1, NDUFS1, NDUFV2, NONO, NR0B2, NR1H4, NR1I3, NR5A2, PCK1, PPARGC1A, RARA, RBM3, RPS5, RPS19, RPSA, RXRA, SAT1, SFPQ, SULT1C3, SYNCRIP, THRB, TUFM, UCP3	15	18	Gene Expression, Lipid Metabolism, Small Molecule Biochemistry
4-G	AIFM1, ATOX1, BAK1, BAX, BMP6, CHMP1A, DMPK, DNM1L, EIF4E, GRB2, LAT, LAX1, LETMD1, LSM1, MAP7, MCM6, MFN1, MFN2, MPV17, NOP10, PA2G4, PARP1, PCNA, POLD1, POLDIP2, PPIA (includes EG:268373), RPL11, RPL30, RPL18A, SOS2, STOML2, TALDO1, TNFSF10, YY1, ZNF143	19	22	Cellular Compromise, Cellular Assembly and Organization, Cell Morphology
4-P	ABCA1, ABCC3, ADIPOQ, AEBP1, AKT2, ALB, BBC3, C5, CCL27, CD36, CETP, CFD, CR1, CX3CL1, Cytochrome c oxidase, DIABLO, FOXC2, HEXA, HEXB, HMGB1L1, IL27, LAMB3, LPL, MFHAS1, MUT, NR1I2, PPARG, PRODH, PSMB10,	13	17	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry

	PTGES, PTPN12, SCNN1G, SLCO1A2, TNF, UCK1			
5-G	APP, ARC, ATP5A1, ATP5C1, ATP5O, CAMK2A, CSNK1G2, DLG4, DSTN, FMR1, GLUL, GRIA1, GRIA2, GRIK5, HTT, INS, LEP, NRXN1, PHB2, PHLPP1, PRDX1, PRDX2, PRDX3, SDHA, SIK1, SLC25A6, SRXN1, SUCLA2, SYT1, VDAC2	14	17	Cell-To-Cell Signaling and Interaction, Nervous System Development and Function, Cell Morphology
5-P	ACTG1, Actin, ADH5 (includes EG:128), APP, AVPI1, BLVRA, C5AR1, CCL5, CD163, CHN2, COPS5, CXCL3, CYBB, DUSP10, FAM38A, FEZ1, FMR1, IFNAR1, IL13, ITM2B, Lamin b, LDL, MIF, MSN, NOS2, PITRM1, PLEC, RAC1, RIPK2, SNCA, SORT1, SPINT2, TNFRSF1A, TPMT, VPS35	13	17	Inflammatory Response, Cell-To-Cell Signaling and Interaction, Hematological System Development and Function
6-G	APH1A (includes EG:51107), BIRC2, CA3, CELA1, CER1, CSNK2A1, CTNNB1, DHX15, DPAGT1, EGF, GNA11, GPI, ICAM2, IFNG, IGBP1, IRAK4, ITGB2, KLK2, NFkB (complex), PCSK6, PI3, PIK3CG, POLR2F, PRPF8, PSEN1, PTEN, RAC1, RPS6KB1, RPS6KB2, SERPINA3, SMNDC1, SOD2, TLR5, TNFRSF1B, TXN2	12	17	Free Radical Scavenging, Cellular Movement, Genetic Disorder
6-P	ABCB1B, BRCA1, CAV1, CCNA2, CCNB1, CCND3, CDK1, CDKN1A, CEBPA, CLTC, DHFR, DNTT, ERCC1, GDF15, GMNN, GSTM1, GSTM5, LETMD1, MCM6, MCM7, NANOG, NR1I2, PHC1, POU5F1, PPARG, RB1CC1, SNX9 (includes EG:51429), STUB1, STX8, TOM1L1, TOPBP1, TP53, TSG101, WWTR1, XRCC5	12	16	Cell Cycle, Cancer, Reproductive System Disease
7-G	AACS, AKT1, ANXA11, ATP5F1, BCL10, CARTPT, CCL3L3, CCND3, CLEC7A, CRYL1, CSNK1E, CYP7A1, DLAT, EGF, EIF6, HTR4, IGFBP2, IKBKG, IL2, IMPDH2, LEP, ME1, MEP1A, MIR122A, N-cor, NDUFS1, NDUFV2, NFATC1, PER1, PER2, PPARGC1A, RPS6KB1, SLC27A5, THRB, THRSP	12	17	Genetic Disorder, Metabolic Disease, Behavior
7-P	ACTL6A, AK3, AKT2, ANXA6, ARID1A, BAF110, CA2, ETS2, FCGR2B, GCAT, GLIPR2, IFI30, IL4, IL5, KLHDC2, MMP1 (includes EG:4312), PDK1, PGM1, PLIN2, PPARG, RHAG, RHCE, SEMA4A, SLC4A1, SMARCA4, SMARCB1, SMARCC1, SMARCC2 (includes EG:6601), SMARCE1, ST7, TADA2B, TEP1, TERT, THY1, TMEM97	12	16	Gene Expression, Cellular Assembly and Organization, Cellular Compromise

Table 1: Top IPA networks for GeneSpring and Partek. Table shows network ID, genes in network from list, number of genes, network score, and top functions of network. Network ID and analysis indicated for GeneSpring (G) and Partek (P).

