ENVIRONMENTAL METALS AND BIRTH DEFECTS: NEW APPROACHES TO UNDERSTANDING THE ROLE OF METALS IN CONGENITAL HEART DEFECTS

Alison P. Sanders

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Environmental Sciences and Engineering.

Chapel Hill
2013

Approved by:

Rebecca C. Fry

Michael Flynn

Amy H. Herring

Andrew F. Olshan

Frederick Pfaender

ABSTRACT

ALISON P. SANDERS: **Environmental metals and birth defects: New approaches to understanding the role of metals in congenital heart defects**
(Under the direction of Rebecca C. Fry)

Birth defects are a leading cause of infant mortality, and cardiac defects are among the most fatal. The majority of birth defects have no known cause whether environmental, genetic, or a combination of these factors. Toxic metals are likely contributors to birth defects in humans, yet the biological pathways that underlie these relationships remain largely unknown. The objective of this research was to assess the contributions of environmental metals to congenital heart defects using **i) an ecologic study to evaluate the relationship between metals and birth defects in North Carolina and ii) an *in vitro* cardiomyocyte model to identify underlying metal-altered molecular pathways**. This research evaluated the associations between arsenic, cadmium, manganese, and lead levels in private well water with the prevalence of specific birth defects using six years of data collected by the North Carolina Birth Defects Monitoring Program. Prevalence ratios (PR) with 95% confidence intervals (CI) were calculated to estimate the association between the prevalence of birth defects in census tracts with the highest average metal levels compared to the lowest average metal levels, adjusted for maternal age, race, and education status. We identified relationships between the highest category of arsenic exposure and the prevalence of conotruncal heart defects (PR: 1.3 95%CI: 0.9-1.8) as well as the highest category of manganese exposure and the

prevalence of atrioventricular septal heart defects (PR: 1.8 95% CI: 1.1-3.1). The findings suggest an ecologic association between concentrations of metals in drinking water and the prevalence of specific birth defects. Next, we applied an *in vitro* approach to identify biological pathways that underlie metal-altered signaling in the heart. We exposed human-derived cardiomyocytes to low-level cadmium chloride (0.5 μM) and assessed changes in genome-wide microRNA (miRNA) and messenger RNA (mRNA) expression levels. We identified 8 miRNAs and 31 genes that were differentially expressed in cardiomyocytes exposed to cadmium compared to unexposed controls. A subset of 8 mRNAs were predicted targets of the cadmium-associated miRNAs. We identified miRNA-dependent signaling pathways that were enriched for gene expression, embryonic development, and organismal development. Taken together, the findings contribute to the understanding of potentially preventable environmentally-mediated birth defects.

To my mother: Kathryn Ann Sanders.

ACKNOWLEDGEMENTS

I am thankful to Rebecca Fry for her mentorship and guidance throughout my time at UNC. I appreciate the dedication she has to her students' success in research and in life. I thank Mike Flynn, Amy Herring, Andy Olshan, and Fred Pfaender for serving as committee members and contributing to this research. I thank our collaborators including Tania Desrosiers, Dianne Enright, Josh Warren, and Bob Meyer for their support and assistance with the analysis of the birth defects data. I especially thank Bob Meyer with the North Carolina Birth Defects Monitoring Program for his guidance and support as well as for providing analysis space in Raleigh, NC. I appreciate the help of Daniel Rojas, Julia Rager, and Kathryn Bailey for their assistance with the *in vitro* cardiomyocyte culture and miRNA analyses. Thanks to Lisa Smeester as well as other past and present Fry Laboratory members who have assisted with projects over the last four years. I am truly grateful to my family and friends for their continued love, support, and encouragement.

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES x

LIST OF ABBREVIATIONS AND SYMBOLS xi

CHAPTER 1: INTRODUCTION 1

The global burden of birth defects 1

Congenital heart defects (CHDs) *1*

Environmental factors that contribute to CHDs *2*

Toxic metals as ubiquitous environmental contaminants 3

Sources and routes of potential toxic metal exposure *4*

Drinking water metal exposure *5*

Prenatal exposure, susceptibility, and development 6

Prenatal biomonitoring of toxic metals *7*

In utero and developmental susceptibility *8*

Relationships between toxic metals and birth defects 9

Epidemiologic evidence of the relationship between drinking water metals and birth defects *10*

Limitations of epidemiologic studies of birth defects *11*

Understanding biological mechanisms that link metals to CHDs 14

Toxicological findings: in vivo and in vitro studies *14*

Biological signaling pathways of CHDs *16*

Epigenetics modifiers: DNA methylation and microRNAs *17*

Research gaps and specific aims 19

References 22

CHAPTER 2: ASSOCIATION BETWEEN ARSENIC, CADMIUM, MANGANESE, AND LEAD LEVELS IN PRIVATE WELLS AND BIRTH DEFECTS PREVALENCE IN NORTH CAROLINA 33

Overview 33

Introduction 34

Methods 36

Results 42

Discussion 46

Conclusions 53

Tables 54

Figures 57

References 64

CHAPTER 3: MICRORNA-DEPENDENT AND INDEPENDENT CADMIUM-INDUCED SIGNLAING IN HUMAN CARDIOMYOCYTES 70

Overview 70

Introduction 71

Methods 73

Results 78

Discussion 81

Conclusions 85

Tables 86

Figures 89

CHAPTER 4: CONCLUSIONS 98

Summary 98

Public health implications 99

Recommendations for future studies 100

APPENDIX ONE: ARSENIC IN NORTH CAROLINA: PUBLIC HEALTH IMPLICATIONS 102

APPENDIX TWO: TOWARDS PRENATAL BIOMONITORING IN NORTH CAROLINA: ASSESSING ARSENIC, CADMIUM, MERCURY, AND LEAD LEVELS IN PREGNANT WOMEN 133

APPENDIX THREE: CADMIUM-ASSOCIATED PATTERNS OF DNA METHYLATION IN LEUKOCYTES FROM MOTHER-BABY PAIRS 160

References 191

LIST OF TABLES

TABLE 2-1. Characteristics of eligible cases and controls 54

TABLE 2-2. Frequency of exposed case and control infants where exposed is defined as living in a census tract with an average metal level greater than or equal to the 90th percentile..... 55

TABLE 2-3. Association [PR (95%CI)] between selected metals and birth defects..... 56

TABLE 3-1. Cadmium exposure to cardiomyocytes significantly increased the expression of nine unique miRNAs corresponding to 56 probesets. 86

TABLE 3-2. Eight miRNA predicted mRNA targets were confirmed experimentally by microarray analysis of mRNA transcript levels in samples exposed cadmium 87

TABLE 3-3. Transcriptional regulators (TR) known to influence the differentially expressed mRNA transcripts were identified for the miRNA-associated genes and miRNA-independent genes..... 88

LIST OF FIGURES

FIGURE 2-1. Average levels of arsenic (A), cadmium (B), manganese (C), and lead (D) within census tracts across North Carolina 57

FIGURE 2-2. Sensitivity analyses plots for selected associations between residence in ecologic units of highest compared to lowest metal levels and prevalence of birth defects. 58

FIGURE 3-1. Significant molecular network of eight miRNAs differentially expressed in cadmium-exposed cardiomyocytes 89

FIGURE 3-2. Heatmap of 31 differentially expressed mRNA transcripts 90

FIGURE 3-3. Significant molecular network of miRNA-associated proteins and miRNA-independent proteins 91

LIST OF ABBREVIATIONS

5-meC	5-methyl-cytosine
As	Arsenic
ASD	Atrial septal defects
AVSD	Atrioventricular septal defect
Cd	Cadmium
CHD	Congenital heart defects
CI	Confidence interval
CL	Cleft lip
CP	Cleft palate
CpG	Cytosine-phosphate-guanine
EA	Esophageal atresia
ECD	Endocardial cushion defect
DHHS	Department of Health and Human Services
HLHS	Hypoplastic left heart syndrome
MCL	Maximum Contaminant Level
miRNA	MicroRNA
Mn	Manganese
miRNA	MicroRNA
mRNA	Messenger RNA
NHANES IV	*Fourth National Report on Human Exposure to Environmental Chemicals*
NTD	Neural tube defect
Pb	Lead

PR	Prevalence ratio
SMCL	Secondary Maximum Contaminant Level
VSD	Ventricular septal defect
TEF	Tracheo-esophageal fistula
TGA	Transposition of the great arteries
TOF	Tetralogy of Fallot
TT	Treatment Technique

CHAPTER 1: INTRODUCTION

The global burden of birth defects

Birth defects affect approximately 3% of all live births and are the leading cause of infant mortality in the US [1]. The estimated cost of US hospital stays due to birth defects was $2.6 billion in 2004 [2], and birth defects are the 10th leading cause of global Disability Adjusted Life Years [3]. Despite the serious physiological and economic impacts of births defects, roughly 60-70% of birth defects have no known environmental or genetic cause [4]. Collectively, birth defects include a spectrum of structural malformations at birth including musculoskeletal, reproductive, gastrointestinal, craniofacial, neural tube, and cardiac anomalies. Each of these broad classifications is further divided into specific phenotypes that range in severity and may be diagnosed from before birth to later in life.

Congenital heart defects (CHDs)

Congenital heart defects (CHDs) are the most common type of birth defects, occurring in nearly 1 in 100 live births [5], and are the leading cause of defect-related deaths [6]. CHDs consist of a heterogeneous group of structural malformations affecting structures of the heart including the valves, chambers and walls, and outflow tracts. The most common CHDs include septal defects such as ventricular septal defects (VSD) and atrial septal defects (ASD), conotruncal defects such as Tetralogy of Fallot (TOF) and

transposition of the great arteries (TGA), and outflow tract obstructions such as hypoplastic left heart syndrome (HLHS) and coarctation of the aorta [7]. Many infants born with CHDs undergo surgery in the first six months of life, which may fully repair a defect; however, diagnosis of a CHD in some cases may require life-long healthcare and multiple reparative surgeries [8]. Approximately 34% of birth defect hospitalizations in the US are due to cardiac and circulatory anomalies, and there is substantial economic burden of CHDs, estimated at nearly $1.4 billion in 2004 [2].

Many CHDs arise from genetic origins such as chromosomal abnormalities or single gene mutations. For example, of infants born with Edwards Syndrome, a genetic disorder that is caused by an additional copy of the 18^{th} chromosome, 90-100% are estimated to develop a CHD [9]. An increasing number of genetic factors or syndromes have been identified for many of the cardiac phenotypes [9, 10]; however, it's estimated that 80% of CHD cases originate from other or unknown causes [11].

Environmental factors that contribute to CHDs

Even with the identification of some genetic risk factors, the biological mechanisms of CHD remains poorly understood. As with other complex diseases, it is hypothesized that a combination of genetic and environmental factors contribute to CHDs, rather than a single component in isolation. Potential environmental exposures of interest in CHD etiology have included maternal diet (including alcohol and tobacco products), maternal health conditions (diabetes, influenza, or rubella), medication use (including retinoic acid and anti-seizure medications), maternal stress, air pollution, pesticides, disinfection byproducts, solvents, and heavy metals [6, 12-14]. Notably, toxic

metal exposure has been associated with birth defects (reviewed by [15]), yet limited studies have directly assessed this relationship or mechanism of action. Given the large economic and societal impacts of birth defects and specifically the contributions of cardiac defects, the studies presented herein examine the links between early life exposure to toxic metals and CHDs at the population and mechanistic levels.

Toxic metals as ubiquitous environmental contaminants

Toxic metals are elements that have no known biological role in human physiology. The research presented in Chapter 2 prioritized four heavy metals for study based on their environmental relevance in North Carolina: arsenic, cadmium, lead, and manganese. Arsenic, cadmium, lead, and manganese are toxic metals ranked by the ATSDR as priority contaminants of concern and are known carcinogens and/or neurodevelopmental toxicants [16-19].

Toxic metals are present in environmental reservoirs including air, water, and soil; and human exposure is widespread in populations worldwide and in the US. Specifically, diet and drinking water are common sources of exposure to each of these metals. Among pregnant women participating in the *Fourth National Report on Human Exposure to Environmental Chemicals* (NHANES IV), 66%, 89%, and 94% had detectable levels of cadmium, mercury, and lead in whole blood [20]. Blood levels of arsenic were not examined. The global prevalence of toxic metals in the environment combined with their well-documented health effects represents a public health concern on a global scale.

Sources and routes of potential toxic metal exposure

Arsenic is the 20th most abundant element in the Earth's crust [21]. Approximately 70% of environmental arsenic is naturally-occurring and 30% is created as a result of anthropogenic sources [22]. Human exposure to arsenic primarily occurs through consumption of contaminated drinking water or food products including rice, apples, [16, 23]. For example, it is estimated that more than 30 million individuals in Bangladesh are potentially exposed to arsenic through contaminated drinking water wells [24].

Cadmium exposure commonly occurs through diet via consumption of shellfish and leafy greens and through exposure to tobacco products [19]. Minor exposure can also occur from byproducts of industrial processes [19]. Blood levels of cadmium among smokers are up to five times higher than blood levels detected in nonsmokers [25], and cadmium levels in pregnant women exposed to second–hand smoke have been shown to be double that of nonsmokers [26].

Exposure to lead can result from various environmental sources such as contaminated dust or soil, drinking water, lead-based paint, cigarette smoke, as well as byproducts of industrial processes [17]. Due to the removal of lead from gasoline in the 1970s and 80s, environmental lead levels have steadily declined. Drinking water is a potential route of lead exposure as a result of lead leaching from lead-soldered pipes [27] and rural counties may experience additional risk factors for environmental lead exposure such as older housing and lower socioeconomic status that may increase an individual's likelihood of exposure [28].

Manganese is typically considered an essential metal, but excess levels of

manganese are known to result in neurotoxic effects in humans [18]. Manganese is typically encountered through diet; however, drinking water may be a significant source of manganese ingestion [29]. Inhalation exposure to manganese aerosols can occur through showering and occupational exposure to welding fumes [18, 30].

Drinking water metal exposure

Arsenic, cadmium, and lead are regulated in public water supplies under the Safe Drinking Water Act at levels of 10 μg/L, 5 μg/L, and 15 μg/L [31]. It should be noted that lead is regulated according to a Treatment Technique (TT) which requires that no more than 10% of samples can exceed 15 μg/L; arsenic and cadmium regulatory standards are set as the Maximum Contaminant Level (MCL) allowed in public distribution systems. Manganese, largely due to issues associated with taste and odor, is regulated at a Secondary Maximum Contaminant Level (SMCL) of 50 μg/L that is not federally enforced in public distribution systems [31]. Detectable levels of metals in public drinking water can range up to the MCL and may potentially contribute to low-level exposure in the general population; however, MCLs in public water distribution systems are largely considered protective of human health. In contrast, millions of US residents rely on drinking water supplies known as private domestic wells that are not regulated by federal government regulations, and there are currently no regulatory standards for domestic well waters. In areas where there a high proportion of private well use is compounded by underlying geology predisposed to elevated metal levels, individuals may unknowingly consume drinking water with contaminant levels in excess of EPA guideline levels.

Approximately 14% (about 42 million people) of the U.S. population and 26% of North Carolinians (2.3 million), use federally unregulated private wells as a primary drinking water source [32]. Private well water consumption in the U.S. is associated with an excess lifetime risk of bladder and lung cancer due to arsenic ingestion that is five times higher than that estimated for public well users [33] (Kumar 2010). Previously, we showed that arsenic levels in North Carolina private wells ranged as high as 806 ppb - a level that is 80 times the current EPA MCL in public distribution systems [34] (see Appendix One) and exceeds a level associated with known cancer risks. This initial study also demonstrated spatial and temporal trends of naturally-occurring arsenic across the state. These data support that ingestion of metal-contaminated drinking water may be a public health concern in North Carolina. Identifying regions of contamination through studies like the one above, can help to shape state and local water monitoring policies [35] as well as assist in developing cost-effective monitoring programs that target at risk populations. Moreover, there is a need to better understand potential health effects associated with private well metal exposure, particularly among susceptible populations, in North Carolina.

Prenatal exposure, susceptibility, and development

The widespread occurrence of environmental metals indicates a public health need to monitor and protect susceptible populations from toxic metal exposures in North Carolina. Pregnant women and children are considered susceptible populations due to their developmental status and/or compromised immune systems. Heavy metals including arsenic, cadmium, lead, and manganese are known human carcinogens and/or

developmental toxicants [16-19]. The known neurotoxic and developmental impacts of toxic metals make these priority contaminants of concern for the protection of population health. Of concern, these metals are able to cross the placental barrier potentially resulting in *in utero* exposure during a critical period of developmental susceptibility [36-38]. Prenatal metal exposure has been implicated in numerous adverse health outcomes in newborns including increased infant mortality, lower birth weight, and intrauterine growth restriction (reviewed by [15]). Prenatal biomonitoring programs would serve to protect pregnant women and the developing fetus, yet these remain limited in number.

Prenatal biomonitoring of toxic metals

Toxic metal exposures *in utero* and during childhood may result in significant health effects to individuals that manifest cumulatively to the detriment of populations. US annual economic losses due to decreased productivity from environmental exposure to mercury and lead were estimated as $8.7 and $43.4 billion respectively [39, 40]. In pregnant women, maternal blood levels of 3.5 μg/L mercury and 5.0 μg/dL lead are recommended for the protection of fetal health [41-43]. However, currently Minnesota, New York City, and New York State are the only jurisdictions that have active guidelines for monitoring maternal blood lead levels in the US [43], in spite of strong evidence to indicate the adverse health effects of metals to susceptible populations such as newborns.

While North Carolina has established a successful program for childhood lead screening [44], we recently showed that in a sample of pregnant women from North Carolina (n=211), more than 57% had a detectable blood levels of more than one toxic metal including arsenic, cadmium, mercury, or lead [45] (see Appendix Two).

Specifically, arsenic (65.7%), cadmium (57.3%), mercury (63.8%), and lead (100%) levels were detected in maternal blood samples. The findings also suggest that factors related to maternal county of residence and race may impact maternal exposure levels. Moreover, some women exceeded the guideline levels of mercury and lead during pregnancy. No reference guideline levels currently exist for prenatal levels of arsenic or cadmium.

Targeted national, statewide, and community level programs promoting mothers' awareness of metals exposure and *in utero* effects would help to protect human health and lessen the economic burden of metal exposure. Biomonitoring programs in North Carolina and other targeted populations can help to protect the health of expectant mothers and their children by reducing prenatal exposures and subsequently the potential for future adverse developmental effects [45].

In utero and developmental susceptibility

Prenatal biomonitoring represents a step towards protecting the health of developing newborns. The first trimester of pregnancy is considered a critical period of fetal susceptibility to maternal exposure to environmental contaminants. Specifically, one month prior to conception and three months after conception is known as the periconceptional period. Environmental exposures during the periconceptional period are of greatest concern to the developing fetus. In birth defect etiology, each major organ system has a more sensitive period during development, which overlaps with the pregnancy/developmental timeline for specific organs. The most sensitive time period for cardiac development occurs in the first through third weeks of pregnancy.

The heart is the first major organ to develop and function in an embryo. The heart arises out of the primary and the secondary heart fields, which originate the left ventricle or the right ventricle, respectively [46]. The two fields converge to form the heart tube, which then loops and bends to begin to form the four-chambered heart [46]. On approximately day 21 after conception, electrophysiological factors induce spontaneous orchestrated contraction of cardiac cells and the human heart begins to beat [47]. The process of heart formation is a highly controlled and time-dependent process [48]. Proper development and function of the cardiovascular system is crucial for embryonic survival and health. Defects in cardiac development may result in fetal loss or congenital heart defects. Physiologically, there are still many unknowns about the processes that control and regulate embryonic heart development. The highly regulated, precise coordination and timing of cardiac development stages in early gestation potentially prior to pregnancy detection means that cardiac organ development is particularly susceptible to environmental influences [46].

Relationships of environmental metals to birth defects

There are many detrimental health outcomes associated with developmental metal exposure including spontaneous abortion, stillbirth, low birth weight, preterm birth, reduced fetal growth, impaired neurodevelopment, and congenital malformation [15, 49-55]. Despite the known developmental effects of metals and susceptible periods of fetal development in pregnancy, few epidemiologic studies have examined the link between maternal metal exposure and the occurrence of structural defects. Exposure to metals, such as arsenic, cadmium, mercury, and lead, is associated with detrimental birth

outcomes in humans [56-61]. Further, epidemiological data suggest there may be links between environmental exposure to metals and birth defects in humans including birth defects [15, 62-69]. These findings are demonstrated in animal models, which clearly show that exposure to toxic metals cause birth defects [70-73], yet epidemiological studies investigating these effects in human populations remain limited in number and study quality.

Epidemiologic evidence of the relationship between drinking water metals and birth defects

This section reviews the literature regarding the four prioritized heavy metals that will be discussed in Chapter 2, namely arsenic, cadmium, lead, and manganese. Metals in drinking water including arsenic and lead have been associated with the occurrence of CHDs [62, 64, 69] as well as NTDs [63, 68]. No studies have examined manganese exposure in drinking water and the relationship to birth defects, and only a single study has analyzed cadmium levels in drinking water with the occurrence of NTDs [63].

Studies that have examined the association between drinking water metal levels and birth defects reported conflicting findings. One study in Bangladesh reported an elevated odds of any birth defect (odds ratio (OR): 1.005, 95% CI: 1.001-1.010) with exposure to environmental arsenic in drinking water, but did not differentiate between types of birth defects [66]. The finding was a small but statistically significant association and the authors recommended future examination of the association with NTDs. A second study investigated the association between arsenic in drinking water as well as occupational metal exposures and the risk of NTDs (lumped spina bifida, anencephaly, and encephalocele) [63]. This case-control study found that environmental metals in

drinking water including arsenic, cadmium, mercury and lead were not strong risk factors for NTDs. Zierler et al. conducted a case-control study in of drinking water arsenic levels and cardiovascular defects in Massachusetts [69]. The authors reported an association between publicly supplied drinking water arsenic levels and the prevalence of congenital heart defects, specifically that of coarctation of the aorta (prevalence odds ratio (POR): 3.4, 95% CI: 1.3-8.9). In an ecologic study of mortality from circulatory diseases in the US, Engel and Smith reported an elevated standardized mortality ratio of 2.0 (95% CI: 1.1-3.4) among female infants with circulatory system anomalies [64]. A case-control study in Massachusetts reported crude and adjusted ORs of 1.8 and 2.2 (no 95% CIs were reported) of grouped cardiovascular defects associated with detectable lead levels in drinking water [62]. A single study, conducted in Scotland, reported no association between lead in drinking water and NTDs and did not assess the relationship with other defects [67].

The association between manganese exposure through drinking water and birth defects has not been previously studied. However, one study reported that exposure to water manganese levels greater than 0.4 mg/L was associated with an elevated risk of all-cause infant mortality compared to unexposed controls (OR: 1.8, 95%CI: 1.2-2.6) [74]. The authors recommended that future studies be conducted to replicate the findings and examine more specifically defined mortality outcomes.

Limitations of epidemiological studies of birth defects

The dearth of epidemiological studies investigating the relationship between environmental metal exposures and birth defects is in part due to methodological hurdles

that impact studies of rare outcomes. Given the nature of birth defect etiology, population-based studies examining environmental causes of disease must be retrospective since it is not known in advance whether a fetus will develop a defect. This results in several major hurdles in studies of birth defects including poor exposure assessment and subsequent issues of temporality, limited sample sizes or poor case ascertainment, and unmeasured confounding.

Among the studies examining drinking water metal exposures, exposure was assigned based on average levels in community or public water supplies. Temporal fluctuations were accounted for in two of the studies by assigning maternal exposure based on public water samples closest to the date of conception [63, 69]. Samples were averaged if the distribution system was unknown or multiple water tests were performed within the periconceptional period. In other instances, samples were obtained from the mother's primary water source up to one year after pregnancy [66]. Ideally, maternal exposure assessment would be performed near the time of conception, however this is not feasible with the preferred case-control design used to study birth defects. Direct biological monitoring of maternal and fetal exposures during the periconceptional period would greatly improve potential exposure misclassification issues in studies of environmental causes of birth defects. This approach is particularly difficult in the study of birth defects due to the low prevalence of defects and the associated cost of monitoring a large number of individuals to detect a small proportion of cases. Moreover, exposure monitoring during the critical period of susceptibility may differ greatly based on the defect of interest. Commonly, information on environmental exposures is obtained through questionnaire or maternal interview. Direct exposure assessment may be less of a

concern for geographically stable contaminants. For instance, in arsenic endemic regions evidence indicates that levels of arsenic in groundwater tend to remain stable for several years [75].

Notwithstanding possible exposure misclassification, these studies were largely hampered by insufficient sample sizes and case enrollment ($n<30$) for analyses. Studies investigating the effects of any exposure on birth defects should separate defect etiologies as outcome as much as possible, preferably to individual phenotype classifications or groups based on similar etiology. Lumping cases into large groups of defects (as by [63, 66] can obscure interpretation of the findings due to outcome misclassification since there are a wide variety of congenital malformation etiologies.

Additional factors that affect metal-defect relationships include prenatal nutrition [76], maternal age [77], race [78], socioeconomic status [68, 79], infant sex [80], and family history [81]. For example, in the U.S. non-Hispanic Black infants compared with non-Hispanic white infants, had a significantly higher birth prevalence of Tetralogy of Fallot [78]. CHDs are also more common among families of lower socioeconomic classes [79] and among families with a previous CHD occurrence [81]. Limited sample sizes or use of retrospective data (e.g. collected through birth defects registries) may prevent adequate adjustment for confounding variables.

The biological mechanism by which toxic metals may induce birth defects is not well understood and is underrepresented in the literature. Studies may be lacking due to the well-documented ability of arsenic and other metals to induce spontaneous abortion [49, 82] that may ultimately prevent the true proportion of incident cases of congenital malformations from being detected. The use of *in vivo* and *in vitro* models has enabled

more direct study of the mechanisms that drive metal-defect associations and largely eliminates the limitations encountered by population-based studies.

Understanding biological mechanisms that link metals to CHDs

As discussed above, CHDs are associated with metal exposures including arsenic and lead [62, 64, 69, 83]. A preponderance of evidence from epidemiological, *in vivo* and *in vitro* studies has provided substantial knowledge of genetic and environmental factors that contribute to CHDs. Toxic metals are known to impact cell function at a molecular level that can result in cancer and non-cancer endpoints [84, 85], however, the role of molecular factors in metal-associated birth defects in humans remain largely unknown. Molecular factors including gene expression and epigenetic processes that ultimately control gene expression may underlie metal-associated CHDs.

Toxicological findings: in vivo and in vitro studies

Given the number of ethical and technical limitations associated with studying environmentally-mediated causes of birth defects in humans, alternative *in vivo* and *in vitro* models provide controlled situations and have enabled substantial advances in the understanding of defect etiology. There is extensive evidence from animal models suggesting that toxic metal exposure is associated with structural malformations [73, 86-94]. For example, metals including arsenic, cadmium, lead, and mercury have been shown to induce structural defects in the chick embryo affecting development of the neural tube, limbs, gut, and heart among other structures [70, 73, 91, 95, 96]. Notably, there is extensive research and debate on the role of arsenic in NTDs as animal models

consistently demonstrate teratogenicity but epidemiologic studies inconsistently report an association [68, 87, 97]. Thus, animal models are useful to better understand molecular mechanisms and pathways involved in the interplay between metal exposure and birth defects [97, 98].

In vivo models present several advantages including that cardiac development is largely conserved cross vertebrate species, animal models provide shorter developmental timelines and the ability to model maternal conditions. *In vivo* models allow researchers to differentiate between the effects of pre- and postnatal exposures and enable direct assessment of exposures during critical windows of susceptibility. Advances in understanding mechanisms of CHD etiology have been made possible through the use of genetically modified animals such as knock-out models. For example, there are a number of candidate genes (e.g, *connexin40* and *connexin45*) that upon deletion from the genome result in septal defects in animal models, including VSDs [99-101]. Further, there is evidence that prenatal metal exposure, for example to cadmium, can lead to long-term alterations in cardiac function in adult rats [102].

The results from *in vivo* experiments are currently curated in publicly accessible databases of genes and environmental contaminants with known associations. Together these have enabled novel computational predictions that allow for relatively quick survey of the role of molecular factors in metal-associated birth defects. We have previously used toxicogenomic approaches to identify biological pathways perturbed by metal exposure using an *in ovo* whole chick embryo model [95]. Specifically, we computationally predicted key biological pathways that were metal-associated and altered during development. Using an *in ovo* approach we validated that blockade of a prioritized

pathway, the glucocorticcoid pathway, reverted arsenic-induced neural tube defects in the chick embryo. The results demonstrated the application of a novel systems biology strategy by which biological pathways can be predicted and subsequently tested to increase our understanding of pathophysiological mechanisms related to birth defects [95].

In vitro models offer the advantage of studying relevant tissue types (eg, human cell types) under highly controlled environments that further reduce the influence of confounding factors in animal studies. *In vitro* studies also benefit from relatively quick experimental time, and ease of collection of biological samples for molecular analyses. Few studies, however, have investigated the toxicogenomic and epigenetic effects of metals in heart cells. Identification of genomic and epigenetic effects of metals levels in metal-perturbed heart function using *in vitro* culture of cardiomyocytes has not been previously described.

Biological signaling pathways related to CHDs

The exact mechanisms contributing to CHDs remain elusive, however key gene families are involved in cardiac development [99, 103-108]. There is growing evidence that prenatal environmental exposure may influence the burden of disease in adult life, and that this relationship is associated with epigenetic modifications altered during the prenatal period [109-112]. It is hypothesized that many key cardiac development genes have upstream regulators that may be modified by metal exposure perhaps through epigenetic processes, leading to altered heart development and function in adulthood [113, 114].

Epigenetic mediators: DNA methylation and microRNAs

Epigenetics is the study of heritable changes that occur "above" the genome that do not directly alter the genetic sequence. Epigenetic factors include DNA methylation, histone modification, microRNAs (miRNA) which have functional consequences in the cell such as modifying gene expression or protein translation without changes in DNA sequence [115, 116]. DNA methylation refers to the addition of a methyl group that can be reversibly added to cytosine bases typically occurring in cytosine-guanine (CpG) rich areas of the DNA. Changes in DNA methylation are potentially transcriptional-regulating and can have profound impacts, particularly during development. The addition or removal of methyl groups from 5-methylcytosine (5-meC), or DNA methylation, is an epigenetic mechanism that may play a key role in mediating biological processes [110], possibly contributing to subsequent health effects resulting from environmental toxicant exposure.

Metals are known modifiers of DNA methylation [84, 117-119]. We recently investigated the effects of cadmium exposure on fetal and maternal patterns of DNA methylation in a cohort of mother-baby pairs from North Carolina. We found that 61 and 92 genes in fetal and maternal DNA were differentially methylated and showed that these genes were enriched for biological pathways involved in transcriptional regulation and apoptosis [120] (See Appendix Three).

miRNAs are short single-stranded RNAs approximately 20 nucleotides in length that post-transcriptionally regulate gene expression through silencing. miRNAs are involved in a number of cell proliferation, apoptosis, and developmental processes [111, 112]. miRNAs are small non-coding RNA segments that selectively bind segments of

messenger RNA (mRNA) and prevent translation of proteins, ultimately resulting in gene silencing. The first miRNA, lin-4, was discovered in 1993 in *C. elegans* and is involved in regulation of key genes in organism development. Currently over 1,100 human miRNAs have been identified and are believed to target approximately 60% of genes [121]. Multiple miRNAs may bind to and/or control downstream translation of proteins. Likewise, a single miRNA can bind multiple mRNA targets. The implications of miRNA post-transcriptional processes include the potential to inhibit biological pathways even in the presence of increased gene expression (mRNA transcripts).

miRNAs are known regulators of gene expression during cardiovascular development (reviewed by [114, 122]). Specifically, miR-133, miR-22, and miR-1 are involved in cell proliferation and maintenance processes that dictate gene expression and cardiac development and/or dysfunction. Of note, miR-1 overexpression negatively regulates cardiac growth, in part by targeting and downregulating heart and neural crest derivatives-expressed protein 2 (Hand2), a transcription factor that promotes cardiomyocyte expansion [123]. It has also be shown that mice deficient in miR-17~92 developed fatal ventricular septal defects [124].

Numerous studies have assessed miRNA expression with respect to arsenic exposure and are reviewed in [125], while only three studies have examined or cadmium-associated miRNA changes in human tissues or cell lines [126-128]. Occupational cadmium exposure was associated with increased expression of miR-146a in peripheral blood leukocytes [126]. In Chinese children, increased levels of urinary arsenic and lead, but not cadmium were associated with decreased levels of miR-21 and miR-221 [127]. Finally, human hepatoblastoma cells exposed to 10 μM cadmium showed decreased

expression levels of 12 miRNAs, four of which belonged to the let-7 family of tumor suppressors [128].

Both DNA methylation and miRNA expression are proposed epigenetic processes that contribute to CHDs [114, 129, 130]. Given that metals alter miRNA expression, it is plausible that metal-induced miRNA alterations in heart cells, may affect a large number of cell division and maintenance processes, which contribute to aberrant gene expression and dysregulated cardiac function and/or development. The research presented in Chapter 3 examines miRNAs as a potential epigenetic mechanism of metal-induced perturbed cardiac gene expression. Epigenetic processes may likely be the "missing link" in previous studies of genetic and environmental factors that influence birth defect etiology.

Research gaps and specific aims

Metal exposure continues to be an important area of public health concern for both maternal and child health. There are knowledge gaps in population-based studies of toxic metals exposure and birth defect outcomes as well as understanding the underlying biological processes that potentially mediate metal-induced defects. The identification of metal-defect relationships as well as the underlying genetic and epigenetic components will lead the way in understanding potentially preventable environmentally-mediated defect etiology.

While studies investigating the relationship between metals in drinking water and the risk of birth defects show putative evidence of a positive association with CHDs, there is room for methodological improvement and need for further studies. Bias due to

possible exposure misclassification, insufficient sample sizes, and issues regarding case identification and ascertainment largely hampered previous studies. In addition, all US studies that have examined the association between drinking water arsenic and birth defect outcomes have referenced populations supplied by public drinking water. Public drinking water is federally regulated for most toxic metals. Since millions of Americans obtain their water from unregulated domestic private wells and private wells and these wells are federally unregulated, studies of private well water and birth defect outcomes are needed. The area of environmental metal exposure and birth defect outcomes has room for methodological improvement and an essential need for additional study to investigate potential associations in potential metal-exposed populations such as North Carolina private well users.

The research presented in Chapter 2 assesses the contributions of environmental metals to the prevalence of CHDs using an ecologic study to evaluate the relationship between private well drinking water metals and CHDs in North Carolina. We analyzed North Carolina private well monitoring data and birth defect registry data to identify associations between environmental metals the prevalence of birth defects. This research investigated epidemiologic relationships between specific birth defects, including CHDs, and a panel of environmental metals (e.g. arsenic, cadmium, manganese, and lead) using statewide birth defect registry data from 2003 to 2008. We hypothesized that private well levels of arsenic, cadmium, manganese, and lead were associated with the prevalence of specific birth defects in North Carolina.

In addition to understanding metal-defect associations at the population level, there is a need to mechanistically understand biological pathways that lead to disease. *In*

vitro models enable highly-controlled experimental conditions to investigate the cellular mechanisms underlying biological responses and pathways in metal-exposed human cells. miRNA expression is clearly involved in embryonic cardiac development and is implicated in response to metal exposure, but is previously unexplored as a potential mechanism of metal-induced birth defects. There is a need to understand the mechanistic impact of metals on cardiomyocyte miRNA expression and subsequent control of gene expression. Identification of genomic and epigenetic modifiers in metal-perturbed heart function using *in vitro* culture of human cardiomyocytes has not been previously described.

The research presented in Chapter 3 assesses the potential contributions of environmental metals to the etiology of CHDs using an *in vitro* cardiomyocyte model to identify underlying metal-altered molecular pathways. We used *in vitro* heart cell culture to assess genome-wide gene expression and miRNA expression levels in human cardiomyocyte cells. This research toxicogenomically identified pathways that are dysregulated in human heart cells as a result of exposure to an environmentally relevant cadmium dose. We hypothesized that *in vitro* metal exposure would alter cardiomyocyte contractility and gene expression through an epigenetic mechanism, namely miRNA expression and key signaling pathways were identified using a systems biology approach.

References

1. Petrini, J., et al., *Contribution of birth defects to infant mortality in the United States*. Teratology, 2002. **66**: p. S3-S6.

2. Russo, C.A. and A. Elixhauser, *Hospitalizations for Birth Defects, 2004: Statistical Brief #24*, in *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*2006: Rockville (MD).

3. Murray, C.J. and A.D. Lopez, *Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study*. Lancet, 1997. **349**(9063): p. 1436-42.

4. Brent, R.L. and D.A. Beckman, *Environmental teratogens*. Bull N Y Acad Med, 1990. **66**(2): p. 123-63.

5. Hoffman, J.I. and S. Kaplan, *The incidence of congenital heart disease*. J Am Coll Cardiol, 2002. **39**(12): p. 1890-900.

6. Jenkins, K.J., et al., *Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics*. Circulation, 2007. **115**(23): p. 2995-3014.

7. Botto, L.D., et al., *Seeking causes: Classifying and evaluating congenital heart defects in etiologic studies*. Birth Defects Res A Clin Mol Teratol, 2007. **79**(10): p. 714-27.

8. Hoffman, J.I., S. Kaplan, and R.R. Liberthson, *Prevalence of congenital heart disease*. Am Heart J, 2004. **147**(3): p. 425-39.

9. Pierpont, M.E., et al., *Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics*. Circulation, 2007. **115**(23): p. 3015-38.

10. Ransom, J. and D. Srivastava, *The genetics of cardiac birth defects*. Semin Cell Dev Biol, 2007. **18**(1): p. 132-9.

11. Bentham, J. and S. Bhattacharya, *Genetic mechanisms controlling cardiovascular development*. Ann N Y Acad Sci, 2008. **1123**: p. 10-9.

12. Cedergren, M.I., et al., *Chlorination byproducts and nitrate in drinking water and risk for congenital cardiac defects*. Environ Res, 2002. **89**(2): p. 124-30.

13. Desrosiers, T.A., et al., *Paternal occupation and birth defects: findings from the National Birth Defects Prevention Study*. Occup Environ Med, 2012. **69**(8): p. 534-42.

14. Hansen, C.A., et al., *Ambient air pollution and birth defects in brisbane, australia*. Plos One, 2009. **4**(4): p. e5408.

15. Wigle, D.T., et al., *Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants*. J Toxicol Environ Health B Crit Rev, 2008. **11**(5-6): p. 373-517.

16. ATSDR. *Toxicological Profile for Arsenic*. 2007 2007 [cited 2011 July 1]; Available from: http://www.atsdr.cdc.gov/ToxProfiles/TP.asp?id=22&tid=3.

17. ATSDR. *Toxicological Profile for Lead*. 2007 [cited 2011 July 1].

18. ATSDR. *Toxicological Profile for Manganese*. 2008 [cited 2011 July 1].

19. ATSDR. *Toxicological Profile for Cadmium*. 2008 [cited 2011 July 1].

20. Woodruff, T.J., A.R. Zota, and J.M. Schwartz, *Environmental chemicals in pregnant women in the United States: NHANES 2003-2004*. Environ Health Perspect, 2011. **119**(6): p. 878-85.