Ingenuity Canonical Pathways	Analysis	Molecules
Oxidative Phosphorylation	GeneSpring	NDUFA4, SDHB, UQCR11, TCIRG1, NDUFB5, COX5B, UQCRB, ATP5G2, NDUFB10, NDUFB9, NDUFA5, NDUFS1, ATP6V1F, NDUFC2, ATP5G1, ATP5H (includes EG:10476), ATP5J2, ATP5F1, ATP5I, COX7C (includes EG:1350), COX4I1, NDUFS4, ETNK2, NDUFA8, SDHA, ATP5J, NDUFV1, NDUFC1, COX7A2, COX6B1, ATP5O, NDUFS7, ATP5A1, COX6C, ATP6V1A, NDUFS3, ATP5C1, UQCR10, NDUFV2, NDUFS8, PPA2, UQCRC2, NDUFA7
Oxidative Phosphorylation	Partek	SDHA, ATP5C1, NDUFS1, NDUFV2, NDUFS8, PPA2, TCIRG1, UQCRC2, ATP6V1G1, NDUFS3
Mitochondrial Dysfunction	GeneSpring	NDUFA4, SDHB, NDUFB5, COX5B, PSENEN, UQCRB, NDUFB10, NDUFB9, NDUFA5, NDUFS1, SOD2, GPX4, COX7C (includes EG:1350), COX4I1, NDUFS4, AIFM1, NDUFA8, SDHA, ATP5J, NDUFV1, COX7A2, COX6B1, NDUFS7, COX6C, ATP5A1, NDUFS3, APH1A (includes EG:51107), ATP5C1, PRDX3, NDUFS8, NDUFV2, TXN2, UQCRC2, TXNRD2, NDUFA7
Mitochondrial Dysfunction	Partek	SDHA, NDUFAF1, NDUFS3, GPX7, APH1A (includes EG:51107), ATP5C1, NDUFS1, PRDX3, NDUFV2, NDUFS8, UQCRC2, MAPK10, GPX4, TXNRD2, AIFM1
Ubiquinone Biosynthesis	GeneSpring	NDUFA4, NDUFC1, NDUFV1, NDUFS7, NDUFB5, NDUFS3, NDUFB10, NDUFS1, NDUFA5, NDUFB9, UFSP2, NDUFV2, NDUFS8, NDUFC2, NDUFA7, NDUFS4, NDUFA8
Ubiquinone Biosynthesis	Partek	NDUFS1, UFSP2, NDUFV2, NDUFS8, AS3MT, NDUFS3
Citrate Cycle	GeneSpring	SDHA, SUCLA2, SDHB, SUCLG2, SUCLG1, DLD, IDH2, MDH1, MDH2, CLYBL, IDH3B
Citrate Cycle	Partek	SDHA, IDH2, MDH2, IDH3B
Protein Ubiquitination Pathway	GeneSpring	PSMA3, PSMA7, UBE2D2, USP48, TCEB2, STUB1, USP47, USP16, UBE2E3, BRCA1, PSMA2, ANAPC11, UCHL3, PSMB5, PSMD13, UBE2R2 (includes EG:54926), PSME2, USP30, USP1, PSMA1, UBE2L6, USP33, PSMD8, USP31, PSMB7, PSMB2, PSMD12, PSMA4
Protein Ubiquitination Pathway	Partek	USP31, PSMA6, PSMA3, PSMB5, PSMB10, PSMA7, STUB1, UBE2R2 (includes EG:54926), USP11, UBE2D2, UBR1, USP48
Butanoate Metabolism	GeneSpring	SDHA, ALDH4A1, SDHB, SUCLG2, AACS, DBT, ALDH3B1, MYO5B, DCXR, HMGCL, ALDH5A1, ALDH9A1
Butanoate Metabolism	Partek	SDHA, ALDH4A1, BDH1, AACS, ALDH3B1, MYO5B, PDHB, ALDH5A1, PRDX6, ACADS
Propanoate Metabolism	GeneSpring	ALDH4A1, HIBCH, SUCLA2, SUCLG2, SUCLG1, ALDH3B1, MUT, ACAD10, UEVLD, ALDH5A1, ALDH9A1
Propanoate Metabolism	Partek	ALDH4A1, HIBCH, ACSM5, ACACB, DHCR24, ALDH3B1, MUT, ACAD10, UEVLD, ALDH5A1, ACADS

Glycine, Serine and Threonine Metabolism	GeneSpring	SARDH, GAMT, DLD, GARS, DBT, ALAS1, PDPR, SMOX, PISD, ABP1, PLCD4, ETNK2
Inositol Metabolism	GeneSpring	COX6B1, AIFM3, OXNAD1, COX6C, BLVRB, TPI1, TM7SF2, NDUFB10, CYB5RL, RP3-402G11.5, IDH3B, NDUFA7, NDUFA8
Glutamate Metabolism	GeneSpring	ALDH4A1, SUCLG2, NADSYN1, GLUL, MYO5B, PPAT, ALDH5A1
Glutamate Metabolism	Partek	ALDH4A1, NAGK, NADSYN1, GPT, MYO5B, GPT2, ALDH5A1
Valine, Leucine and Isoleucine Degradation	GeneSpring	ALDH4A1, HIBCH, DBT, ALDH3B1, AOX1, MUT, HMGCL, ACAD10, ALDH5A1, ALDH9A1
Valine, Leucine and Isoleucine Degradation	Partek	ALDH4A1, HIBCH, MCCC1, ALDH3B1, MUT, ACAD10, ALDH5A1, ACADS
LPS/IL-1 Mediated Inhibition of RXR Function	GeneSpring	ALDH4A1, GSTM3 (includes EG:2947), GSTA5, ALDH9A1, IRAK1, SLC27A5, TRAF2, GSTA4, GSTM4, ALAS1, ACSL4, SMOX, ALDH3B1, PLTP, FABP3, SULT1C3, NDST1, SULT1B1, ALDH5A1
LPS/IL-1 Mediated Inhibition of RXR Function	Partek	ALDH4A1, SULT1C2, GSTM1, GSTM3 (includes EG:2947), GSTA5, SLCO1A2, GSTA4, CYP7A1, GSTM4, ALDH3B1, FABP7, PLTP, SULT1C3, LBP, ABCC3, NDST1, ABCC4, CYP2A6, SULT1B1, ALDH5A1
Pyruvate Metabolism	GeneSpring	ALDH4A1, LDHD, DLAT, DLD, ALDH3B1, ME1, MDH1, MDH2, UEVLD, ALDH5A1, ALDH9A1
Pyruvate Metabolism	Partek	ALDH4A1, AKR1A1, ACSM5, ACACB, LDHD, DLAT, ALDH3B1, ME1, MDH2, UEVLD, PDHB, ALDH5A1
Glycolysis/Gluconeogenesis	GeneSpring	ALDH4A1, PGM1, GAPDH (includes EG:14433), TPI1, ALDH9A1, ADH5 (includes EG:128), GPI, ENO1, DLAT, GALK1, DLD, ALDH3B1, UEVLD, ALDH5A1
Glycolysis/Gluconeogenesis	Partek	ADH5 (includes EG:128), ALDH4A1, AKR1A1, ACSM5, DLAT, PGM1, ALDH3B1, TPI1, UEVLD, PDHB, ALDH5A1

Table 2: Top IPA canonical pathways. Table shows Ingenuity Canonical Pathways, the analysis that produced it, and the molecules that enrich the pathway.