21. NRC, *Arsenic in drinking water: 2001 Update*, ed. N.R. Council2001, Washington, DC: National Academy Press.

22. Nriagu, J.O. and J.M. Pacyna, *Quantitative assessment of worldwide contamination of air, water and soils by trace metals*. Nature, 1988. **333**(6169): p. 134-9.

23. Gilbert-Diamond, D., et al., *Rice consumption contributes to arsenic exposure in US women*. Proc Natl Acad Sci U S A, 2011. **108**(51): p. 20656-60.

24. Mukherjee, A., et al., *Arsenic contamination in groundwater: A global perspective with emphasis on the Asian scenario*. Journal of Health Population and Nutrition, 2006. **24**(2): p. 142-163.

25. Jarup, L., et al., *Health effects of cadmium exposure--a review of the literature and a risk estimate*. Scand J Work Environ Health, 1998. **24 Suppl 1**: p. 1-51.

26. Guan, H., et al., *Maternal and fetal exposure to four carcinogenic environmental metals*. Biomed Environ Sci, 2010. **23**(6): p. 458-65.

27. EPA. *Actions You Can Take to Reduce Lead in Drinking Water*. 1993 [cited 2013 June 2]; Available from: http://water.epa.gov/drink/info/lead/lead1.cfm.

28. Norman, E.H., et al., *Rural-urban blood lead differences in North Carolina children*. Pediatrics, 1994. **94**(1): p. 59-64.

29. Bouchard, M.F., et al., *Intellectual impairment in school-age children exposed to manganese from drinking water*. Environ Health Perspect, 2011. **119**(1): p. 138-43.

30. Elsner, R.J. and J.G. Spangler, *Neurotoxicity of inhaled manganese: public health danger in the shower?* Med Hypotheses, 2005. **65**(3): p. 607-16.

31. EPA *Drinking Water Contaminants. Available online at:* http://water.epa.gov/drink/contaminants/index.cfm. 2012.

32. Kenny, J.F., et al., *Estimated use of water use in the United States in 2005*, 2009, U.S. Geological Survey. p. 1-60.

33. Kumar, A., et al., *Arsenic exposure in US public and domestic drinking water supplies: A comparative risk assessment*. Journal of Exposure Science and Environmental Epidemiology, 2010. **20**(3): p. 245-254.

34. Sanders, A.P., et al., *Arsenic in North Carolina: public health implications*. Environ Int, 2012. **38**(1): p. 10-6.

35. Miranda, M.L., D.C. Dolinoy, and M.A. Overstreet, *Mapping for prevention: GIS models for directing childhood lead poisoning prevention programs*. Environmental Health Perspectives, 2002. **110**(9): p. 947-953.

36. Rudge, C.V., et al., *The placenta as a barrier for toxic and essential elements in paired maternal and cord blood samples of South African delivering women*. Journal of Environmental Monitoring, 2009. **11**(7): p. 1322-1330.

37. Salpietro, C.D., et al., *Cadmium concentration in maternal and cord blood and infant birth weight: a study on healthy non-smoking women*. Journal of Perinatal Medicine, 2002. **30**(5): p. 395-399.

38. Fechter, L.D., *Distribution of manganese in development*. Neurotoxicology, 1999. **20**(2-3): p. 197-201.

39. Landrigan, P.J., et al., *Environmental pollutants and disease in American children: Estimates of morbidity, mortality, and costs for lead poisoning, asthma, cancer, and developmental disabilities*. Environmental Health Perspectives, 2002. **110**(7): p. 721-728.

40. Trasande, L., P.J. Landrigan, and C. Schechter, *Public health and economic consequences of methyl mercury toxicity to the developing brain*. Environmental Health Perspectives, 2005. **113**(5): p. 590-596.

41. Mahaffey, K.R., R.P. Clickner, and R.A. Jeffries, *Adult Women's Blood Mercury Concentrations Vary Regionally in the United States: Association with Patterns of Fish Consumption (NHANES 1999-2004)*. Environmental Health Perspectives, 2009. **117**(1): p. 47-53.

42. Hightower, J.M., A. O'Hare, and G.T. Hernandez, *Blood mercury reporting in NHANES: Identifying Asian, Pacific Islander, Native American, and multiracial groups*. Environmental Health Perspectives, 2006. **114**(2): p. 173-175.

43. CDC (Eds.: Ettinger, A.S., Wengrovitz, A.G., Portier, C., Brown, M.J.), *Guidelines for the identification and management of lead exposure in pregnant and lactating women*, D.o.E.a.E.H.S. National Center for Environmental Health, Editor 2010, DHHS: Atlanta, GA.

44. NCDHHS, *North Carolina Strategic Plan to Eliminate Childhood Lead Poisoning: North Carolina Childhood Lead Poisoning Prevention Program (NCCLPPP)*, in *Children's Environmental Health Branch*2009.

45. Sanders, A.P., et al., *Toward prenatal biomonitoring in North Carolina: Assessing arsenic, cadmium, mercury, and lead levels in pregnant women* Plos One, 2012. **in press.**

46. Black, B.L., *Transcriptional pathways in second heart field development*. Semin Cell Dev Biol, 2007. **18**(1): p. 67-76.

47. Mahler, G.J. and J.T. Butcher, *Cardiac developmental toxicity*. Birth Defects Res C Embryo Today, 2011. **93**(4): p. 291-7.

48. Barnett, P., M. van den Boogaard, and V. Christoffels, *Localized and temporal gene regulation in heart development*. Curr Top Dev Biol, 2012. **100**: p. 171-201.

49. Vahter, M., *Effects of arsenic on maternal and fetal health*. Annu Rev Nutr, 2009. **29**: p. 381-99.

50. Bellinger, D.C., *Teratogen update: lead and pregnancy*. Birth Defects Res A Clin Mol Teratol, 2005. **73**(6): p. 409-20.

51. Grazuleviciene, R., et al., *Effects of Elevated Levels of Manganese and Iron in Drinking Water on Birth Outcomes*. Polish Journal of Environmental Studies, 2009. **18**(5): p. 819-825.

52. Axelrad, D.A., et al., *Dose-response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data*. Environ Health Perspect, 2007. **115**(4): p. 609-15.

53. Ericson, J.E., et al., *Prenatal manganese levels linked to childhood behavioral disinhibition*. Neurotoxicol Teratol, 2007. **29**(2): p. 181-7.

54. Kim, Y., et al., *Prenatal lead and cadmium co-exposure and infant neurodevelopment at 6 months of age: The Mothers and Children's Environmental Health (MOCEH) study*. Neurotoxicology, 2012. **35C**: p. 15-22.

55. Yu, X.D., et al., *Prenatal exposure to multiple toxic heavy metals and neonatal neurobehavioral development in Shanghai, China*. Neurotoxicol Teratol, 2011. **33**(4): p. 437-43.

56. Nishijo, M., et al., *Effects of maternal exposure to cadmium on pregnancy outcome and breast milk*. Occupational and Environmental Medicine, 2002. **59**(6): p. 394-396.

57. Kuhnert, B.R., et al., *The relationship between cadmium, zinc, and birth weight in pregnant women who smoke*. Am J Obstet Gynecol, 1987. **157**(5): p. 1247-51.

58. Sikorski, R., et al., *Smoking during pregnancy and the perinatal cadmium burden*. J Perinat Med, 1988. **16**(3): p. 225-31.

59. Ahmad, S.A., et al., *Arsenic in drinking water and pregnancy outcomes*. Environ Health Perspect, 2001. **109**(6): p. 629-31.

60. von Ehrenstein, O.S., et al., *Pregnancy outcomes, infant mortality, and arsenic in drinking water in West Bengal, India*. Am J Epidemiol, 2006. **163**(7): p. 662-9.

61. Dorea, J.G. and C.M. Donangelo, *Early (in uterus and infant) exposure to mercury and lead*. Clinical Nutrition, 2006. **25**(3): p. 369-376.

62. Aschengrau, A., S. Zierler, and A. Cohen, *Quality of community drinking water and the occurrence of late adverse pregnancy outcomes*. Arch Environ Health, 1993. **48**(2): p. 105-13.

63. Brender, J.D., et al., *Maternal exposure to arsenic, cadmium, lead, and mercury and neural tube defects in offspring*. Environmental Research, 2006. **101**(1): p. 132-139.

64. 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.

65. Huang, J., et al., *Effect of exposure to trace elements in the soil on the prevalence of neural tube defects in a high-risk area of China*. Biomed Environ Sci. **24**(2): p. 94-101.

66. Kwok, R.K., R.B. Kaufmann, and M. Jakariya, *Arsenic in drinking-water and reproductive health outcomes: A study of participants in the Bangladesh integrated nutrition programme*. Journal of Health Population and Nutrition, 2006. **24**(2): p. 190-205.

67. Macdonell, J.E., H. Campbell, and D.H. Stone, *Lead levels in domestic water supplies and neural tube defects in Glasgow*. Archives of Disease in Childhood, 2000. **82**(1): p. 50-53.

68. Shalat, S.L., D.B. Walker, and R.H. Finnell, *Role of arsenic as a reproductive toxin with particular attention to neural tube defects*. J Toxicol Environ Health, 1996. **48**(3): p. 253-72.

69. Zierler, S., et al., *Chemical quality of maternal drinking water and congenital heart disease*. Int J Epidemiol, 1988. **17**(3): p. 589-94.

70. Song, G., et al., *Effects of choline on sodium arsenite-induced neural tube defects in chick embryos*. Food and Chemical Toxicology, 2011. **In Press, Uncorrected Proof**.

71. Machado, A.F., et al., *Teratogenic response to arsenite during neurulation: relative sensitivities of C57BL/6J and SWV/Fnn mice and impact of the splotch allele*. Toxicol Sci, 1999. **51**(1): p. 98-107.

72. Wang, A., et al., *Reproductive and developmental toxicity of arsenic in rodents: a review*. Int J Toxicol, 2006. **25**(5): p. 319-31.

73. Doi, T., et al., *Epigenetic effect of cadmium on global de novo DNA hypomethylation in the cadmium-induced ventral body wall defect (VBWD) in the chick model*. Toxicol Sci, 2011. **120**(2): p. 475-80.

74. Hafeman, D., et al., *Association between manganese exposure through drinking water and infant mortality in Bangladesh*. Environ Health Perspect, 2007. **115**(7): p. 1107-12.

75. Cheng, Z., et al., *Limited temporal variability of arsenic concentrations in 20 wells monitored for 3 years in Araihazar, Bangladesh*. Environmental Science & Technology, 2005. **39**(13): p. 4759-66.

76. Blom, H.J., et al., *Neural tube defects and folate: case far from closed*. Nature Reviews Neuroscience, 2006. **7**(9): p. 724-731.

77. Reefhuis, J. and M.A. Honein, *Maternal age and non-chromosomal birth defects, Atlanta--1968-2000: teenager or thirty-something, who is at risk?* Birth Defects Res A Clin Mol Teratol, 2004. **70**(9): p. 572-9.

78. Canfield, M.A., et al., *National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999-2001*. Birth Defects Research Part a-Clinical and Molecular Teratology, 2006. **76**(11): p. 747-756.

79. Agha, M.M., et al., *Socioeconomic status and prevalence of congenital heart defects: does universal access to health care system eliminate the gap?* Birth Defects Res A Clin Mol Teratol, 2011. **91**(12): p. 1011-8.

80. Freeman, S.B., et al., *Ethnicity, sex, and the incidence of congenital heart defects: a report from the National Down Syndrome Project*. Genet Med, 2008. **10**(3): p. 173-80.

81. Oyen, N., et al., *Recurrence of congenital heart defects in families*. Circulation, 2009. **120**(4): p. 295-301.

82. Vahter, M., *Health effects of early life exposure to arsenic*. Basic Clin Pharmacol Toxicol, 2008. **102**(2): p. 204-11.

83. Jackson, L.W., et al., *Parental lead exposure and total anomalous pulmonary venous return*. Birth Defects Res A Clin Mol Teratol, 2004. **70**(4): p. 185-93.

84. Martinez-Zamudio, R. and H.C. Ha, *Environmental epigenetics in metal exposure*. Epigenetics, 2011. **6**(7): p. 820-827.

85. Salnikow, K. and A. Zhitkovich, *Genetic and epigenetic mechanisms in metal carcinogenesis and cocarcinogenesis: nickel, arsenic, and chromium*. Chem Res Toxicol, 2008. **21**(1): p. 28-44.

86. Robinson, J.F., et al., *Methylmercury induced toxicogenomic response in C57 and SWV mouse embryos undergoing neural tube closure*. Reprod Toxicol, 2010. **30**(2): p. 284-91.

87. Holson, J.F., et al., *Appropriate use of animal models in the assessment of risk during prenatal development: an illustration using inorganic arsenic*. Teratology, 2000. **62**(1): p. 51-71.

88. Robinson, J.F., et al., *Arsenic- and cadmium-induced toxicogenomic response in mouse embryos undergoing neurulation*. Toxicol Appl Pharmacol, 2011. **250**(2): p. 117-29.

89. Domingo, J.L., *Metal-induced developmental toxicity in mammals: a review*. J Toxicol Environ Health, 1994. **42**(2): p. 123-41.

90. Birge, W.J. and O.W. Roberts, *Toxicity of metals to chick embryos*. Bull Environ Contam Toxicol, 1976. **16**(3): p. 319-24.

91. Gilani, S.H. and Y. Alibhai, *Teratogenicity of metals to chick embryos*. J Toxicol Environ Health, 1990. **30**(1): p. 23-31.

92. Hovland, D.N., Jr., et al., *Differential sensitivity of the SWV and C57BL/6 mouse strains to the teratogenic action of single administrations of cadmium given throughout the period of anterior neuropore closure*. Teratology, 1999. **60**(1): p. 13-21.

93. Thompson, J. and J. Bannigan, *Effects of cadmium on formation of the ventral body wall in chick embryos and their prevention by zinc pretreatment*. Teratology, 2001. **64**(2): p. 87-97.

94. Gerber, G.B., A. Leonard, and P. Hantson, *Carcinogenicity, mutagenicity and teratogenicity of manganese compounds*. Crit Rev Oncol Hematol, 2002. **42**(1): p. 25-34.

95. Ahir, B.K., et al., *Systems biology and birth defects prevention: blockade of the glucocorticoid receptor prevents arsenic-induced birth defects*. Environ Health Perspect, 2013. **121**(3): p. 332-8.

96. Papaconstantinou, A.D., et al., *Mercury, cadmium, and arsenite enhance heat shock protein synthesis in chick embryos prior to embryotoxicity*. Birth Defects Res B Dev Reprod Toxicol, 2003. **68**(6): p. 456-64.

97. Wlodarczyk, B.J., et al., *Arsenic-induced neural tube defects in mice: alterations in cell cycle gene expression*. Reprod Toxicol, 1996. **10**(6): p. 447-54.

98. DeLaughter, D.M., et al., *What chick and mouse models have taught us about the role of the endocardium in congenital heart disease*. Birth Defects Res A Clin Mol Teratol, 2011. **91**(6): p. 511-25.

99. Kruger, O., et al., *Cardiac morphogenetic defects and conduction abnormalities in mice homozygously deficient for connexin40 and heterozygously deficient for connexin45*. J Mol Cell Cardiol, 2006. **41**(5): p. 787-97.

100. Kumai, M., et al., *Loss of connexin45 causes a cushion defect in early cardiogenesis*. Development, 2000. **127**(16): p. 3501-12.

101. Simon, A.M., et al., *Heart and head defects in mice lacking pairs of connexins*. Dev Biol, 2004. **265**(2): p. 369-83.

102. Zepeda, R., et al., *Cardiac tissue injury resistance during myocardial infarction at adulthood by developmental exposure to cadmium*. Cardiovasc Toxicol, 2012. **12**(1): p. 64-72.

103. Matsson, H., et al., *Alpha-cardiac actin mutations produce atrial septal defects*. Hum Mol Genet, 2008. **17**(2): p. 256-65.

104. Ching, Y.H., et al., *Mutation in myosin heavy chain 6 causes atrial septal defect*. Nat Genet, 2005. **37**(4): p. 423-8.

105. Schott, J.J., et al., *Congenital heart disease caused by mutations in the transcription factor NKX2-5*. Science, 1998. **281**(5373): p. 108-11.

106. Garg, V., et al., *GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5*. Nature, 2003. **424**(6947): p. 443-7.

107. Allen, S.P., et al., *Misexpression of noggin leads to septal defects in the outflow tract of the chick heart*. Dev Biol, 2001. **235**(1): p. 98-109.

108. Bartlett, H., G.J. Veenstra, and D.L. Weeks, *Examining the cardiac NK-2 genes in early heart development*. Pediatr Cardiol, 2010. **31**(3): p. 335-41.

109. Bollati, V. and A. Baccarelli, *Environmental epigenetics*. Heredity, 2010. **105**(1): p. 105-112.

110. Wilson, A.G., *Epigenetic Regulation of Gene Expression in the Inflammatory Response and Relevance to Common Diseases*. Journal of Periodontology, 2008. **79**(8): p. 1514-1519.

111. Huang, Y., et al., *Biological functions of microRNAs: a review*. J Physiol Biochem, 2011. **67**(1): p. 129-39.

112. Ying, S.Y., D.C. Chang, and S.L. Lin, *The microRNA (miRNA): overview of the RNA genes that modulate gene function*. Mol Biotechnol, 2008. **38**(3): p. 257-68.

113. Kou, C.Y., et al., *Epigenetic regulation of neonatal cardiomyocytes differentiation*. Biochem Biophys Res Commun, 2010. **400**(2): p. 278-83.

114. Cordes, K.R. and D. Srivastava, *MicroRNA regulation of cardiovascular development*. Circ Res, 2009. **104**(6): p. 724-32.

115. Jirtle, R.L. and M.K. Skinner, *Environmental epigenomics and disease susceptibility*. Nat Rev Genet, 2007. **8**(4): p. 253-62.

116. Reamon-Buettner, S.M., V. Mutschler, and J. Borlak, *The next innovation cycle in toxicogenomics: environmental epigenetics*. Mutat Res, 2008. **659**(1-2): p. 158-65.

117. Fragou, D., et al., *Epigenetic mechanisms in metal toxicity*. Toxicology Mechanisms and Methods, 2011. **21**(4): p. 343-352.

118. Baccarelli, A. and V. Bollati, *Epigenetics and environmental chemicals*. Curr Opin Pediatr, 2009. **21**(2): p. 243-51.

119. Cheng, T.F., S. Choudhuri, and K. Muldoon-Jacobs, *Epigenetic targets of some toxicologically relevant metals: a review of the literature*. J Appl Toxicol, 2012.

120. Sanders, A.P., et al., *Cadmium-associated patterns of DNA methylation in leukocytes from mother-baby pairs* in revision.

121. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res, 2009. **19**(1): p. 92-105.

122. Liu, N. and E.N. Olson, *MicroRNA regulatory networks in cardiovascular development*. Dev Cell, 2010. **18**(4): p. 510-25.

123. Zhao, Y., E. Samal, and D. Srivastava, *Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis*. Nature, 2005. **436**(7048): p. 214-20.

124. Ventura, A., et al., *Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters*. Cell, 2008. **132**(5): p. 875-86.

125. Hou, L., D. Wang, and A. Baccarelli, *Environmental chemicals and microRNAs*. Mutat Res, 2011. **714**(1-2): p. 105-12.

126. Bollati, V., et al., *Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes*. Environ Health Perspect, 2010. **118**(6): p. 763-8.

127. Kong, A.P., et al., *Associations between microRNA (miR-21, 126, 155 and 221), albuminuria and heavy metals in Hong Kong Chinese adolescents*. Clin Chim Acta, 2012. **413**(13-14): p. 1053-7.

128. Fabbri, M., et al., *Whole genome analysis and microRNAs regulation in HepG2 cells exposed to cadmium*. ALTEX, 2012. **29**(2): p. 173-82.

129. Chowdhury, S., et al., *Maternal Genome-Wide DNA Methylation Patterns and Congenital Heart Defects*. Plos One, 2011. **6**(1): p. 11.

130. Obermann-Borst, S.A., et al., *Congenital heart defects and biomarkers of methylation in children: a case-control study*. European Journal of Clinical Investigation, 2011. **41**(2): p. 143-150.

CHAPTER 2: ASSOCIATION BETWEEN ARSENIC, CADMIUM, MANGANESE, AND LEAD LEVELS IN PRIVATE WELLS AND BIRTH DEFECTS PREVALENCE IN NORTH CAROLINA: AN ECOLOGIC STUDY

Overview

Heavy metals including arsenic, cadmium, lead, and manganese are known human carcinogens and/or developmental toxicants that are able to cross the placental barrier from mother to fetus. In this population-based study, we assess the association between metal concentrations in private well water and birth defect prevalence in North Carolina. We conducted a semi-ecologic study including 20,151 infants born between 2003 and 2008 with selected birth defects (cases) identified by the North Carolina Birth Defects Monitoring Program, and 668,371 non-malformed infants (controls). We geocoded maternal residences at delivery and over 20,000 well locations measured for metals by the North Carolina Division of Public Health. The average level of each metal was calculated among wells sampled within North Carolina census tracts. Individual exposure was assigned as the average metal level of the census tract that contained the geocoded maternal residence. Prevalence ratios (PR) with 95% confidence intervals (CI) were calculated to estimate the association between the prevalence of birth defects in the highest category (≥90th percentile) of average census tracts metal level and compared to the lowest category (≤50th percentile). Statewide, private well metal levels exceeded the EPA Maximum Contaminant Level in 2.2, 0.1, 3.9, and 21 percent of wells tested for arsenic, cadmium, lead, and manganese, respectively. The highest category of arsenic exposure was weakly associated with a higher prevalence of conotruncal heart defects

(PR: 1.3 95% CI: 0.9-1.8). The highest category of manganese exposure was significantly associated with a higher prevalence of atrioventricular septal heart defects (PR: 1.8 95% CI: 1.1-3.1) and was modestly associated with conotruncal heart defects (PR: 1.3 95% CI: 1.0-1.9). These findings suggest an ecologic association between concentrations of certain metals in drinking water and specific birth defects.

Introduction

Heavy metals including arsenic, cadmium, lead, and manganese are known human carcinogens and/or developmental toxicants [1-4]. These metals are able to cross the placental barrier potentially resulting in in utero exposure during a critical period of developmental susceptibility [5-7]. There are many detrimental developmental health outcomes potentially associated with metal exposure including spontaneous abortion, stillbirth, low birth weight, preterm birth, reduced fetal growth, impaired neurodevelopment, and congenital malformation [8-15].

Birth defects are a leading cause of infant mortality in the U.S. [16, 17], yet 60-70% have no known cause [18]. In epidemiologic studies, toxic metals including arsenic, cadmium, and lead have been associated with congenital structural malformations (reviewed by [9]). Metal exposure from arsenic- and lead-contaminated drinking water has been associated with an increased occurrence of congenital heart defects (CHDs) [19-21] as well as neural tube defects (NTDs) in human populations [22, 23]. The teratogenicity of arsenic, cadmium, manganese, and lead has been demonstrated in animal models, specifically affecting cardiac, limb, musculoskeletal, craniofacial, and central nervous system development [10, 24-26]. It is clear from animal studies that metal

exposure can result in structural malformation; however, limited population-based human studies have examined such relationships.

Diet and drinking water are common sources of metal exposure in non-occupationally exposed individuals. In North Carolina, 2.3 million residents use private wells for drinking water [27], and the water quality of these wells is not federally regulated [28]. Notably, several statewide assessments of North Carolina private wells have indicated geographical regions with naturally-occurring arsenic [29-31], with levels in some wells as high as 806 ppb - a level that is 80 times the current EPA Maximum Contaminant Level (MCL) standard (10 ppb) in public distribution systems [29]. These data support that ingestion of metal-contaminated drinking water may be a public health concern in North Carolina. To date, no studies that examined the association between metals in private wells and birth defect prevalence have been carried out in North Carolina.

In the present study we examined private well water levels of arsenic, cadmium, manganese, and lead across North Carolina, and used an ecologic study design to assess the association between metal levels and specific birth defect phenotypes. Individual level data on maternal water consumption were not available; therefore, we conducted a semi-ecologic study by assigning maternal drinking water metal exposure based on geocoded residence within a census tract. We compared the prevalence of twelve specific defect types or groups of defects in census tracts with the highest average metal levels to the prevalence in census tracts with the lowest average metal levels. This work represents the first effort to assess the association between statewide levels of metals in North Carolina wells with birth defects.

Methods

Study design and study population

We conducted a statewide semi-ecologic study including 20,151 eligible cases and 668,371 non-malformed controls in North Carolina born between 2003 and 2008. To be eligible for this study, participants were liveborn singleton infants born between 2003 and 2008, with a geocoded residence at delivery within a North Carolina census tract. In addition, eligible case infants could not have any known chromosomal abnormalities. Infants from non-singleton births (n=25,069), without a geocoded residence at delivery (n=38,216), or case infants with known chromosomal abnormalities (n=1,305) were excluded from this study.

Individual-level data regarding maternal source of drinking water and frequency of consumption were not available; therefore our ecologic study used the census tract as the unit of analysis. Our analysis included 797 census tracts across North Carolina. Private well water data were provided by the North Carolina Department of Health and Human Services within the Division of Public Health. The well water database contained metal concentrations collected from private wells across the state between 1998 and 2010 measured by inductively coupled plasma-mass spectrometry. This database contains historical records of well water samples collected by and analyzed at the State Laboratory of Public Health in Raleigh, North Carolina as we have previously described [29].

Outcome classification

The following twelve structural defects or groups of defects were included in this study: (1) spina bifida without anencephaly (n=218); (2) anotia and microtia (n=94); (3) conotruncal defects including common truncus, Tetralogy of Fallot (TOF), and transposition of the great arteries (TGA) (n=435); (4) atrioventricular septal defects (AVSD) and endocardial cushion defects (ECD) (n=150); (5) hypoplastic left heart syndrome (HLHS) (n=142); (6) cleft palate (CP) (n=351); (7) cleft lip (CL) with or without CP (n=516); (8) esophageal atresia (EA) and tracheo-esophageal fistula (TEF) (n=140); (9) pyloric stenosis (n=1,204); (10) reduction defects of the upper and lower limbs (n=255); (11) gastroschisis (n=215); and (12) hypospadias (n=1,994).

Case infants with selected birth defects were identified by the North Carolina Birth Defects Monitoring Program (BDMP). The BDMP is a statewide active surveillance program covering all 100 North Carolina counties [32]. Cases were liveborn infants identified through systematic review and abstraction of medical records by BDMP field staff. Diagnoses were confirmed by supporting documentation in the medical record, such as surgical or autopsy reports, medical imaging, and physical exams. Selected birth defects in this analysis included a spectrum of phenotypes comprising NTDs, CHDs, gastrointestinal, genitourinary, and musculoskeletal defects. Non-malformed controls infants were identified from birth certificate records during the same years.

Exposure assessment

Environmental metals of interest included arsenic, cadmium, manganese, and lead. We analyzed over 20,000 measurements of metal concentrations in private wells collected by the North Carolina Department of Health and Human Services across the state between 1998 and 2010. Only private wells with a corresponding GPS coordinate or with a complete geocoded address were considered in this study (~67% of wells sampled statewide). GPS coordinates were standardized to a decimal degree format and private well locations were geocoded according to previously described methods [29]. The number of geocoded wells varied by metal where 46,286 had arsenic measures, 22,915 had cadmium measures, 70,671 had manganese measures, and 76,387 had lead measurements. Geocoded well locations were then assigned to North Carolina census tracts corresponding to the 2000 census using ESRI ArcGISTM 9.3 (Redlands, CA).

The average level of each of the four metals (arsenic, cadmium, manganese, and lead) was calculated within each census tract across the state. The distributions of average metal levels were mapped using ESRI ArcGISTM 9.3 software (Redlands, CA). Census tracts with fewer than ten well measures for each contaminant over the total time period were excluded from further analysis (n=352, 44%).

Maternal residence at delivery for cases and controls was geocoded by the North Carolina State Center for Health Statistics. Cases and controls were assigned to census tracts in North Carolina based on geocoded maternal residence at delivery using ESRI ArcGISTM 9.3 (Redlands, CA).

A dichotomous exposure contrast was developed to compare exposed and non-exposed Census tracts at delivery. First, we examined the distribution of average metal

concentrations across all tracts with at least ten well water measures, and calculated the corresponding 50th and 90th percentiles of average concentration for each metal. Then, we assigned an exposure status to each Census tract for each metal based on these percentiles as follows: residences were considered "exposed" for a given metal if the average concentration of the metal within the corresponding tract was equal to or greater than the 90th percentile of metal concentration across all tracts combined; tracts were considered "non-exposed" for a given metal if the average concentration of the metal within the corresponding tract was less than or equal to the 50th percentile of metal concentration across all tracts combined. Thus, the final dichotomous exposure contrast compared "exposed" census tracts (a tract where metal concentrations exceeded the 90th percentile of the distribution of metal concentrations across the state) and "non-exposed" census tract. The resulting 50th and 90th percentiles were as follows: 1.27 and 2.32 ppb for arsenic, 0.54 and 1.82 ppb for cadmium, 47.25 and 133.33 ppb for manganese, and 3.65 and 7.80 ppb for lead.

Covariates

The following maternal and infant characteristics were obtained for cases and controls from the birth certificate: residence at delivery, maternal race/ethnicity, age, education, marital status, parity, diabetes and smoking status as well as infant birth weight, gestational age, gender, and date of birth. Subjects were excluded from this dataset if they were missing data on any covariates including maternal race, age, or education (n=2,092).

Statistical analysis

The final cohort included in the analysis comprised 686,435 births (20,094 cases and 666,341 controls). Estimates of the association between metal concentrations in drinking water and the prevalence of each birth defect within census tracts were calculated by log-linear regression using SAS 9.3 (SAS Institute Inc., Cary, North Carolina). Prevalence ratios (PRs) with 95% confidence intervals (CIs) were calculated to estimate the association between average concentration of each metal within census tracts and the prevalence of birth defects, adjusted for maternal age at delivery (continuous), race (categorized as non-Hispanic white, non-Hispanic Black, or Other) and education status (categorized as mother's highest level of completed education as less high school, high school, or greater than high school). To account for multiple testing, Bonferroni-corrected PRs with 99.9% CI estimates (alpha = 0.001) were also calculated.

To examine potential metal-by-metal interaction, Spearman rank correlation was used to examine the relationship between each pair-wise metal across the ecologic unit. The corresponding correlation coefficient (r) and p-values were calculated.

A multiplicative interaction effect was estimated for areas of both higher arsenic and manganese levels and the prevalence of each birth defect category. Specifically, the prevalence of birth defects among infants residing in census tracts where both arsenic and manganese levels were greater than or equal the 90th percentile were compared to areas where both arsenic and manganese were less than or equal to the 50th percentile.

Sensitivity analyses

A sensitivity analysis was conducted using the census block group as the ecologic unit of analysis to determine whether the pattern of findings observed at the later tract level was robust. The analysis included 5,264 block groups across North Carolina. Using the smaller block group level as the unit of exposure resulted in a finer geographical exposure assessment, but fewer study subjects (roughly 40% of the cohort included with the tract ecologic unit analysis) since a larger proportion of the block groups did not have adequate well sampling data (>75%). Maternal residence at delivery and the private well locations were assigned to North Carolina block groups corresponding to the 2000 census. Exposure levels for each metal were calculated for each block group using the same methods as the tract level analysis. The resulting 50th and 90th percentiles of block group averages were 1.24, 2.80 for arsenic, 0.5 and 1.5 for cadmium, 41.62 and 155.28 for manganese, and 3.32 and 7.43 for lead. PRs and corresponding 95% CIs were calculated using log-linear regression as described in the main analysis conducted with the census tract ecologic unit.

To further refine the exposure assessment, a secondary sensitivity analysis was conducted using a nested cohort of individuals residing outside of public water supply distribution areas. Using North Carolina public water supply area service data (EPA, 2007) and the spatial overlay tool in ArcGIS, the nested analysis was limited to only those subjects with a geocoded maternal residence at delivery outside of a public water supply area. This restricted our analysis to individuals who are more likely to use private drinking water wells (n=117,530, ~17% of the cohort in the main analysis) infants

included in the analysis. PRs and corresponding 95% CIs were calculated identically to the main analysis using the tract and block group ecologic units.

Results

Study population characteristics

Among all eligible infants identified by the BDMP and live birth certificates, 20,151 cases and 668,371 non-malformed control infants were included in this study. Table 1 shows the demographic characteristics of the mothers and infants. Case and control subjects were similar with respect to age, race, and sex. Case infants were more likely to have mothers with educational attainment greater than high school.

Metal levels elevated in some private wells

From the well water database records collected between 1998 and 2010, we analyzed between 22,000 and 72,000 wells with geocoded locations (Supplemental Material, Table S1). Notably, private well metal levels that exceeded the EPA regulatory standard for public drinking water were detected in 1,436 (2.25%) of arsenic, 24 (0.1%) of cadmium, 2,563 (3.9%) of lead, and 15,031 (21.3%) of manganese samples.

The geocoded levels of arsenic, cadmium, manganese, and lead in private wells were summarized to average census tract levels. The average census tract levels of each metal are presented in Figure 1. The range of average tract levels was 0.5 to 21.72 ppb for arsenic, 0.5 to 8.75 ppb for cadmium, 2.5 to 689.66 ppb for lead, and 15 to 730.91 ppb for manganese. The average level of each metal exceeded EPA drinking water guideline

levels in at least one census tract (Supplemental Material, Table S1). Notable trends of differential metal levels are evident across the state. Both arsenic and manganese levels were elevated around the central area of the state. A Spearman rank correlation showed a positive relationship between tract-level arsenic and manganese levels ($r=0.31$ $p<0.001$) (Supplemental Material, Table S2a). Additionally, cadmium levels in wells were elevated in areas near cities (e.g. Raleigh, Charlotte) whereas lead levels were randomly spatially distributed and were not strongly correlated with other metals (Supplemental Table 2a).

To examine the spatial occurrence of each birth defect category within census tracts, we assigned a dichotomous exposure status based on 50th and 90th percentiles of average concentration for each metal. The patterns of these levels are represented in Figure 1, where the dark brown category represents "exposed" and the light yellow category represents "unexposed". This dichotomous exposure assignment was developed to sharpen the exposure contrast by emphasizing the tails of the distribution of metal concentrations in our study area and was developed independently from EPA MCL guidelines for metals in public drinking water supplies. Table 2 reports the frequency of exposed case and control infants among the highest exposure category for each metal. A lower proportion of case infants were exposed to cadmium (n=288, 5.0%), compared to case infants exposed to arsenic (n=531, 9.3%), manganese (443, 7.8%) or lead (n=467, 8.2%).

Crude and adjusted PRs and 95% CIs were calculated to estimate the association between average concentration of each metal within census tracts and the prevalence of birth defects (Supplemental Material, Table S3; Table 3). Table 3 reports the adjusted PR estimates comparing the prevalence among the metal- "exposed" versus "unexposed"

groups. An elevated PR was observed for higher arsenic exposure and the prevalence of conotruncal heart defects (PR: 1.3 95% CI: 0.9-1.8), but the observed 95% CI was consistent with unity. The highest category of cadmium exposure was associated with a significantly decreased prevalence of spina bifida (PR: 0.4 95% CI: 0.2-1.0) and pyloric stenosis (PR: 0.6 95% CI: 0.4-0.8). Notably, the highest category of manganese exposure was significantly associated with a higher prevalence AVSD/ECD (PR: 1.8 95% CI: 1.1-3.1). An elevated PR was also observed for higher manganese exposure and the prevalence of conotruncal heart defects (PR: 1.4 95% CI: 1.0-1.9). Since manganese is an important essential metal, it is noteworthy that no statistically significant inverse relationships were observed between prevalence of any defect and elevated manganese. The highest quartile of lead exposure was significantly associated with a decreased prevalence of AVSD/ECD (PR: 0.4 95% CI: 0.1-0.9). The findings did not suggest patterns of association with the highest categories of cadmium or lead levels and an increased prevalence of any birth defects analyzed in this study. To account for multiple comparisons, Bonferroni corrected 99.9% CIs were calculated (Supplemental Material, Table S4). The protective association between cadmium exposure and pyloric stenosis was the only observed association that remained statistically significant after Bonferroni correction (PR: 0.6 99.9% CI: 0.3-1.0).

We examined the potential interaction between joint exposure to arsenic and manganese and the prevalence of each birth defect. Adjusted PR estimates and 95% CIs revealed elevated point estimates for the interaction effect of arsenic and manganese exposure with defect outcomes including spina bifida, AVSD/ECD, cleft lip with or

without cleft palate, and gastroschisis. However, no statistically significant associations were observed (Supplemental Material, Table S9).

Sensitivity analyses results

To assess the impact of potential exposure misclassification in our ecologic analysis, we conducted sensitivity analyses by i) changing the ecologic unit to census block groups to achieve a geographically finer resolution of exposure, and ii) including only individuals whose maternal residence at delivery was outside public water utility distribution areas. The estimated PRs and 95% CIs for each sensitivity analysis and metal-by-outcome pair are provided (Supplemental Material, Tables S5-S8). Overall, similar trends were observed for analyses excluding individuals living outside of public distribution areas when compared to the main analyses. For both types of sensitivity analyses, the PR estimates had wider confidence intervals due to the reduction in sample size.

The PRs and 95% CIs for selected relationships assessed in the sensitivity analyses are shown in Figure 2. The findings suggest patterns of association consistent with the highest categories of arsenic or manganese levels and conotruncal heart defects as well as manganese levels and spina bifida. For example, the association between manganese and spina bifida was estimated under six conditions including crude and adjusted PRs (95% CI) of 1.1 (0.6-1.8), which increased to 1.9 (0.8-4.6) when restricted to individuals outside of public supply areas. The same estimates using the block group ecologic unit estimated a crude PR (95% CI) of 2.2 (1.1-4.4), an adjusted PR (95% CI) of

2.1 (1.1-4.2), which increased to 2.7 (1.1-7.0) when restricted to individuals outside of public supply areas.

Discussion

We conducted a statewide semi-ecologic study to investigate the potential association between heavy metal concentrations in private well water and birth defects prevalence. Specifically, we examined the association between prevalence of twelve birth defect phenotypes and levels of arsenic, cadmium, manganese, and lead summarized at the census tract ecologic unit. Taken together, our analyses identified three notable findings: i) metal levels are elevated in private wells across the state and in some cases co-occur, ii) increased arsenic levels were modestly associated with conotruncal heart defects, and iii) increased manganese levels were modestly associated with conotruncal heart defects as well as spina bifida.

Elevated environmental metals are spatially correlated in private wells

Spatial trends of private well metal contamination were identified across North Carolina, and some concentrations exceed the EPA standard for public drinking water systems. The MCL for arsenic in public water supplies is 10 μg/L, and 5 μg/L for cadmium [33]. Lead is regulated according to a Treatment Technique (TT) which requires that no more than 10% of wells can exceed 15 μg/L [33]. Manganese has a recommended level in public water supplies as a Secondary Maximum Contaminant Level (SMCL) of 50 μg/L based on taste and is not federally enforceable [33]. Based on these standards, 1,436 (2.25%) of arsenic, 24 (0.1%) of cadmium, 2,563 (3.9%) of lead,

and 15,031 (21.3%) of manganese well samples exceeded guideline levels, respectively. The extent of samples exceeding the SMCL for manganese is substantial and supported by a nationwide study that demonstrated a similar proportion of wells (~21%) exceeding 50 ppb [34].