A) GeneSpring

Functional Cluster	Enrichment Score	Category	Term	%	p-value
1	44.4	GOTERM_CC_FAT	GO:0005739~mitochondrion	21.8	2.39E-57
2	20.8	KEGG_PATHWAY	mmu00190:Oxidative phosphorylation	4.2	7.02E-21
2	20.8	GOTERM_BP_FAT	GO:0022900~electron transport chain	3.0	1.16E-14
3	9.9	GOTERM_BP_FAT	GO:0006412~translation	5.5	1.85E-16
3	9.9	GOTERM_CC_FAT	GO:0005840~ribosome	4.0	1.66E-13
4	9.0	GOTERM_BP_FAT	GO:0051186~cofactor metabolic process	3.8	2.25E-14

B) Partek

Functional Cluster	Enrichment Score	Category	Term	%	p-value
1	13.6	GOTERM_CC_FAT	GO:0005739~mitochondrion	15.0	1.42E-20
2	11.1	GOTERM_BP_FAT	GO:0006412~translation	5.2	9.03E-14
2	11.1	GOTERM_CC_FAT	GO:0005840~ribosome	4.0	2.81E-13
4	2.3	INTERPRO	IPR012335:Thioredoxin fold	1.7	1.07E-04

Table 3: DAVID analysis on gene expression. Tables show enriched terms using lists from (A) GeneSpring and (B) Partek where $p \leq 0.01$ with any FC indicate similar terms are enriched using DAVID. Listed are one (or more) representative terms for a functional cluster, along with the category that term came from, the percent (%) enrichment that is the number of genes from the input list vs. all the genes in the term, and p-value of enrichment.

Gene Symbol	p-value	fold-change	Description
AIFM1	8.53E-04	1.403	apoptosis-inducing factor, mitochondrion-associated, 1
ATP5A1	3.74E-03	1.214	ATP synthase, H+ transporting, mitochondrial
ATP5A1	3.74E-03	1.214	
ATP5C1	1.68E-04	1.349	
ATP5C1	1.68E-04	1.349	
ATP5F1	1.97E-04	1.183	
ATP5G1	1.85E-03	1.138	
ATP5G2	2.22E-03	1.175	
ATP5H	2.33E-04	1.199	
ATP5I	1.83E-03	1.152	
ATP5J	2.12E-03	1.103	
ATP5J	2.12E-03	1.103	
ATP5J2	6.82E-03	1.163	
ATP5O	1.40E-04	1.222	
ATP6V1A	5.31E-03	1.152	
COX4I1	2.13E-03	1.23	cytochrome c oxidase subunit
COX5B	6.77E-04	1.225	
COX5B	6.77E-04	1.225	
COX6B1	1.47E-03	1.236	
COX6B1	1.47E-03	1.236	
COX6C	5.21E-03	1.162	
COX6C	5.21E-03	1.162	
COX7A2	7.75E-03	1.192	
COX7A2	7.75E-03	1.192	
COX7C	1.43E-03	1.238	
COX7C	1.43E-03	1.238	
GPX4	2.15E-04	1.396	glutathione peroxidase 4 (phospholipid hydroperoxidase)
NDUFA4	8.20E-03	1.266	NADH dehydrogenase (ubiquinone)
NDUFA4	8.20E-03	1.266	
NDUFA5	3.32E-03	1.248	
NDUFA5	3.32E-03	1.248	
NDUFA7	7.16E-03	1.137	
NDUFA7	7.16E-03	1.137	
NDUFA8	5.52E-03	1.22	
NDUFA8	5.52E-03	1.22	
NDUFB10	1.09E-03	1.266	
NDUFB10	1.09E-03	1.266	
NDUFB5	1.92E-03	1.224	

NDUFB5	1.92E-03	1.224	
NDUFB9	3.22E-03	1.127	
NDUFB9	3.22E-03	1.127	
NDUFC1	3.08E-03	1.127	
NDUFC2	2.48E-03	1.373	
NDUFS1	7.45E-04	1.359	
NDUFS1	7.45E-04	1.359	
NDUFS3	6.59E-04	1.256	
NDUFS3	6.59E-04	1.256	
NDUFS4	6.04E-03	1.238	
NDUFS4	6.04E-03	1.238	
NDUFS7	4.55E-03	1.286	
NDUFS7	4.55E-03	1.286	
NDUFS8	2.76E-04	1.332	
NDUFS8	2.76E-04	1.332	
NDUFV1	2.13E-03	1.214	
NDUFV1	2.13E-03	1.214	
NDUFV2	9.24E-03	1.29	
NDUFV2	9.24E-03	1.29	
PPA2	8.23E-04	1.26	pyrophosphatase (inorganic) 2
PRDX3	1.71E-03	1.381	peroxiredoxin 3
SDHA	1.16E-03	1.269	succinate dehydrogenase complex, subunit
SDHA	1.16E-03	1.269	
SDHB	2.43E-04	1.239	
SDHB	2.43E-04	1.239	
SOD2	1.75E-03	1.169	superoxide dismutase 2, mitochondrial
TCIRG1	9.47E-04	1.303	T-cell, immune regulator 1, ATPase, H+ transporting
TXN2	6.60E-03	1.247	thioredoxin 2
TXNRD2	2.51E-03	1.368	thioredoxin reductase 2
UQCR10	1.16E-03	1.185	ubiquinol-cytochrome c reductase
UQCR11	4.17E-03	1.176	
UQCRB	2.95E-03	1.175	
UQCRB	2.95E-03	1.175	
UQCRC2	1.76E-03	1.274	
UQCRC2	1.76E-03	1.274	