North Carolina has the fourth largest state population relying on private wells for drinking water; the total number served is surpassed by populations in only Pennsylvania, California, and Michigan [27]. Arsenic has been previously recognized as a contaminant of concern in certain regions of North Carolina [29-31]. Notably, this study identifies at least one census tract and block group where the average metal levels exceed respective EPA standards, and we mapped spatial trends of heavy metals in wells across the state. These data are supported by our previous results that demonstrated arsenic contamination in areas along a geographic area called the Carolina terrane [29]. Here we find that manganese correlates spatially with arsenic. While the co-occurrence of arsenic and manganese has been described in North Carolina geologic conditions, a spatial correlation has not been previously reported with respect to private well levels. Elevated cadmium levels were evident near major cities including Raleigh and Charlotte. Urbanization and anthropogenic activities are contributing factors to groundwater contamination from metals including cadmium and arsenic [35]. Lead can occur at elevated levels in drinking water by leaching from certain faucets, fittings, or water systems [36], which may explain the random spatial appearance of levels across the state.

Arsenic levels modestly associated with heart defect prevalence

Our findings show that residence within regions of higher arsenic levels was modestly associated with a higher prevalence of conotruncal heart defects at the census tract ecologic unit, and showed statistically significant association at the block group ecologic unit. While there is strong evidence of the teratogenic effects of arsenic compounds in animal models, currently inorganic arsenic teratogenicity in humans remains debated [25, 37]. Fetotoxic effects of arsenic exposure include an increase in spontaneous abortion and stillbirth [8, 38]. Perhaps due to early fetal loss, there are limited human epidemiologic studies that demonstrate a positive association between arsenic levels and birth defects. Previous population-based studies have shown associations between drinking water arsenic levels and CHDs [19-21], NTDs [22, 37], and increased prevalence of all birth defects [39]. Arsenic has been implicated in numerous studies of cardiovascular diseases in adult populations [40, 41]; however, its role in abnormal prenatal heart cardiac development requires further study.

Manganese levels associated with heart defects and spina bifida prevalence

Our findings support a potential association between residence in areas of highest manganese levels and the prevalence of conotruncal heart defects. While not statistically significant at the tract ecologic unit, residence in areas of highest manganese levels was significantly associated with increased prevalence of spina bifida at the census block group ecologic unit. While manganese is an essential nutrient at low doses, lifetime exposure to drinking water levels above the EPA health benchmark of 0.3 mg/L may have detrimental effects on the central nervous system [3]. No studies to date have

examined the relationship between drinking water manganese and birth defects in humans. In animals, both increased and decreased levels of manganese been associated with teratogenicity (reviewed in [42]). Furthermore, manganese exposure has been associated with several adverse newborn outcomes including infant mortality [43], intrauterine growth restriction [44] and lower birth weight [11, 45]. Given that manganese is able to cross the placental barrier and potentially bioaccumulate in fetal brain tissue [7], the potential association between manganese and spina bifida is biologically plausible. Dietary exposures and inhalation of manganese aerosols during showering remain important considerations exposure in addition to drinking water [46, 47]. These data suggest that additional studies examining the health effects of manganese in drinking water should be carried out.

Arsenic-by-manganese interaction

We assessed a potential interaction effect between regions of higher arsenic and manganese with the prevalence of each defect category. The results show a potential interactive effect on the association with spina bifida, conotruncal heart defects, atrioventricular septal heart defects, and cleft lip with or without cleft palate. Specifically, the estimates reveal that the prevalence of these defects among infants residing in census tracts where both arsenic and manganese levels were greater than or equal the 90th percentile compared to areas where both arsenic and manganese were less than or equal to the 50th percentile were increased when compared to the PR estimates observed for either metal alone. The findings suggest a potential joint effect in areas of arsenic and manganese co-occurrence, but these analyses had limited power as reflected in the width

of the reported CIs because the total sample size was considerably reduced. This association has not been previously examined in studies of environmental metal exposures and birth defects.

Cadmium and lead levels revealed reduced or null associations

Our results suggest some associations between residence in areas of the highest cadmium and lead levels and a reduced birth defects prevalence. A decreased prevalence in pyloric stenosis was associated with residence in areas of the highest cadmium exposure, a relationship that remained statistically significant after Bonferroni correction. We also observed a protective effect of lead exposure with atrioventricular septal heart defects. There were no statistically significant positive relationships between birth defect prevalence and residence in areas of the highest cadmium or lead levels. At the time of this study, we are aware of only three studies that examined the relationship between drinking water levels of cadmium and/or lead and birth defects in humans [21, 22, 48]. Two studies reported null associations between drinking water levels and NTDs [22, 48], and the other reported an increased odds of heart defects associated with detectable lead levels in drinking water [21]. Blood lead levels or occupational lead exposure have been associated with CHDs [49, 50], NTDs [51, 52], and cleft lip [52, 53]. Further, animal studies confirm cadmium- and lead-induced malformations in animal models [54-57]. While counterintuitive, there is evidence for protective effects of maternal cigarette smoking, a major source of cadmium, and NTDs and conotruncal heart defects [58].

Study limitations and strengths

This study had several limitations that are important when interpreting the results. Data on individual use of well water were unavailable; therefore, any observed associations are limited to the ecologic unit of analysis. Thus, the findings from this study cannot be applied at an individual level, but are suggestive associations that warrant additional follow-up. Similar ecologic approaches have been used to examine associations between metals in private wells and health outcomes [59, 60]. In this study, maternal residence at delivery rather than conception was linked to the ecologic unit for exposure assignment. Maternal mobility during pregnancy may introduce nondifferential exposure classification during pregnancy [61]; however, the effect may be tempered in this study as some maternal mobility could have occurred within the same ecologic unit. This study did not evaluate the relationship between personal water consumption patterns and the risk of defects. These findings require cautious interpretation and additional studies are needed to evaluate the link between metals and specific birth defects.

Exposure assessment remains a challenge among epidemiological studies of birth defects [62, 63]. For example, there are no existing biomonitoring programs in North Carolina to actively assess environmental exposures during the prenatal period. As a result, research paradigms usually require retrospective exposure assessment that is clearly limited by issues of temporality. Specifically, in this study the assigned exposures represent a census tract average over a period of 12 years and BDMP data represent the location of cases at delivery. This study is clearly limited by issues of temporality and does not examine exposure particular to the critical periconceptional period during early pregnancy. Our recent study suggests that prenatal biomonitoring may be warranted in

North Carolina [64], especially among at-risk populations. Such a biomonitoring program would enable a direct assessment of the relationship between environmental exposures and children born with birth defects.

A large proportion of census tracts and block groups were not assigned exposure status due to limited sampling (fewer than 10 wells). Clearly there may be effects of additional confounders that are related to geographical location such as socioeconomic status and/or residence in rural versus urban areas. We attempt to address this by refining our exposure assessment, at the expense of estimate precision, by applying two sensitivity analyses. A sensitivity analysis removed individuals residing in tracts and block groups where a proportion of the geographic area is served by public water. Given the lack of individual-level data, this ecologic method served as an imperfect proxy for type of water supply as it potentially excludes a fraction of individuals who use private wells but may be within the public distribution service area. The results of the sensitivity analyses are encouraging because for some metal-defect trends, PRs estimates using the block group unit and when restricted to individuals outside of public water utility areas suggest that exposure misclassification at the tract level may have biased associations towards the null. In both sensitivity analyses, the CIs reflect a loss of precision due to decreased sample sizes.

This study has several strengths including an active statewide case ascertainment program, which provided a large sample size of cases and controls representing six years of birth defects surveillance in North Carolina. The outcomes we investigated included specific phenotypes or groups of phenotypes based on similar etiologic mechanisms. The

statewide surveillance program allowed exclusion of case infants born with chromosomal abnormalities, as this is a known risk factor for some structural defects [65].

Conclusions

Overall, our findings suggest evidence of a possible relationship between concentrations of arsenic and manganese in drinking water and specific birth defects. Given that our results support previous evidence of an association between metal exposure and conotruncal heart defects as well as spina bifida, these relationships should be prioritized for future studies. Ongoing research in our laboratory is directed at understanding the biological mechanisms that underlie metal-defect associations using animal and cell culture models [66]. Further studies are necessary to validate the potential associations observed in this study and to better understand the complex environmental and genetic mechanisms involved in potentially preventable environmentally-mediated birth defects.

Tables

Table 2-1. Characteristics of eligible case and control infants in the study.

Characteristic	Case infants (n=20,151)	Control infants (n=668,371)
Infant Sex n, (%)		
Female	7,684 (38.1%)	327,514 (49.0%)
Male	12,463 (61.8%)	340,854 (51.0%)
Missing	4 (<0.1%)	3 (<0.1%)
Maternal age (ave ± stdev)	26.8 ± 6.2	26.9 ± 6.1
Missing	0	3 (<0.1%)
Maternal Race n, (%)		
NHW	11,658 (57.9%)	381,252 (57.0%)
NHB	4,751 (23.6%)	152,371 (22.8%)
Other[a]	3,742 (18.6%)	134,748 (20.2%)
Maternal Education[b] n (%)		
< high school	5,035 (25.0%)	151,727 (22.7%)
High school	5,934 (29.4%)	190,391 (28.5%)
> high school	9,125 (45.3%)	324,225 (48.5%)
Missing	57 (0.3%)	2,028 (0.3%)

[a] Other race/ethnicity included individuals reporting Hispanic, Asian, Pacific Islander/Native American and Unknown race.

[b] Highest year of education completed.

Table 2-2. Frequency of exposed case and control infants where exposed is defined as living in a census tract with an average metal level greater than or equal to the 90th percentile.

	Arsenic	Cadmium	Manganese	Lead
1. Spina bifida (n=218)	20	5	17	21
2. Anotia/microtia (n=94)	9	3	7	8
3. Conotruncals (n=435)	42	28	47	37
4. AVSD/ECD (n=150)	14	8	19	5
5. HLHS (n=142)	15	8	8	15
6. Cleft palate (n=351)	40	17	26	25
7. Cleft lip ± CP (n=516)	46	25	42	38
8. EA/TEF (n=140)	10	6	6	9
9. Pyloric stenosis (n=1,204)	127	40	98	106
10. Limb reduction (n=255)	22	13	17	22
11. Gastroschisis (n=215)	15	7	11	17
12. Hypospadias (n=1,994)	171	128	145	164
Total Cases (n=5,714)	531 (9.3%)	288 (5.0%)	443 (7.8%)	467 (8.2%)
Total Controls	57,377	40,294	51,186	52,691

Abbreviations: 'Conotruncals' consists of a group of heart defects including common truncus, Tetralogy of Fallot, and transposition of the great arteries; Atrioventricular septal/Endocardial cushion defects (AVSD/ECD); Hypoplastic left heart syndrome (HLHS); Esophageal atresia also called tracheo-esophageal fistula (EA/TEF); 'Limb reduction' includes reduction defects of the upper and lower limbs.

Table 2-3. Association[a] [PR (95%CI)] between selected metals and birth defects.

Defect	As	Cd	Mn	Pb
1. Spina bifida	1.2 (0.7-1.9)	0.4 (0.2-1.0)*	1.1 (0.6-1.8)	1.3 (0.8-2.1)
2. Anotia/microtia	1.2 (0.6-2.7)	0.5 (0.2-1.7)	0.9 (0.4-2.1)	1.0 (0.5-2.2)
3. Conotruncals	1.3 (0.9-1.8)	1.1 (0.7-1.7)	1.4 (1.0-1.9)	1.0 (0.7-1.5)
4. AVSD/ECD	1.0 (0.6-1.9)	0.7 (0.3-1.5)	1.8 (1.1-3.1)*	0.4 (0.1-0.9)*
5. HLHS	0.9 (0.5-1.6)	0.8 (0.4-1.7)	0.6 (0.3-1.2)	1.4 (0.8-2.5)
6. Cleft palate (CP)	1.1 (0.8-1.6)	0.7 (0.4-1.2)	0.8 (0.6-1.3)	0.7 (0.5-1.1)
7. Cleft lip ± CP	1.0 (0.7-1.4)	0.8 (0.5-1.3)	1.0 (0.7-1.4)	0.9 (0.6-1.3)
8. EA/TEF	1.0 (0.5-2.0)	0.6 (0.3-1.5)	0.6 (0.3-1.5)	1.0 (0.5-2.0)
9. Pyloric stenosis	1.1 (0.9-1.3)	0.6 (0.4-0.8)*	1.0 (0.8-1.2)	1.0 (0.8-1.3)
10. Limb reduction	0.8 (0.5-1.4)	0.8 (0.4-1.4)	0.8 (0.4-1.3)	1.1 (0.7-1.8)
11. Gastroschisis	0.7 (0.4-1.2)	0.8 (0.4-1.7)	0.6 (0.3-1.1)	1.0 (0.6-1.6)
12. Hypospadias	1.0 (0.8-1.2)	1.0 (0.9-1.3)	0.9 (0.8-1.1)	1.1 (0.9-1.3)

[a] Prevalence ratios were adjusted for maternal age, race, and education status.

*$p<0.05$

Abbreviations: 'Conotruncals' consists of a group of heart defects including common truncus, Tetralogy of Fallot, and transposition of the great arteries; Atrioventricular septal/Endocardial cushion defects (AVSD/ECD); Hypoplastic left heart syndrome (HLHS); Esophageal atresia also called tracheo-esophageal fistula (EA/TEF); 'Limb reduction' includes reduction defects of the upper and lower limbs.

Figures

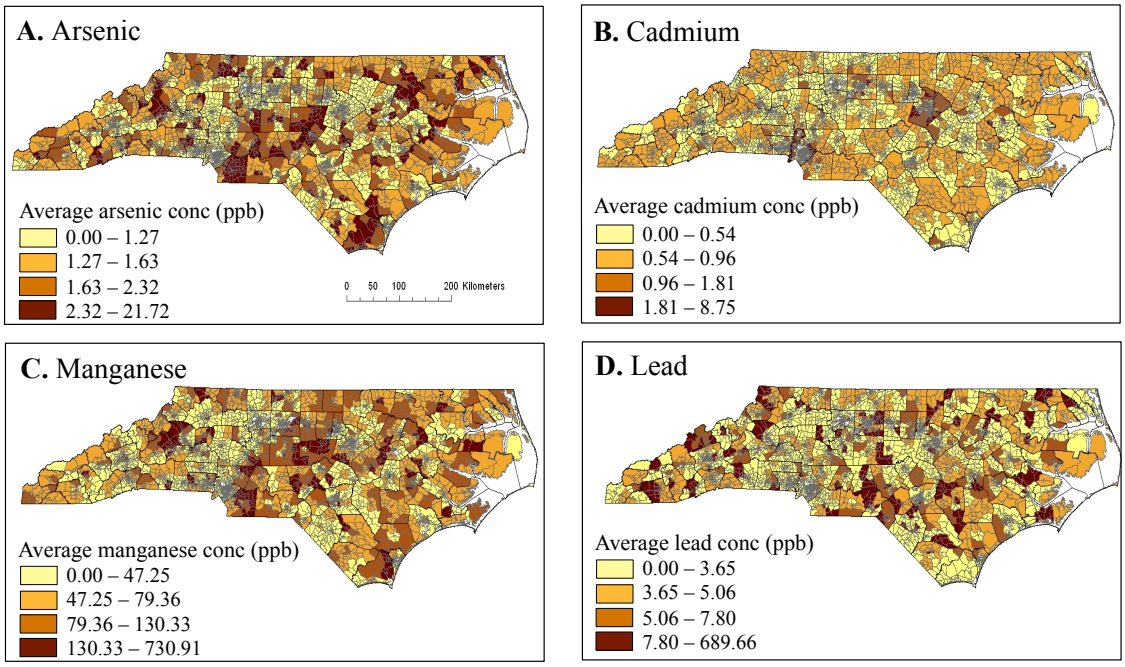

Figure 2-1. Average levels of arsenic (A), cadmium (B), manganese (C), and lead (D) within census tracts across North Carolina. The color gradient represents the calculated exposure percentile categories ($\leq 50^{th}$; $50\text{-}75^{th}$; $75\text{-}90^{th}$; $\geq 90^{th}$ percentiles). The association between birth defects prevalence was tested comparing areas with average metal levels greater than or equal to the 90^{th} percentile (dark brown) versus areas with areas less than or equal to the 50^{th} percentile (light yellow).

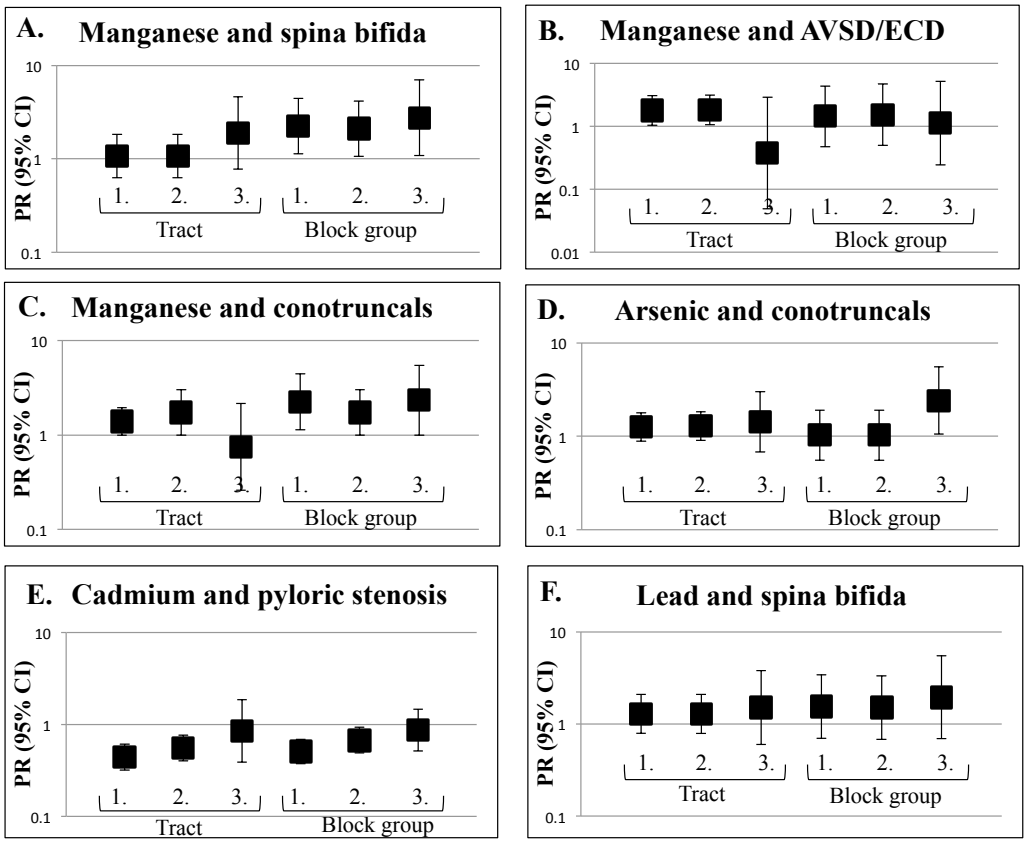

Figure 2-2. Sensitivity analyses plots for selected associations between residence in ecologic units of highest compared to lowest metal levels and prevalence of birth defects including the following metal-defect pairs: A) manganese and spina bifida, B) manganese and atrioventricular septal defects/endocardial cushion defects (AVSD/ECD), C) manganese and conotruncal heart defects, D) arsenic and conotruncal heart defects, E cadmium and pyloric stenosis, and F) lead and spina bifida. Prevalence ratios (PRs) and 95% confidence intervals (CIs) correspond as follows: 1. Crude, 2. Adjusted for maternal age, race and education status, or 3. Restricted to individuals outside public service areas and adjusted for maternal age, race and education status.

Supplemental Tables

Table 2-S1. Number of geocoded wells for each metal as well as number of census tract and block groups with average metal levels above EPA drinking water standards for public distribution systems.

	Total number of geocoded wells	Tracts (n_{Total}=797)	Block Groups (n_{Total}=5,264)	EPA reference level
As	46,286	3 (0.4%)	42 (0.8%)	10 ppb
Cd	22,915	1 (0.1%)	2 (<0.1%)	5 ppb
Mn	70,671	209 (26.2%)	531 (10.1%)	50 ppb
Pb	76,387	19 (2.4%)	47 (0.9%)	15 ppb

Table 2-S2. Spearman rank correlations comparing average (ppb) census tract metal levels (Table 2a) and average (ppb) census block group metal levels (Table 2b). *p<0.05

Table 2-S2a. Tract Spearman correlations

	Cd	Mn	Pb
As	-0.11*	0.31*	-0.02
Cd		0.03	0.06
Mn			0.18*

Table 2-S2b. Block group Spearman correlations

	Cd	Mn	Pb
As	-0.03	0.25*	-0.03
Cd		0.02	0.02
Mn			0.05

Table 2-S3. Crude PRs and 95% CIs using the census tract ecologic unit.

Defect	As	Cd	Mn	Pb
1. Spina bifida	1.2 (0.7, 2.0)	0.4 (0.2, 1.0)	1.1 (0.6, 1.8)	1.3 (0.8, 2.1)
2. Anotia/microtia	1.3 (0.6, 2.7)	0.5 (0.2, 1.6)	0.9 (0.4, 2.2)	1.0 (0.5, 2.2)
3. Conotruncals	1.3 (0.9, 1.8)	1.1 (0.7, 1.7)	1.4 (1.0, 1.9)	1.0 (0.7, 1.5)
4. AVSD/ECD	1.0 (0.6, 1.8)	0.7 (0.3, 1.5)	1.8 (1.0, 3.0)*	0.4 (0.1, 0.9)*
5. HLHS	0.9 (0.5, 1.6)	0.7 (0.4, 1.6)	0.6 (0.3, 1.2)	1.4 (0.8, 2.4)
6. Cleft palate	1.2 (0.8, 1.7)	0.7 (0.4, 1.2)	0.8 (0.6, 1.3)	0.7 (0.5, 1.1)
7. Cleft lip ± CP	1.0 (0.8, 1.4)	0.8 (0.5, 1.2)	1.0 (0.7, 1.4)	0.9 (0.6, 1.2)
8. EA/TEF	1.0 (0.5, 2.0)	0.7 (0.3, 1.6)	0.6 (0.3, 1.5)	1.0 (0.5, 2.0)
9. Pyloric stenosis	1.2 (1.0, 1.5)	0.4 (0.3, 0.6)*	1.0 (0.8, 1.2)	1.0 (0.8, 1.3)
10. Limb reduction	0.9 (0.6, 1.4)	0.7 (0.4, 1.3)	0.8 (0.5, 1.4)	1.1 (0.7, 1.8)
11. Gastroschisis	0.8 (0.5, 1.4)	0.5 (0.2, 1.1)	0.6 (0.3, 1.1)	1.0 (0.6, 1.7)
12. Hypospadias	0.9 (0.8, 1.1)	1.1 (0.9, 1.3)	1.0 (0.8, 1.1)	1.0 (0.9, 1.2)

*p<0.05

Table 2-S4. Bonferroni corrected adjusted PRs and 99.9% CIs using the census tract ecologic unit. PRs were adjusted for maternal age, race, and education status.

Defect	As	Cd	Mn	Pb
1. Spina bifida	1.2 (0.5-2.7)	0.4 (0.1-1.8)	1.1 (0.4-2.6)	1.3 (0.6-2.9)
2. Anotia/microtia	1.2 (0.3-4.5)	0.5 (0.1-3.9)	0.9 (0.2-3.7)	1.0 (0.3-3.7)
3. Conotruncals	1.3 (0.7-2.3)	1.1 (0.5-2.2)	1.4 (0.8-2.4)	1.0 (0.6-1.9)
4. AVSD/ECD	1.0 (0.4-2.8)	0.7 (0.2-2.5)	1.8 (0.7-4.5)	0.4 (0.1-1.6)
5. HLHS	0.9 (0.3-2.4)	0.8 (0.2-2.8)	0.6 (0.2-1.9)	1.4 (0.5-3.7)
6. Cleft palate	1.1 (0.6-2.1)	0.7 (0.3-1.7)	0.8 (0.4-1.7)	0.7 (0.4-1.5)
7. Cleft lip ± CP	1.0 (0.6-1.7)	0.8 (0.4-1.7)	1.0 (0.6-1.8)	0.9 (0.5-1.6)
8. EA/TEF	1.0 (0.3-3.2)	0.6 (0.2-2.7)	0.6 (0.1-2.6)	1.0 (0.3-3.3)
9. Pyloric stenosis	1.1 (0.8-1.5)	0.6 (0.3-1.0)*	1.0 (0.7-1.4)	1.0 (0.7-1.5)
10. Limb reduction	0.8 (0.4-1.9)	0.8 (0.3-2.1)	0.8 (0.3-1.9)	1.1 (0.5-2.4)
11. Gastroschisis	0.7 (0.3-1.8)	0.8 (0.2-3.0)	0.6 (0.2-1.7)	1.0 (0.4-2.3)
12. Hypospadias	1.0 (0.7-1.3)	1.0 (0.8-1.5)	0.9 (0.7-1.3)	1.1 (0.8-1.4)

*p<0.05

Table 2-S5. Adjusted PRs and 95% CIs using the census tract ecologic unit and excluding individuals living in public distribution areas. PRs were adjusted for maternal age, race, and education status.

Defect	As	Cd	Mn	Pb
1. Spina bifida	1.4 (0.6-3.5)	1.4 (0.2-12.2)	1.9 (0.8-4.6)	1.5 (0.6-3.8)
2. Anotia/microtia	1.1 (0.3-3.9)	NE	0.4 (0.1-3.1)	1.5 (0.4-5.4)
3. Conotruncals	1.4 (0.7-3.0)	0.4 (0.0-2.8)	0.7 (0.3-2.2)	1.2 (0.5-2.6)
4. AVSD/ECD	1.1 (0.4-3.5)	0.7 (0.1-5.9)	0.4 (0.0-2.9)	0.3 (0.0-2.2)
5. HLHS	1.2 (0.4-3.4)	0.7 (0.1-5.5)	0.8 (0.2-2.8)	0.7 (0.2-3.1)
6. Cleft palate	1.1 (0.6-2.3)	0.6 (0.1-2.5)	0.3 (0.1-1.0)	0.5 (0.2-1.3)
7. Cleft lip ± CP	0.9 (0.5-1.8)	1.3 (0.5-3.2)	0.7 (0.3-1.8)	0.7 (0.3-1.6)
8. EA/TEF	1.3 (0.5-3.8)	2.4 (0.6-9.2)	0.6 (0.1-4.5)	1.4 (0.5-4.4)
9. Pyloric stenosis	1.2 (0.8-1.7)	0.9 (0.4-1.9)	1.1 (0.7-1.7)	1.1 (0.7-1.7)
10. Limb reduction	0.3 (0.1-1.3)	1.0 (0.2-4.3)	0.5 (0.1-2.0)	0.8 (0.2-2.6)
11. Gastroschisis	1.0 (0.5-2.3)	0.6 (0.1-4.3)	0.6 (0.2-1.9)	1.1 (0.4-2.6)
12. Hypospadias	0.9 (0.7-1.3)	1.5 (1.0-2.3)	0.7 (0.4-1.1)	0.9 (0.6-1.3)

NE -Not estimated.

Table 2-S6. Crude PRs and 95% CIs using the block group ecologic unit.

Defect	As	Cd	Mn	Pb
1. Spina bifida	0.9 (0.4-2.3)	1.1 (0.6-2.1)	2.2 (1.1-4.4)*	1.5 (0.7-3.4)
2. Anotia/microtia	1.1 (0.3-3.7)	0.9 (0.4-2.0)	0.7 (0.2-2.9)	0.4 (0.0-2.7)
3. Conotruncals	1.0 (0.6-1.9)	1.2 (0.8-1.8)	1.7 (1.0-3.0)*	1.1 (0.6-2.1)
4. AVSD/ECD	1.9 (0.7-5.3)	0.4 (0.2-1.3)	1.4 (0.5-4.4)	1.2 (0.3-4.0)
5. HLHS	1.0 (0.3-2.8)	0.7 (0.3-1.7)	0.5 (0.1-1.9)	2.0 (0.9-4.4)
6. Cleft palate	0.8 (0.4-1.7)	0.9 (0.5-1.4)	0.5 (0.2-1.2)	0.8 (0.4-1.5)
7. Cleft lip ± CP	1.4 (0.8-2.4)	0.9 (0.6-1.3)	0.7 (0.3-1.4)	1.2 (0.7-2.1)
8. EA/TEF	1.3 (0.5-3.3)	0.7 (0.3-1.6)	1.3 (0.5-3.4)	1.1 (0.4-3.3)
9. Pyloric stenosis	1.2 (0.8-1.7)	0.5 (0.4-0.7)*	1.0 (0.7-1.5)	0.9 (0.6-1.3)
10. Limb reduction	0.7 (0.2-2.0)	0.7 (0.4-1.4)	1.3 (0.6-2.6)	1.2 (0.5-2.8)
11. Gastroschisis	0.3 (0.1-1.2)	0.4 (0.2-1.0)	0.8 (0.3-2.2)	1.2 (0.5-2.8)
12. Hypospadias	0.9 (0.6-1.2)	1.1 (0.9-1.3)	0.9 (0.7-1.2)	0.8 (0.6-1.1)

Table 2-S7. Adjusted PRs and 95% CIs using the block group ecologic unit. PRs were adjusted for maternal age, race, and education status.

Defect	As	Cd	Mn	Pb
1. Spina bifida	0.9 (0.3-2.2)	1.2 (0.6-2.3)	2.1 (1.1-4.2)*	1.5 (0.7-3.3)
2. Anotia/microtia	1.0 (0.3-3.4)	0.9 (0.4-2.2)	0.7 (0.2-2.9)	0.4 (0.0-2.7)
3. Conotruncals	1.0 (0.6-1.9)	1.2 (0.8-1.8)	1.7 (1.0-3.0)	1.1 (0.6-2.0)
4. AVSD/ECD	1.8 (0.7-5.0)	0.5 (0.2-1.5)	1.5 (0.5-4.7)	1.3 (0.4-4.5)
5. HLHS	0.9 (0.3-2.6)	0.8 (0.3-1.7)	0.5 (0.1-2.0)	2.1 (0.9-4.5)
6. Cleft palate	0.8 (0.4-1.6)	0.9 (0.6-1.5)	0.5 (0.2-1.2)	0.8 (0.4-1.6)
7. Cleft lip ± CP	1.3 (0.8-2.3)	0.9 (0.6-1.4)	0.7 (0.3-1.4)	1.3 (0.7-2.2)
8. EA/TEF	1.3 (0.5-3.4)	0.7 (0.3-1.5)	1.3 (0.5-3.4)	1.1 (0.4-3.3)
9. Pyloric stenosis	1.1 (0.7-1.5)	0.7 (0.5-0.9)*	1.0 (0.7-1.4)	0.9 (0.6-1.3)
10. Limb reduction	0.7 (0.2-1.9)	0.8 (0.4-1.7)	1.2 (0.6-2.5)	1.0 (0.4-2.4)
11. Gastroschisis	0.3 (0.1-1.0)	0.7 (0.3-1.7)	0.8 (0.3-2.1)	1.2 (0.5-2.7)
12. Hypospadias	0.9 (0.7-1.2)	1.1 (0.9-1.3)	0.9 (0.7-1.2)	0.9 (0.6-1.2)

*p<0.05

Table 2-S8. Adjusted PRs and CIs at the block group level excluding individuals living in public distribution areas. PRs were adjusted for maternal age, race, and education status.

Defect	As	Cd	Mn	Pb
1. Spina bifida	0.9 (0.2-4.2)	2.0 (0.7-5.6)	2.7 (1.1-7.0)*	1.9 (0.7-5.5)
2. Anotia/microtia	NE	0.9 (0.2-4.3)	NE	NE
3. Conotruncals	2.4 (1.1-5.5)*	0.8 (0.3-2.4)	2.3 (1.0-5.5)	0.8 (0.3-2.4)
4. AVSD/ECD	1.2 (0.3-5.7)	0.4 (0.1-3.4)	1.1 (0.2-5.2)	NE
5. HLHS	2.2 (0.7-7.0)	1.2 (0.3-4.1)	0.6 (0.1-4.4)	1.6 (0.5-4.9)
6. Cleft palate	0.8 (0.2-2.6)	0.7 (0.2-2.0)	0.6 (0.2-1.9)	1.1 (0.4-2.6)
7. Cleft lip ± CP	1.1 (0.5-2.5)	1.3 (0.6-2.8)	0.4 (0.1-1.5)	0.7 (0.2-1.9)
8. EA/TEF	1.9 (0.5-7.2)	1.4 (0.4-5.2)	0.5 (0.1-4.1)	1.0 (0.2-4.7)
9. Pyloric stenosis	1.0 (0.6-1.7)	0.9 (0.5-1.5)	1.1 (0.6-1.8)	1.0 (0.6-1.7)
10. Limb reduction	1.2 (0.3-4.0)	0.8 (0.2-2.9)	0.9 (0.3-2.7)	1.2 (0.3-4.1)
11. Gastroschisis	0.5 (0.1-2.0)	0.5 (0.1-2.2)	1.0 (0.3-2.9)	0.9 (0.3-2.7)
12. Hypospadias	0.9 (0.5-1.3)	1.1 (0.8-1.5)	0.8 (0.5-1.3)	0.8 (0.5-1.2)

NE -Not estimated.

*p<0.05

Table 2-S9. Crude and adjusted PRs and 95% CIs estimate arsenic-by-manganese interaction using the tract level ecologic unit. PRs were adjusted for maternal age, race, and education status.

Defect	As*Mn Crude	As*Mn Adjusted
1. Spina bifida	2.1 (0.3-16.1)	2.0 (0.3-15.1)
2. Anotia/microtia	0.4 (0.0-4.1)	0.4 (0.0-3.8)
3. Conotruncals	1.1 (0.3-3.5)	1.2 (0.4-3.7)
4. AVSD/ECD	1.3 (0.2-8.3)	1.3 (0.2-8.2)
5. HLHS	4.6 (0.3-66.0)	4.5 (0.3-65.8)
6. Cleft palate	1.1 (0.3-4.6)	1.1 (0.3-4.5)
7. Cleft lip ± CP	3.0 (0.8-11.5)	2.9 (0.8-11.0)
8. EA/TEF	NE	NE
9. Pyloric stenosis	1.1 (0.5-2.3)	1.0 (0.5-2.1)
10. Limb reduction	1.2 (0.2-5.6)	1.5 (0.3-8.0)
11. Gastroschisis	6.4 (0.3-144.1)	6.0 (0.3-134.6)
12. Hypospadias	0.7 (0.4-1.3)	0.7 (0.4-1.4)

NE -Not estimated.

References

1. ATSDR. Toxicological Profile for Arsenic. 2007 2007 [cited 2011 July 1]; Available from: http://www.atsdr.cdc.gov/ToxProfiles/TP.asp?id=22&tid=3.

2. ATSDR. Toxicological Profile for Lead. 2007 [cited 2011 July 1].

3. ATSDR. Toxicological Profile for Manganese. 2008 [cited 2011 July 1].

4. ATSDR. Toxicological Profile for Cadmium. 2008 [cited 2011 July 1].

5. Rudge, C.V., et al., The placenta as a barrier for toxic and essential elements in paired maternal and cord blood samples of South African delivering women. Journal of Environmental Monitoring, 2009. 11(7): p. 1322-1330.

6. Salpietro, C.D., et al., Cadmium concentration in maternal and cord blood and infant birth weight: a study on healthy non-smoking women. Journal of Perinatal Medicine, 2002. 30(5): p. 395-399.

7. Fechter, L.D., Distribution of manganese in development. Neurotoxicology, 1999. 20(2-3): p. 197-201.

8. Vahter, M., Effects of arsenic on maternal and fetal health. Annu Rev Nutr, 2009. 29: p. 381-99.

9. Wigle, D.T., et al., Epidemiologic evidence of relationships between reproductive and child health outcomes and environmental chemical contaminants. J Toxicol Environ Health B Crit Rev, 2008. 11(5-6): p. 373-517.

10. Bellinger, D.C., Teratogen update: lead and pregnancy. Birth Defects Res A Clin Mol Teratol, 2005. 73(6): p. 409-20.

11. Grazuleviciene, R., et al., Effects of Elevated Levels of Manganese and Iron in Drinking Water on Birth Outcomes. Polish Journal of Environmental Studies, 2009. 18(5): p. 819-825.

12. Axelrad, D.A., et al., Dose-response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data. Environ Health Perspect, 2007. 115(4): p. 609-15.

13. Ericson, J.E., et al., Prenatal manganese levels linked to childhood behavioral disinhibition. Neurotoxicol Teratol, 2007. 29(2): p. 181-7.

14. Kim, Y., et al., Prenatal lead and cadmium co-exposure and infant neurodevelopment at 6 months of age: The Mothers and Children's Environmental Health (MOCEH) study. Neurotoxicology, 2012. 35C: p. 15-22.

15. Yu, X.D., et al., Prenatal exposure to multiple toxic heavy metals and neonatal neurobehavioral development in Shanghai, China. Neurotoxicol Teratol, 2011. 33(4): p. 437-43.

16. Petrini, J., et al., Contribution of birth defects to infant mortality in the United States. Teratology, 2002. 66 Suppl 1: p. S3-6.

17. Mathews, T.J. and M.F. MacDorman, Infant mortality statistics from the 2005 period linked birth/infant death data set. National Vital Statistics Report, 2008. 57: p. 1-32.

18. Brent, R.L., Environmental causes of human congenital malformations: the pediatrician's role in dealing with these complex clinical problems caused by a multiplicity of environmental and genetic factors. Pediatrics, 2004. 113(4 Suppl): p. 957-68.

19. Zierler, S., et al., Chemical quality of maternal drinking water and congenital heart disease. Int J Epidemiol, 1988. 17(3): p. 589-94.

20. 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.

21. Aschengrau, A., S. Zierler, and A. Cohen, Quality of community drinking water and the occurrence of late adverse pregnancy outcomes. Arch Environ Health, 1993. 48(2): p. 105-13.

22. Brender, J.D., et al., Maternal exposure to arsenic, cadmium, lead, and mercury and neural tube defects in offspring. Environmental Research, 2006. 101(1): p. 132-139.

23. Shalat, S.L., D.B. Walker, and R.H. Finnell, Role of arsenic as a reproductive toxin with particular attention to neural tube defects. J Toxicol Environ Health, 1996. 48(3): p. 253-72.

24. Thompson, J. and J. Bannigan, Cadmium: Toxic effects on the reproductive system and the embryo. Reproductive Toxicology, 2008. 25(3): p. 304-315.

25. Holson, J.F., et al., Appropriate use of animal models in the assessment of risk during prenatal development: an illustration using inorganic arsenic. Teratology, 2000. 62(1): p. 51-71.

26. Gilani, S.H. and Y. Alibhai, Teratogenicity of metals to chick embryos. J Toxicol Environ Health, 1990. 30(1): p. 23-31.

27. Kenny, J.F., et al., Estimated use of water use in the United States in 2005, 2009, U.S. Geological Survey. p. 1-60.

28. EPA. Arsenic in Drinking Water. 2006 [cited 2010 May 24]; Available from: http://epa.gov/safewater/arsenic/index.html.