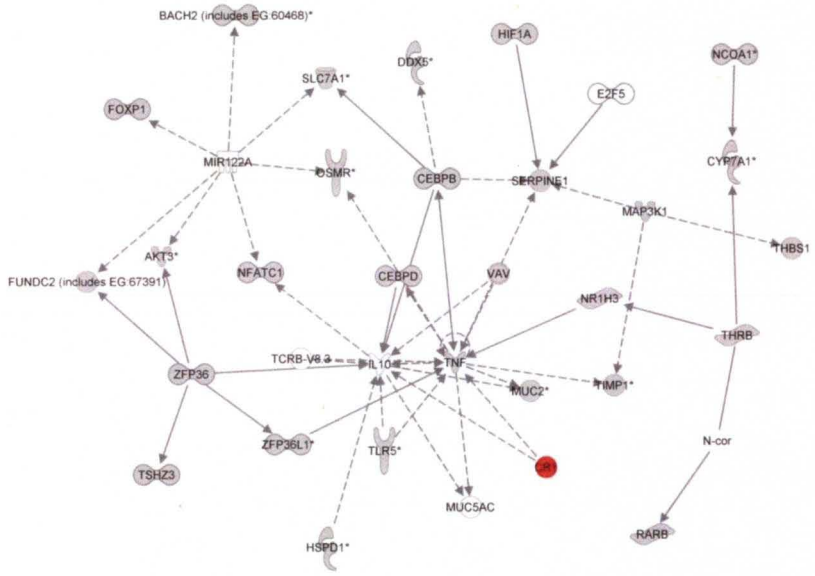
Table 4: DAVID and IPA mitochondrion analysis. Table shows genes that are common between DAVID GO term mitochondrion and IPA pathways

mitochondrial dysfunction and oxidative phosphorylation with transcript p-value and fold-change. Of the 78 genes between the 2 IPA pathways, 72 genes appear in the DAVID GO term mitochondrion.

	No. Unique Genes			
MicroRNA	**TargetScan**	**miRBase**	**Intersection**	**Notes**
mmu-miR-140	158	1000	43	140/876-3p
mmu-miR-18	125	2051	71	18ab
mmu-miR-206	478	861	61	1/206
mmu-miR-337	0	0	0	
Totals	761	3912	175	

Table 5: MicroRNA database comparison. Table shows number of predicted microRNA gene targets found in TargetScan and miRBase. MicroRNA was input into each database and predicted gene targets were collected. MiRBase reported many more predicted gene targets than TargetScan. Notes indicate TargetScan microRNA that have the same predicted targets; i.e. entering miR-1 or miR-206 in TargetScan will result in the same gene list.

Path Designer Network 1



Path Designer Network 2

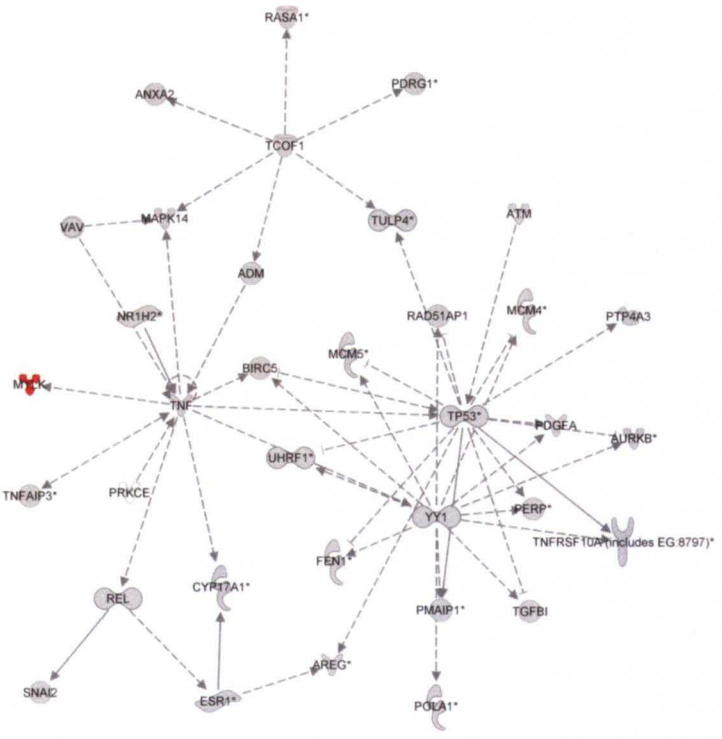
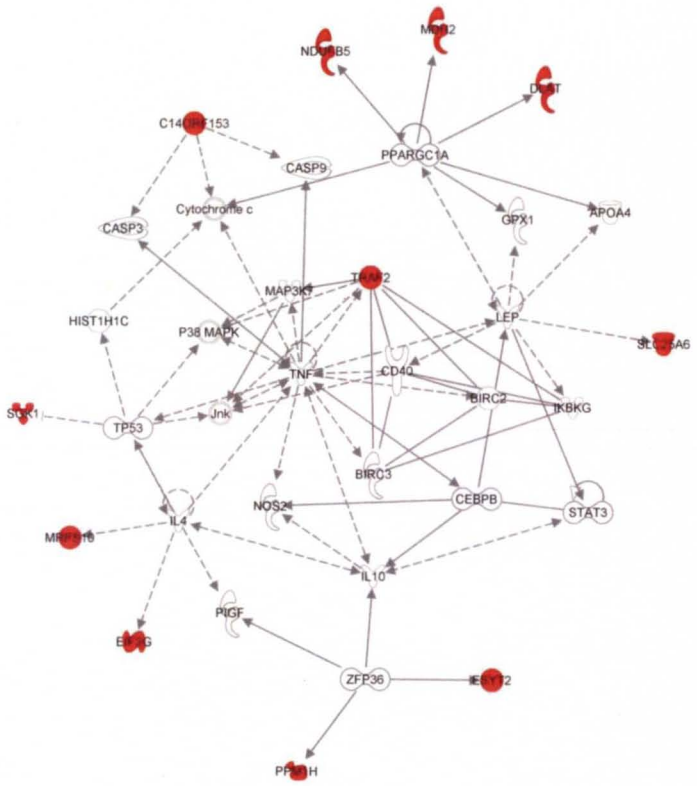