29. Sanders, A.P., et al., Arsenic in North Carolina: public health implications. Environ Int, 2012. 38(1): p. 10-6.

30. Pippin, C.G., M. Butczynski, and J. Clayton, Distribution of Total Arsenic in Groundwater in the North Carolina Piedmont, D.o.E.a.N.R.-G. Section, Editor 2003, NCDENR: Mooresville, NC.

31. Kim, D., et al., Spatial Modeling for Groundwater Arsenic Levels in North Carolina. Environmental Science & Technology, 2011. 45(11): p. 4824-4831.

32. SCHS. North Carolina Birth Defects Monitoring Program. 2013; Available from: http://www.schs.state.nc.us/schs/bdmp/. .

33. EPA. Drinking water contaminants: National primary drinking water regulations. 2012 [cited 2013 February 14]; Available from: http://water.epa.gov/drink/contaminants/index.cfm.

34. DeSimone, L.A., P.A. Hamilton, and R.J. Gilliom, The quality of our nation's waters - Quality of water from domestic wells in principal aquifers of the United States, 1991-2004- Overview and major findings U.S. Geological Survey Circular, 2009. 1332: p. 48.

35. Bhagure, G.R. and S.R. Mirgane, Heavy metal concentrations in groundwaters and soils of Thane Region of Maharashtra, India. Environ Monit Assess, 2011. 173(1-4): p. 643-52.

36. EPA. Actions You Can Take to Reduce Lead in Drinking Water. 1993 [cited 2013 June 2]; Available from: http://water.epa.gov/drink/info/lead/lead1.cfm.

37. Shalat, S.L., D.B. Walker, and R.H. Finnell, Role of arsenic as a reproductive toxin with particular attention to neural tube defects. Journal of Toxicology and Environmental Health, 1996. 48(3): p. 253-272.

38. Rahman, A., et al., Association of arsenic exposure during pregnancy with fetal loss and infant death: a cohort study in Bangladesh. Am J Epidemiol, 2007. 165(12): p. 1389-96.

39. Kwok, R.K., R.B. Kaufmann, and M. Jakariya, Arsenic in drinking-water and reproductive health outcomes: A study of participants in the Bangladesh integrated nutrition programme. Journal of Health Population and Nutrition, 2006. 24(2): p. 190-205.

40. Tseng, W.P., Effects and dose--response relationships of skin cancer and blackfoot disease with arsenic. Environ Health Perspect, 1977. 19: p. 109-19.

41. Navas-Acien, A., et al., Arsenic exposure and cardiovascular disease: a systematic review of the epidemiologic evidence. Am J Epidemiol, 2005. 162(11): p. 1037-49.

42. Gerber, G.B., A. Leonard, and P. Hantson, Carcinogenicity, mutagenicity and teratogenicity of manganese compounds. Crit Rev Oncol Hematol, 2002. 42(1): p. 25-34.

43. Hafeman, D., et al., Association between manganese exposure through drinking water and infant mortality in Bangladesh. Environ Health Perspect, 2007. 115(7): p. 1107-12.

44. Vigeh, M., et al., Blood manganese concentrations and intrauterine growth restriction. Reprod Toxicol, 2008. 25(2): p. 219-23.

45. Zota, A.R., et al., Maternal blood manganese levels and infant birth weight. Epidemiology, 2009. 20(3): p. 367-73.

46. Bouchard, M.F., Manganese in Drinking Water: Bouchard Responds. Environmental Health Perspectives, 2011. 119(6): p. A241.

47. Elsner, R.J. and J.G. Spangler, Neurotoxicity of inhaled manganese: public health danger in the shower? Med Hypotheses, 2005. 65(3): p. 607-16.

48. Macdonell, J.E., H. Campbell, and D.H. Stone, Lead levels in domestic water supplies and neural tube defects in Glasgow. Archives of Disease in Childhood, 2000. 82(1): p. 50-53.

49. Jackson, L.W., et al., Parental lead exposure and total anomalous pulmonary venous return. Birth Defects Res A Clin Mol Teratol, 2004. 70(4): p. 185-93.

50. Bound, J.P., et al., Involvement of deprivation and environmental lead in neural tube defects: a matched case-control study. Arch Dis Child, 1997. 76(2): p. 107-12.

51. Cengiz, B., et al., Serum zinc, selenium, copper, and lead levels in women with second-trimester induced abortion resulting from neural tube defects: a preliminary study. Biol Trace Elem Res, 2004. 97(3): p. 225-35.

52. Irgens, A., et al., Reproductive outcome in offspring of parents occupationally exposed to lead in Norway. Am J Ind Med, 1998. 34(5): p. 431-7.

53. Kristensen, P., et al., Perinatal outcome among children of men exposed to lead and organic solvents in the printing industry. Am J Epidemiol, 1993. 137(2): p. 134-44.

54. Webster, W.S. and K. Messerle, Changes in the mouse neuroepithelium associated with cadmium-induced neural tube defects. Teratology, 1980. 21(1): p. 79-88.

55. McClain, R.M. and B.A. Becker, Teratogenicity, fetal toxicity, and placental transfer of lead nitrate in rats. Toxicol Appl Pharmacol, 1975. 31(1): p. 72-82.

56. Thompson, J. and J. Bannigan, Effects of cadmium on formation of the ventral body wall in chick embryos and their prevention by zinc pretreatment. Teratology, 2001. 64(2): p. 87-97.

57. Thompson, J.M. and J.G. Bannigan, Omphalocele induction in the chick embryo by administration of cadmium. J Pediatr Surg, 2007. 42(10): p. 1703-9.

58. Grewal, J., et al., Maternal periconceptional smoking and alcohol consumption and risk for select congenital anomalies. Birth Defects Res A Clin Mol Teratol, 2008. 82(7): p. 519-26.

59. Liu-Mares, W., et al., Pancreatic cancer clusters and arsenic-contaminated drinking water wells in Florida. BMC Cancer, 2013. 13: p. 111.

60. Nieder, A.M., et al., Bladder cancer clusters in Florida: identifying populations at risk. J Urol, 2009. 182(1): p. 46-50; discussion 51.

61. Canfield, M.A., et al., Residential mobility patterns and exposure misclassification in epidemiologic studies of birth defects. J Expo Sci Environ Epidemiol, 2006. 16(6): p. 538-43.

62. Katz, E.A., G.M. Shaw, and D.M. Schaffer, Exposure assessment in epidemiologic studies of birth defects by industrial hygiene review of maternal interviews. Am J Ind Med, 1994. 26(1): p. 1-11.

63. Borja-Aburto, V.H., et al., [Difficulties of the methods for studying environmental exposure and neural tube defects]. Salud Publica Mex, 1999. 41 Suppl 2: p. S124-31.

64. Sanders, A.P., et al., Towards prenatal biomonitoring in North Carolina: assessing arsenic, cadmium, mercury, and lead levels in pregnant women. Plos One, 2012. 7(3): p. e31354.

65. Forrester, M.B. and R.D. Merz, Impact of excluding cases with known chromosomal abnormalities on the prevalence of structural birth defects, Hawaii, 1986-1999. Am J Med Genet A, 2004. 128A(4): p. 383-8.

66. Ahir, B.K., et al., Systems biology and birth defects prevention: blockade of the glucocorticoid receptor prevents arsenic-induced birth defects. Environ Health Perspect, 2013. 121(3): p. 332-8.

CHAPTER 3: miRNA-DEPENDENT AND INDEPENDENT CADMIUM-INDUCED SIGNALING IN HUMAN CARDIOMYOCYTES

Overview

Cadmium is a toxic metal with widespread human exposure resulting from diet, air, and drinking water. Despite the strong evidence of cadmium-associated cardiovascular disease, the mechanisms that underlie metal-induced cardiovascular endpoints remain largely unknown. MicroRNAs (miRNAs) are post-transcriptional regulators of gene expression that influence heart development and function, but the role of miRNAs remains unstudied as a potential mechanism of cadmium-associated disease. This is the first study to assess cadmium-induced changes in miRNA levels in human cardiomyocytes. The objective of this study was to assess miRNAs as potential epigenetic modifiers of gene expression in human cardiomyocytes exposed to cadmium. We exposed human-derived induced pluripotent stem cell cardiomyocytes to 0.5 μM of cadmium chloride (56.3 ppb) for 24 hours. Micro (mi) RNAs and messenger (m) RNAs were isolated and analyzed for differential expression analysis using genome-wide microarrays. A non-cytotoxic dose of 0.5 μM cadmium chloride increased the expression levels of eight miRNAs. Network analysis of the miRNAs revealed that cadmium potentially alters signaling pathways associated with cancer, connective tissue disorders, and inflammatory disease. Analysis of messenger RNA (mRNA) expression showed 31 genes were upregulated in cadmium-exposed cardiomyocytes, and eight (26%) matched to predicted targets of the identified differentially expressed miRNAs. The miRNA-

associated mRNAs mapped to signaling pathways including gene expression, embryonic development, and organismal development. miRNA-independent mRNA changes were seen for 21 genes including a family of metallothionein genes showing increased transcript levels. Cadmium alters miRNA patterns in cardiomyocytes that regulate gene expression and potentially play a role in metal-induced cardiovascular disease.

Introduction

Cadmium is a carcinogen of the liver, kidney, and pancreas as well as a known risk factor for cardiovascular disease [1, 2]. Numerous poor cardiovascular endpoints are associated with cadmium exposure including coronary heart disease [3], peripheral artery disease [4, 5], myocardial infarction [6], as well as stroke and heart failure [7, 8]. Cadmium is a cardiovascular risk factor even at low chronic exposure levels [9, 10]. It is estimated that the Disability Adjusted Life Years (DALYs) due to cardiovascular disease will be 169 million (11% of estimated global DALYs) by 2020 [11]. Cadmium is therefore an environmental risk factor of considerable potential global disease burden [12, 13].

Human exposure to cadmium occurs primarily through diet and cigarette smoke [2, 14]. There is evidence that cadmium can accumulate in the human heart at levels higher than observed in many other organs [15]. In addition to cadmium-associated heart disease later in life, early life exposure to cadmium is associated with altered cardiac development in animal models [16-19], but to date this has not been examined in humans. Despite the strong evidence of cadmium-associated adverse cardiovascular endpoints, an

exact molecular mechanism that links cadmium to heart disease is currently unknown.

The suggested impacts of cadmium exposure include inhibition of DNA repair, generation of reactive oxygen species, and perturbation of cell cycle progression or apoptosis (reviewed in [20]). Cadmium is believed to act as an epigenetic or indirectly genotoxic metal [21]. The relationship between transcriptional mediators and gene expression is understudied and few studies have investigated the complex interplay between the toxicogenomic and potential epigenetic mediators of cadmium-associated heart disease. Indeed, animal studies confirm that cadmium exposure alters cardiac gene expression [16, 17, 22], and we have shown that low dose cadmium exposure in human lymphoblast cells alters genes enriched for biological functions including cancer and cardiovascular disease [23]. Expanding this work, here we examine the transcriptional effects in cardiomyoctyes, with a particular focus on microRNA (miRNAs) as mediators. We hypothesize that metal-induced miRNA changes are a potential epigenetic mechanism that may control downstream gene expression that leads to altered cardiac function or disease.

miRNAs are short single-stranded RNAs approximately 20 nucleotides in length that post-transcriptionally regulate gene expression through silencing. miRNAs are involved in a number of cell proliferation, apoptosis, and developmental processes [24-26]. Currently over 1,100 human miRNAs have been identified [27], and are believed to target approximately 60% of genes [28]. miRNA control of gene expression in heart cells is involved in both heart development and failure (reviewed in [29, 30]). Specifically, miR-133 and miR-1 are miRNAs involved in the regulation of a large number of cell division and maintenance processes, which contribute to altered gene expression and

dysregulated cardiac function and/or development. No studies to date have examined cadmium-induced changes in miRNA levels in human cardiomyocytes.

In the present study, we examine the effects of cadmium at the genomic (mRNA) and epigenetic (miRNA) levels to identify potential mechanisms of metal-perturbed heart function using *in vitro* cardiomyocyte culture. Given that alterations in cardiomyocyte signaling can lead to abnormal heart development or function, this work represents novel insight into miRNA-mediated metal-induced disease.

Methods

Cardiomyocyte culture

iCell® cardiomyocytes derived from human induced-pluripotent stem cells were purchased from Cellular Dynamics International (Madison, WI). Cells were stored and thawed according to the manufacturer's protocol. The iCell® cardiomyocytes were plated at a density of 15,000 cells per well (~47,000 cells/cm^2) in a 24-well plate that was pre-treated with 0.1% gelatin. Cells were initially maintained in iCell Cardiomyocytes Plating Medium then transferred into iCell Cardiomyocytes Maintenance Medium every 48 hours after. Cells were incubated at 37C and 7% CO_2 for 144 hours until they reached a state of synchronous beating. After 144 hours of incubation, cells were exposed to cadmium or received a mock control. Treatment groups included the untreated control and cadmium chloride treatments selected based on environmental relevance. Cadmium chloride solutions were prepared fresh from stock at the following final test concentrations: 0.05 μM (5.6 ppb), 0.1 μM (11.3 ppb), and 0.5 μM (56.3 ppb).

Cytotoxicity analysis

A 96-well plate seeded with iCell Cardiomyocytes at the same density (~47,000 cells/cm^2) was examined for cytotoxicity. After 120 hours incubation, cells received culture medium with or without cadmium chloride treatment for 24 hrs. Treatment groups were divided as follows: the (**i**) untreated control, **(ii)** 0.1 mM staurosporin positive control, and (**iii**) cadmium chloride treatment. Cadmium chloride solutions were prepared fresh from stock at the following final test concentrations: 0.01 µM, 0.05 µM, 0.1 µM, 0.5 µM, 1.0µM, 5.0 µM, 10.0µM, and 50.0µM.

After 24-hour exposure to cadmium chloride, cardiomyocyte viability was assessed using the CellTiter-Glo Luminescent Cell Viability Assay (Promega, G7570; Madison, WI). The assay quantifies adenosine triphosphate (ATP), which signals the presence of metabolically active cells. Luminescence was measured in relative light units (RLU) on the GloMax®-96 Microplate Luminometer (Promega, Madison, WI). Luminescence fold change was calculated as the mean luminescence for cadmium- or staurosporin- exposed samples divided by untreated control values. Mean luminescence was calculated from six replicates under each treatment condition after subtraction of the average background luminescence.

RNA and miRNA isolation

After 24-hour exposure to cadmium chloride, cells were lysed and stored at -80 using the Qiagen AllPrep® DNA/RNA/Protein Mini Kit for simultaneous purification of genomic DNA, total RNA, and total protein from human cells according to kit specifications (Qiagen, Valencia, CA). Simultaneous mRNA (>200 nucleotides) and

miRNA (18-200 nucleotide length) extraction was performed using the supplementary protocol for purification of miRNA and the RNeasy® MinElute® Cleanup Kit (Qiagen, Valencia, CA). miRNA and mRNA were quantified using a NanoDrop spectrophotometer 2000c (Thermo Scientific, Wilmington, DE). mRNA quality was assessed using the 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Samples used for mRNA microarray analysis were required to have an RNA Integrity Number (RIN) value above 9.6.

Microarray analysis

A starting material of 20 ng of the collected small RNA fraction (18-200 nucleotides) was labeled and hybridized to the Agilent Human miRNA Microarray release 16.0, which measures the expression levels of 1,348 human miRNAs across 53,880 probesets (Agilent Technologies, Santa Clara. CA). Three samples the 0.5 μM cadmium chloride-exposed and three unexposed control samples were hybridized using 20 ng input small RNA per sample.

Three samples exposed to 0.5 μM cadmium chloride and three unexposed control samples were amplified, fragmented, and hybridized to Affymetrix GeneChip Human Gene 2.0 ST Arrays (Affymetrix, Santa Clara, CA) using 50 ng input RNA per sample. The Nugen Ovation Pico WTA system V2 and the Nugen Encore™ Biotin module (Nugen, San Carlos, CA) were used to prepare RNA for hybridization. Briefly 50 ng of total RNA was converted into amplified SPIA cDNA using the Nugen Ovation Pico WTA system V2 and accompanying user guide. The amplified SPIA cDNA was purified using the Qiagen QIAquick PCR purification kit. Five μg of SPIA cDNA was fragmented

and labeled using the Nugen Encore Biotin module and accompanying user guide. The hybridization cocktail for hybridization onto Affymetrix GeneChip Human Gene 2.0 ST arrays was prepared using the Nugen Encore Biotin module user guide's cocktail assembly table. The hybridization cocktails were denatured and hybridized onto Affymetrix GeneChip Human Gene 2.0 ST arrays for 18 hours at 45°C in an Affymetrix GeneChip Hybridization Oven 640. After 18 hours, the arrays were washed and stained in the Affymetrix Fluidics Station 450. The arrays were scanned using the Affymetrix GeneChip Scanner 3000 7G Plus with autoloader. The Affymetrix GeneChip Human Gene 2.0 ST Array contains 53,617 probesets corresponding to over 25,000 RefSeq (Entrez) genes.

Statistical analysis

Microarray data were normalized using Robust Multi-Chip Average (RMA) [31]. miRNA data were extracted using Agilent Feature Extraction software and probes with signal intensities with an absolute expression level above a background level (background was set to 26.5, the median signal per array) in all conditions were removed from the analysis to eliminate background noise. Differential expression was defined as a significant difference in mRNA or miRNA expression levels between exposed and unexposed samples, where the following two statistical requirements were met: (i) fold change of ≥ 1.5 or ≤ -1.5 (average exposed versus average unexposed); (ii) p-value < 0.01 (ANOVA). Analysis of variance (ANOVA) p-values were calculated using Partek® Genomics Suite™ software (St. Louis, MO). To control the rate of false positives, false discovery rate q-values were calculated [32]. Data will be submitted to the National

Center for Biotechnology Information's Gene Expression Omnibus repository [33] and the accession numbers will be provided here for the mRNA and miRNA datasets.

Pathway enrichment analysis of cadmium-altered miRNA and mRNA transcripts

We performed enriched biological functions and network analyses using the Ingenuity Pathway Analysis (IPA) tool (Ingenuity® Systems, Redwood City, CA). IPA is a proprietary database that curates gene-phenotype associations, molecular interactions, regulatory events, and chemical knowledge to provide a global molecular network. The set of differentially expressed miRNAs were uploaded into the IPA core pathway analysis module and overlaid onto this global molecular network. Related networks were algorithmically constructed based on connectivity.

Putative miRNA-mRNAs relationships were identified using the IPA microRNA Target Filter. Given the list of cadmium-altered miRNAs, mRNA targets were identified based on a knowledgebase of predicted and experimentally observed relationships. An independent miRNA database was used to verify the predicted mRNA targets [34].

We compared our experimentally identified differentially expressed mRNAs with the list of putative mRNA targets matching the set of differentially expressed miRNAs. This resulted in classification of two distinct subsets of differentially expressed mRNA in samples exposed to cadmium. The miRNA-associated gene set as well as the miRNA-independent geneset was overlaid onto the global molecular network.

Functional analysis was carried out to identify biological functions and disease signatures most significantly associated with the observed differentially expressed miRNAs, as well as the miRNA-associated and miRNA-independent genesets. Statistical

significance of each biological function or disease was calculated using Fisher's exact text with an alpha set at 0.05.

Upstream transcriptional regulators were identified using IPA for each differentially expressed mRNA gene list. The upstream regulator analysis uses prior knowledge of to identify potential regulating molecules involved in signaling cascades that may explain observed gene expression changes. Statistically significant overlap p-values between the uploaded geneset and genes known to be modified by a transcriptional regulator were calculated using Fisher's exact text with an alpha set at 0.05. Each list of transcriptional regulator was filtered for molecules that correspond to known transcription factors.

Results

Cadmium exposure modulates miRNAs in human cardiomyocytes

In this study, we set out to determine whether cadmium exposure alters miRNA expression levels in human cardiomyocytes. Human-derived iCell cardiomyocytes were exposed to 0, 0.05, 0.1 or 0.5 μM cadmium chloride. None of the doses tested resulted in significant cytotoxicity to the cardiomyocytes (Supplemental Material, Table S1). After exposure, we collected small RNAs (miRNAs) and measured their relative abundances using the Agilent Human miRNA Microarray (version 16.0). A total of 33,573 of 53,880 total probesets were detectable above background levels in these cells. A total of 56 probesets corresponding to nine unique miRNAs were upregulated in cadmium-exposed samples compared with control samples (Table 1; Supplemental Material, Table S2).

There were no miRNAs with significantly decreased expression levels in response to cadmium. The nine significantly differentially expressed miRNAs included miR-1, miR-125b-5p, miR-130a-3p, miR-20a-5p, miR-21-5p, miR-22-3p, miR-27b-3p, miR-30e-5p, and miR-1274a (Table 1).

Upregulated miRNAs integrated into biological networks

To identify potential biological pathways impacted by the cadmium-altered miRNAs, we overlaid the nine differentially expressed miRNAs onto molecular pathway maps using IPA. Eight miRNAs constructed a single network that was enriched for functions including cancer, connective tissue disorders, and inflammatory disease ($p<1x10^{-23}$) (Figure 1). These miRNAs have known links to tumor suppressor gene p53 (*TP53*). To note, miR-1274a was not recognized by IPA and was excluded from further analysis.

Cadmium exposure modulates gene expression in human cardiomyocytes

To test whether cadmium alters mRNA expression levels, we assessed mRNA abundance in cardiomyocytes. A total of 83 probesets corresponding to 31 unique genes were differentially expressed in cadmium-exposed samples compared with control samples (Figure 2; Supplemental Material, Table S3). Only five of the 31 genes (16%) showed decreased expression among the cadmium-exposed samples. Among the remaining 26 genes showing increased expression in the cadmium-exposed samples, a family of ten metallothionein (MT) genes was significantly upregulated including *MT1A*, *MT1B*, *MT1CP*, *MT1E*, *MT1F*, *MT1G*, *MT1H*, *MT1L*, *MT1M*, *MT1L*, and *MT1X*

(Supplemental Material, Table S3). Taken together, all 31 significantly differentially expressed genes were mapped to signaling pathways including renal and urological disease, cellular compromise, and psychological disorders ($p<1\times10^{-26}$) (Supplemental Material, Table S4).

Conservation of predicted and observed mRNA targets

To understand downstream genomic changes regulated by the eight cadmium-altered miRNAs, we computationally predicted putative mRNA targets. We identified 6,735 predicted mRNA targets. We compared the list of 6,735 predicted mRNA targets with the 31 significantly differentially expressed genes. There was an overlap between eight (26%) of our experimentally observed differentially expressed genes and the predicted miRNA targets (Table 2). Specifically, the microarray analysis confirmed differential expression of ALX homeobox 4 (*ALX4*), malic enzyme 1 (*ME1*), *MT1F*, *MT1G*, *MT1H*, DNA-directed RNA polymerase III subunit G (*POLR3G*), solute carrier family 30 member 2 (*SLC30A2*), and sushi domain containing 5 (*SUSD5*). Among these modulated mRNAs, multiple miRNAs were common regulators of transcription. *POLR3G* was predicted to be the target of miR-17-5p, miR-27a-3p, and miR-30c-5p. *ALX4* was a predicted target of both miR-17-5p and miR-21-5p (Table 2).

The eight miRNA-associated mRNA transcripts were mapped in IPA to signaling pathways including gene expression, embryonic development, and organismal development ($p<1\times10^{-16}$) (Supplemental Material, Table S4). The remaining miRNA-independent genes (n=23) were mapped to signaling pathways including cell signaling, nucleic acid metabolism and small molecule biochemistry ($p<1\times10^{-41}$) (Supplemental

Material, Table S4).

Enrichment of transcription factor binding sites in cadmium-associated genes

The mRNA transcripts modified by cadmium independently of miRNAs (n=23) as well as those potentially under miRNA control (n=8) were examined for enrichment of upstream transcriptional regulators. This analysis demonstrated enrichment for a set of eight and 16 upstream regulators, respectively between the two lists (Table 3). Notably, three transcription factors were common to both lists including metal regulatory transcription factor 1 (MTF1), lysine (K)-specific demethylase 5B (KDM5B), and histone deacetylase (hdac).

Discussion

Cadmium is a known risk factor for cardiovascular disease [1, 2]; however, the mechanism by which cadmium impacts cardiac function is understudied. This is the first study to assess miRNA expression as a key transcriptional regulator in cardiomyocyte response to cadmium. miRNAs regulate gene expression by post-transcriptional degradation of mRNA or reducing translational efficiency. Each miRNA can regulate the expression of several hundred target genes [28]. Thus, miRNAs represent potential epigenetic modulators that control the downstream expression of genes that may lead to altered cardiac function or disease state. Here, we exposed human cardiomyocytes to 0.5 μM cadmium chloride for 24 hours and assessed genome-wide changes in miRNA and mRNA expression profiles.

The expression levels of over 1,000 miRNA were measured and compared

between cadmium-exposed and unexposed cardiomyocytes. A total of nine miRNAs were identified with altered expression. Network analysis identified that eight of the miRNAs are linked to the *p53* tumor suppressor gene. The *p53* gene encodes the p53 protein which is a critical transcription factor involved in tumor suppression and under normal conditions is protective of cancer progression [35]. Increased levels of p53 are linked to cardiomyocyte apoptosis, which may lead to fatal ischemic heart injury [36, 37].

The individual identified miRNAs included miR-1, miR-20a-5p, miR-21-5p, miR-22-3p, miR-27b-3p, miR-30e-5p, miR-125b-5p, miR-130a-3p, and miR-1274a. Many of the cadmium-responsive miRNAs have known links to cardiac development or function. More specifically, miR-1, miR-21, miR-22, miR-27b, and miR-125b have been shown to have altered expression in tissues exhibiting cardiac conditions such as arrhythmia and hypertrophy compared to normal tissues [38-44]. Of note, miR-1 overexpression negatively regulates cardiac growth, in part by targeting and downregulating heart and neural crest derivatives-expressed protein 2 (Hand2), a transcription factor that promotes cardiomyocyte expansion [45]. In addition, miR-20a belongs to a cluster of miRNAs known as miR-17~92 that are oncogenic [46]; mice deficient in miR-17~92 developed fatal ventricular septal heart defects [47]. Both miR-130a and miR-30e are known regulators of cardiac development [48, 49], but have not been previously associated with abnormal cardiac function or disease states.

To date, only three studies have examined cadmium-associated miRNA changes in human tissues or cell lines [50-52]. Increased expression of miR-146a in peripheral blood leukocytes was associated with occupational cadmium exposure [50]. Decreased expression of miR-21 in urine samples was associated with increased levels of urinary

arsenic and lead, but not cadmium in a cohort of children [52]. A previous *in vitro* study, noted twelve miRNAs down regulated in human hepatoblastoma cells exposed to 10 μM cadmium [51]. The authors identified differential expression of miR-130a in cadmium-treated cells [51], however expression levels were decreased in contrast to the increased miR-130a expression observed in the present study. No previously known relationship exists between cadmium and the other miRNAs identified in this study.

miRNAs are potential regulators of mRNA abundance. To examine the potential impact of the modulation of the eight miRNAs, mRNA expression levels were examined in the cardiomyocytes. A total of 31 genes showed altered expression among the cadmium-exposed samples compared to controls. Among these, a family of ten MT genes was identified. Induction of MT gene expression by cadmium has been shown *in vitro* [51], *in vivo* [53], and in human samples [54, 55]. MTs are involved in cellular protection against metal toxicity as well as the homeostasis of essential metals [56].

We predicted miRNA-mRNA interactions then compared the predicted targets to our list of 31 differentially expressed genes. Eight of these were predicted to be regulated by the cadmium-modulated miRNAs. These miRNA-associated mRNAs were enriched for signaling pathways involved in gene expression, embryonic development, and organismal development. Paradoxically, all eight of the miRNA-associated mRNAs showed increased expression levels in the cadmium-exposed samples. Although binding of miRNA typically negatively regulates gene expression, recent evidence demonstrates that some miRNA act by increasing target gene expression [57]. It is suggested that factors other than miRNA binding might promote or inhibit miRNA-mRNA interactions [24]. miRNA-mediated upregulation of target genes can occur by upregulation of

previously suppressed target genes or by decreased expression of other inhibitory proteins and/or transcription factors.

This study also identified upstream regulators of the differentially expressed genes. Notably, the MT genes were enriched for upstream regulation by transcription factors such as MTF1. MTF1 is a transcription factor associated with metal detoxification that coordinates induction of MT gene expression in multiple species. Andrews *et al.* suggested the other signal transduction cascades might be involved in MT gene expression activated by MTF-1 in cadmium-treated cells [58]. The present study provides evidence that miRNAs, such as miR-1, may partially play a role as mediators of signal transduction.

Cardiomyocytes are the dominant cell type in the heart, including the septa [59], which are the most common site of cardiac defects in humans. Pathways that regulate early cardiomyocyte differentiation are also involved in processes that lead to adult heart disease [30], and animal models have shown that embryonic or early life cadmium exposure may influence cardiomyocyte cell fate and function in postnatal life [16, 17]. In general, cultured induced pluripotent stem cells mimic phenotypic properties of embryonic cells and are similar to embryo-derived stem cells [60, 61]. Human induced pluripotent stem cell-derived cardiomyocytes provide a superior *in vitro* model for cadmium toxicity studies over previous heart cell lineages because they exhibit spontaneous contractility and electrophysiological properties of main cardiac cell types [62, 63]. These properties enable *in vitro* culture of cells capable of mimicking acute cardiac dysfunction including arrhythmias [63]. Despite these advantages, it remains unclear how miRNA behavior in induced pluripotent stem derived cardiomyocytes

compares to that of primary cardiomyocytes [61, 64]. The future characterization of miRNA control of gene expression in primary cardiomyocytes and *in vivo* models will provide information on how normal heart cells respond to cadmium.

In this study, we examined human cardiomyocytes miRNA and mRNA responses to acute environmentally relevant levels of cadmium. Further studies comparing the acute versus chronic miRNA response to exposure, as has been shown for other environmental exposures [65], will be useful in understanding these relevant differences as population tend to experience low-level chronic cadmium exposure. These miRNAs may potentially serve as novel biomarkers of exposure or metal-associated disease in future studies.

Conclusions

This study demonstrates that miRNAs are altered in cardiomyocytes exposed to cadmium. miRNAs are predicted to control mRNA transcript levels, and we demonstrate that these potentially regulate a set of eight differentially expressed mRNAs in human cardiomyocytes. Taken together, these findings support a role of miRNAs as a potential mechanism of cadmium-associated functional changes in human heart cells. More research is needed to examine the biological effects of cadmium exposure *in vivo* and among populations suffering from cadmium-associated cardiovascular disease. Understanding the pathophysiological role of cadmium in the heart can inform the mechanism of cadmium-associated cardiovascular disease and aid in the potential treatment of metal-associated disease.

Tables

Table 3-1. Cadmium exposure to cardiomyocytes significantly increased the expression of nine unique miRNAs corresponding to 56 probesets. Significant differential expression between exposed and unexposed samples was defined by a fold change (FC) ≥ |1.5| and a p-value < 0.01.

miRNA	N Probesets	p-value range	Ave FC
miR-1	17	9.97E-04 – 9.43E-03	2.03
miR-20a-5p	1	4.37E-03	1.54
miR-21-5p	13	1.81E-03 – 9.95E-03	1.98
miR-22-3p	7	1.00E-03 – 7.34E-03	1.69
miR-27b-3p	1	3.67E-03	1.51
miR-30e-5p	1	9.85E-03	1.58
miR-125b-5p	10	1.37E-03 – 9.74E-03	1.62
miR-1274a_v16.0	3	5.77E-03 – 5.77E-03	1.56
miR-130a-3p	3	1.11E-05 – 3.60E-03	1.55

Table 3-2. Eight miRNA predicted mRNA targets were confirmed experimentally by microarray analysis of mRNA transcript levels in samples exposed cadmium. Significant differential mRNA transcript expression between exposed and unexposed samples was defined by a fold change (FC) ≥ |1.5| and a p-value < 0.01.

miRNA	Confidence	Gene Symbol	mRNA FC	p-value
miR-1	Moderate (predicted)	*MT1F*	4.47	2.08E-07
miR-1	Moderate (predicted)	*MT1G*	68.99	2.10E-06
miR-1	Moderate (predicted)	*MT1H*	148.52	2.69E-05
miR-17-5p	High (predicted)	*ALX4*	1.51	4.68E-03
miR-17-5p	High (predicted)	*POLR3G*	1.57	3.04E-03
miR-21-5p	High (predicted)	*ALX4*	1.51	4.68E-03
miR-27a-3p	Moderate (predicted)	*POLR3G*	1.57	3.04E-03
miR-27a-3p	High (predicted)	*SUSD5*	1.63	2.57E-03
miR-30c-5p	High (predicted)	*ME1*	1.51	3.56E-03
miR-30c-5p	High (predicted)	*POLR3G*	1.57	3.04E-03
miR-30c-5p	Moderate (predicted)	*SLC30A2*	2.04	1.65E-03

Table 3-3. Transcriptional regulators (TR) known to influence the differentially expressed mRNA transcripts were identified for the miRNA-associated genes and miRNA-independent genes. P-values represent a statistically significant overlap between the cadmium-associated geneset and genes known to be modified by a upstream transcriptional regulator ($p<0.05$).

Upstream Regulator miRNA-target associated	**Molecule Type**	**Target molecules**	**p-value**
MUC1	TR	*MT1G, MT1H*	5.26E-05
KDM5B	TR	*MT1F, MT1H*	8.96E-04
Hdac	group	*MT1F, MT1H*	1.14E-03
HDAC1	TR	*MT1G, MT1H*	1.32E-03
MTF1	TR	*MT1G*	2.00E-03
HAND2	TR	*ALX4*	5.98E-03
BACH1	TR	*ME1*	7.97E-03
E2F5	TR	*MT1G*	8.76E-03
SUPT16H	TR	*POLR3G*	8.76E-03
ELF4	TR	*MT1G*	1.23E-02
NCOR1	TR	*ME1*	1.79E-02
E2F2	TR	*MT1G*	2.41E-02
HDAC2	TR	*MT1H*	2.49E-02
RBL1	TR	*MT1G*	2.57E-02
E2F3	TR	*MT1G*	2.80E-02
NR3C2	ligand-dependent nuclear receptor	*MT1H*	3.50E-02
Upstream Regulator miRNA-independent	**Molecule Type**	**Target molecules**	**p-value**
MTF1	TR	*MT1E, SLC30A1*	3.89E-06
KDM5B	TR	*MT1E, MT1X*	2.45E-03
Hdac	group	*MT1B, MT1E*	3.12E-03
TP73	TR	*MT1B, MT1L*	5.75E-03
SMARCA4	TR	*MT1E, MT1L*	1.52E-02
NRF1	TR	*MT1E*	2.95E-02
CTNNB1	TR	*ASCL2, MT1L*	4.58E-02
USF1	TR	*MT1E*	4.70E-02

Figures

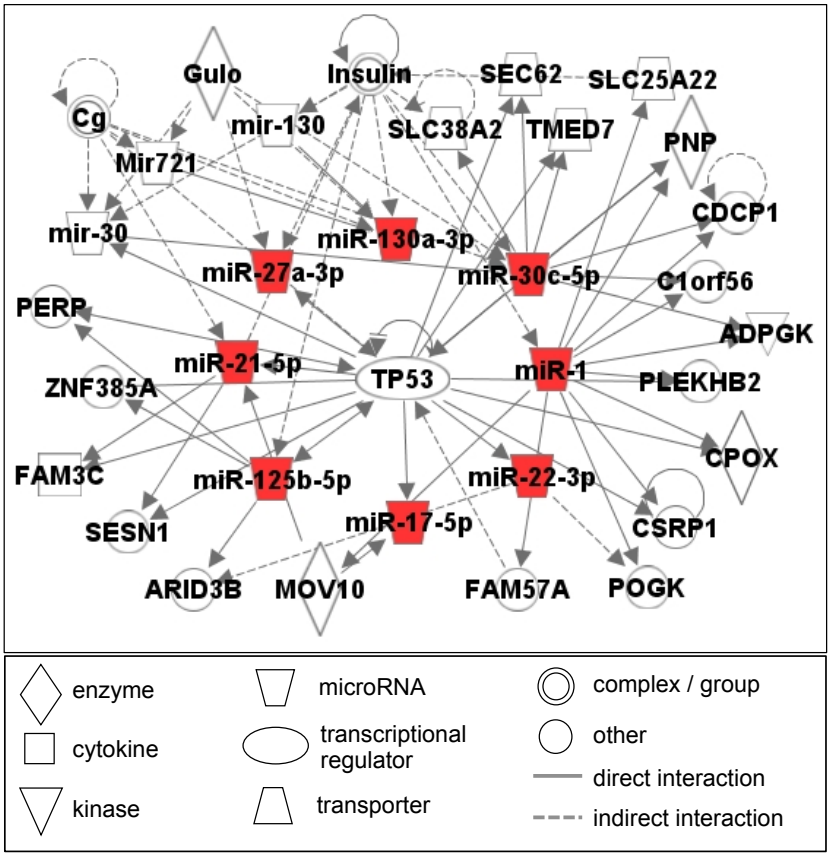

Figure 3-1. Significant molecular network of eight miRNAs differentially expressed in cadmium-exposed cardiomyocytes. Tumor suppressor protein p53 (TP53) appears as a central node. The network was enriched for signaling pathways associated with cancer, connective tissue disorders, and inflammatory disease ($p<1x10^{-23}$). Colors represent increased expression of miRNA (red symbols) or associated proteins (clear symbols).

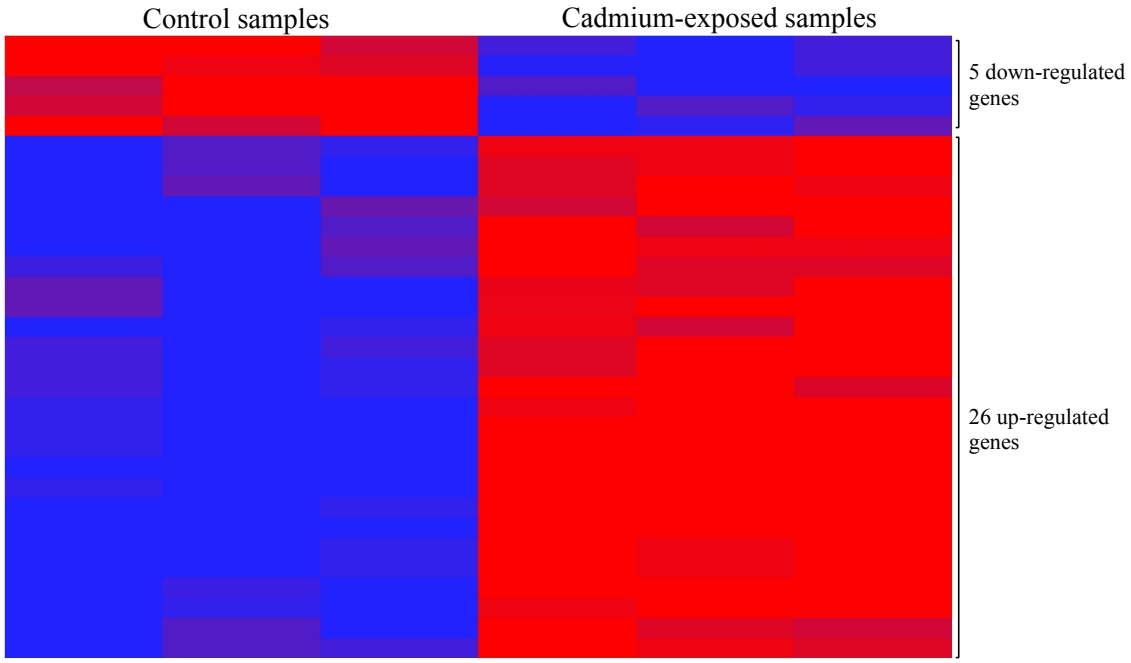

Figure 3-2. Heatmap of 31 differentially expressed mRNA transcripts. Data were mean standardized and hierarchical clustering was performed. Blue indicates relative low expression, and red indicates relative high expression.

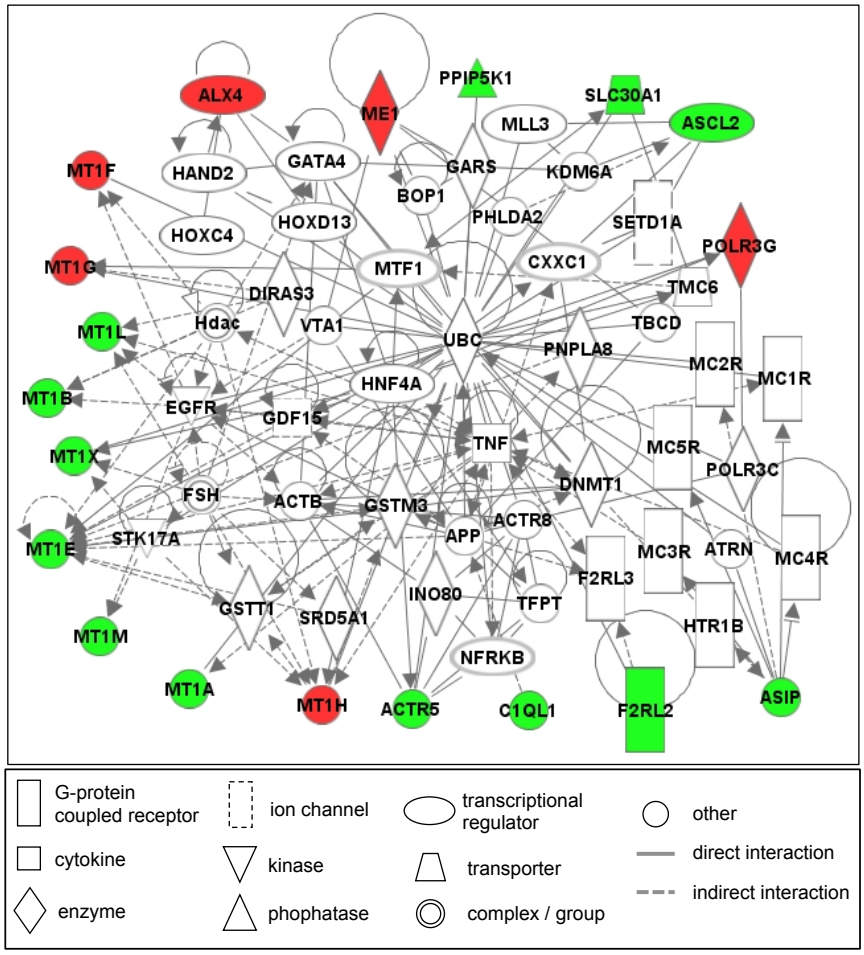

Figure 3-3. Significant molecular network of miRNA-associated proteins (red color) and miRNA-independent proteins (green color). Networks display interactions of proteins encoded by genes that were differentially expressed in cadmium-exposed cardiomyocytes.

Supplemental Tables

Table 3-S1. Cadmium chloride was not cytotoxic to cardiomyocytes after 24-hour exposure. Luminescence fold change was calculated as the mean luminescence for cadmium- or staurosporin- exposed samples divided by untreated control values. Each treatment group included six replicates.

Treatment Condition	**Fold change**
Staurosporin (positive control)	0.01
0.01 μM $CdCl_2$	1.20
0.05 μM $CdCl_2$	1.13
0.1 μM $CdCl_2$	1.20
0.5 μM $CdCl_2$	1.25
1.0 μM $CdCl_2$	1.34
5.0 μM $CdCl_2$	1.18
10.0 μM $CdCl_2$	1.08
50.0 μM $CdCl_2$	1.07

Table 3-S2. Cadmium exposure to cardiomyocytes significantly increased the expression of nine unique miRNAs corresponding to 56 probesets. Significant differential expression between exposed and unexposed samples was defined by a fold change (FC) ≥ |1.5| and a p-value < 0.01.

miRNA	Probeset ID	p-value	q-value	FC
miR-1	44458	9.97E-04	0.49	1.75
miR-1	19282	1.02E-03	0.49	1.72
miR-1	4289	1.13E-03	0.50	1.91
miR-1	28368	2.28E-03	0.60	1.87
miR-1	8669	3.28E-03	0.65	2.17
miR-1	61032	3.61E-03	0.65	2.10
miR-1	30149	4.10E-03	0.67	2.29
miR-1	53139	4.26E-03	0.67	1.96
miR-1	46440	4.38E-03	0.67	2.27
miR-1	39526	4.51E-03	0.67	2.01
miR-1	41053	5.18E-03	0.69	1.99
miR-1	14899	5.45E-03	0.70	2.13
miR-1	33250	6.13E-03	0.70	1.87
miR-1	37885	7.77E-03	0.73	2.07
miR-1	20269	7.83E-03	0.73	1.88
miR-1	57664	8.50E-03	0.74	2.15
miR-1	17168	9.43E-03	0.75	2.32
miR-20a-5p	58184	4.37E-03	0.67	1.54
miR-21-5p	5799	1.81E-03	0.57	2.10
miR-21-5p	28018	2.97E-03	0.65	2.09
miR-21-5p	59575	3.10E-03	0.65	2.02
miR-21-5p	18483	3.34E-03	0.65	2.16
miR-21-5p	35083	3.68E-03	0.66	2.06
miR-21-5p	25570	4.82E-03	0.67	2.15
miR-21-5p	20794	5.46E-03	0.70	2.03
miR-21-5p	9456	6.29E-03	0.71	1.87
miR-21-5p	26951	6.53E-03	0.71	1.92
miR-21-5p	10775	6.70E-03	0.71	1.85
miR-21-5p	9322	8.66E-03	0.74	1.90
miR-21-5p	51597	9.89E-03	0.75	1.87
miR-21-5p	43392	9.95E-03	0.75	1.73
miR-22-3p	28021	1.00E-03	0.49	1.60
miR-22-3p	32003	1.04E-03	0.49	1.73
miR-22-3p	2399	2.69E-03	0.63	1.67
miR-22-3p	11673	3.39E-03	0.65	1.69
miR-22-3p	9068	4.15E-03	0.67	1.67
miR-22-3p	10971	5.70E-03	0.70	1.74

miR-22-3p	56058	7.34E-03	0.72	1.71
miR-27b-3p	23533	3.67E-03	0.66	1.51
miR-30e-5p	25735	9.85E-03	0.75	1.58
miR-125b-5p	15935	1.37E-03	0.53	1.73
miR-125b-5p	37771	4.37E-03	0.67	1.57
miR-125b-5p	43022	4.38E-03	0.67	1.65
miR-125b-5p	17949	5.53E-03	0.70	1.65
miR-125b-5p	41051	5.67E-03	0.70	1.68
miR-125b-5p	20012	6.90E-03	0.71	1.51
miR-125b-5p	55379	7.54E-03	0.72	1.62
miR-125b-5p	28571	8.12E-03	0.74	1.53
miR-125b-5p	53702	8.56E-03	0.74	1.60
miR-125b-5p	43869	9.74E-03	0.75	1.64
miR-1274a_v16.0	23273	5.77E-03	0.70	1.55
miR-1274a_v16.0	2371	5.81E-03	0.70	1.61
miR-1274a_v16.0	21226	9.43E-03	0.75	1.53
miR-130a-3p	3655	1.11E-05	0.10	1.51
miR-130a-3p	1878	2.07E-04	0.35	1.52
miR-130a-3p	20288	3.60E-03	0.65	1.62

Table 3-S3. Cadmium exposure to cardiomyocytes significantly altered the expression of 31 mRNA transcripts (FC ≥|1.5| and p-value<0.01). Five of the transcripts showed decreased expression in samples exposed to 0.5 μM cadmium chloride.

Gene Symbol	p-value	FDR q-value	Fold Change
MT1H	2.69E-05	2.24E-01	148.52
MT1A	1.39E-04	4.29E-01	93.61
MT1L	8.25E-06	8.25E-02	87.13
MT1M	7.28E-06	8.25E-02	85.76
MT1G	2.10E-06	5.26E-02	68.99
MT1E	2.58E-04	5.38E-01	7.64
MT1B	7.14E-05	3.57E-01	7.55
MT1F	2.08E-07	1.04E-02	4.47
MT1CP	1.44E-04	4.29E-01	3.95
SLC30A1	9.48E-05	4.29E-01	2.81
MT1X	1.42E-04	4.29E-01	2.48
F2RL2	5.50E-03	8.56E-01	2.26
SLC30A2	1.65E-03	8.56E-01	2.04
SUSD5	2.57E-03	8.56E-01	1.63
ASCL2	9.69E-03	8.60E-01	1.60
ACTR5	8.73E-03	8.56E-01	1.59
POLR3G	3.04E-03	8.56E-01	1.57
LOC100132472	8.16E-03	8.56E-01	1.54
OR10AD1	7.89E-03	8.56E-01	1.53
FAM27L	4.55E-03	8.56E-01	1.53
OR1E2	4.77E-03	8.56E-01	1.52
C1QL1	8.31E-03	8.56E-01	1.51
ME1	3.56E-03	8.56E-01	1.51
PPIP5K1	4.58E-03	8.56E-01	1.51
ALX4	4.68E-03	8.56E-01	1.51
ASIP	9.61E-03	8.60E-01	1.51
LOC440895	6.64E-03	8.56E-01	-1.55
LOC441728	3.58E-03	8.56E-01	-1.57
C12orf36	6.40E-03	8.56E-01	-1.64
MIR570	7.26E-03	8.56E-01	-1.64
SNORD63	8.08E-03	8.56E-01	-2.11

Table 3-S4. Differentially expressed mRNA transcripts that were i) miRNA-associated, ii) mRNA-independent, or iii) both were mapped to signaling pathways. Gene names in bold correspond to the 31 mRNA genes with altered expression in cadmium-exposed samples compared to unexposed controls.

ID	Molecules in miRNA-associated networks	P-value	Top Functions
1	**ALX4**, ATP7B, BACH1, BOP1, DLX2, DLX5, E2F5, ELF4, ERCC8, GARS, GATA4, GORASP2, GSTM3, GTF2A1, GTF3C1, HAND2, HOXA11, HOXD13, LDB1, **ME1**, **MT1F, MT1G, MT1H**, POLR3C, POLR3D, **POLR3G**, POLR3H, SOX10, SPINK7, STK17A, TBCD, TOM1, UBC, UGDH, VTA1	10^{-16}	Gene Expression, Embryonic Development, Organismal Development
2	CREB3, **SLC30A2**	10^{-3}	Molecular Transport, Cancer, Cellular Development
3	ACVR1, **SUSD5**, TGFB1	10^{-3}	Inflammatory Disease, Inflammatory Response, Skeletal and Muscular Disorders
ID	**Molecules in miRNA-independent networks**	**P-value**	**Top Functions**
1	**ACTR5**, ACTR8, **ASCL2**, **ASIP**, ATRN, beta-estradiol, **C1QL1**, EGFR, eumelanin, **F2RL2**, FAM168A, FOS, GPR50, HTR1B, INO80C, MC1R, MC2R, MC3R, MC4R, MC5R, mir-548, **MT1A, MT1B, MT1E, MT1L, MT1M, MT1X,** MTF1, NFRKB, POMC, **PPIP5K1**, **SLC30A1**, STK17A, TMC6, UBC	10^{-41}	Cell Signaling, Nucleic Acid Metabolism, Small Molecule Biochemistry
2	**OR1E2**, RAD21	10^{-3}	Cellular Assembly and Organization, DNA Replication, Recombination, and Repair, Cellular Function and Maintenance

ID	All 31 mRNA molecules in networks	P-value	Top Functions
1	ACTB, **ACTR5**, ACTR8, APP, **C1QL1**, CXXC1, DIRAS3, EGFR, **F2RL3**, FSH, GDF15, glutathione disulfide, GSTM3, GSTT1, Hdac, HNF4A, INO80, **MT1A, MT1B, MT1E, MT1F, MT1H, MT1L, MT1M, MT1X,** MTF1, NFRKB, nitrogen, PNPLA8, POLR3C, **POLR3G**, SRD5A1, STK17A, TFPT, TNF	10^{-26}	Renal and Urological Disease, Cellular Compromise, Psychological Disorders
2	**ALX4**, **ASCL2**, **ASIP**, ATRN, beta-estradiol, BOP1, CXXC1, DNMT1, **F2RL2**, GARS, GATA4, HAND2, HOXC4, HOXD13, HTR1B, KDM6A, MC1R, MC2R, MC3R, MC4R, MC5R, **ME1**, mir-548, MLL3, **MT1G,** MTF1, NFRKB, PHLDA2, **PPIP5K1**, SETD1A, **SLC30A1**, TBCD, TMC6, UBC, VTA1	10^{-21}	Cell Signaling, Nucleic Acid Metabolism, Small Molecule Biochemistry
3	CREB3, **SLC30A2**	10^{-3}	Molecular Transport, Cancer, Cellular Development
4	**OR1E2**, RAD21	10^{-3}	Cellular Assembly and Organization, DNA Replication, Recombination, and Repair, Cellular Function and Maintenance
5	ACVR1, **SUSD5**, TGFB1	10^{-3}	Inflammatory Disease, Inflammatory Response, Skeletal and Muscular Disorders

Note: References for Chapter 3 appear after the appendices.

CHAPTER 4: CONCLUSIONS AND IMPLICATIONS

Summary

The majority of birth defects have no known cause whether environmental, genetic, or combination of these factors. Toxic metals are likely contributors to birth defects in humans, yet the biological pathways that underlie these relationships remain largely unknown. The studies presented herein examined i) the epidemiologic association between toxic metal exposure and CHDs at the population and ii) transcriptional regulation and epigenetic modifiers using cardiomyocyte cultures exposed to cadmium *in vitro*.

In the second chapter of this dissertation, we evaluated the associations between arsenic, cadmium, manganese, and lead levels in private well water with the prevalence of specific birth defects using six years of data collected by the North Carolina Birth Defects Monitoring Program. We estimated the association between prevalence of birth defects and residence in census tracts with the highest average metal levels and compared to residence in areas with the lowest average metal levels. Our exploratory study suggests ecologic associations between concentrations of toxic metals in drinking water and the prevalence of specific birth defects.

Next, we applied an *in vitro* approach to identify biological pathways that underlie exposure to metals and their potential association with altered signaling in the heart. We exposed human-derived cardiomyocytes to low-levels of cadmium and

assessed genome-wide miRNA and mRNA expression levels. We identified sets of miRNA-dependent and -independent genes that were differentially expressed in cardiomyocytes exposed to cadmium compared to unexposed controls.

Public health implications

Taken together the body of my doctoral research examines the potential health effects of metals in North Carolina and contributes to the overarching goal of protecting the health of susceptible populations by raising public health awareness of the global burden and potential impacts of toxic metals. Our studies have demonstrated that detectable metal levels were present in private well across the state and showed that further monitoring of levels of toxic metals North Carolina populations was prudent particularly for pregnant women. The epidemiologic study in Chapter Two contributes to the body of literature that describes the potential relationship between metal leveling in drinking water and CHDs. The toxicological study presented in Chapter Three describes a potential epigenetic mechanism, namely miRNA regulation of gene expression, which may contribute to metal-induced heart defects.

In addition to targeted biomonitoring programs, public health strategies to prevent exposure such as raising awareness among pregnant mothers may be vital to protecting the health of newborns. Informing pregnant women of the sources of toxic metals such as diet (including drinking water), cigarette smoke, and other environmental sources will help to reduce exposures. Increased education about the potential risks of environmental contaminants including toxic metals will be key to reducing and preventing harm to

infants. The evidence provided by the environmental and biomonitoring studies herein have provided impetus for targeted biomonitoring programs and public health campaigns in North Carolina, specifically related to well water testing and monitoring prenatal lead exposure.

Recommendations for future studies

Our study contributes to the growing body of epidemiological literature on metal exposure and the relationship to birth defects. Direct biological monitoring of maternal and fetal exposures during the periconceptional period would greatly improve potential exposure misclassification issues in studies of environmental causes of birth defects. Further epidemiologic investigations should incorporate individual-level assessment of maternal drinking water quality and consumption. The results from this study suggest that research aimed at identifying metal-induced health outcomes should consider the likely presence of multiple metals. Moreover, studies examining the relationship between metal exposures in North Carolina and other metal-mediated health end points are needed.

The research presented here used acute exposure and exposure response to a specific metal. Moving forward it will be important to understand the effects of chronic exposure and to assess long term alterations in miRNA expression levels. Moreover, studies examining the mechanistic responses to metal mixtures will be of interest since metal exposure rarely occurs in isolation. Further, use of an *in vivo* model or direct human biological sampling would allow increased translatability of this work the understanding of metal-induced disease. There is growing evidence that links miRNAs

and environmental exposures, and miRNAs have potential use as diagnostic biomarkers. Understanding the mechanisms that underlie metal-induced miRNA alteration are important for the development of more effective preventions strategies for environmentally-mediated diseases.

It is likely a complex interplay of transcriptional regulators, including multiple epigenetic and genetic factors that will ultimately help researchers achieve a better understanding of metal-mediated disease. Taken together, the findings presented in this dissertation contribute to the understanding of potentially preventable environmentally-mediated birth defects. Identifying environmental causes of adverse birth outcomes is important for the prevention of birth defects.

APPENDIX ONE: ARSENIC IN NORTH CAROLINA: PUBLIC HEALTH IMPLICATIONS

Alison P. Sanders[1], Kyle P. Messier[1], Mina Shehee[2], Kenneth Rudo[2], Marc L. Serre[1], and Rebecca C. Fry[1*]

Authors' Affiliations:

[1] Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, 1213 Michael Hooker Research Building, Chapel Hill, North Carolina 27599

[2] Medical Evaluation and Response Assessment Unit, North Carolina Department of Health and Human Services, 1912 Mail Service Center, Raleigh, North Carolina 27699

ABSTRACT

Arsenic is a known human carcinogen and relevant environmental contaminant in drinking water systems. We set out to comprehensively examine statewide arsenic trends and identify areas of public health concern. Specifically, arsenic trends in North Carolina private wells were evaluated over an eleven-year period using the North Carolina Department of Health and Human Services database for private domestic well waters. We geocoded over 63,000 domestic well measurements by applying a novel geocoding algorithm and error validation scheme. Arsenic measurements and geographical coordinates for database entries were mapped using Geographic Information System techniques. Furthermore, we employed a Bayesian Maximum Entropy (BME) geostatistical framework, which accounts for geocoding error to better estimate arsenic values across the state and identify trends for unmonitored locations. Of the approximately 63,000 monitored wells, 7,712 showed detectable arsenic concentrations that ranged between 1 and 806 μg/L. Additionally, 1,436 well samples exceeded the EPA drinking water standard. We reveal counties of concern and demonstrate a historical pattern of elevated arsenic in some counties, particularly those located along the Carolina terrane. We analyzed these data in the context of populations using private well water and identify counties for targeted monitoring, such as Stanly and Union Counties. By spatiotemporally mapping these data, our BME estimate revealed arsenic trends at unmonitored locations within counties and better predicted well concentrations when compared to the classical kriging method. This study reveals relevant information on the location of arsenic-contaminated private domestic wells in North Carolina and indicates potential areas at increased risk for adverse health outcomes.

INTRODUCTION

Ingestion of arsenic through drinking water is implicated in heart disease, neurological abnormalities, as well as cancers of the skin, lung, kidney, and bladder (NRC 2001). The United States Environmental Protection Agency (EPA) regulates arsenic in public drinking water supplies at a maximum contaminant level (MCL) of 10 μg/L (EPA 2010). Although this standard is enforceable in public water systems via the Safe Drinking Water Act, there is no federal regulatory standard for domestic well waters. Approximately 14 percent (about 42 million people) of the U.S. population obtains water from unregulated private domestic wells (Kenny et al. 2009). It is estimated that domestic well users in the U.S. carry an excess lifetime risk of bladder and lung cancer of 66 people per million people, almost five times higher than that estimated for public well users (Kumar et al. 2010).

Although arsenic exposure through drinking water is documented worldwide (Mukherjee et al. 2006), there is a paucity of data in the U.S.. For example, in a survey of U.S. drinking water supplies, many of the Mid-Atlantic states had insufficient data with less than 10% of counties represented (Welch et al. 1999). To help fill this void, regional evaluations of groundwater arsenic in Appalachia (Shiber 2005), Idaho (Hagan 2004), Maine (Yang et al. 2009), Michigan (Kim et al. 2002), Nevada (Shaw et al. 2005; Walker et al. 2006), New England (Ayotte et al. 2003), and New Hampshire (Peters et al. 1999) have provided data beyond those collected in nationwide studies and have demonstrated contamination of drinking sources at spatial scales finer than the county level. In

addition, while the USGS provides routine ambient well monitoring nationwide, data gathered currently represent only a small fraction of groundwater sources. Domestic wells are not often monitored in such nationwide programs and may more adequately reflect exposure to unregulated contaminated water.

It is known that areas of North Carolina contain naturally occurring arsenic deposits including the geological region of Carolina terrane (or Carolina slate belt) making it an ideal area for further investigation and public health efforts (Foley et al. 2001). An initial study by Pippin et al., (2005), characterized arsenic occurrence in North Carolina groundwater between 1996-2004 and employed classic kriging techniques to map arsenic probabilities in the state. More recently, Kim et al. (*in press*), characterized geologic determinants of arsenic in Orange County, North Carolina. Our work builds upon earlier characterization of arsenic in North Carolina wells by developing methods to map historical contamination for the purposes of protecting public health. Importantly, the number of individuals in North Carolina currently using domestic wells for drinking water is estimated at 2.3 million (Kenny et al. 2009). North Carolina has the fourth-largest state population in the U.S. using private groundwater wells as a drinking source and is exceeded only by Michigan, California, and Pennsylvania (Kenny et al. 2009). Many states, such as North Carolina, still remain understudied for the presence of arsenic in drinking water at a statewide level. Here, we present results obtained from a statewide program to monitor unregulated private domestic wells.

Given the known health risks and occurrence of arsenic in drinking water, we set out to identify areas of concern and quantitatively assess concentrations in domestic wells throughout North Carolina. To assess arsenic trends in monitored private domestic wells, we applied a novel geocoding scheme and mapped arsenic levels in wells using standard Geographical Information System (GIS) techniques (Nuckols et al. 2004; Miranda et al. 2002; Holton 2002; Bellander et al. 2001). To estimate arsenic values at unmonitored locations, we then applied a novel Bayesian Maximum Entropy (BME) framework (Christakos 1990, 2000; Serre and Christakos 1999; De Nazelle et al. 2010) to predict arsenic contamination across the state and to examine areas of interest. These analytical techniques were applied to more than 60,000 domestic well water arsenic measures collected by the North Carolina Department of Health and Human Services (NCDHHS) dating back to 1998. In this work, we identify spatial and temporal arsenic trends in North Carolina domestic wells and indicate specific locations and populations of concern. Importantly, the geocoding and geostatistical methods presented here can be applied to track contaminant trends in other states. Our results indicate a large number of contaminated wells in North Carolina and suggest that ongoing monitoring of well water contaminants is prudent. Moreover, these data provide new information of specific areas in North Carolina where targeted well monitoring programs can be used in a cost-effective manner.

MATERIALS AND METHODS

Data collection

Domestic well water samples were collected by the NCDHHS Division of Public Health (DPH) State Laboratory of Public Health and Epidemiology Section which provides groundwater monitoring assistance to North Carolina homeowners. Following DPH guidelines, local health department officials collected water from homeowners' indoor, outdoor, or well tap after allowing the water to run for 5-10 minutes. Arsenic analyses were performed by the NCDHHS Laboratory for Environmental Inorganic Chemistry. Samples were transported to the DPH State Laboratory of Public Health and analyzed within 48 hours as per EPA Method 200.8 Revision 5.4 via inductively coupled plasma mass spectrometry (ICP-MS) with adherence to formal quality assurance/quality control (QA/QC) protocols (EPA 1994). Sample aliquots were acidified with nitric acid to below a pH of 2.0 for at least sixteen hours prior to ICP-MS analysis. A 50-mL subsample was then digested at 95ºC for at least 2 h. For samples with high amounts of undissolved particulates, the digestate was filtered through a 0.45 μm filter to prevent damage to the analytical instrumentation. This method detects for total arsenic; species of As(III) and As(V) were not differentiated. The NCDHHS detection limit for arsenic had shifted in the past decade. In early 2000, the detection limit decreased from 10.0 to 1.0 μg/L. More recently, the detection limit was raised to 5.0 μg/L – the level at which the laboratory presently operates. The laboratory maintains QC requirements of Fortified Blank recoveries of 95 – 105% and reagent blanks must show no contamination. In all lab analyses yttrium, rhodium, lutetium were included as internal standards to account for instrument drift and physical interferences. Every QC requirement must be met for the

sample analysis to be approved by the laboratory manager. Analyses were then entered into an extensive electronic database maintained by the State Laboratory of Public Health.

Electronic database formatting

Results from arsenic well water analyses were housed in an electronic database containing informational fields for arsenic concentration, well location ID, county name, global positioning system (GPS) location (if available), well owner address, and date collected. We analyzed 68,836 electronically available well water records of measured arsenic concentrations collected between October 19, 1998 and February 25, 2010. Data cleaning of the arsenic database excluded entries with insufficient information on location or those with improper/incomplete chemical analysis. We required that entries included in this study provided the following minimum information: a county name, a valid sampling date, and an approved laboratory chemical analysis for arsenic. The resulting comprehensive database of 63,856 well measures was used for all subsequent analyses.

Descriptive statistics, heat map generation, and county ranking

For all analyses, arsenic measurements below the detection limit (DL) were treated as 0.5 times the DL (0.5 × DL). Descriptive statistics were calculated using Spotfire DecisionSite 8.1 (TIBCO, Palo Alto, CA) for each of the 100 North Carolina counties. Heat maps to visualize temporal county trends were prepared using Partek Genomics Suite 6.4 (St. Louis, Missouri). Hierarchical cluster analysis using Euclidean

distance as a measure was used to identify relationships over time among the top 35 counties exceeding the standard (Eisen et al. 1998). Furthermore, counties were ranked by 1) the percentage of wells exceeding the EPA standard over the full time period examined, and 2) the percentage of wells exceeding the EPA standard multiplied by the percentage of county population using self-supplied domestic wells (data reported by Kenny et al. 2009). We considered the current EPA MCL of 10 μg/L as the threshold, although the regulatory standard (originally 50 μg/L) was lowered over the course of data collection in this study.

Four-class geocoding algorithm and error validation scheme

A geocoding algorithm was developed to extract spatial coordinates and associated location error for the 63,856 private well measurements contained in the database. A four-class strategy, detailed below, was used to assign each data entry, *l,* with a spatial coordinate $\boldsymbol{s}_l=(s_1,s_2)_l$, where s_1 and s_2 were the longitude and latitude best-representing the well location given the level of recorded information: GPS, street address, zip code, or county. The first of the four classes (Class I) was represented by data entries with available GPS coordinates, $\boldsymbol{s}_{l(\mathrm{GPS})}$. For wells with reported GPS locations, geographical coordinates were standardized to decimal degrees format. Standardized GPS coordinates were visualized in ESRI ArcGIS^{TM} software Version 9.0 (Redlands, CA). Well locations were classified as Class I when GPS coordinates were available and within the longitude/latitude boundaries of North Carolina. The second class (Class II) assigned geocoded owner address (GOA) coordinates, $\boldsymbol{s}_{l(\mathrm{GOA})}$, to data entries based on street

address. To geocode the data, we applied a multi-stage geocoding process in which a series of local and national reference files were used sequentially in order from most comprehensive spatial detail to least as follows: a North Carolina point reference file (courtesy of NCDHHS Spatial Analysis Group), followed by a North Carolina Department of Transportation line reference file, then followed by a U.S. Street Address line reference file (Tele Atlas Dynamap Transportation, 2003). Resulting from this process, match scores of 0-100 were associated with each successfully geocoded address coordinate, $\boldsymbol{s}_{l(\mathrm{GOA})}$. To determine a match score threshold at which each reference file provided acceptable coordinates, we developed a method using one-way analysis of variance (ANOVA) to select acceptable match scores based on the distance between GPS and geocoded coordinates, $d_l = \|\boldsymbol{s}_{l(\mathrm{GPS})} - \boldsymbol{s}_{l(\mathrm{GOA})}\|$. We calculated the average distance, d_l, for wells with match score of 100 (perfectly matched addresses) to serve as the referent group. The remaining wells with a geocoded location were binned into deciles according to match score (e.g. 99-90, 89-80, 79-70 and so on) and the average distance d_l was also calculated for each match score decile. We used ANOVA to compare the average distance d_l to identify which match score deciles had an average distance d_l that was not statistically significantly different ($=0.05$) from the average distance d_l obtained with a match score of 100. Using this criterion, point file match scores of 70 and higher and line file match scores of 60 and above were considered acceptable for describing geocoded well locations. A data point was classified as Class II when it was not a Class I, was successfully geocoded by a given reference file with a match score above the match score threshold, and the county name included in the owner address matched the recorded county of sampling location. Class II data entries were assigned a single coordinate pair

representing the geocoded owner address coordinates $\boldsymbol{s}_{l(\mathrm{GOA})}$ resulting from the multi-stage geocode process. The third class (Class III) included zip code centroid coordinates calculated and assigned using ArcGIS™. A well location was categorized as Class III when it was neither a Class I or Class II, was inside the North Carolina boundary and county of well location and a zip code was available to geocode. Class III data entries were assigned zip code centroid coordinates. The fourth class (Class IV) included county centroid coordinates for each of the 100 counties in North Carolina. A well location was considered Class IV when it did not meet the requirements of any of the previous classes, but a county centroid coordinate was available. Class IV data entries were assigned county centroid coordinates. Each entry in the database was categorized as one of the four aforementioned classes and the resulting four-class geocoded data was used in all subsequent analyses. In summary, the geocoding process assigned one of each of the following four classes to every arsenic measure in the database: Class I (GPS location), Class II (street address), Class III (zip code centroid), Class IV (county centroid). Wells with assigned geographic coordinates and corresponding arsenic concentration data were visualized using ArcGIS™ ESRI version 9.0 software.

The four-class geocoding scheme enabled maximum incorporation of spatial information contained in the private well database. Next, we assessed error associated with each of the four classes to account for uncertainty introduced by the geocoding process. For Class I, GPS device positional error resulted from inaccuracies in satellite triangulation. Positional error associated with GPS instrumentation found by others was between 5 and 25 meters (Wing and Eklund 2007; Hulbert and French 2001). Without

rigorous quantification of GPS device error by the NCDHHS, a conservative Class I error of 50 m was estimated for the GPS coordinates $\boldsymbol{s}_{l(\mathrm{GPS})}$. For Class II, the location error of geocoded street address coordinates $\boldsymbol{s}_{l(\mathrm{GOA})}$ was considered to be a combined effect of two error sources: character-matching error as captured by the match score and inaccuracy in reference file coordinates. For Class II data entries we estimated the location error as the distance $d_l=\|\boldsymbol{s}_{l(\mathrm{GPS})}-\boldsymbol{s}_{l(\mathrm{GOA})}\|$ between the well GPS location and the street address geocoded location. Class II location error was described as a function of match score grouped into deciles (e.g. 100, 90-99, 80-89, 70-79). The location error for a geocoded owner address coordinate $\boldsymbol{s}_{l(\mathrm{GOA})}$ corresponding to a given match score was assigned the median of the distances d_l in the corresponding match score bin (Supplemental Figure S3). For Classes III and IV, entries provided limited locational information, thus data entries were assigned an error estimate of the median radius of zip code or county, respectively.

Spatiotemporal geostatistical estimation of arsenic concentrations

In addition to mapping the actual historical arsenic measures contained in the database, a BME geostatistical framework was applied to estimate concentrations for locations at which no data were available. We let $X(\boldsymbol{s}',t)$ be a space/time random field (S/TRF) representing the yearly arsenic concentration at location $\boldsymbol{s}'$ and year t. We defined the yearly county arsenic concentration at location $\boldsymbol{s}$ and time t as the areal average of $X(\boldsymbol{s}',t)$ over an area the size of a county around $\boldsymbol{s}$, i.e.

$$Z_R(\mathbf{s},t) = \frac{1}{\|A_R\|}\int_{A_R(s)} d\mathbf{s}' X(\mathbf{s}',t) \qquad \textbf{(1)}$$

where $A_R(\boldsymbol{s})$ was an area of radius R around $\boldsymbol{s}$, and the subscript R in Z_R emphasized the county level observation scale over which arsenic was estimated, which in this work corresponded to the median county radius in North Carolina (approximately 11 km).

First, kriging principles were applied to estimate arsenic concentrations across space and time using only the county averages. The average arsenic concentrations were calculated within the boundary of any county i. This county average provided an exact (hard) value $z_{hard(i,t)}$ for the S/TRF $Z(\boldsymbol{s}_i,t)$ at the centroid $\boldsymbol{s}_i$ of county i. These hard data were processed with the well-documented kriging method (Cressie 1990) to model the mean trend (assumed constant) and covariance $c_Z(\boldsymbol{p},\boldsymbol{p}')$ of the S/TRF $Z(\boldsymbol{s},t)$ where $\boldsymbol{p}=(\boldsymbol{s},t)$ was the space/time coordinate and obtain kriging estimates of county arsenic concentrations at a grid of unmonitored locations.

To incorporate the geocoded information and refine the spatial resolution of our arsenic estimate, we developed a BME framework to account for errors associated with geocoded classes by generating soft data. The soft data for yearly arsenic county concentrations were constructed at soft data points located on a static fine resolution grid across the state. The mean μ_j (Equation 2) and variance σ_j^2 (Equation 3) for the yearly county arsenic concentration incorporated geocoding distance error and assigned weight w_l (Equation 4). As such, the arsenic value at a given soft data point $\boldsymbol{s}_j$ was assigned a mean and variance calculated as the weighted sample average and sample variance,

respectively, of the geocoded private well data that were within a distance R of $\boldsymbol{s}_j$, where the weights decreased as a function of increased geocoding location error, i.e.

$$\mu_j = \frac{\sum w_l As_l}{\sum w_l} \quad \textbf{(2)}$$

$$\sigma_j^2 = \frac{\dfrac{\sum w_l (As_l - \mu_j)^2}{\sum w_l}}{n} \quad \textbf{(3)}$$

where As_l and n served as the l-th and total number of measured arsenic values within R of $\boldsymbol{s}_j$, respectively, and w_l described a weight that was inversely related to the location error $_l$ of the l-th arsenic data point. Weights were calculated as

$$w_l = \frac{R - \varepsilon_l}{R} \quad \textbf{(4)}$$

which captured the chance that well l was correctly placed within a circle of radius R despite its location error ε_l. To prevent the probability of estimating a negative concentration, we defined the soft data using a Gaussian probability distribution function truncated below zero, with mean and variance calculated from Equations 2 and 3. Equations (2-4) provided values for the soft data points $\boldsymbol{s}_j$, which together with the hard data $z_{hard(i,t)}$ at the county centroids $\boldsymbol{s}_i$, constituted the site specific knowledge S that was incorporated to produce refined estimates of yearly county arsenic concentrations using the BME method (Christakos 1990, 2000; De Nazelle et al. 2010) and its *BMElib* numerical implementation (Serre and Christakos 1999; Christakos et al. 2002). The estimated values were mapped using ArcGIS™ ESRI version 9.0 software.

Lastly, to quantitatively evaluate the difference in performance of the kriging and BME methods for accurate estimation of arsenic concentrations, we applied a cross-validation approach. Each data point was sequentially removed from the estimation scheme and then re-estimated using the remaining space/time data points (Money et al. 2009). Mean square error (MSE) was then derived from this cross-validation process and calculated as the sum of the squared differences between the re-estimated and original values.

USGS and EPA database retrieval

Archived arsenic monitoring data from the online Water Quality section of the National Water Information System (NWIS) and STORET database were obtained electronically (USGS 2010). Basic statistics were compared from the NCDHHS database with the field sample dataset from the NWIS in North Carolina.

RESULTS

Arsenic levels exceed EPA standard in North Carolina monitored wells

The historical database of monitored well data revealed increased well sampling in recent years. Prior to 2008 the NCDHHS sampled over 4,000 wells per year with a notable increase to over 10,000 wells per year from 2008 to present (Supporting Information Table S1). Across the state, a total of 1,436 well measurements (2.25%) exceeded the current standard of 10 μg/L and 233 exceeded 50 μg/L (Table 1; Supporting Information Table S1 and Table S2). Over the 11-year period, 7,713 samples measured above the detection limit, representing approximately 12% of all private wells tested (Table 1). The remaining domestic well water records were below the detection limit (Supporting Information Table S2).

To systematically determine counties with elevated arsenic levels, counties were ranked by the percentage of wells that exceeded the EPA standard across the eleven-year period (Figure 1; Table 1; Supporting Information Table S1). The top ten counties with the highest percentage of wells containing elevated arsenic levels were in order: Stanly, Union, Anson, Montgomery, Dare, Randolph, Davidson, Alexander, Cleveland, and Currituck (Supporting Information Table S1). Of the measured wells that exceeded the EPA standard, nearly 70% were within these ten counties. The remaining 30% of wells that exceeded the EPA standard were collected from the remaining ninety North Carolina counties (Table 1). To identify counties that might exhibit similar arsenic trends over time, cluster analysis was performed. Members of the top ten counties (e.g. Union,

Stanly, Alexander) had statistical temporal relationships in arsenic levels across the 11-year period (Supporting Information Figure S1). This comprehensive temporal assessment revealed a historical pattern of arsenic levels in counties along the Carolina terrane, demonstrating that some counties appear to have been high for over a decade.

Table 1 also reports the demographics of county and state population (and percent of total) using private domestic well water (Kenny et al. 2009). Notably, in some of the highly ranked counties, such as Randolph County, nearly 48% of the county population uses private domestic wells as a primary water source. Counties were assigned a second ranking based on the percentage of population using self-supplied domestic wells multiplied by the percentage of wells exceeding the EPA standard. The following top ten counties were identified: Stanly, Union, Montgomery, Randolph, Lincoln, Pender, Dare, Moore, Person, and Transylvania.

A four-class geocoding algorithm increased spatial information

The geocoding scheme developed in this work categorized the data into four classes (Supporting Information Figure S2). Approximately 3.6% (2,295 well measures) of the database had original GPS coordinates available and represent Class I. Geocoded well locations representing Class II comprised 68.9% (43,991 well measures) of the database. The remaining well locations were categorized as Class III (13.3%) and Class IV (14.2%) by assigning zip code and county centroid coordinates, respectively.