Figure 4: MicroRNA database comparison network. The top 2 networks from the IPA analysis of the intersection of miRBase and TargetScan for microRNA predicted target genes indicate enrichment of different regulatory processes. The top functions of network 1 (top) are tissue morphology, cellular growth and proliferation, and gene expression and network 2 (bottom) are cancer, cell morphology, and DNA replication, recombination, and repair. Gene symbols are located in the center of each box. Colors represent those genes significantly expressed and upregulated (red) in the GeneSpring dataset. A legend can be found in Figure 3.

genespring-microrna_mrna-2009-12-01-n1



genespring-microrna_mna-2009-12-01-n2

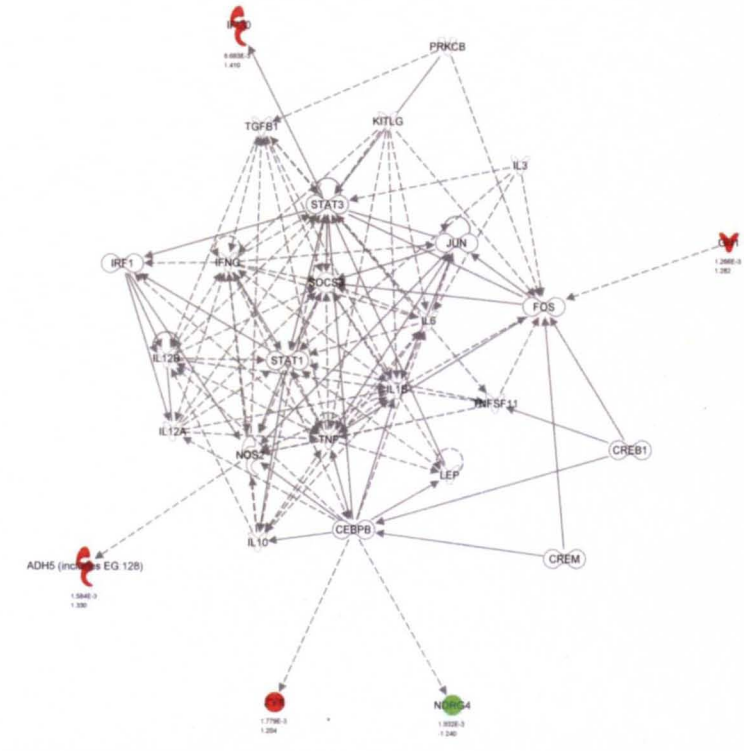


Figure 5: MicroRNA and transcript intersection comparison network. The top 2 networks from the IPA analysis of the intersection of miRBase predicted gene targets and transcript genes where $p \leq 0.01$ with any fold-change indicates networks similar in function to the transcript gene set. The top functions associated with each of the networks follows: cell death, hematological system development and function, and inflammatory disease (top) and cellular development, cell death, and cell signaling (bottom). A legend can be found in Figure 3.

Functional Cluster	Category	Term	% Enrichment	*p*-value
1	GOTERM_BP_ALL	GO:0065007~ biological regulation	44.3	1.60E-08
1	GOTERM_MF_ALL	GO:0030528~ transcription regulator activity	18.2	6.39E-07
1	GOTERM_MF_ALL	GO:0003677~ DNA binding	23.9	8.09E-07
2	GOTERM_BP_ALL	GO:0009790~ embryonic development	10.8	4.42E-07
3	GOTERM_BP_ALL	GO:0006366~ transcription from RNA polymerase II promoter	9.0	4.41E-05
4	GOTERM_BP_ALL	GO:0048514~ blood vessel morphogenesis	5.7	1.14E-04
4	GOTERM_BP_ALL	GO:0001525~ angiogenesis	4.6	6.94E-04
5	SP_PIR_KEYWORDS	zinc-finger	15.3	9.19E-05
5	GOTERM_MF_ALL	GO:0046872~ metal ion binding	29.6	2.28E-03

Table 6: MicroRNA and transcript intersection DAVID analysis. Listed in this table are enriched terms for the microRNA predicted gene targets, implicating many transcriptional regulatory networks. Listed are one (or more) representative terms for a functional cluster, along with the category that term came from, the percent (%) enrichment that is the number of genes from the input list vs. all the genes in the term, and *p*-value of enrichment.

Enriched with	No. genes	Enrichment factor	TF name	Fold-Change*
M00292[Freac-4]	23	1.779	Foxd1 forkhead box D1	NC
M00997[DEC]	19	1.974	Bhlhe40 basic helix-loop-helix family, member e40	1.5 fc
M01066[BLIMP1]	23	1.453	Prdm1 PR domain containing 1, with ZNF domain	(1.2)-fc
M00134[HNF-4]	18	1.347	Pparg peroxisome proliferator activated receptor gamma	1.6 fc
M00026[RSRFC4]	21	1.338	Same as Mef2a ??	NC
M00146[HSF1]	12	1.527	Heat shock factor protein 1	NC
M00626[RFX1_(EF-C)]	23	1.797	Rfx1 regulatory factor X, 1 (influences HLA class II expression)	NC
M00055[N-Myc]	20	1.262	Mycn v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	NC
M00403[aMEF-2]	17	1.348	Mef2a myocyte enhancer factor 2A	NC
M00186[SRF]	8	1.391	Srf serum response factor	NC
M00191[ER]	17	1.441	Esr1 estrogen receptor 1 (alpha)	1.3 fc
M01004[Helios_A]	20	1.741	Ikzf2 IKAROS family zinc finger 2	NC
M00279[MIF-1]	21	1.502	Gm4924 predicted gene 4924	
M00114[Tax/CREB]	21	1.673	tax p40 [Human T-lymphotropic virus 1] and Creb1 cAMP responsive element binding protein 1	NC
M00196[Sp1]	51	1.236	Sp1 trans-acting transcription factor 1	NC
M00447[AR]	17	1.642	Ar androgen receptor	NC
M00963[T3R]	20	1.636	Rxra retinoid X receptor alpha	NC
M00005[AP-4]	18	1.381	Tcfap4 transcription factor AP4	NC
M00322[c-Myc:Max]	25	1.343	Myc myelocytomatosis oncogene	1.6 fc myc

*Fold-change in arsenic exposed vs. not exposed if FC ≥ |1.1|, otherwise no change (NC). Parentheses represents negative value.