A geocoding location error was assessed for each geocoding class (Supporting Information Table S3). A conservative estimate of 50 m location error for Class I was established based on previously reported quantification of GPS error (Wing and Eklund 2007; Hulbert and French 2001). Class II location errors were determined as a function of match score. Acceptable geocoding match scores were established at 60-100 and the corresponding median location error ranged between 78 meters and 758 meters (Supporting Information Figure S3, Table S3). Location error for Class III and Class IV was approximated as half the median radius of a zip code or county, 3,500 and 11,000 m, respectively.

Mapping of arsenic in monitored private domestic wells

We applied the results of our four-class geocoding process to map arsenic levels in monitored wells and identify regions of arsenic contamination in North Carolina (Figure 2). As an example, we show locations of the geocoded wells and those that exceeded the EPA standard in 2009 (Figure 2A). Notably, wells exceeding the EPA standard were located primarily in the south-central region of the state. The calculated county averages for 2009 are also provided (Figure 2B). The highest county average was observed in Stanly County, where the average of 89 domestic well records approached the 10 μg/L EPA standard.

Spatiotemporal modeling of estimated arsenic concentrations

Next, we set out to refine the spatial scale and apply the results of the geocoding process using two geostatistical estimation methods, namely 1) the classical kriging method and 2) our novel BME framework. Using the classical method that incorporates no distance error information, the spatial distribution of kriging estimates of county arsenic concentrations across North Carolina is represented (Figure 2C). The kriging estimates correspond to the spatial interpolation of county averages assigned to their county centroid. In comparison, the BME framework with a county-level observation scale (Figure 2D) interpolated data in between county averages. The county observation scale enabled estimates of aggregated arsenic concentrations across a ~11 km radius. This map incorporated estimated distance errors associated with the geocoded data to obtain less biased predicted arsenic values. From this BME map we identified regions within counties where elevated arsenic is endemic. For example, southeastern Union County and the border between Stanly and Montgomery Counties are areas of special concern (Figure 2-D1). The BME estimates reveal that in these areas the arsenic concentration may reach 18 μg/L. Cross-validation analysis indicated that the BME framework better estimated arsenic concentrations than the kriging method. The MSE for the BME method was 41% lower than that of kriging (Supporting Information Table S4). In total, the geocoded data and the rigorous processing of location errors through the BME method significantly improved our understanding of arsenic distributions across unsampled areas of North Carolina.

DISCUSSION

Arsenic is a known human carcinogen and relevant environmental contaminant in drinking water systems. Publicly available data at the NCDHHS represent a volume of historical arsenic analyses in North Carolina domestic well waters performed under stringent EPA guidelines that remain largely unanalyzed. A primary goal of this research was to identify trends in areas of North Carolina with elevated arsenic concentrations in monitored domestic well waters. We revealed arsenic trends by county in monitored wells over time as well as estimated concentrations in unmonitored locations. The geocoding methods developed in this study data enabled a comprehensive report of over 4,000 yearly arsenic measurements with geographical coordinates from 1998-2007 and over 10,000 from 2008 to present, a substantial increase relative to the USGS and EPA ambient monitoring systems. Specifically, the number of records analyzed represents a 600-fold increase from samples collected by the USGS (USGS 2010) and more than three times the number of records analyzed in other studies North Carolina wells (Pippin 2005; Kim et al. *in press*). The substantial increase in recent well sampling is likely due to state legislation adopted in 2008 that requires every newly constructed well be tested. These types of monitoring programs, as evidenced here, are successful to ensuring increased awareness of well water contaminants and protected health of individual homeowners.

A notable result of this study is the surprisingly high levels of arsenic (up to 806 μg/L) that were detected in some homeowners' domestic wells. In addition, more than 1,436 (2.25%) of wells exceeded the EPA standard. Some of the top-ranked counties

identified here as most frequently exceeding the EPA MCL have not previously been highlighted in nation- or statewide studies (Welch et al. 1999) including Anson, Montgomery, Dare, Alexander, Cleveland, and Currituck Counties. We identify historical trends in counties along the Carolina terrane and through comprehensive temporal assessment reveal that arsenic levels have been elevated for over a decade. In some of these counties, greater than 50% of the population use domestic wells (Kenny et al. 2009). Importantly, some of the identified counties of concern have rapidly growing populations (US Census Bureau 2006), which compounded by an arsenic-prone geology may have public health implications for residents. Simultaneously, rural areas are more likely to lack connection to a municipal drinking water system. By ranking based on percentage of population at risk we identify counties where county-level well monitoring programs may be cost-effective. By integrating information on population size in these counties, we show Union and Stanly continue to rank among the most at-risk county populations. Our data confirm increased concentrations of arsenic in counties located along the Carolina terrane and highlight elevated levels over a decade-long period. In addition, the counties of Stanly and Union have large populations (roughly 26,000 and 49,000 individuals) relying on private well water sources. Currently, no epidemiologic literature has investigated the health impact of arsenic in these populations.

This study presents a new approach to assigning geographical coordinates when sample locations are described by inconsistent recorded information. It was evident from the spatial locations of GPS data that GPS device use was not uniform across the state. It was necessary, therefore, to increase the number of geocoded locations using additional

information (Classes II-IV) to avoid bias and provide more accurate spatial representations. As such, we applied a four-step geocoding process and error estimation scheme that increased the available geographic coordinates of ~3,000 to more than 63,000 well analyses. We increased the knowledgebase using available locational information to assign geographic coordinates of domestic wells based on four spatial classes: GPS, street address, zip code, and county. As an example, we tripled the number of successfully geocoded points used in previous analyses over comparable geographic areas and timeframes (Pippin 2005; Kim et al. *in press*). Additionally, while others have shown that multi-stage geocoding methods improved the match rate compared to single-step methods (Lovasi et al. 2007), to our knowledge, the present study is one of the first to use GPS locations to systematically quantify and account for geocoding location error. We present a widely applicable method of systematically determining acceptable match scores resulting from the multi-stage address geocoding procedure that serves as an alternative to a minimum match score determined *a priori* (Yang et al. 2004).

The general BME framework has been applied to arsenic levels in Bangladesh groundwaters to estimate aqueous concentrations where data are not available (Serre et al. 2003). In this study, we apply the BME framework to U.S. groundwaters and incorporate location error into a novel arsenic estimation process, which aggregates monitored arsenic levels to a county level observation scale (~11km). To the best of our knowledge, this is the first study that accounts for locational error. The cross-validation analysis shows that the BME approach presented in this work more accurately predicts arsenic than conventional modeling approaches. Within this framework, the county

observation scale refines well-to-well variation and we would not expect to find the high concentrations seen in individual monitored wells (e.g. up to 806 μg/L). By narrowing the scope to counties of interest in the southwestern Piedmont we identified southeastern Union County and the border between Stanly and Montgomery Counties as areas of special concern. The levels documented in this study indicate areas of arsenic contamination at nearly twice the EPA standard-a level estimated to double the risk of bladder and lung cancer (NRC 2001). The local observation scale enables our predictive maps to be useful for public health screening purposes by identifying areas where the risk of arsenic exposure is high and by providing a science-based criterion for cost-effective monitoring of wells.

Through our analyses of over 60,000 geocoded well locations we were able to more accurately assess spatial and temporal arsenic trends in both monitored wells and estimates at unmonitored locations across North Carolina. We found that areas near the eastern coast and along the Carolina terrane have high prevalence of arsenic contamination in private wells. The presence of arsenic in the Carolina terrane has been confirmed by geological studies in this area (Foley et al. 2001) and is supported by models that incorporate geologic determinants (Kim et al. *in press*). Abundant literature details the sources of groundwater arsenic contamination through anthropogenic factors including agricultural or industrial practices (Foley and Ayuso 2008; Jackson et al. 2006; Appleyard et al. 2004; Embrick et al. 2005), yet, much of arsenic contamination highlighted in this study is thought to be naturally occurring due to the underlying geology (Foley et al. 2001; Welch et al. 2000; Duker et al. 2005). The Carolina terrane,

however, does not underlie each of the top ten counties (Pender, Dare, and Currituck Counties for instance) and it is possible that a combination of anthropogenic and natural sources may contribute to arsenic contamination, however additional studies are warranted. Currently, no EPA Superfund National Priorities List or Toxic Release Inventory sites are listed in these three counties that would indicate substantial anthropogenic contribution to environmental arsenic concentrations. Moreover, while an EPA superfund site is located in Haywood County with reported historical use of arsenical pesticides, no wells in that county that exceeded the EPA standard.

Studies like this one represent a major step towards arsenic surveillance in contaminated areas. To lessen the risk of exposure to arsenic in drinking water, recommended preventative measures include point-of-use removal, modification of well depth, and/or use of an alternate water source (Alaerts and Khouri 2004; Pratson et al. 2010). These solutions are rarely cost-effective, however, and may not be feasible in rural areas. Simple cost-effective technologies for mitigating arsenic are not widely available, and, in lieu of federal or state water quality regulation of domestic wells, the best mitigation is in the form of well water testing programs. Trivalent (As3+: arsenite) and pentavalent (As5+: arsenate) arsenic most commonly occur in natural waters (Duker et al. 2005; Feng et al. 2001). Due to the variable toxicity of arsenic species (As3+ being more toxic), additional studies are warranted to determine the distribution of individual arsenic species in drinking water. Targeted monitoring is crucial in reducing the financial cost of testing for speciated arsenic in every monitored well and the methods developed here can be applied to this end towards arsenic in other regions as well as to other contaminants of

concern to public health.

The compounding of environmental and social factors means residents could be at increased risk for health effects from arsenic. Additional studies are warranted to further ascertain the sources of and biogeochemical behavior of arsenic in unstudied regions of North Carolina and to help reduce exposures among at risk populations. By identifying regions of contamination through studies such as this one, cost-effective monitoring programs can target at risk populations to protect public health and help to shape state and local water monitoring policies (Miranda et al. 2011). Moving forward we anticipate research that will integrate these findings with biomonitoring and health outcome data to substantiate risks posed to populations in arsenic endemic areas. The next steps to protecting individuals include community education in at risk areas and biomonitoring of those populations most at risk including children and pregnant women.

FIGURES

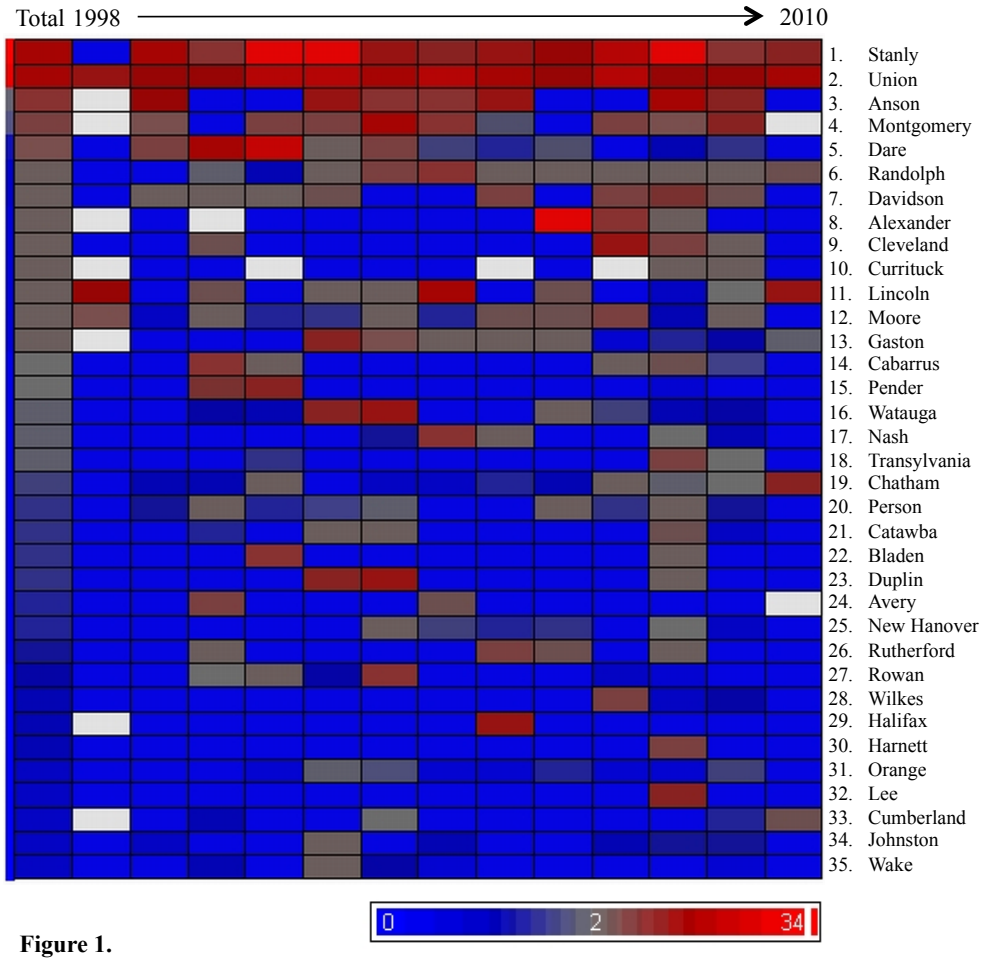

Figure 1.

Figure 1. The top thirty-five counties that exceed the EPA standard (10 μg/L). Counties are ranked by the percent of wells that exceed the EPA standard and are represented in a heat map. Counties with percent of wells exceeding the statewide 2.25% appear in red-scale, while those below the statewide percent appear in blue-scale. Counties with no information available appear in white.

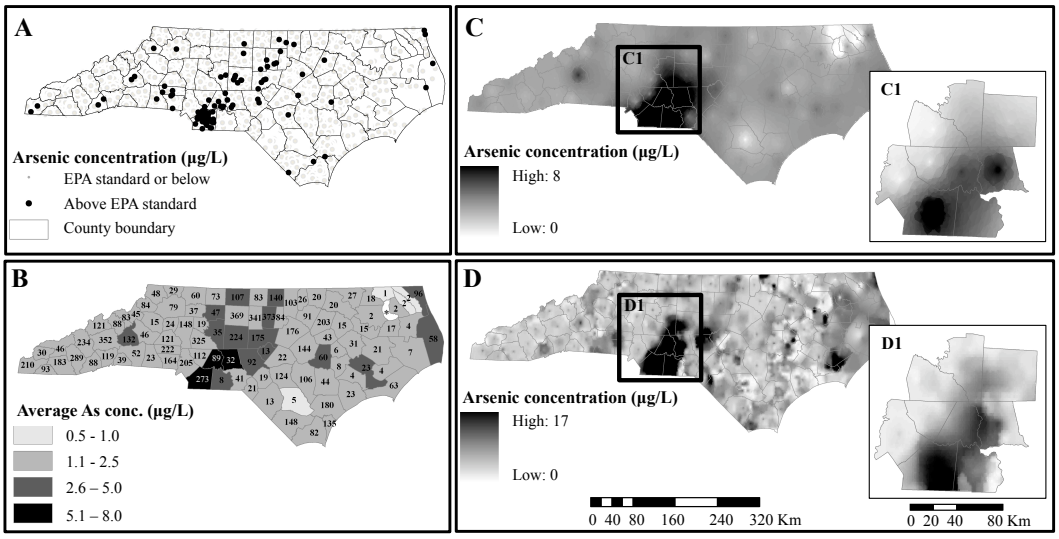

Figure 2. Geocoded arsenic concentrations in 2009. (**A**) Samples exceeding the EPA standard are shown in black. Well locations of samples below the standard appear in gray. (**B**) County averages are displayed in grayscale and the number of arsenic analyses in 2009 appear within each county. *No wells were sampled in Chowan County in 2009. (**C**) A classical kriging method estimated arsenic distribution across the state at unmonitored locations. (**D**) The Bayesian Maximum Entropy framework estimated arsenic distribution across the state at unmonitored locations.

TABLES

Table 1. Top 25-ranked[a] North Carolina counties.

Rank[a]	County	Total no. of wells sampled	No. of wells that exceed EPA standard (%)	No. of wells above detect (%)	Pop. using domestic wells, in thousands (%)[b]	Pop. at risk rank[c]
1	Stanly	849	176 (20.73)	485 (57.13)	25.947 (44.00)	1
2	Union	3250	634 (19.51)	1454 (44.74)	49.197 (30.20)	2
3	Anson	98	10 (10.20)	34 (34.69)	2.704 (10.60)	11
4	Montgomery	372	34 (9.14)	120 (32.26)	8.213 (30.06)	3
5	Dare	572	36 (6.29)	137 (23.95)	8.091 (23.87)	7
6	Randolph	1595	72 (4.51)	394 (24.70)	66.845 (48.31)	4
7	Davidson	552	23 (4.17)	106 (19.20)	36.893 (23.86)	13
8	Alexander	128	5 (3.91)	30 (23.44)	6.818 (19.21)	15
9	Cleveland	269	9 (3.35)	37 (13.75)	9.234 (9.39)	29
10	Currituck	153	5 (3.27)	24 (15.69)	3.492 (15.11)	23
11	Lincoln	990	31 (3.13)	139 (14.04)	39.858 (57.06)	5
12	Moore	1093	33 (3.02)	206 (18.85)	34.920 (42.75)	8
13	Gaston	1697	47 (2.77)	278 (16.38)	57.915 (29.53)	14
14	Cabarrus	626	15 (2.40)	98 (15.65)	37.367 (24.87)	21
15	Pender	800	19 (2.38)	115 (14.38)	33.179 (71.46)	6
16	Watauga	671	15 (2.24)	77 (11.48)	11.970 (28.18)	20
17	Nash	1137	25 (2.20)	145 (12.75)	4.870 (5.33)	43
18	Transylvania	424	9 (2.12)	24 (5.66)	15.156 (51.16)	10
19	Chatham	1404	26 (1.85)	455 (32.41)	32.080 (55.31)	12
20	Person	847	15 (1.77)	165 (19.48)	25.605 (68.80)	9
21	Catawba	454	8 (1.76)	42 (9.25)	62.709 (41.35)	17
22	Bladen	114	2 (1.75)	7 (6.14)	13.898 (42.19)	16
23	Duplin	240	4 (1.67)	13 (5.42)	15.305 (29.44)	24
24	Avery	261	4 (1.53)	28 (10.73)	8.291 (47.00)	18
25	New Hanover	1449	22 (1.52)	248 (17.12)	16.371 (9.12)	41
--	Other counties	43811	157 (0.36)	2852 (6.51)	1668.398 (24.95)	--
Total	--	63856	1436 (2.25)	7713 (12.08)	2295.33 (26.43)	--

[a]Rank based on tendency of wells to exceed the EPA standard throughout an 11-year period; [b]Data from (Kenny et al. 2009); [c]Rank based on composite of percentage of wells that exceed the EPA standard through the 11-year period and percentage of county residents using private domestic well water.

REFERENCES

Alaerts GJ, Khouri N. 2004. Arsenic contamination of groundwater: Mitigation strategies and policies. Hydrogeology Journal 12(1): 103-114.

Appleyard S, Wong S, Willis-Jones B, Angeloni J, Watkins R. 2004. Groundwater acidification caused by urban development in Perth, Western Australia: source, distribution, and implications for management. Australian Journal of Soil Research 42 (5-6): 579-585.

Ayotte JD, Montgomery DL, Flanagan SM, Robinson KW. 2003. Arsenic in groundwater in eastern New England: Occurrence, controls, and human health implications. Environ Sci Technol 37(10): 2075-2083.

Bellander T, Berglind N, Gustavsson P, Jonson T, Nyberg F, Pershagen G, et al. 2001. Using geographic information systems to assess individual historical exposure to air pollution from traffic and house heating in Stockholm. Environ Health Perspect 109(6): 633-639.

Christakos G. 1990. A Bayesian Maximum-Entropy view to the spatial estimation problem. Mathematical Geology 22(7): 763-777.

Christakos G. 2000. Modern Spatiotemporal Geostatistics: Oxford University Press.

Christakos G, Bogaert P, Serre ML. 2002. Temporal GIS: Advanced Functions for Field-Based Applications. New York, NY: Springer-Verlag.

Cressie N. 1990. THE ORIGINS OF KRIGING. Mathematical Geology 22(3): 239-252.

De Nazelle A, Arunachalam S, Serre ML. 2010. Bayesian Maximum Entropy Integration of Ozone Observations and Model Predictions: An Application for Attainment Demonstration in North Carolina. Environ Sci Technol 44(15): 5707-5713.

Duker AA, Carranza EJM, Hale M. 2005. Arsenic geochemistry and health. Environ Int 31(5): 631-641.

Embrick LL, Porter KM, Pendergrass A, Butcher DJ. 2005. Characterization of lead and arsenic contamination at Barber Orchard, Haywood County, NC. Microchemical Journal 81 (1): 117-121.

EPA. 1994. Method 200.8 Revision 5.4 - Determination of Trace Elements in Water and Wastes by Inductively Coupled Plasma-Mass Spectrometry. Cincinnati, OH.

Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95(25): 14863-14868.

EPA. 2010. Arsenic in Drinking Water. Available: http://epa.gov/safewater/arsenic/index.html [accessed May 24 2010].

Feng ZM, Xia YJ, Tian DF, Wu KK, Schmitt M, Kwok RK, et al. 2001. DNA damage in buccal epithelial cells from individuals chronically exposed to arsenic via drinking water in Inner Mongolia, China. Anticancer Res 21(1A): 51-57.

Foley N, Ayuso RA, Seal R. 2001. Remnant colloform pyrite at the Haile gold deposit, South Carolina: A textural key to genesis. Econ Geol Bull Soc Econ Geol 96(4): 891-902.

Foley NK, Ayuso RA. 2008. Mineral sources and transport pathways for arsenic release in a coastal watershed, USA. Geochem.-Explor. Environ. Anal. 8: 59-75.

Hagan EF. 2004. Ground Water Quality Technical Brief: Statewide ambient groundwater quality monitoring program arsenic speciation results (2002 & 2003).

Holton WC. 2002. Locating lead - Mapping leads to intervention. Environ Health Perspect 110(9): A533-A533.

Hulbert IAR, French J. 2001. The accuracy of GPS for wildlife telemetry and habitat mapping. Journal of Applied Ecology 38(4): 869-878.

Jackson BP, Seaman JC, Bertsch PM. 2006. Fate of arsenic compounds in poultry litter upon land application. Chemosphere 65 (11): 2028-2034.

Kim D, Miranda ML, Tootoo J, Bradley P, Gelfand AE. 2011 (*in press)*. Spatial modeling for groundwater arsenic levels in North Carolina. Environ Sci Technol *in press*.

Kim MJ, Nriagu J, Haack S. 2002. Arsenic species and chemistry in groundwater of southeast Michigan. Environ Pollut 120(2): 379-390.

Kenny JF, Barber NL, Hutson SS, Linsey KS, Lovelace JK, Maupin MA. 2009. Estimated use of water use in the United States in 2005.

Kumar A, Adak P, Gurian PL, Lockwood JR. 2010. Arsenic exposure in US public and domestic drinking water supplies: A comparative risk assessment. J Expo Sci Environ Epidemiol 20(3): 245-254.

Lovasi GS, Weiss JC, Hoskins R, Whitsel EA, Rice K, Erickson CF, et al. 2007. Comparing a single-stage geocoding method to a multi-stage geocoding method: how much and where do they disagree? International Journal of Health Geographics 6.

Money E, Carter GP, Serre ML. 2009. Using river distances in the space/time estimation of dissolved oxygen along two impaired river networks in New Jersey. Water Research 43(7): 1948-1958.

Miranda ML, Dolinoy DC, Overstreet MA. 2002. Mapping for prevention: GIS models for directing childhood lead poisoning prevention programs. Environ Health Perspect 110(9): 947-953.

Miranda ML, Edwards SE. 2011. Use of spatial analysis to support environmental health research and practice. North Carolina Medical Journal 72 (2): 132-135.

Mukherjee A, Sengupta MK, Hossain MA, Ahamed S, Das B, Nayak B, et al. 2006. Arsenic contamination in groundwater: A global perspective with emphasis on the Asian scenario. Journal of Health Population and Nutrition 24(2): 142-163.

National Research Council (NRC). 2001. Arsenic in drinking water: 2001 Update. Washington, DC: National Academy Press.

Nuckols JR, Ward MH, Jarup L. 2004. Using geographic information systems for exposure assessment in environmental epidemiology studies. Environ Health Perspect 112(9): 1007-1015.

Peters SC, Blum JD, Klaue B, Karagas MR. 1999. Arsenic occurrence in New Hampshire drinking water. Environ Sci Technol 33(9): 1328-1333.

Pippin CG. 2005. Distribution of total arsenic in groundwater in the North Carolina Piedmont. NGWA Naturally Occurring Contaminants Conference: Arsenic, Radium, Radon, and Uranium, 89-102.

Pratson E, Vengosh A, Dwyer G, Pratson L, Klein E. 2010. The Effectiveness of Arsenic Remediation from Groundwater in a Private Home. Ground Water Monitoring and Remediation 30(1): 87-93.

Serre ML, Christakos G. 1999. Modern geostatistics: computational BME analysis in the light of uncertain physical knowledge - the Equus Beds study. Stochastic Environmental Research and Risk Assessment 13(1-2): 1-26.

Serre ML, Kolovos A, Christakos G, Modis K. 2003. An application of the holistochastic human exposure methodology to naturally occurring arsenic in Bangladesh drinking water. Risk Anal 23(3): 515-528.

Shaw WD, Walker M, Benson M. 2005. Treating and drinking well water in the presence of health risks from arsenic contamination: Results from a US hot spot. Risk Anal 25(6): 1531-1543.

Shiber JG. 2005. Arsenic in domestic well water and health in Central Appalachia, USA. Water Air Soil Pollut 160(1-4): 327-341.

U.S. Census Bureau. 2006. Population Estimates for the 100 Fastest-Growing U.S. Counties by Percentage Growth from July 1, 2004 to July 1, 2005. Washington, DC.

USGS. 2010. National Water Information System. Available: http://waterdata/usgs.gov/nc/nwis/qwdata 2010].

Walker M, Shaw WD, Benson M. 2006. Arsenic consumption and health risk perceptions in a rural western US area. J Am Water Resour Assoc 42(5): 1363-1370.

Welch AH, Helsel DR, Focazio MJ, Watkins SA. 1999. Arsenic in ground water supplies of the United States. New York:Elsevier Science.

Welch AH, Westjohn DB, Helsel DR, Wanty RB. 2000. Arsenic in ground water of the United States: Occurrence and geochemistry. Ground Water 38(4): 589-604.

Wing MG, Eklund A. 2007. Performance comparison of a low-cost mapping grade global positioning systems (GPS) receiver and consumer grade GPS receiver under dense forest canopy. Journal of Forestry 105(1): 9-14.

Yang D, Bilaver L, Hayes O, Goerge R. 2004. Improving geocoding practices: evaluation of geocoding tools. Journal of Medical Systems 28(4): 361-370.

Yang Q, Jung HB, Culbertson CW, Marvinney RG, Loiselle MC, Locke DB, et al. 2009. Spatial Pattern of Groundwater Arsenic Occurrence and Association with Bedrock Geology in Greater Augusta, Maine. Environ Sci Technol 43(8): 2714-2719.

APPENDIX TWO: TOWARDS PRENATAL BIOMONITORING IN NORTH CAROLINA: ASSESSING ARSENIC, CADMIUM, MERCURY, AND LEAD LEVELS IN PREGNANT WOMEN

Alison P. Sanders[1], Kaye Flood[2], Shu Chiang[2], Amy H. Herring[3], Leslie Wolf[2], and Rebecca C. Fry[1*]

Authors' Affiliations:

[1] Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina.

[2] State Laboratory of Public Health, North Carolina Department of Health and Human Services, Raleigh, North Carolina.

[3] Department of Biostatistics, Gillings School of Global Public Health, and Carolina Population Center, University of North Carolina, Chapel Hill, North Carolina.

ABSTRACT

Exposure to toxic metals during the prenatal period carries the potential for adverse developmental effects to the fetus, yet such exposure remains largely unmonitored in the United States. The aim of this study was to assess maternal exposure to four toxic metals (arsenic (As), cadmium (Cd), mercury (Hg), and lead (Pb)) in a cohort of pregnant women in North Carolina. We analyzed blood samples submitted to the North Carolina Department of Health and Human Services for blood typing to assess toxic metal levels in pregnant women (n=211) across six North Carolina counties. Whole blood metal concentrations were measured by inductively coupled plasma mass spectrometry. The association between maternal characteristics, including county of residence, age, and race, and metal exposure was analyzed using multiple linear regression analysis. A large fraction of the blood samples showed detectable levels for each of the four metals. Specifically, As (65.7%), Cd (57.3%), Hg (63.8%), and Pb (100%) were detected in blood samples. Moreover, compared with adult females participating in the *Fourth National Report on Human Exposure to Environmental Chemicals* and guidelines for pregnant women, some women in the sample population exceeded benchmark levels of Cd, Hg, and Pb. Evidence from this pilot study indicates that pregnant women in North Carolina are exposed to As, Cd, Hg, and Pb and suggests that factors related to maternal county of residence and race may impact maternal exposure levels. As increased levels of one or more of these metals *in utero* have been associated with detrimental developmental and reproductive outcomes, further study is clearly warranted to establish the impacts to newborns.

INTRODUCTION

Arsenic (As), cadmium (Cd), mercury (Hg), and lead (Pb) are metals ranked among the top ten most toxic substances by the Agency for Toxic Substances and Disease Registry [1]. Maternal exposures to toxic metals may result from diet, air, drinking water, occupational exposures, and/or tobacco use. Evidence suggests that each of these metals is able to cross the placental barrier resulting in prenatal exposure [2,3,4,5] and maternal blood levels of these contaminants correlate with umbilical cord blood levels [5,6]. Biologically, prenatal exposure to metals is of concern for a variety of reasons. For example, prenatal exposure to arsenic is associated with later life health effects in adults [7,8,9] including increased mortality and increased risk of lung and liver cancer [10,11,12]. In addition, *in utero* exposure to arsenic has been shown to alter genomic signaling of key biological pathways [13]. Furthermore, maternal exposures to toxic metals can increase the risk for poor birth outcomes, including low birth weight [3,14,15], reduced fetal growth [16], and reproductive and cognitive deficits in adolescents [17,18,19].

Toxic metal exposures *in utero* and during childhood may result in significant health effects to individuals that manifest cumulatively to the detriment of populations. U.S. annual economic losses due to decreased productivity from environmental exposure to Hg and Pb were estimated as $8.7 and $43.4 billion respectively [20,21]. In spite of strong evidence to indicate the adverse health effects of metals to susceptible populations such as newborns, few statewide biomonitoring programs are in place. For example, while North Carolina has established a successful program for childhood lead screening [22], there remains no prenatal assessment of environmental contaminants. This is of

concern in North Carolina because the occurrence of metals including As, Pb, and Hg has been reported in environmental reservoirs with the potential for human exposure [23,24,25].

Known health risks are associated with exposure to heavy metals, particularly during the periconceptional and prenatal periods. Established recommended advisory blood levels of 5.8 μg/L Hg and 10.0 μg/dL Pb are considered protective of human health for the general population [26,27]. However, because cord blood levels of Hg and Pb have been found to exceed maternal blood levels, the maternal blood levels of 3.5 μg/L Hg and 5.0 μg/dL Pb are recommended for the protection of fetal health [26,28,29]. Blood levels in individuals exceeding these reference concentrations are associated with decreased IQ and cognitive function [20,26]. Despite evidence of associated detrimental health effects, there are currently no biological advisory levels pertaining to As or Cd for pregnant women or the general population.

The major aim of this pilot study was to assess blood levels of As, Cd, Hg, and Pb in pregnant women residing in six North Carolina counties and to examine associations of metal levels according to maternal age, race, and county of residence. Evidence gathered in this study will be used to further assess maternal and prenatal exposures in future studies and to provide impetus for targeted biomonitoring programs or public health campaigns in North Carolina. Biomonitoring of maternal levels of toxic metals in targeted populations may help reduce prenatal exposures and subsequently the potential for future adverse developmental effects.

METHODS

Study Design

To assess blood levels of toxic metals in pregnant women, we analyzed blood samples submitted to the North Carolina Department of Health and Human Services (NCDHHS) State Laboratory of Public Health for blood group typing. Residual volumes of blood originally submitted for group typing were analyzed for As, Cd, Hg, and Pb. Data were provided under the context of a data use agreement for a limited data set using a data custodian as an intermediary to protect identifiable information. This study was reviewed by the UNC Public Health-Nursing IRB (#09-1399) and was determined to be exempt from review according to regulatory category 4 under the National Institutes of Health description on Research on Human Specimens (44 CFR 46.101(b)). Similar strategies that assess metal levels in samples collected under the auspices of existing public health monitoring programs have been used to establish environmental contaminant levels in individuals [30].

Women in this study were North Carolina residents in their third trimester of pregnancy and receiving prenatal care at local health departments. The mothers' venous blood samples were collected between October 2009 and January 2011. Bladen, Cumberland, Richmond, Stanly, Union, and Wake Counties were selected for analysis based on trends in private well water metal levels reported in publicly available data [31]. Throughout the remainder of the manuscript county identities are coded and designated as A-F, using random assignment. Data collected included: (1) pregnant women's (n=211) whole blood level of As, Cd, Hg, and Pb, (2) county of residence, (3) age, and

(4) race. No additional demographic information on subjects was requested in accordance with the limited data use agreement.

Blood Analysis

The North Carolina Department of Health and Human Services (NCDHHS) for blood typing group is certified by the U.S. DHHS as compliant with the 1988 Clinical Laboratory Improvement Amendment (CLIA). In accordance with CLIA certification, the laboratory adheres to quality assessment/quality control (QA/QC) requirements with oversight by the Centers for Medicare and Medicaid Services. Blood samples from non-fasting pregnant women were collected into phlebotomy tubes containing ethylenediaminetetraacetic acid (EDTA) as an anticoagulant and stored at 4°C prior to processing at the Division of Public Health State Laboratory. Whole blood analysis was performed via Inductively Coupled Plasma Dynamic Reaction Cell Mass Spectrometry (ICP/DRC/MS) according to the US CDC laboratory methods for metals in whole blood, DLS Method code CTL-TMS 3.01 [32,33]. The samples were analyzed on a Perkin Elmer Elan DRC II ICP/MS instrument. Arsenic was measured in blood using the DRC mode to remove isobaric interference from Argon Chloride at mass 75. Iridium was used as an internal standard for all elements. The limits of detection (LODs) were 0.23 μg-As/L, 0.11 μg-Cd/L, 0.23 μg-Hg/L, and 0.13 μg-Pb/dL. The results represent levels of total metal concentrations; speciation of metals was not performed. The ICP-MS analyses for Cd, Hg, and Pb were similar to those used in the Fourth National Report on Human Exposure to Environmental Chemicals (NHANES IV).

Statistical analysis

Statistical analysis was conducted with the statistical package SAS 9.2 (SAS Institute Inc., Cary, North Carolina). Blood samples with metal levels below the LOD were assigned an imputed value equal to LOD/√2, as in NHANES IV [34]. The data were log-transformed to address the non-normal distribution of metal concentrations. Spearman's correlation coefficient estimates among metal pairs and corresponding p-values were calculated. Linear regression was performed on every pairwise metal combination by adjusting for maternal age to test for associations between blood metal levels and race (NHW served as the referent group). Additional multivariate analyses were performed adjusting for maternal age, race, and categorized county of residence to assess the relationship between metal levels and residence in individual counties. Furthermore, differences between the mean metal levels of individual counties were compared using Scheffe's test (α=0.05).

RESULTS

Here we monitored metal levels in blood samples collected from pregnant women (n=211) residing in North Carolina between October 2009 and January 2011. Study participant characteristics are presented in Table 1. The average woman's age was 25 years and ranged from 15 to 43 years. The study population was 2% Asian, 38% non-Hispanic Black (NHB), 6% Hispanic, and 55% non-Hispanic White (NHW). The monitored women were residents of County A (n=50), County B (n=28), County C (n=50), County D (n=13), County E (n=25), and County F (n=45) in North Carolina.

From the 211 individuals in the sample population, blood metal levels were detected among 138 for As (65.7%), 121 for Cd (57.3%), 134 for Hg (63.8%), and 211 for Pb (100%) samples. Descriptive statistics of metal levels are presented in Table 2. The 95th percentile whole blood Cd, Hg, and Pb levels among women of childbearing age (age 16-49) participating in NHANES IV (n=4,241) were reported as 1.60 μg/L, 4.40 μg/L, and 3.50 μg/dL, respectively [34]. A total of three (1.4%), five (2.4%), and four (1.9%) women in the study population exceeded these levels for Cd, Hg and Pb, respectively. Arsenic in whole blood was not measured in NHANES IV and therefore could not be compared to these data. However, three (1.4%) women here had levels of As comparable to those reported for an environmentally exposed cohort of pregnant women [2]. Blood Cd levels were within previously reported ranges for pregnant women, however, some Hg and Pb levels exceeded reported levels [35]. Specifically, of the pregnant women sampled, four women (1.9%) had Hg concentrations above the EPA blood guideline of 5.8 μg/L and two women (0.94%) had Pb levels above 5.0 μg/dL (the

CDC blood lead advisory level for pregnant women) [26,27] (Figure 1). Additionally, five women (2.3%) had blood Hg levels above the 3.5 μg/L level of concern during pregnancy [28,29].

Our analysis revealed relationships among some sample population characteristics and maternal metal levels. Table 3 shows the unadjusted Spearman correlation coefficients computed for each pairwise combination of metals. Notably, maternal levels of As and Hg positively covaried in the study population ($r=0.45$, $p<0.001$). This suggests that for some pregnant mothers, As and Hg are elevated in tandem. Specifically samples collected from Counties A, C, and F demonstrated a significant positive correlation between maternal As and Hg ($p<0.05$, data not shown). No other metal pair correlations were statistically significant at the $\alpha=0.001$ significance level.

Our data also indicate that maternal race and residence in certain counties were associated with elevated body burden of some toxic metals. Linear regression of maternal race on metal levels adjusted for maternal age is presented in Table 4. The results reveal that NHB maternal race was associated with increased blood As ($p<0.05$) and Cd levels ($p<0.001$) when compared to NHW mothers. Asian women in this study also appear to have elevated levels of As, Cd, and Hg, although the sample size was small and additional study is needed to verify this trend.

Table 5 shows the geometric mean metal levels by coded maternal county of residence. The compared differences among average metal levels in individual counties using Scheffe's test after adjusting for maternal age and race are also displayed. The highest geometric mean As level (0.87 μg/L) was in County F. The As levels in samples

collected from Counties C and F were significantly elevated when compared to at least one other county (Scheffe's $p<0.05$; Table 5). Cd levels in samples collected from Counties A and C were significantly elevated when compared to at least one other county (Scheffe's $p<0.05$), and the highest geometric mean level of Cd (0.27 μg/L) was in County A. Among Hg levels, samples from County F had the highest geometric mean Hg (0.87 μg/L), and average levels in Counties D and F were significantly elevated when compared to one or more other counties (Scheffe's $p<0.05$; Table 5). Finally, maternal blood levels of Pb in County C had the highest geometric mean (1.28 μg/L), and were significantly elevated compared individually to those from Counties A, E, and F (Scheffe's $p<0.05$). Notably, samples collected from County C had the highest levels of As, Cd, and Pb. Samples collected from County F had elevated average levels of As and Hg. Counties A and D had elevated average levels of Hg and Cd, respectively. Average Pb and Hg levels exceeded those reported for women participating in NHANES IV in County C and County F, respectively (Table 5) [34].