Table 7: TFBS analysis on gene expression. Listing of significantly enriched Transcription Factor Binding Sites (TFBS), from significantly differentially expressed genes in GeneSpring dataset where $p \geq 0.05$ and FC≥|1.5|. All TFBS where p≤0.05 are presented along with the transcription factor that binds the site, number of genes involved in our list (out of 94 genes), and enrichment factor, the fold-change (FC) of the TF was provided if it was available in our original gene

data if the FC was more than 1.1 fold. TFs were resolved from the TFBS ID using MAPPER (http://snpper.chip.org/mapper/mapper-main).

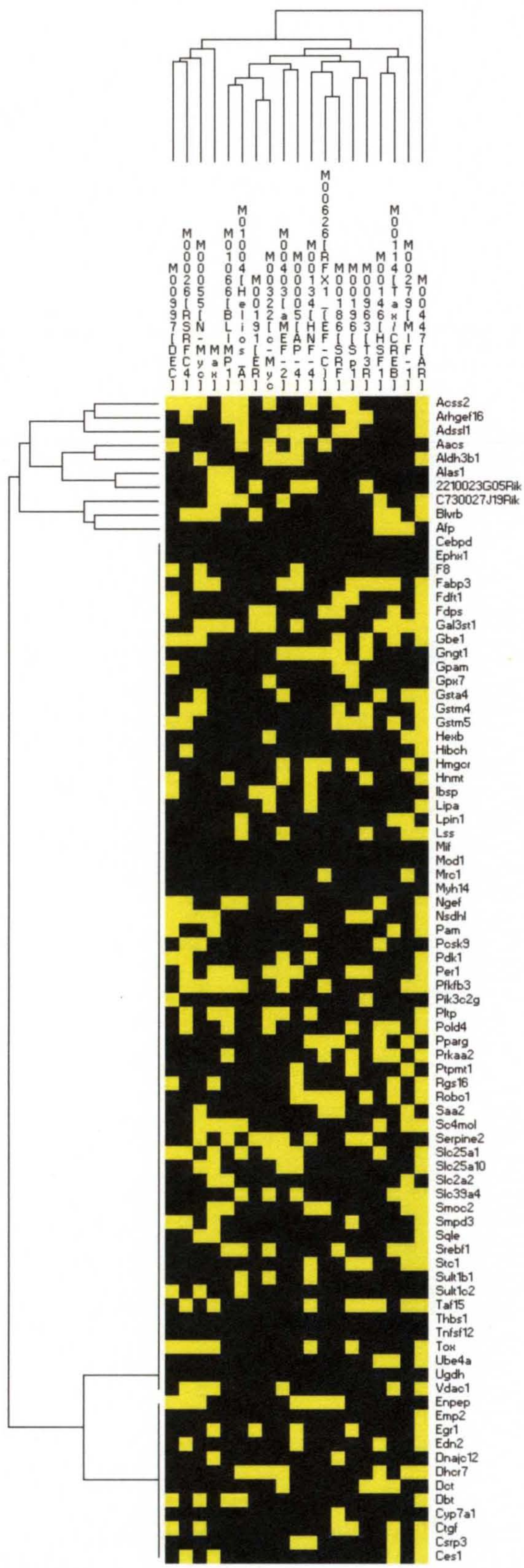

M00097[DEC]
M00026[RSRFC4]
M00055[N-Myc]
Max]
M00106[BLIMP1]
M01004[Helios_A]
M00191[ER]
M00322[c-Myc]
M00403[aMEF-2]
M00005[AP-4]
M00134[HNF-4]
M00626[RFX1_[EF-C]]
M00186[SRF]
M00196[Sp1]
M00963[T3R]
M00146[HSF1]
M00114[Tax/CREB]
M00279[MIF-1]
M00447[AR]
Acss2
Arhgef16
Adssl1
Aacs
Aldh3b1
Alas1
2210023G05Rik
C730027J19Rik
Blvrb
Afp
Cebpd
Ephx1
F8
Fabp3
Fdft1
Fdps
Gal3st1
Gbe1
Gngt1
Gpam
Gpx7
Gsta4
Gstm4
Gstm5
Hexb
Hibch
Hmgcr
Hnmt
Ibsp
Lipa
Lpin1
Lss
Mif
Mod1
Mrc1
Myh14
Ngef
Nsdhl
Pam
Pcsk9
Pdk1
Per1
Pfkfb3
Pik3c2g
Pltp
Pold4
Pparg
Prkaa2
Ptpmt1
Rgs16
Robo1
Saa2
Sc4mol
Serpine2
Slc25a1
Slc25a10
Slc2a2
Slc39a4
Smoc2
Smpd3
Sqle
Srebf1
Stc1
Sult1b1
Sult1c2
Taf15
Thbs1
Tnfsf12
Tox
Ube4a
Ugdh
Vdac1
Enpep
Emp2
Egr1
Edn2
Dnajc12
Dhcr7
Dct
Dbt
Cyp7a1
Ctgf
Csrp3
Ces1

Figure 6: TFBS on gene list clustering. Transcription factor binding sites and genes used to find enriched TFBS (filtered where $p \leq 0.05$ and $FC > |\pm 1.5|$). A yellow box indicates the gene (on the rows) was found to contain a TFBS (on the columns). Clusters are formed on columns and rows using McQuitty's Criteria in PermutMatrix (http://www.lirmm.fr/~caraux/PermutMatrix/).

REFERENCES

1. IARC, *Some drinking-water disinfectants and contaminants, including arsenic.* 2004: World Health Organization.
2. National Research Council (U.S.). Subcommittee on Arsenic in Drinking Water., *Arsenic in drinking water : 2001 update*. 2001, Washington, DC: National Academy Press. xiv, 225 p.
3. Nordstrom, D.K., *Public health. Worldwide occurrences of arsenic in ground water.* Science, 2002. **296**(5576): p. 2143-5.
4. Ryker, S., *Mapping arsenic in groundwater.* Geotimes, 2001. **46**: p. 34-38.
5. Meliker, J.R., et al., *Arsenic in drinking water and cerebrovascular disease, diabetes mellitus, and kidney disease in Michigan: a standardized mortality ratio analysis.* Environ Health, 2007. **6**: p. 4.
6. ISSI, *Arsenic Occurrence in Public Drinking Water Supplies*. 2000, EPA: Washington, D.C.
7. Amini, M., et al., *Statistical modeling of global geogenic arsenic contamination in groundwater.* Environ Sci Technol, 2008. **42**(10): p. 3669-75.
8. Polizzotto, M.L., et al., *Near-surface wetland sediments as a source of arsenic release to ground water in Asia.* Nature, 2008. **454**(7203): p. 505-8.
9. Yu, W.H., C. M. Harvey, and C. F. Harvey, *Arsenic in groundwater in Bangladesh: A geostatistical and epidemiological framework for evaluating health effects and potential remedies.* Water Resour. Res., 2003. **39**(6).
10. EPA, U.S. *Arsenic in Drinking Water.* 2006; Available from: http://www.epa.gov/safewater/arsenic/index.html.
11. Chen, C.L., et al., *Arsenic in drinking water and risk of urinary tract cancer: a follow-up study from northeastern Taiwan.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(1): p. 101-10.
12. Rossman, T.G., *Mechanism of arsenic carcinogenesis: an integrated approach.* Mutat Res, 2003. **533**(1-2): p. 37-65.
13. Applebaum, K.M., et al., *Polymorphisms in nucleotide excision repair genes, arsenic exposure, and non-melanoma skin cancer in New Hampshire.* Environ Health Perspect, 2007. **115**(8): p. 1231-6.
14. Lubin, J.H., et al., *Respiratory cancer and inhaled inorganic arsenic in copper smelters workers: a linear relationship with cumulative exposure that increases with concentration.* Environ Health Perspect, 2008. **116**(12): p. 1661-5.

15. Lamm, S.H., et al., *Arsenic cancer risk confounder in southwest Taiwan data set.* Environ Health Perspect, 2006. **114**(7): p. 1077-82.
16. Engel, R.R. and A.H. Smith, *Arsenic in drinking water and mortality from vascular disease: an ecologic analysis in 30 counties in the United States.* Arch Environ Health, 1994. **49**(5): p. 418-27.
17. Tseng, C.H., *Blackfoot disease and arsenic: a never-ending story.* J Environ Sci Health C Environ Carcinog Ecotoxicol Rev, 2005. **23**(1): p. 55-74.
18. Navas-Acien, A., et al., *Arsenic exposure and prevalence of type 2 diabetes in US adults.* JAMA, 2008. **300**(7): p. 814-22.
19. Lloyd-Jones, D., et al., *Executive summary: heart disease and stroke statistics--2010 update: a report from the American Heart Association.* Circulation, 2010. **121**(7): p. 948-54.
20. Lusis, A.J., *Atherosclerosis.* Nature, 2000. **407**(6801): p. 233-41.
21. Fauci, A.S., *Harrison's principles of internal medicine / editors, Anthony S. Fauci ... [et al.].* 17th ed. 2008, New York: McGraw-Hill Medical. v. <1-2 >.
22. Libby, P., P.M. Ridker, and A. Maseri, *Inflammation and atherosclerosis.* Circulation, 2002. **105**(9): p. 1135-43.
23. Gibson, F.C., 3rd, et al., *Innate immune signaling and Porphyromonas gingivalis-accelerated atherosclerosis.* J Dent Res, 2006. **85**(2): p. 106-21.
24. Hektoen, L., *The Vascular Changes of Tuberculous Meningitis, Especially the Tuberculous Endarterities.* J Exp Med, 1896. **1**(1): p. 112-63.
25. Ford, P.J., et al., *Anti-P. gingivalis response correlates with atherosclerosis.* J Dent Res, 2007. **86**(1): p. 35-40.
26. Brea, A., et al., *Nonalcoholic fatty liver disease is associated with carotid atherosclerosis: a case-control study.* Arterioscler Thromb Vasc Biol, 2005. **25**(5): p. 1045-50.
27. Berg, A.H. and P.E. Scherer, *Adipose tissue, inflammation, and cardiovascular disease.* Circ Res, 2005. **96**(9): p. 939-49.
28. Hotamisligil, G.S., *Inflammation and metabolic disorders.* Nature, 2006. **444**(7121): p. 860-7.
29. Libby, P., et al., *Inflammation in atherosclerosis: transition from theory to practice.* Circ J, 2010. **74**(2): p. 213-20.
30. Ghosh, J., et al., *Taurine prevents arsenic-induced cardiac oxidative stress and apoptotic damage: role of NF-kappa B, p38 and JNK MAPK pathway.* Toxicol Appl Pharmacol, 2009. **240**(1): p. 73-87.
31. Plump, A.S. and J.L. Breslow, *Apolipoprotein E and the apolipoprotein E-deficient mouse.* Annu Rev Nutr, 1995. **15**: p. 495-518.
32. Simeonova, P.P., et al., *Arsenic exposure accelerates atherogenesis in apolipoprotein E(-/-) mice.* Environ Health Perspect, 2003. **111**(14): p. 1744-8.
33. Bunderson, M., et al., *Arsenic exposure exacerbates atherosclerotic plaque formation and increases nitrotyrosine and leukotriene biosynthesis.* Toxicol Appl Pharmacol, 2004. **201**(1): p. 32-9.
34. Simeonova, P.P. and M.I. Luster, *Arsenic and atherosclerosis.* Toxicol Appl Pharmacol, 2004. **198**(3): p. 444-9.