DISCUSSION

Pregnant women are exposed to toxic metals

The aim of this study was to assess toxic metal exposure in pregnant mothers in North Carolina. The data suggest that indeed pregnant women in North Carolina are exposed to toxic metals. For each of the four toxic metals studied, more than 50% of the women had detectable blood levels. Specifically, depending on the metal detectable blood levels were found in 57-100% of the samples. Most of the metal measurements revealed levels of exposure that did not fall outside regulated limits. For some individuals, however, blood levels of toxic metals exceeded the 95th percentile of Cd, Hg, and Pb reported in the U.S. population. Moreover, a few women had metal levels above current guideline levels adopted for health protection. Specifically, some women had blood metal levels above reference values indicative of a threshold for adverse health effects: five women (2.3%) had blood Hg levels above 3.5 μg/L and two women (0.94%) had Pb levels greater than 5.0 μg/dL (the CDC blood lead advisory). Since no current threshold levels exist for As in blood we compared our findings with a previous study by Concha et al [2]. Three women in our study showed levels similar to those reported among As-exposed pregnant women served by contaminated public water supplies containing up to 200 μg/L As (blood range: 5.6-13 μg/L) [2]. In addition, thirty-seven women had blood levels greater than 1 μg/L, a level previously associated with known exposure to As in drinking water [36].

As and Hg are positively correlated in pregnant mothers' blood

Environmental exposures in the general public more commonly occur as mixtures than as single-contaminant exposures. Contaminant mixtures can contribute to an individuals' susceptibility and potentially have synergistic effects resulting in disease manifestation [37,38]. Here we examined the correlation of the four toxic metals among the pregnant women. Notably, we report a relationship between maternal As and Hg blood levels. An implication of this finding is that maternal and fetal body burden may result from combined or interactive effects of more than one metal. Thus, concomitant exposure is an important consideration in future studies that assess *in utero* health outcomes from metal exposure.

Non-Hispanic Black maternal race is associated with increased Cd

The findings also highlight associations between blood metal levels and maternal race. Our data suggest that across the sample population NHB maternal race was associated with increased Cd exposure compared to NHW women. To our knowledge this association has not been previously reported. Poor birth outcomes in North Carolina and across the U.S. disproportionately affect infants born to NHB mothers [39,40], despite the lack of a definitive causal mechanism. Our findings suggest that the role of maternal race in birth outcome etiology may be confounded by metal exposure, and further investigation of social, environmental, and genetic factors disproportionately affecting NHB mothers is needed.

County of residence is associated with mothers' metal levels

This initial study suggests that on average, women residing in various North Carolina counties experience differential exposure to toxic metals during pregnancy. We show that on average, pregnant women residing in Counties C and F were likely to have higher levels of As when compared to women residing in the other counties. Exposure to As primarily occurs through ingestion of food and drinking water [41]. In public drinking water supplies, allowable levels of As can range up to the EPA Maximum Contaminant Level (MCL) of 10 μg/L. As such, individuals can be exposed to detectable levels of As that are still within regulatory guidelines. We have shown recently that in North Carolina, some counties have elevated levels of As in unregulated private well water [23]. Here we show that women living in Counties C and F have elevated blood As, two counties that were not identified as particular outliers in the well water analysis. This could suggest an alternate route of exposure among residents, perhaps through diet. In addition, this study did not select for individuals based on their type of drinking water source. Taken together, these results suggest that a future study of women supplied by private drinking water sources may be prudent.

We also found that on average, pregnant women residing in Counties D and F also had elevated levels of blood Hg. Most Hg exposures in the U.S. are attributable to diet, particularly fish consumption [42]; however, dental amalgams [4] and some industrial processes also contribute to human exposure [43]. In North Carolina, a statewide freshwater fish advisory for mercury (specifically in largemouth bass) has been in effect since 2009. A recent study in Durham County, North Carolina found that the type of fish consumption and corresponding Hg content varied among different

demographic groups of pregnant women [24]. There are no reports of public or private well drinking water Hg levels above acceptable EPA MCL standard of 2 μg/L in either of these counties.

Cadmium exposure commonly occurs through food consumption or inhalation of cigarette smoke, although minor exposure can occur from byproducts of industrial processes [44]. Previous studies have shown Cd levels in maternal and cord blood are higher in smokers [45], and that Cd levels among smokers were reportedly 4-5 times higher than nonsmokers [46]. Moreover, Cd levels in pregnant women exposed to second–hand smoke were double that of nonsmokers [6]. In our study, residence in County A was associated with increased maternal blood Cd levels; yet, the smoking rate in this county is lower than the statewide rate (20.9% of adults) [47]. Detectable levels of Cd in public drinking water can range up to 5.0 μg/L (the EPA MCL) and may potentially contribute to low-level exposure in the general population. Further research is needed to establish the source of maternal Cd exposure.

Exposure to Pb can result from various environmental sources such as contaminated dust or soil, drinking water, lead-based paint, cigarette smoke, as well as byproducts of industrial processes [48]. Elevated maternal blood levels of Pb were associated with residence in County C. Similar to the other metals, Pb levels in public drinking water supplies can range up to the 15 μg/L and up to 10% of samples can exceed this level under EPA guidelines. Currently, no evidence indicates that North Carolina public water sources contain endemic levels of lead. While drinking water is a potential route of lead exposure, rural counties may experience additional risk factors for Pb exposure such as older housing and low socioeconomic status that may put mothers at

risk [49].

Although the design of this study did not allow us to ascertain exposure sources, both anthropogenic and environmental sources can contribute to an individuals' toxic metal exposure. Body weight, race, occupational factors, duration of residence, socioeconomic status, smoking status, gestational time during pregnancy, and genetic factors can affect the body burden of metals. It is likely that multiple exposure routes and sources may have contributed to the metal levels assessed in this study and warrant future follow-up.

Biomonitoring programs may reduce the risk of prenatal metal exposure

Despite evidence that adverse health effects are attributable to *in utero* metal exposure, currently Minnesota, New York City, and New York State are the only jurisdictions that have active guidelines for monitoring maternal blood lead levels in the U.S. [26]. The initiation of statewide metal biomonitoring programs in the U.S. may be accomplished through collaborations with state programs to better protect public health, particularly in susceptible populations such as pregnant women and children. This study demonstrates that detectable metal levels were present in maternal blood and suggests that further biomonitoring of blood levels of toxic metals in susceptible North Carolina populations is prudent. Furthermore, *in lieu* of existing targeted biomonitoring programs, public health strategies to prevent exposure such as raising awareness among pregnant mothers may be vital to protecting the health of newborns.

In conclusion, this study demonstrates that the pregnant women under study were exposed to four toxic metals at detectable levels. Specifically, we show evidence of elevated levels of As, Hg, and Pb among some women in the third trimester of pregnancy. In addition, we identified statistically significant relationships between maternal As and Hg in blood levels. Furthermore, we report associations between maternal blood metal levels and maternal race as well as county of residence. The increased levels of one or more of these metals are associated with detrimental neurological and physiological developmental outcomes. Informing pregnant women of the sources of toxic metals such as diet (including drinking water), cigarette smoke, and other environmental sources may help to reduce exposures. Targeted national, statewide, and community level programs promoting mothers' awareness of metals exposure and *in utero* effects would each help to protect human health and lessen the economic burden of metal exposure. Biomonitoring programs in North Carolina and other at-risk areas can help to protect the health of expectant mothers and their children.

REFERENCES

1. ATSDR (2010) CERCLA Priority list of hazardous substances for 2007. Atlanta, GA: Centers for Disease Control.

2. Concha G, Vogler G, Lezcano D, Nermell B, Vahter M (1998) Exposure to inorganic arsenic metabolites during early human development. Toxicological Sciences 44: 185-190.

3. Salpietro CD, Gangemi S, Minciuollo PL, Briuglia S, Merlino MV, et al. (2002) Cadmium concentration in maternal and cord blood and infant birth weight: a study on healthy non-smoking women. Journal of Perinatal Medicine 30: 395-399.

4. Clifton JC (2007) Mercury exposure and public health. Pediatric Clinics of North America 54: 237-269+.

5. Rudge CV, Rollin HB, Nogueira CM, Thomassen Y, Rudge MC, et al. (2009) The placenta as a barrier for toxic and essential elements in paired maternal and cord blood samples of South African delivering women. Journal of Environmental Monitoring 11: 1322-1330.

6. Guan HA, Piao FY, Li XW, Li QJ, Xu L, et al. (2010) Maternal and Fetal Exposure to Four Carcinogenic Environmental Metals. Biomedical and Environmental Sciences 23: 458-465.

7. Vahter M (2008) Health effects of early life exposure to arsenic. Basic & Clinical Pharmacology & Toxicology 102: 204-211.

8. Tokar EJ, Qu W, Waalkes MP (2011) Arsenic, Stem Cells, and the Developmental Basis of Adult Cancer. Toxicological Sciences 120: S192-S203.

9. Waalkes MP, Liu J (2008) Early-life arsenic exposure: Methylation capacity and beyond. Environmental Health Perspectives 116: A104-A104.

10. Smith AH, Marshall G, Yuan Y, Ferreccio C, Liaw J, et al. (2006) Increased mortality from lung cancer and bronchiectasis in young adults after exposure to arsenic in utero and in early childhood. Environmental Health Perspectives 114: 1293-1296.

11. Liaw J, Marshall G, Yuan Y, Ferreccio C, Steinmaus C, et al. (2008) Increased childhood liver cancer mortality and arsenic in drinking water in northern Chile. Cancer Epidemiology Biomarkers & Prevention 17: 1982-1987.

12. Liu J, Waalkes MP (2008) Liver is a target of arsenic carcinogenesis. Toxicological Sciences 105: 24-32.

13. Fry RC, Navasumrit P, Valiathan C, Svensson JP, Hogan BJ, et al. (2007) Activation of inflammation/NF-kappa B signaling in infants born to arsenic-exposed mothers. Plos Genetics 3: 2180-2189.

14. Huyck KL, Kile ML, Mahiuddin G, Quamruzzaman Q, Rahman M, et al. (2007) Maternal arsenic exposure associated with low birth weight in Bangladesh. Journal of Occupational and Environmental Medicine 49: 1097-1104.

15. Frery N, Nessmann C, Girard F, Lafond J, Moreau T, et al. (1993) Environmental exposure to cadmium and human birth-weight. Toxicology 79: 109-118.

16. Quansah R, Jaakkola JJK (2009) Paternal and maternal exposure to welding fumes and metal dusts or fumes and adverse pregnancy outcomes. International Archives of Occupational and Environmental Health 82: 529-537.

17. Ronchetti R, van den Hazel P, Schoeters G, Hanke W, Rennezova Z, et al. (2006) Lead neurotoxicity in children: Is prenatal exposure more important than postnatal exposure? Acta Paediatrica 95: 45-49.

18. Thompson J, Bannigan J (2008) Cadmium: Toxic effects on the reproductive system and the embryo. Reproductive Toxicology 25: 304-315.

19. Jedrychowski W, Jankowski J, Flak E, Skarupa A, Mroz E, et al. (2006) Effects of prenatal exposure to mercury on cognitive and psychomotor function in one-year-old infants: Epidemiologic cohort study in Poland. Annals of Epidemiology 16: 439-447.

20. Trasande L, Landrigan PJ, Schechter C (2005) Public health and economic consequences of methyl mercury toxicity to the developing brain. Environmental Health Perspectives 113: 590-596.

21. Landrigan PJ, Schechter CB, Lipton JM, Fahs MC, Schwartz J (2002) Environmental pollutants and disease in American children: Estimates of morbidity, mortality, and costs for lead poisoning, asthma, cancer, and developmental disabilities. Environmental Health Perspectives 110: 721-728.

22. NCDHHS (2009) North Carolina Strategic Plan to Eliminate Childhood Lead Poisoning: North Carolina Childhood Lead Poisoning Prevention Program (NCCLPPP). Children's Environmental Health Branch. Raleigh, NC.

23. Sanders AP, Messier KP, Shehee M, Rudo K, Serre ML, et al. (2011) Arsenic in North Carolina: Public Health Implications. Environment International 38: 10-16.

24. Miranda ML, Edwards S, Maxson PJ (2011) Mercury Levels in an Urban Pregnant Population in Durham County, North Carolina. International Journal of Environmental Research and Public Health 8: 698-712.

25. Miranda ML, Dolinoy DC, Overstreet MA (2002) Mapping for prevention: GIS models for directing childhood lead poisoning prevention programs. Environmental Health Perspectives 110: 947-953.

26. CDC (Eds.: Ettinger AS, Wengrovitz, AG, Portier, C, Brown, MJ) (2010) Guidelines for the identification and management of lead exposure in pregnant and lactating women. Atlanta, GA: DHHS.

27. EPA (2010) Methylmercury exposure. Available online: http://www.epa.gov/mercury/exposure.htm. Accessed July 21, 2011.

28. Mahaffey KR, Clickner RP, Jeffries RA (2009) Adult Women's Blood Mercury Concentrations Vary Regionally in the United States: Association with Patterns of Fish Consumption (NHANES 1999-2004). Environmental Health Perspectives 117: 47-53.

29. Hightower JM, O'Hare A, Hernandez GT (2006) Blood mercury reporting in NHANES: Identifying Asian, Pacific Islander, Native American, and multiracial groups. Environmental Health Perspectives 114: 173-175.

30. Rice DC, Lincoln R, Martha J, Parker L, Pote K, et al. (2010) Concentration of metals in blood of Maine children 1-6 years old. Journal of Exposure Science and Environmental Epidemiology 20: 634-643.

31. NCDHHS (2011) North Carolina State Laboratory of Public Health: Inorganic Sample Results. Available online: http://slphreporting.ncpublichealth.com/EnvironmentalSciences/InOrganicChemistry/.

32. CDC (2008) Laboratory Procedure Manual: Lead and Cadmium in Whole Blood. Division of Laboratory Sciences National Center for Environmental Health: Atlanta, GA.

33. McShane WJ, Pappas RS, Wilson-McElprang V, Paschal D (2008) A rugged and transferable method for determining blood cadmium, mercury, and lead with inductively coupled plasma-mass spectrometry. Spectrochimica Acta Part B-Atomic Spectroscopy 63: 638-644.

34. CDC (2009) Fourth National Report on Human Exposure to Environmental Chemicals (NHANES). Department of Health and Human Services: Atlanta, GA.

35. Gerhardsson L, Lundh T (2010) Metal Concentrations in Blood and Hair in Pregnant Females in Southern Sweden. Journal of Environmental Health 72: 37-41.

36. Concha G, Nermell B, Vahter M (1998) Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina. Environmental Health Perspectives 106: 355-359.

37. Sexton K, Hattis D (2007) Assessing cumulative health risks from exposure to environmental mixtures - Three fundamental questions. Environmental Health Perspectives 115: 825-832.

38. Nordberg GF, Jin T, Hong F, Zhang A, Buchet JP, et al. (2005) Biomarkers of cadmium and arsenic interactions. Toxicology and Applied Pharmacology 206: 191-197.

39. NCDHHS (2010) Racial and Ethnic Health Disparities in North Carolina: Report Card 2010. Raleigh, NC.

40. Alexander GR, Kogan M, Bader D, Carlo W, Allen M, et al. (2003) US birth weight/gestational age-specific neonatal mortality: 1995-1997 rates for whites, Hispanics, and blacks. Pediatrics 111: 6.

41. ATSDR (2007) Toxicological Profile for Arsenic. Centers for Disease Control.

42. Ramon R, Murcia M, Aguinagalde X, Amurrio A, Llop S, et al. (2011) Prenatal mercury exposure in a multicenter cohort study in Spain. Environment International 37: 597-604.

43. ATSDR (1999) Toxicological Profile for Mercury Atlanta, GA: Centers for Disease Control. Available online: http://www.epa.gov/region4/waste/npl/nplnc/capefrnc.htm. Accessed June 30, 2011.

44. ATSDR (2008) Toxicological Profile for Cadmium. Atlanta, GA: Centers for Disease Control.

45. Rollin HB, Rudge CVC, Thomassen Y, Mathee A, Odland JO (2009) Levels of toxic and essential metals in maternal and umbilical cord blood from selected areas of South Africa-results of a pilot study. Journal of Environmental Monitoring 11: 618-627.

46. Jarup L, Berglund M, Elinder CG, Nordberg G, Vahter M (1998) Health effects of cadmium exposure - a review of the literature and a risk estimate - Preface. Scandinavian Journal of Work Environment & Health 24 (Supplemental 1): 1-51.

47. BRFSS (2005). Results from the 2004 N.C. Behavioral Risk Factor Surveillance System Survey: BRFSS. Available online: http://www.schs.nc.us/SCHS/brfss/2004/highlights.html. Accessed July 1, 2011.

48. ATSDR (2007) Toxicological Profile for Lead. Atlanta, GA: Centers for Disease Control.

49. Norman EH, Bordley WC, Hertzpicciotto I, Newton DA (1994) Rural-urban blood lead differences in North-Carolina children. Pediatrics 94: 59-64.

FIGURES

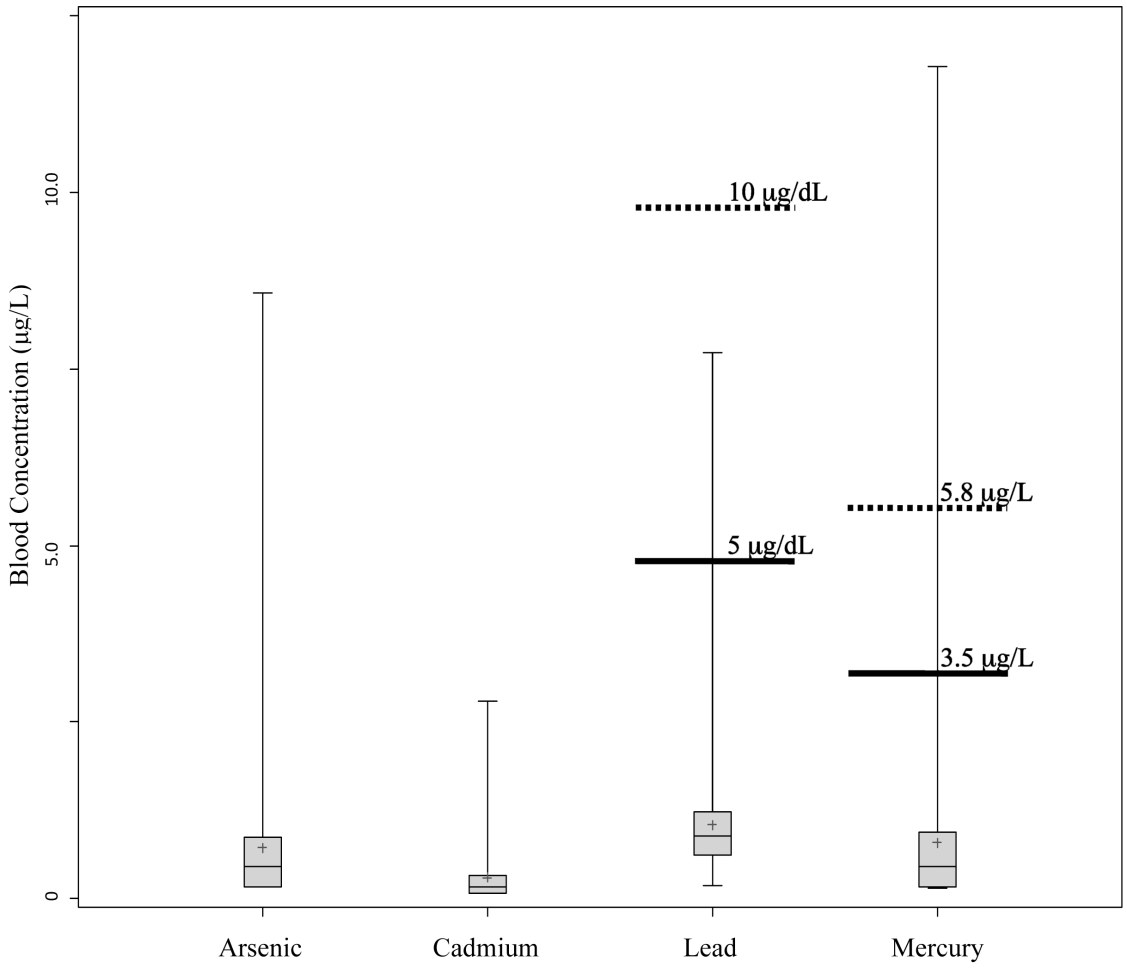

Figure 1. Boxplots of blood levels of As, Cd, Pb, and Hg in 211 pregnant women residing in North Carolina. Solid horizontal lines indicate guideline values believed to be protective of fetal health. Dashed lines represent established guideline levels for the general public. No values have been established for As or Cd. The center horizontal line in each box blot corresponds to the sample median and the central plus sign corresponds to the sample mean. Horizontal lines also represent the 25th and 75th percentiles as well as minimum and maximum measured values. *Lead units are expressed in µg/dL.

TABLES

Table 1. Characteristics of the pregnant women (n=211) in the third trimester of pregnancy.

Maternal age (years), mean±SD (range)	25.0±5.7 (15-43)
Race/ethnicity, n (%)	
Asian	4 (2%)
Non-Hispanic Black	79 (38%)
Hispanic	13 (6%)
Non-Hispanic White	115 (55%)
County of residence, n (%)	
County A	50 (24%)
County B	28 (13%)
County C	50 (24%)
County D	13 (6%)
County E	25 (12%)
County F	45 (22%)

Table 2. Detectable levels and geometric averages of the four toxic metals in women.

Blood Metals	n	% Detected	Geometric Mean[a] (range)
Arsenic	210	65.7	0.445 (<0.23 - 8.58) μg/L
Cadmium	211	57.3	0.181 (<0.11 - 2.79) μg/L
Mercury	210	63.8	0.453 (<0.23 - 11.78) μg/L
Lead	211	100	0.890 (0.19 - 7.72) μg/dL

[a]The geometric mean was calculated with measures <LOD assigned values of LOD/√2 [35].

Table 3. Spearman's rank correlation coefficients for metals in blood of North Carolina pregnant women (n=211).

	As	Cd	Hg	Pb
As	1.	-0.14*	0.45**	0.12
Cd		1	-0.21*	0.07
Hg			1	0.19*
Pb				1

*P<0.05; **P<0.001; Data below detect were assigned imputed values of LOD/√2

Table 4. Linear regression of age-adjusted maternal race on blood metal levels (Beta coefficient and 95% CI). NHW served as the referent group.

	Asian	Hispanic	NHB	NHW (ref)
As	0.51 (0.14-0.88)*	-0.13 (-0.35-0.08)	0.12 (0.02-0.23)*	--
Cd	0.39 (0.01-0.77)*	0.20 (-0.02-0.41)	0.25 (0.15-0.36)**	--
Hg	0.64 (0.23-1.05)*	-0.07 (-0.31-0.16)	0.01 (-0.11-0.12)	--
Pb	-0.01 (-0.24-0.22)	-0.08 (-0.21-0.05)	-0.07 (-0.13- -0.00)*	--

* $p<0.05$; ** $P<0.001$; Metal levels were log-transformed.

Table 5. Concentration of As, Cd, Hg, and Pb in maternal blood collected from women residing in selected North Carolina counties – reported as the geometric mean.

	Maternal blood concentration[a] (μg/L)[*]			
	As	Cd	Hg	Pb
County A (n=50)	0.29	0.27[B]	0.28	0.70
County B (n=28)	0.36	0.14	0.39	0.89
County C (n=50)	0.49[A]	0.23[B]	0.40	1.28[D]
County D (n=13)	0.33	0.09	0.70[C]	1.02
County E (n=25)	0.39	0.11	0.45	0.69
County F (n=45)	0.87[A]	0.17	0.87[C]	0.86
Total (n=211)	0.44	0.18	0.45	0.89
NHANES IV	NA	0.33	0.78	1.22

Abbreviations: NA not available; NHANES: Geometric mean levels (2003-2004) reported for women (n=4,241) participating in NHANES IV [35].

[a] Concentration reported as the geometric mean blood levels collected from women in the third trimester of pregnancy.

[*]Units of Pb are expressed in μg/dL

[A]Arsenic levels in County C were significantly elevated compared individually to County A (Scheffe's $p<0.05$). As levels in County F were significantly elevated compared to each of the other five counties individually (Scheffe's $p<0.05$).

[B] Cadmium levels in County A were significantly elevated compared individually to County B, D, and E (Scheffe's $p<0.05$). Cd levels in County C were significantly elevated compared individually to County D and E (Scheffe's $p<0.05$)

[C] Mercury levels in County F were significantly elevated compared individually to County A, B, and C (Scheffe's $p<0.05$). Hg levels in County D were significantly elevated compared individually to County A (Scheffe's $p<0.05$).

[D] Lead levels in County C were significantly elevated compared individually to County A, E, and F (Scheffe's $p<0.05$).

APPENDIX THREE: CADMIUM-ASSOCIATED PATTERNS OF DNA METHYLATION IN LEUKOCYTES FROM MOTHER-BABY PAIRS

Alison P. Sanders[1], Lisa Smeester[1], Daniel Rojas[2], Tristan DeBussycher[3], Michael C. Wu[4], Fred A. Wright[4], Yi-Hui Zhou[4], Jessica E. Laine[5], Julia E. Rager[1], Geeta K. Swamy[6], Allison Ashley-Koch[7], Marie Lynn Miranda[8], and Rebecca C. Fry[1,2]

Authors' affiliations:

[1] Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA.

[2] Curriculum in Toxicology, School of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA.

[3] Center for Bioinformatics, UNC-Chapel Hill, Chapel Hill, North Carolina, USA.

[4] Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA.

[5] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, USA.

[6] Department of Obstetrics and Gynecology, Duke University, Durham, North Carolina, USA.

[7] Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, USA.

[8] School of Natural Resources and Environment and Department of Pediatrics, University of Michigan, Ann Arbor, MI, USA.

ABSTRACT

Cadmium is prevalent in the environment and understudied as a developmental toxicant. It can cross the placental barrier from mother to fetus and is linked to detrimental effects in newborns. This study examines the relationship between maternal exposure to cadmium and leukocyte DNA methylation patterns in seventeen mother-newborn pairs. A methylated cytosine-guanine (CpG) island recovery assay was used to assess over 4.6 million sites spanning 16,421 CpG islands. Exposure to cadmium was classified for each mother-newborn pair according to maternal blood levels and compared to levels of cotinine. Comparative methylation analysis was performed to identify genes with differential methylation levels. DNA motifs that were overrepresented among the differentially methylated genes were identified. Subsets of genes were identified that showed altered DNA methylation levels in fetal DNA associated with exposure to cadmium (n=61), cotinine (n=366), or both (n=30). In maternal DNA, several differentially methylated genes were associated with cadmium (n=92) and cotinine (n=134) exposure. Relative differences in DNA hyper- or hypo- methylation in selected maternal samples were reproducible using an alternate technology. While the gene sets were largely distinct between mothers and newborns, functional similarities at the biological pathway level were identified including transcriptional regulation and apoptosis. Furthermore, conserved DNA motifs with sequence similarity to specific transcription factor binding sites were identified within the CpG islands of the gene sets. This pilot study provides evidence for distinct patterns of DNA methylation alterations or "footprints" in fetal and maternal DNA associated with exposure to cadmium.

INTRODUCTION

Cadmium (Cd) is a heavy metal that ranks among the top ten chemicals in the Agency for Toxic Substances and Disease Registry priority list of hazardous substances [1]. Cd is a known lung carcinogen and a putative carcinogen in other tissues including the liver, prostate, kidney, bladder, stomach, and pancreas [2-4]. Cd is widespread in the environment and is found in byproducts of industrial processes, contaminated water or soil, certain foods such as shellfish, and tobacco products [5]. Cd in the blood has a half-life of approximately 3-4 months and can range up to 10 years [6].

Environmental and low level health effects of Cd are of growing interest in the environmental health field, particularly among susceptible populations such as pregnant women and children [7]. Of concern, among pregnant women (n=253) studied in the Fourth National Report on Human Exposure to Environmental Chemicals (NHANES IV), 66% had detectable blood Cd levels with an average level of 0.22 μg/L [8]. Given that fetal levels correlate with maternal levels [9, 10], we sought to determine the impact of Cd exposure on DNA methylation patterns in a cohort of mother-baby pairs in Durham County, North Carolina.

Cd is a component of cigarette smoke and concurrent exposure is common. In utero exposure to both Cd and cigarette smoke is associated with lower newborn birth weight [9, 11, 12]. Further, Cd has been suggested to be the component of cigarette smoke that affects fetal skeletal growth [13]. Prenatal tobacco smoke exposure is a known modifier of DNA methylation patterns [14-16] and is therefore an important

consideration in this study. Cotinine, the primary metabolite of nicotine and biomarker of cigarette smoke exposure, is a reliable measure of actual dose received with a half-life of less than one day [17, 18]. Thus, cotinine is a general measure of recent exposure to tobacco products, whereas Cd can represent long term and historic tobacco exposure. In addition, diet is the major source of Cd exposure among non-smokers [6]. Therefore, in this study we measured maternal cotinine and Cd levels to compare and contrast the methylation levels associated with either contaminant and to differentiate Cd-specific patterns in DNA methylation from those associated with cotinine.

Prenatal Cd exposure has been inversely associated with fetal growth parameters such as newborn length, weight, height, and head circumference [9, 12, 19-24], as well as adverse cognitive developmental effects later in life [23, 25]. The modes of action of health effects related to prenatal Cd exposure are not well established. An epigenetic mechanism has been hypothesized [26], but is understudied in human populations. DNA methylation is an epigenetic mechanism that may play a key role in mediating biological processes [27], possibly contributing to subsequent health effects resulting from environmental toxicant exposure. Indeed there is evidence from animal and cell culture studies that Cd alters DNA methyltransferase activity and subsequently DNA methylation [2, 26, 28-33]. Thus, in an effort to understand potential impacts of prenatal Cd exposure, we assess DNA methylation in blood leukocytes and distinguish between Cd- and cotinine-associated changes. Here we examine genome-wide site-specific DNA methylation changes associated with in utero and maternal Cd exposure.

RESULTS

This study consisted of 34 subjects, 17 mother-newborn pairs, selected as a nested cohort from the CEHI Healthy Pregnancy, Healthy Baby study in Durham, North Carolina. Maternal and infant characteristics are presented in Table 1. The average maternal age was 28 years. Most of the women had more than one child (n=13; 76.5%), including the infants described in this study. There were similar proportions of male (n=9; 52.9%) and female children (n=8; 47.1%). All newborns, with one exception, had a birth weight greater than 2500 g with a range of 2495 to 3740 g. Levels of maternal Cd ranged from below detect to 1.05 μg/L with an average maternal blood concentration of 0.44 μg/L. Ten mothers in the study had Cd levels above the NHANES median level in pregnant women and were classified as the "higher cadmium-exposed" group (See Methods and Materials). Maternal cotinine levels ranged from below detect to 166.96 μg/L with an average of 14.5 μg/L. Eleven mothers had a detectable level of cotinine and were classified as "cotinine-exposed". Of the mothers with detectable blood levels of cotinine, only two women had levels above 10 μg/L (a level associated with active or passive smoking activity [17]).

Linear regression analysis revealed that maternal Cd or cotinine levels did not vary significantly with respect to maternal age, race, parity, child's sex or birth weight. Exclusion of the low birth weight infant from the dataset did not significantly affect regression analyses (data not shown). Generally, younger mothers had higher levels of Cd, but this finding was not statistically significant ($p>0.05$). A linear relationship between maternal serum cotinine and Cd levels was not observed ($p=0.6$); however, Spearman rank correlation revealed a positive association ($r=0.62$, $p=0.008$).

Cd-associated gene-specific DNA methylation

Methylated DNA collected from each newborn and mother was isolated using Methyl-CpG-binding domain protein 2 (MBD2) and hybridized onto Affymetrix Human Promoter 1.0R arrays. This resulted in the assessment of 4.6 million sites that were bioinformatically analyzed in the context of human CpG islands enriched in the promoter regions of 16,421 genes as in our recent publication [34] (see Methods and Materials).

Prior to analysis for the exposures of interest, differences in average methylation levels associated with maternal age, race, and infant sex were identified. Analysis of DNA methylation patterns from maternal DNA revealed there was a significant difference in gene specific methylation levels associated with mothers' age (n= 596 genes) and race (n=83 genes). In fetal DNA, there was significant difference in gene specific DNA methylation levels associated with maternal age (n=39 genes), race (n=949 genes), and infant sex (n=176 genes) (Supplemental Material, Table S1).

In relation to newborn environmental exposures of Cd and cotinine, two gene sets were identified with significantly different average methylation abundances in fetal DNA that were Cd-associated (n=61 genes) or cotinine-associated (n=366 genes) (Figure 1A-B, Supplemental Material, Table S2). For the fetal DNA assessment, one gene was hypomethylated with increasing Cd, and five were hypomethylated with respect to increasing cotinine (Figure 1C). A total of 30 hypermethylated genes overlapped between the Cd and cotinine gene sets (Figure 1C, Supplemental Material, Table S2).

In maternal DNA, distinct sets of Cd-associated (n=92) or cotinine-associated genes (n=134) were identified (Figure 1A-B, Supplemental Material, Table S3). For the

maternal DNA assessment, eleven of the 92 Cd-associated differentially methylated genes were hypomethylated, whereas four of 134 cotinine-associated genes were hypomethylated. There were no overlapping genes between differentially methylated Cd- or cotinine-associated genes in maternal DNA (Figure 1D). None of the Cd-associated genes were the same between mothers and newborns. Very few (n=12) of the cotinine-associated genes were differentially methylated in both newborns and mothers. The majority of differentially methylated genes had increasing promoter methylation with increasing Cd or cotinine level for both fetal and maternal DNA.

An independent technology was used for validation analyses to assess methylation levels, namely, the Illumina HumanMethylation450 BeadChip (450K). Maternal DNA from two mothers was selected based on stratified Cd-exposure status. Eighty-three of the 92 Cd-associated differentially methylated islands identified in maternal DNA with the MIRA assay were also contained on the Illumina 450K platform. A total of 68 out of the 83 CpG islands (82%) showed concordant hyper- or hypo- methylation between the two platforms (Supplemental Material, Table S3). The data also showed a statistically significant Spearman rank correlation ($r=0.35$; $p=0.0009$) comparing average methylation levels at CpG islands contained on both platforms.

In addition to identifying genes that are differentially methylated and associated with Cd or cotinine levels, a comparison was also performed between the DNA methylomes of mothers and newborns. Regardless of environmental exposure, this comparison of the 34 individuals' methylomes showed that there were 12,820 genes significantly differentially methylated between fetal and maternal DNA that were not due to differences in maternal age, race, or infant sex (Supplemental Material, Figure S1). All

of the 12,820 genes showed lower average methylation abundance in maternal DNA when compared to fetal DNA.

Biological functions are enriched among differentially methylated genes

The Cd- and cotinine-associated gene sets in mother-baby pairs were analyzed to determine whether they encode proteins that play a similar role in the cell. Each of the gene sets was analyzed for enriched biological functions using two independent methods (see Materials and Methods). The most significant enriched biological functional categories were gene expression, cell cycle, cell death, and nervous system development (Table 2). Interestingly, while there was no overlap between the individual gene sets, genes that encode proteins that play a role in regulation of transcription were enriched among Cd-associated genes in both fetal and maternal DNA ($p<0.001$). Apoptosis was an enriched biological process in mothers and newborns for genes associated with either Cd or cotinine levels ($p<0.05$).

Common DNA motifs identified in gene sets

We hypothesized that the identified genes may contain common underlying sequences or motifs. To examine this, the individual promoter regions of CpG islands for each differentially methylated gene were analyzed for statistically enriched common sequence patterns (e.g. motifs). The motifs were then compared with known transcription factor binding sites. Among the Cd- and cotinine-associated gene sets, significantly enriched motifs of conserved DNA sequences were identified (Figure 2). The conserved motifs showed sequence similarity to binding sites for several transcription factors

including: transcription factor 7-like 1 (TCF7L1 also known as TCF3), metal-responsive transcription factor-1 (MTF-1), transcription factor AP-2-eplison (TCFAP2E) and serum response factor (SRF) (Table 3). Notably, for the Cd-associated gene sets identified in maternal and fetal DNA, the enriched motifs had significant sequence similarity to the binding sites of MTF-1 and TCF7L1.

DISCUSSION

There is growing evidence that the prenatal environment may influence the burden of disease in adult life, and that this relationship is strongly associated with epigenetic modifications induced during the prenatal period [27, 35]. Given the relationship of Cd with birth outcomes in humans [9, 12] and its known role as a mediator of DNA methyltransferase activity [26], we set out to identify patterns of differential methylation in newborn and maternal leukocyte DNA associated with exposure to Cd. The data were compared to cotinine-associated patterns of differential methylation.

The data from this study support that there are marked differences between the leukocyte methylomes of mothers and their newborns, and that the methylation patterns are largely distinct based on the environmental exposure to Cd or cotinine for both mothers and their babies. Notably, the genes that are identified as differentially methylated for both of the environmental exposures contain common DNA motifs/sequence and the CpG islands in their promoter regions encode proteins with conserved biological processes in the cell.

The suggested impacts of Cd exposure include enzyme inhibition, generation of reactive oxygen species, and perturbation of apoptosis or cell cycle [36]; however, the ability to create genomic instability without genotoxic action has strongly suggested an epigenetic mechanism of Cd toxicity [26]. Cd-associated genome-wide DNA methylation has not been previously assessed in human samples. Here sets of genes with altered methylation levels that were associated with in utero exposure to Cd- (n=61) and

cotinine- (n=366) were identified. Only thirty of the genes overlapped between the Cd- or cotinine associated gene sets in newborns. Among mothers, the Cd- (n=92) and cotinine-associated (n=134) gene sets were identified. There was no overlap between the Cd- or cotinine associated gene sets in the maternal DNA. When the Cd- and cotinine-specific gene sets were contrasted between mothers and newborns, we found that these sets were largely distinct; none of the Cd-associated genes, and very few (n=12) of the cotinine-associated genes, were similar between mothers and newborns. Overall, the majority of the differentially methylated genes showed increased average methylation levels with increasing exposure. Similar to the findings here, previous *in vitro* and *ex vivo* experiments showed that Cd exposure led to hypermethylation of DNA after prolonged chronic exposures, however hypomethylation was present after acute Cd exposure [26, 29, 30]. A study of adult non-smoking women in Argentina found no significant association between blood Cd and global or site-specific methylation levels within promoter regions of two candidate genes; however, global hypomethylation was associated with urinary Cd levels [37].

The Illumina 450K platform was used as an alternate technology to substantiate methylation levels across the Cd-associated gene set (n=92) for which 83 CpG islands were contained on both arrays. Validation was performed on two maternal DNA samples stratified by Cd-exposure status. A significant Spearman rank trend was observed between the MIRA and 450K platform methylation levels (R=0.35; p=0.0009). A majority of the CpG islands (82%) had concordant relative hyper- or hypo- methylation on the 450K platform in the Cd-exposed versus unexposed mothers.