35. States, J.C., et al., *Arsenic and cardiovascular disease.* Toxicol Sci, 2009. **107**(2): p. 312-23.
36. Wu, J., et al., *High dietary fat exacerbates arsenic-induced liver fibrosis in mice.* Exp Biol Med (Maywood), 2008. **233**(3): p. 377-84.
37. Arteel, G.E., et al., *Subhepatotoxic exposure to arsenic enhances lipopolysaccharide-induced liver injury in mice.* Toxicol Appl Pharmacol, 2008. **226**(2): p. 128-39.
38. Liu, J. and M.P. Waalkes, *Liver is a target of arsenic carcinogenesis.* Toxicol Sci, 2008. **105**(1): p. 24-32.
39. Srivastava, S., et al., *Arsenic exacerbates atherosclerotic lesion formation and inflammation in ApoE-/- mice.* Toxicol Appl Pharmacol, 2009. **241**(1): p. 90-100.
40. Qian, Y., V. Castranova, and X. Shi, *New perspectives in arsenic-induced cell signal transduction.* J Inorg Biochem, 2003. **96**(2-3): p. 271-8.
41. Bode, A.M. and Z. Dong, *The paradox of arsenic: molecular mechanisms of cell transformation and chemotherapeutic effects.* Crit Rev Oncol Hematol, 2002. **42**(1): p. 5-24.
42. Srivastava, S., et al., *In utero arsenic exposure induces early onset of atherosclerosis in ApoE-/- mice.* Reprod Toxicol, 2007. **23**(3): p. 449-56.
43. Fry, R.C., et al., *Activation of inflammation/NF-kappaB signaling in infants born to arsenic-exposed mothers.* PLoS Genet, 2007. **3**(11): p. e207.
44. Kozul, C.D., et al., *Chronic exposure to arsenic in the drinking water alters the expression of immune response genes in mouse lung.* Environ Health Perspect, 2009. **117**(7): p. 1108-15.
45. Liu, J., et al., *Global gene expression associated with hepatocarcinogenesis in adult male mice induced by in utero arsenic exposure.* Environ Health Perspect, 2006. **114**(3): p. 404-11.
46. Kozul, C.D., et al., *Laboratory diet profoundly alters gene expression and confounds genomic analysis in mouse liver and lung.* Chem Biol Interact, 2008. **173**(2): p. 129-40.
47. Carter, M.G., et al., *Transcript copy number estimation using a mouse whole-genome oligonucleotide microarray.* Genome Biol, 2005. **6**(7): p. R61.
48. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
49. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery.* Genome Biol, 2003. **4**(5): p. P3.
50. Ulitsky, I., et al., *Expander: from expression microarrays to networks and functions.* Nat Protoc, 2010. **5**(2): p. 303-22.
51. Lubin, J.H., L.E. Beane Freeman, and K.P. Cantor, *Inorganic arsenic in drinking water: an evolving public health concern.* J Natl Cancer Inst, 2007. **99**(12): p. 906-7.
52. Argos, M., et al., *Gene expression profiles in peripheral lymphocytes by arsenic exposure and skin lesion status in a Bangladeshi population.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(7): p. 1367-75.

53. Van Gaal, L.F., I.L. Mertens, and C.E. De Block, *Mechanisms linking obesity with cardiovascular disease.* Nature, 2006. **444**(7121): p. 875-80.
54. Fosslien, E., *Cardiovascular complications of non-steroidal anti-inflammatory drugs.* Ann Clin Lab Sci, 2005. **35**(4): p. 347-85.
55. Lin, M.T. and M.F. Beal, *Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases.* Nature, 2006. **443**(7113): p. 787-95.
56. Folden, D.V., et al., *Malondialdehyde inhibits cardiac contractile function in ventricular myocytes via a p38 mitogen-activated protein kinase-dependent mechanism.* Br J Pharmacol, 2003. **139**(7): p. 1310-6.
57. Lane, N., *Mitochondrial disease: powerhouse of disease.* Nature, 2006. **440**(7084): p. 600-2.
58. Alexiou, P., et al., *Lost in translation: an assessment and perspective for computational microRNA target identification.* Bioinformatics, 2009. **25**(23): p. 3049-55.

CURRICULUM VITAE

NAME: Matthew R. Zajack

UPDATED: 05 August, 2010

ADDRESS: 1040 Cherokee Rd APT D2
Louisville, KY 40204

CONTACT: 715-456-2485
matt.z@centurytel.net

EDUCATION: B.S. Mathematics and Computer Science
University of Wisconsin-Stout, Menomonie, WI
2007

M.S Pharmacology and Toxicology
University of Louisville, Louisville, KY
2008-present

EMPLOYMENT: IPIBS Fellow, Research Assistant
Dept. of Pharmacology and Toxicology
University of Louisville, Louisville, KY
2008-2010

SOCIETIES: National Strength and Conditioning Association (NSCA)
American College of Sports and Medicine (ACSM)

ABSTRACTS: Zajack, M.R., Micciche, A.F., Taylor, B.F., and States J.C. Development of Fusion Vectors for Anaphase Promoting Complex: The Framework to an Arsenic Trioxide Binding Experiment. *Research!Louisville*, 2008.

Zajack, M.R., Micciche, A.F., Taylor, B.F., and States J.C. Development of Fusion Vectors for Anaphase Promoting Complex: The Framework to an Arsenic Trioxide Binding Experiment. *Brown Cancer Center Retreat*, 2008.

Zajack, M.R., Micciche, A.F., Taylor, B.F., and States J.C. Development of Fusion Vectors for Anaphase Promoting

Complex: The Framework to an Arsenic Trioxide Binding Experiment. *Ohio Valley Society of Toxicology*, 2008.

Zajack, Matt. Arsenic trioxide may inhibit the chemotherapeutic target Anaphase Promoting Complex/Cyclosome subunit, APC11. *First Year Presentation in Department of Pharmacology and Toxicology*, 2009.

M. Zajack, Y.Piao, M.S.Ko, E. Rouchka, and J.C. States. Stimulation of Inflammation and Mitochondrial Dysfunction Pathways by Arsenic Exposure in Livers of ApoE-Knockout Mice. *Ohio Valley Society of Toxicology*, summer session, 2009.

M. Zajack, Y.Piao, M.S.Ko, E. Rouchka, and J.C. States. Induction of Inflammation and Mitochondrial Dysfunction Pathways in Livers of Arsenic-Exposed ApoE-Knockout Mice. *Research!Louisville*, 2009.

M. Zajack, Y.Piao, M.S.Ko, E. Rouchka, and J.C. States. Stimulation of Inflammation and Mitochondrial Dysfunction Pathways by Arsenic Exposure in Livers of ApoE-Knockout Mice. *Society of Toxicology (SOT)*, 2010.