A larger number of genes with differential DNA methylation were associated with cotinine exposure in maternal DNA (n=134 genes) and fetal DNA (n=366 genes), compared to the number of Cd-associated genes. Three previous studies of prenatal cigarette smoke exposure examined gene-specific DNA methylation [14-16]. One study reported global hypomethylation with gene-specific hypermethylation of eight genes investigated [14] while the other found largely gene-specific hypomethylation among 38 total genes [15]. The third study reported 26 significant CpGs mapped to 10 genes with both hyper- and hypomethylation associated with plasma cotinine levels [16]. None of the genes identified were common between the three studies and none are similarly reported in the present study. There are several factors that could potentially influence differences in observed methylation patterns that include exposure type, exposure duration, tissue type (i.e. buccal vs. placental vs. cord blood), subject's age at sampling, type of assay used or other unaccounted for co-exposures.

These data also highlight that regardless of exposure, there were significantly different methylation profiles between mothers and newborns that were not due to differences in maternal age, race, or infant sex. This is supported by a recent study that compared to newborn DNA, less methylation is observed among CpG island promoters in older individuals [7]. The increased methylation levels in newborns has relevance for developmental biology, and further evaluation of these basic processes may increase our understanding of how early life exposure resulting in epigenetic shifts can have long-term health effects.

To classify the differentially methylated genes by known ontologies, two independent enrichment analyses were used. Transcription regulatory processes emerged

as significantly enriched among the Cd gene lists in both fetal and maternal DNA. Cell death, specifically apoptosis, was a significantly enriched function for each gene list. Cd has been shown in vitro to perturb pathways involved in inflammatory response, cell survival, apoptosis, tumorigenesis, and oxidative stress [26, 38-41]. In animal models, in utero exposure to Cd is associated with a wide range of cell cycle and proliferative genomic responses [42]. These data support an epigenetic mechanism, namely DNA methylation, by which genes involved in transcriptional regulation and apoptosis may be altered by Cd exposure. RNA is not available for the specific study subjects analyzed here for gene expression analysis at a functional level. Therefore, these data are interpreted with caution, as it is not anticipated that all of the changes in DNA methylation will impact gene expression.

Of interest, many of the genes contained in the Cd and cotinine gene sets showed common patterns in DNA sequence in the CpG islands within promoters. Among these conserved motif regions, binding sites for a common set of transcription factors were identified. The specific binding of transcription factors to target sites is a recognized mechanism that protects CpG islands from methylation [43, 44]. Recent studies have demonstrated that transcription factor binding results in local regions of low methylation and in contrast, absence of DNA-binding factors triggers the remethylation of local promoter regions [45, 46]. The identified conserved motifs showed sequence similarity to several binding sites including TCF7L1, MTF-1, TCFAP2E and SRF. Notably, MTF-1 is known to respond to changes in cellular concentrations of multiple metals and coordinate expression of genes protective against metal toxicity [47, 48]. The motifs with sequence similarity to TCFAP2E and SRF binding sites were found only in gene sets identified in

newborns, but not mothers. All of the identified transcription factors are known to regulate developmental processes within cells and are worthy of future investigation [49-52]. Taken together, these results suggest that patterns of DNA methylation that are associated with Cd may represent "footprints" indicating transcription factor presence or absence that occur during periods of DNA methylation. Further research is needed to validate this hypothesis.

The cord blood sampling done here is representative of newborn leukocyte DNA rather than potential target organ systems such as the kidney, liver, or bone. There are obvious ethical and technical reasons to use leukocyte DNA as a proxy for target tissue analysis. Peripheral blood leukocyte patterns of DNA methylation have not been shown to differ greatly among cell types [16, 53]. Moreover, the use of circulating white blood cells as proxies for disease has been shown [54, 55]. In addition, there is increasing evidence that there are contaminant-specific changes to DNA methylation associated with various environmental contaminant exposures [34, 56, 57]. Future studies should aim to investigate tissue-specific differences as they relate to contaminants of interest. In this study, RNA was not available to confirm gene expression levels in the cord blood. In future studies, genomic examination of the epigenetic gene targets identified here will aid understanding of how methylation may alter gene expression and subsequent disease manifestation.

This study provides evidence of Cd- and cotinine-associated DNA methylation changes in the leukocyte DNA of newborns and their mothers. This study identifies genome-wide site-specific DNA methylation changes associated with in utero Cd exposure, and distinguishes these from methylation changes attributable to cotinine

exposure, a general proxy measure for exposure to tobacco products. These distinct patterns of environmentally-associated DNA methylation alterations or footprints in fetal and maternal DNA may have functional consequences in the cell and warrant further research. Maternal metal exposure continues to be an important area of public health concern for both maternal and child health. Increased education about the potential risks of environmental contaminants including Cd and cigarette smoke will be key to reducing and preventing harm to infants.

MATERIALS AND METHODS

Study participants

The Children's Environmental Health Initiative (CEHI) conducted a prospective cohort study of pregnant women living in Durham County, North Carolina from 2005-2011. This study is a key component of the Southern Center on Environmentally-Driven Disparities in Birth Outcomes (SCEDDBO), an interdisciplinary center aimed at understanding how environmental, social, and host factors jointly contribute to health disparities (http://cehi.snre.umich.edu/projects/sceddbo/). The CEHI Healthy Pregnancy, Healthy Baby study was reviewed and approved by the Institutional Review Boards at Duke University (Pro00007633) and the University of North Carolina (#09-0866). All women participating in this study consented for maternal venous and newborn cord blood collection for chemical and genetic analysis.

Women receiving prenatal care at either the Duke Obstetrics Clinic or the Durham County Health Department Prenatal Clinic were eligible to participate if they planned to deliver at Duke University Medical Center, were at least 18 years of age, were English-literate, lived in Durham County, and did not have a multi-fetal gestation or any known fetal genetic or congenital anomalies. Additional methods on subject recruitment, enrollment, and data collection have been described previously [58]. Women were enrolled between 18 and 28 weeks of pregnancy, and demographic data were collected including maternal age, race, and parity, as well as child's sex and birthweight. At delivery, biological samples including maternal venous blood and newborn cord blood

were collected. From the CEHI study, we selected a nested subcohort of 17 mother-infant pairs stratified by maternal Cd level (see Supplemental Material).

Methylated CpG island recovery assay (MIRA)

Venous maternal blood and newborn cord blood samples were obtained at delivery. DNA was extracted using Qiagen's PAXgene Blood DNA kit (Qiagen, Valencia, CA) according to manufacturer's protocol. DNA was re-suspended in nuclease-free water and stored at -80C prior to DNA methylation assessment. CpG methylated DNA was collected using the MethylCollector Ultra Kit (Active Motif Carlsbad, CA) and enriched DNA was amplified using the WGA2 kit (Sigma, St Louis, MO) according to manufacturer instructions. Amplified DNA was then hybridized to the Affymetrix Human Promoter 1.0R arrays (Affymetrix, Santa Clara, CA) which quantitates methylation at over 4.6 million probe sites.

Statistical analyses

Linear regression analyses were performed using the statistical package SAS 9.3 (SAS Institute Inc, Cary, NC) to examine relationships between maternal blood Cd levels, serum cotinine levels and demographic characteristics for women and children. Maternal demographics included age, race, and parity, and children's demographics included sex and birthweight. The relationship between Cd and cotinine was assessed

with linear regression as well as Spearman rank correlation. Cd and cotinine levels below detect were treated as zero.

The study subjects' data for each of the 4.6 million probe sets resulting from the MIRA assay were normalized using robust multi-chip average then summarized at a gene-specific CpG island level based on the reference Human Genome 18 (HG18) [59]. The resulting average methylation abundances for 16,421 CpG islands were compared using ANOVA (Partek Genomic Suite 6.4, St Louis Missouri). Differential methylation was assessed for each island and statistically defined as: (i) average island promoter methylation with a minimum absolute change of 30%; and (ii) a p-value < 0.05. Additionally, a false discovery rate (FDR) corrected q-value estimate was calculated and is reported. After identifying covariate-associated DNA methylation within CpG islands (see Supplemental Material), differential methylation was also examined according to sample type (maternal versus newborn), higher versus lower Cd-exposed, and cotinine-exposed versus cotinine-unexposed (see Supplemental Material). The same statistical requirements were applied to all analyses.

Gene ontology/pathway enrichment analysis

Gene ontology/pathway enrichment analysis was performed using two independent methodologies. The differentially methylated genes were analyzed in the context of interacting networks using Ingenuity Pathway Analysis Software (Ingenuity Systems, Inc., Redwood City, CA) and functional clustering using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (available

online: http://david.abcc.ncifcrf.gov/tools.jsp). Methods are provided in Supplemental Material. Pathway enrichment analysis was performed for the exposure-specific gene lists for both mothers and newborns to identify biological pathways significantly associated with the differentially methylated gene sets.

Enriched DNA motif identification

CpG island sequences were retrieved from the UCSC genome browser website [60] for each exposure-associated methylated gene set. These sequences are representative of CpG islands within promoter regions. Position specific letter probability matrices, also known as motifs, were identified using Multiple Em for Motif Elicitation (MEME) version 4.8.1 [61]. A first order Hidden Markov Model and negative position specific priors were calculated from a background set of 200 randomly selected CpG island sequences that were not differentially methylated. Additional parameter specifications are reported in the Supplemental Material. The motif with the highest significance (p-value) for each group of differentially methylated genes was compared to known transcription factor binding sites using TOMTOM [61].

Illumina Methylation 450K BeadChip Validation

Additional DNA methylation analysis was performed on two maternal DNA samples stratified by maternal blood Cd levels (0.8 μg/L vs. below detect) and matched

on maternal race, age and insurance status. Bisulfite conversion was performed using the Zymo EZ DNA Methylation kit (Zymo Research, Irvine, CA) according to manufacturers instructions. Methylation was assessed at 485,577 CpGs in maternal DNA using the Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, CA). BeadChip processing was performed at Expression Analysis Inc. (Durham, NC; www.expressionanalysis.com) and processed with Illumina's GenomeStudio Methylation module Version 1.8 (Illumina Inc., San Diego, CA). The proportion of methylation (β) for each CpG was calculated as the ratio of methylated signal intensity divided by the sum of both methylated and unmethylated signals.

For statistical comparison, the 92 differentially methylated Cd-associated genes in maternal DNA were extracted from the Illumina 450K results. Because the MIRA technology enriches for promoter regions, the 450K data were filtered for probe sets positioned in islands. For quality control, probes with a detection p-value greater than 0.0001 were removed prior to analysis. This resulted in 83 genes for comparison between the two platforms. Average β values for island-associated CpG sites were calculated and a β value ratio equal to the average β for the Cd-exposed mother divided by the unexposed mother was calculated for each gene. To compare relative differences in methylation across the platforms, the β value ratio was converted to a log2 fold change. A Spearman rank correlation and corresponding p-value were calculated to compare the 450K β ratio to the MIRA fold changes across the two platforms.

REFERENCES

1. ATSDR. Priority List of Hazardous Substances. Online at: http://wwwatsdrcdcgov/spl/ 2011.

2. Arita A, Costa M. Epigenetics in metal carcinogenesis: nickel, arsenic, chromium and cadmium. Metallomics 2009; 1:222-8.

3. Waalkes MP. Cadmium carcinogenesis in review. Journal of Inorganic Biochemistry 2000; 79:241-4.

4. IARC. Volume 58. Cadmium. 1993:119-238.

5. ATSDR. Toxicological Profile for Cadmium. Atlanta, GA: Centers for Diseaes Control, 2008.

6. Jarup L, Akesson A. Current status of cadmium as an environmental health problem. Toxicol Appl Pharmacol 2009; 238:201-8.

7. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, et al. Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci U S A 2012; 109:10522-7.

8. Woodruff TJ, Zota AR, Schwartz JM. Environmental Chemicals in Pregnant Women in the United States: NHANES 2003-2004. Environmental health perspectives 2011; 119:878-85.

9. Salpietro CD, Gangemi S, Minciuollo PL, Briuglia S, Merlino MV, Stelitano A, et al. Cadmium concentration in maternal and cord blood and infant birth weight: a study on healthy non-smoking women. Journal of Perinatal Medicine 2002; 30:395-9.

10. Galicia-Garcia V, Rojas-Lopez M, Rojas R, Olaiz G, Rios C. Cadmium levels in maternal, cord and newborn blood in Mexico City. Toxicology Letters 1997; 91:57-61.

11. Yu XD, Yan CH, Shen XM, Tian Y, Cao LL, Yu XG, et al. Prenatal exposure to multiple toxic heavy metals and neonatal neurobehavioral development in Shanghai, China. Neurotoxicol Teratol 2011; 33:437-43.

12. Frery N, Nessmann C, Girard F, Lafond J, Moreau T, Blot P, et al. ENVIRONMENTAL EXPOSURE TO CADMIUM AND HUMAN BIRTH-WEIGHT. Toxicology 1993; 79:109-18.

13. Godfrey K, Walker-Bone K, Robinson S, Taylor P, Shore S, Wheeler T, et al. Neonatal bone mass: influence of parental birthweight, maternal smoking, body composition, and activity during pregnancy. J Bone Miner Res 2001; 16:1694-703.

14. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal Tobacco Smoke Exposure Affects Global and Gene-specific DNA Methylation. American Journal of Respiratory and Critical Care Medicine 2009; 180:462-7.

15. Suter M, Ma J, Harris AS, Patterson L, Brown KA, Shope C, et al. Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. Epigenetics 2011; 6.

16. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking During Pregnancy. Environmental health perspectives 2012.

17. Klebanoff MA, Levine RJ, Clemens JD, DerSimonian R, Wilkins DG. Serum cotinine concentration and self-reported smoking during pregnancy. Am J Epidemiol 1998; 148:259-62.

18. Jarvis MJ, Russell MA, Benowitz NL, Feyerabend C. Elimination of cotinine from body fluids: implications for noninvasive measurement of tobacco smoke exposure. American journal of public health 1988; 78:696-8.

19. Lin CM, Doyle P, Wang DL, Hwang YH, Chen PC. Does prenatal cadmium exposure affect fetal and child growth? Occupational and Environmental Medicine 2011; 68:641-6.

20. Nishijo M, Tawara K, Honda R, Nakagawa H, Tanebe K, Saito S. Relationship between newborn size and mother's blood cadmium levels, Toyama, Japan. Archives of Environmental Health 2004; 59:22-5.

21. Ronco AM, Arguello G, Munoz L, Gras N, Llanos M. Metals content in placentas from moderate cigarette consumers: correlation with newborn birth weight. Biometals 2005; 18:233-41.

22. Llanos MN, Ronco AM. Fetal growth restriction is related to placental levels of cadmium, lead and arsenic but not with antioxidant activities. Reproductive Toxicology 2009; 27:88-92.

23. Tian LL, Zhao YC, Wang XC, Gu JL, Sun ZJ, Zhang YL, et al. Effects of Gestational Cadmium Exposure on Pregnancy Outcome and Development in the Offspring at Age 4.5 Years. Biological Trace Element Research 2009; 132:51-9.

24. Shirai S, Suzuki Y, Yoshinaga J, Mizumoto Y. Maternal exposure to low-level heavy metals during pregnancy and birth size. Journal of Environmental Science and

Health Part a-Toxic/Hazardous Substances & Environmental Engineering 2010; 45:1468-74.

25. Bonithonkopp C, Huel G, Moreau T, Wendling R. PRENATAL EXPOSURE TO LEAD AND CADMIUM AND PSYCHOMOTOR DEVELOPMENT OF THE CHILD AT 6 YEARS. Neurobehavioral Toxicology and Teratology 1986; 8:307-10.

26. Takiguchi M, Achanzar WE, Qu W, Li GY, Waalkes MP. Effects of cadmium on DNA-(Cytosine-5) methyltransferase activity and DNA methylation status during cadmium-induced cellular transformation. Experimental Cell Research 2003; 286:355-65.

27. Wilson AG. Epigenetic Regulation of Gene Expression in the Inflammatory Response and Relevance to Common Diseases. Journal of Periodontology 2008; 79:1514-9.

28. Martinez-Zamudio R, Ha HC. Environmental epigenetics in metal exposure. Epigenetics 2011; 6:820-7.

29. Benbrahim-Tallaa L, Waterlandz RA, Dill AL, Webber MM, Waalkes MP. Tumor suppressor gene inactivation during cadmium-induced malignant transformation of human prostate cells correlates with overexprression of de Novo DNA methyltransferase. Environmental health perspectives 2007; 115:1454-9.

30. Huang DJ, Zhang YM, Qi YM, Chen C, Ji WH. Global DNA hypomethylation, rather than reactive oxygen species (ROS), a potential facilitator of cadmium-stimulated K562 cell proliferation. Toxicology Letters 2008; 179:43-7.

31. Benbrahim-Tallaa L, Tokar EJ, Diwan BA, Dill AL, Coppin JF, Waalkes MP. Cadmium Malignantly Transforms Normal Human Breast Epithelial Cells into a Basal-like Phenotype. Environmental health perspectives 2009; 117:1847-52.

32. Doi T, Puri P, McCann A, Bannigan J, Thompson J. Epigenetic Effect of Cadmium on Global De Novo DNA Hypomethylation in the Cadmium-Induced Ventral Body Wall Defect (VBWD) in the Chick Model. Toxicological Sciences 2011; 120:475-80.

33. Jiang GF, Xu L, Song SZ, Zhu CC, Wu Q, Zhang L. Effects of long-term low-dose cadmium exposure on genomic DNA methylation in human embryo lung fibroblast cells. Toxicology 2008; 244:49-55.

34. Smeester L, Rager JE, Bailey KA, Guan XJ, Smith N, Garcia-Vargas G, et al. Epigenetic Changes in Individuals with Arsenicosis. Chemical Research in Toxicology 2011; 24:165-7.

35. Bollati V, Baccarelli A. Environmental epigenetics. Heredity 2010; 105:105-12.

36. Waisberg M, Joseph P, Hale B, Beyersmann D. Molecular and cellular mechanisms of cadmium carcinogenesis. Toxicology 2003; 192:95-117.

37. Hossain MB, Vahter M, Concha G, Broberg K. Low-level environmental cadmium exposure is associated with DNA hypomethylation in Argentinean women. Environmental health perspectives 2012; 120:879-84.

38. Benton MA, Rager JE, Smeester L, Fry RC. Comparative genomic analyses identify common molecular pathways modulated upon exposure to low doses of arsenic and cadmium. Bmc Genomics 2011; 12:10.

39. Andrews GK. Regulation of metallothionein gene expression by oxidative stress and metal ions. Biochemical Pharmacology 2000; 59:95-104.

40. Shin HJ, Park KK, Lee BH, Moon CK, Lee MO. Identification of genes that are induced after cadmium exposure by suppression subtractive hybridization. Toxicology 2003; 191:121-31.

41. Robinson JF, Yu XZ, Moreira EG, Hong SW, Faustman EM. Arsenic- and cadmium-induced toxicogenomic response in mouse embryos undergoing neurulation. Toxicology and Applied Pharmacology 2011; 250:117-29.

42. Robinson JF, Yu X, Hong S, Griffith WC, Beyer R, Kim E, et al. Cadmium-induced differential toxicogenomic response in resistant and sensitive mouse strains undergoing neurulation. Toxicological sciences : an official journal of the Society of Toxicology 2009; 107:206-19.

43. Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, et al. Sp1 elements protect a CpG island from de novo methylation. Nature 1994; 371:435-8.

44. Macleod D, Charlton J, Mullins J, Bird AP. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev 1994; 8:2282-92.

45. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D. Identification of genetic elements that autonomously determine DNA methylation states. Nat Genet 2011; 43:1091-7.

46. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 2011; 480:490-5.

47. Selvaraj A, Balamurugan K, Yepiskoposyan H, Zhou H, Egli D, Georgiev O, et al. Metal-responsive transcription factor (MTF-1) handles both extremes, copper load and copper starvation, by activating different genes. Genes Dev 2005; 19:891-6.

48. Andrews GK. Cellular zinc sensors: MTF-1 regulation of gene expression. Biometals 2001; 14:223-37.

49. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev 2008; 22:746-55.

50. Merrill BJ, Gat U, DasGupta R, Fuchs E. Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. Genes Dev 2001; 15:1688-705.

51. Winger Q, Huang J, Auman HJ, Lewandoski M, Williams T. Analysis of transcription factor AP-2 expression and function during mouse preimplantation development. Biol Reprod 2006; 75:324-33.

52. Nelson TJ, Balza R, Jr., Xiao Q, Misra RP. SRF-dependent gene expression in isolated cardiomyocytes: regulation of genes involved in cardiac hypertrophy. J Mol Cell Cardiol 2005; 39:479-89.

53. Deaton AM, Webb S, Kerr AR, Illingworth RS, Guy J, Andrews R, et al. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. Genome Res 2011; 21:1074-86.

54. Anglim PP, Alonzo TA, Laird-Offringa IA. DNA methylation-based biomarkers for early detection of non-small cell lung cancer: an update. Molecular Cancer 2008; 7.

55. Sinnaeve PR, Donahue MP, Grass P, Seo D, Vonderscher J, Chibout SD, et al. Gene Expression Patterns in Peripheral Blood Correlate with the Extent of Coronary Artery Disease. Plos One 2009; 4.

56. Pilsner JR, Hu H, Ettinger A, Sanchez BN, Wright RO, Cantonwine D, et al. Influence of prenatal lead exposure on genomic methylation of cord blood DNA. Environmental health perspectives 2009; 117:1466-71.

57. Perera F, Tang WY, Herbstman J, Tang DL, Levin L, Miller R, et al. Relation of DNA Methylation of 5 '-CpG Island of ACSL3 to Transplacental Exposure to Airborne Polycyclic Aromatic Hydrocarbons and Childhood Asthma. Plos One 2009; 4.

58. Miranda ML, Edwards S, Maxson PJ. Mercury Levels in an Urban Pregnant Population in Durham County, North Carolina. International Journal of Environmental Research and Public Health 2011; 8:698-712.

59. Zhou X, Maricque B, Xie M, Li D, Sundaram V, Martin EA, et al. The Human Epigenome Browser at Washington University. Nat Methods 2011; 8:989-90.

60. Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, et al. The UCSC Table Browser data retrieval tool. Nucleic Acids Res 2004; 32:D493-6.

61. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 2009; 37:W202-8.

FIGURES

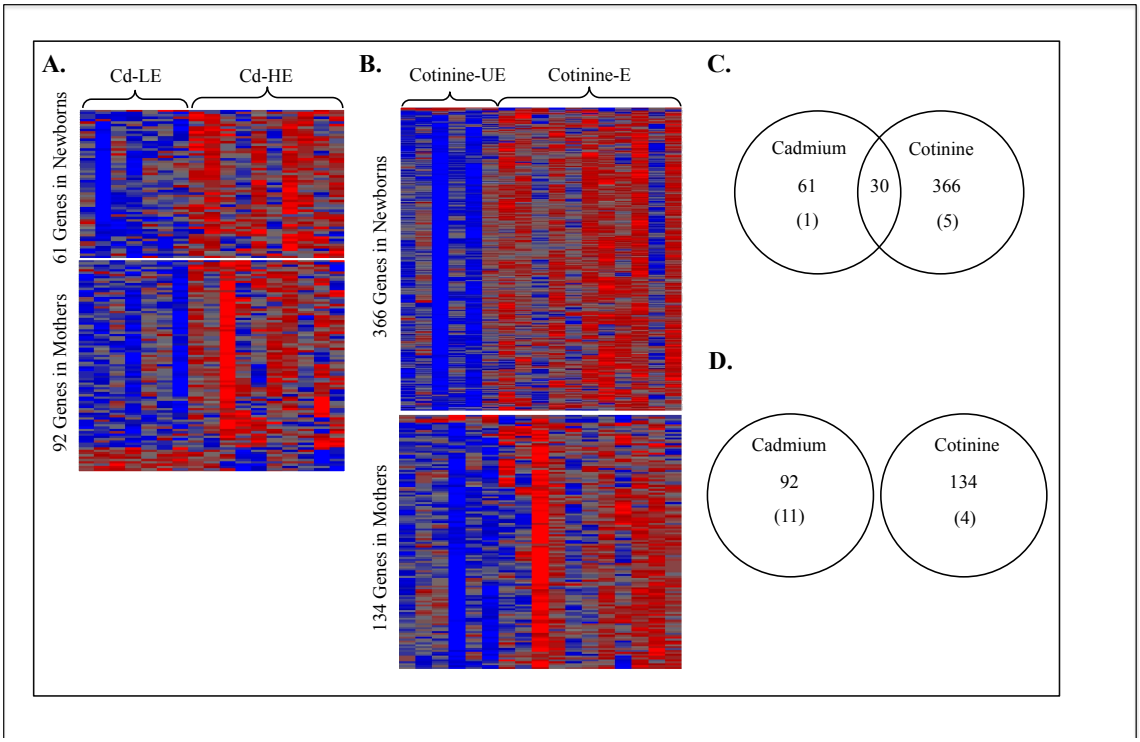

Figure 1. Heat map of differentially methylated genes in fetal and maternal DNA associated with Cd (A), or cotinine (B). Venn diagram representing the total number of Cd-or cotinine-associated differentially methylated genes among newborns (C) or mothers (D). Heat maps represent average DNA methylation levels of exposure-associated gene sets ($p<0.05$). Data are z-score normalized for each gene. Individuals are ordered form left to right based on increasing level of exposure. Red color indicates a relative increase in average DNA methylation and blue color represents a relative decrease in average DNA methylation. In Venn diagrams, number in parentheses indicates the number of hypomethylated genes contained in each set. Abbreviations: LE – lower exposed, HE – higher exposed, UE – unexposed, E – exposed.

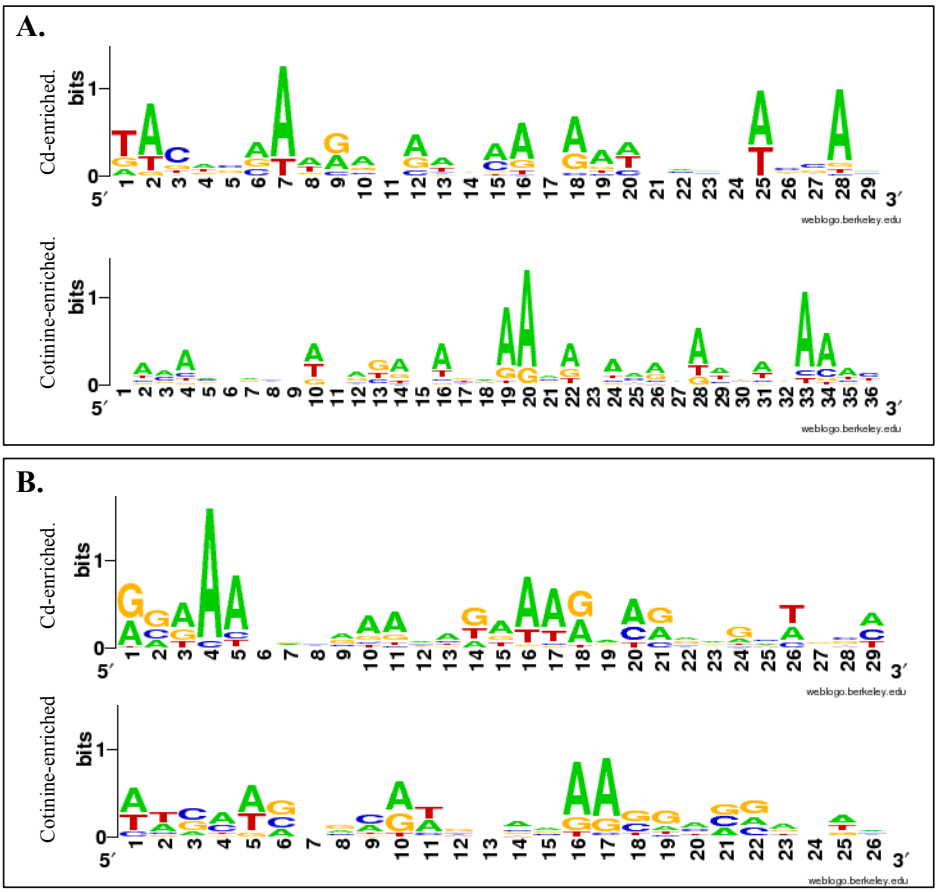

Figure 2. Significantly enriched DNA motifs identified within the Cd and cotinine-associated gene lists in fetal DNA (A) and maternal DNA (B) as identified by MEME. The motifs shown have standard orientation 5' to 3'. Estimates of sequence conservation at each nucleotide position within the motif are shown where a higher bit score is represented by increased font size.

TABLES

Table 1. Characteristics of mother-newborn pairs (n=34 subjects) from Durham, North Carolina.

	Mean ± SD (range) / N (%)
Maternal cadmium (μg/L)	0.44 ± 0.31 (0 - 1.05)
Maternal cotinine (μg/L)	14.5 ± 42.23 (0 - 166.96)
Maternal age (years)	28 ± 7 (19-42)
Maternal race*	
NHB	/ 12 (70.6)
NHW	/ 4 (23.5)
Other	/ 1 (5.8)
Parity	
First	/ 4 (23.5)
Second	/ 5 (29.4)
Third or higher	/ 8 (47.1)
Child's Sex	
M	/ 9 (52.9)
F	/ 8 (47.1)
Birth weight (g)	3210 ± 377 (2495-3740)

*This study intentionally over-sampled NHB mothers.

Table 2. Enriched biological functions within the differentially methylated gene sets ($p<0.05$).

		Fetal DNA		Maternal DNA	
Category	**Function Annotation**	Cd [a]	Cotinine [a]	Cd [a]	Cotinine [a]
Gene expression	Regulation of transcription	0.001* [0.01]	--	0.001* [0.001]	--
Tissue morphology	Adipose tissue quantity	0.004	--	--	--
Cancer	Hyperplasia	0.004	--	--	--
Lipid metabolism[b]	Lipid accumulation	0.005	--	--	--
Cell death	Apoptosis	0.02 [0.05]	0.001*	1.2×10^{-4}* [0.04]	0.05
Cell cycle	Delay in G1, interphase	--	3.0×10^{-5}* [0.003]	--	--
Nervous system[c]	Neuronal quantity	--	0.002 [0.03]	5.2×10^{-5}* [0.03]	--
Cell proliferation	Tumor proliferation	--	0.002	--	0.03
Protein degradation	Proteolysis	--	0.003 [0.003]	--	0.01
Infectious disease	Replication of virus	--	--	--	0.01
Neurological disease	Movement disorder	--	--	--	0.02

*$p<0.001$; [ ] = p-value of gene ontology terms in DAVID.
[a] Associated contaminant exposure in fetal or maternal DNA
[b] Lipid metabolism, molecular transport, small molecule biochemistry
[c] Nervous system development and function

Table 3. Transcription factor binding sites showed sequence similarity to the identified motifs within each Cd- or cotinine-associated gene set.

	Fetal DNA		Maternal DNA	
Transcription Factor	Cd [a]	Cotinine [a]	Cd [a]	Cotinine [a]
TCFAP2E	2.3×10^{-5}*	9.2×10^{-4}*	--	4.9×10^{-3}
TCF7L1	6.7×10^{-5}*	2.0×10^{-5}*	5.1×10^{-4}*	1.5×10^{-4}*
SRF	2.2×10^{-4}*	3.2×10^{-4}*	1.5×10^{-3}	5.0×10^{-3}
MTF1	3.1×10^{-4}*	1.6×10^{-4}*	3.5×10^{-4}*	2.0×10^{-3}

* $p< 0.001$

[a] Associated contaminant exposure in fetal or maternal DNA

References

1. Tellez-Plaza, M., et al., *Cadmium exposure and all-cause and cardiovascular mortality in the U.S. general population*. Environ Health Perspect, 2012. 120(7): p. 1017-22.

2. ATSDR. *Toxicological Profile for Cadmium*. 2008 [cited 2011 July 1].

3. Tellez-Plaza, M., et al., *Cadmium exposure and incident cardiovascular disease*. Epidemiology, 2013. 24(3): p. 421-9.

4. Navas-Acien, A., et al., *Lead, cadmium, smoking, and increased risk of peripheral arterial disease*. Circulation, 2004. 109(25): p. 3196-201.

5. Fagerberg, B., et al., *Cadmium exposure, intercellular adhesion molecule-1 and peripheral artery disease: a cohort and an experimental study*. BMJ Open, 2013. 3(3).

6. Everett, C.J. and I.L. Frithsen, *Association of urinary cadmium and myocardial infarction*. Environ Res, 2008. 106(2): p. 284-6.

7. Agarwal, S., et al., *Heavy metals and cardiovascular disease: results from the National Health and Nutrition Examination Survey (NHANES) 1999-2006*. Angiology, 2011. 62(5): p. 422-9.

8. Peters, J.L., et al., *Cadmium exposure in association with history of stroke and heart failure*. Environ Res, 2010. 110(2): p. 199-206.

9. Tellez-Plaza, M., A. Navas-Acien, and E. Guallar, *Cadmium as a novel cardiovascular risk factor: Supportive evidence and future directions*. Nature Reviews Cardiology (CNIC Edition), 2010. 7: p. 41-46.

10. Messner, B. and D. Bernhard, *Cadmium and cardiovascular diseases: cell biology, pathophysiology, and epidemiological relevance*. Biometals, 2010. 23(5): p. 811-22.

11. Mackay, J. and G.A. Mensah, *The Atlas of Heart Disease and Stroke*2004, Geneva, Switzerland: WHO.

12. Bhatnagar, A., *Environmental cardiology: studying mechanistic links between pollution and heart disease*. Circ Res, 2006. 99(7): p. 692-705.

13. Weinhold, B., *Environmental cardiology: getting to the heart of the matter*. Environ Health Perspect, 2004. 112(15): p. A880-7.

14. Jarup, L. and A. Akesson, *Current status of cadmium as an environmental health*

problem. Toxicol Appl Pharmacol, 2009. 238(3): p. 201-8.

15. Rahil-Khazen, R., et al., *Multi-element analysis of trace element levels in human autopsy tissues by using inductively coupled atomic emission spectrometry technique (ICP-AES)*. J Trace Elem Med Biol, 2002. 16(1): p. 15-25.

16. Zepeda, R., et al., *Cardiac tissue injury resistance during myocardial infarction at adulthood by developmental exposure to cadmium*. Cardiovasc Toxicol, 2012. 12(1): p. 64-72.

17. Ronco, A.M., et al., *Maternal exposure to cadmium during gestation perturbs the vascular system of the adult rat offspring*. Toxicol Appl Pharmacol, 2011. 251(2): p. 137-45.

18. Hallare, A.V., et al., *Combined effects of temperature and cadmium on developmental parameters and biomarker responses in zebrafish (Danio rerio) embryos*. Journal of Themal Biology, 2005. 30: p. 7-17.

19. Cheng, S.H., P.K. Chan, and R.S. Wu, *The use of microangiography in detecting aberrant vasculature in zebrafish embryos exposed to cadmium*. Aquat Toxicol, 2001. 52(1): p. 61-71.

20. Waisberg, M., et al., *Molecular and cellular mechanisms of cadmium carcinogenesis*. Toxicology, 2003. 192(2-3): p. 95-117.

21. Takiguchi, M., et al., *Effects of cadmium on DNA-(Cytosine-5) methyltransferase activity and DNA methylation status during cadmium-induced cellular transformation*. Experimental Cell Research, 2003. 286(2): p. 355-365.

22. Yazihan, N., et al., *Involvement of galectin-3 in cadmium-induced cardiac toxicity*. Anadolu Kardiyol Derg, 2011. 11(6): p. 479-84.

23. Benton, M.A., et al., *Comparative genomic analyses identify common molecular pathways modulated upon exposure to low doses of arsenic and cadmium*. BMC Genomics, 2011. 12: p. 173.

24. Zhao, Y. and D. Srivastava, *A developmental view of microRNA function*. Trends Biochem Sci, 2007. 32(4): p. 189-97.

25. Jovanovic, M. and M.O. Hengartner, *miRNAs and apoptosis: RNAs to die for*. Oncogene, 2006. 25(46): p. 6176-87.

26. Johnson, C.D., et al., *The let-7 microRNA represses cell proliferation pathways in human cells*. Cancer Res, 2007. 67(16): p. 7713-22.

27. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation*

and deep-sequencing data. Nucleic Acids Res, 2011. 39(Database issue): p. D152-7.

28. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res, 2009. 19(1): p. 92-105.

29. Divakaran, V. and D.L. Mann, *The emerging role of microRNAs in cardiac remodeling and heart failure*. Circ Res, 2008. 103(10): p. 1072-83.

30. Cordes, K.R. and D. Srivastava, *MicroRNA regulation of cardiovascular development*. Circ Res, 2009. 104(6): p. 724-32.

31. Irizarry, R.A., et al., *Summaries of Affymetrix GeneChip probe level data*. Nucleic Acids Res, 2003. 31(4): p. e15.

32. Storey, J.D., *The positive false discovery rate: a Bayesian interpretation and the q-value*. The Annals of Statistics, 2003. 31: p. 2013-2035.

33. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository*. Nucleic Acids Res, 2002. 30(1): p. 207-10.

34. Elefant, N., et al., *RepTar: a database of predicted cellular targets of host and viral miRNAs*. Nucleic Acids Res, 2011. 39(Database issue): p. D188-94.

35. Levine, A.J., *p53, the cellular gatekeeper for growth and division*. Cell, 1997. 88(3): p. 323-31.

36. Crow, M.T., *Revisiting p53 and its effectors in ischemic heart injury*. Cardiovasc Res, 2006. 70(3): p. 401-3.

37. Matsusaka, H., et al., *Targeted deletion of p53 prevents cardiac rupture after myocardial infarction in mice*. Cardiovasc Res, 2006. 70(3): p. 457-65.

38. Yang, B., et al., *The muscle-specific microRNA miR-1 regulates cardiac arrhythmogenic potential by targeting GJA1 and KCNJ2*. Nat Med, 2007. 13(4): p. 486-91.

39. Thum, T., et al., *MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts*. Nature, 2008. 456(7224): p. 980-4.

40. Zhang, X., et al., *Regulation of cardiac microRNAs by serum response factor*. J Biomed Sci, 2011. 18: p. 15.

41. Busk, P.K. and S. Cirera, *MicroRNA profiling in early hypertrophic growth of the left ventricle in rats*. Biochem Biophys Res Commun, 2010. 396(4): p. 989-93.

42. Jentzsch, C., et al., *A phenotypic screen to identify hypertrophy-modulating microRNAs in primary cardiomyocytes*. J Mol Cell Cardiol, 2012. 52(1): p. 13-20.

43. Huang, Z.P., et al., *MicroRNA-22 Regulates Cardiac Hypertrophy and Remodeling in Response to Stress*. Circ Res, 2013. 112(9): p. 1234-43.

44. Xu, X.D., et al., *Attenuation of microRNA-22 derepressed PTEN to effectively protect rat cardiomyocytes from hypertrophy*. J Cell Physiol, 2012. 227(4): p. 1391-8.

45. Zhao, Y., E. Samal, and D. Srivastava, *Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis*. Nature, 2005. 436(7048): p. 214-20.

46. Kasinski, A.L. and F.J. Slack, *Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy*. Nat Rev Cancer, 2011. 11(12): p. 849-64.

47. Ventura, A., et al., *Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters*. Cell, 2008. 132(5): p. 875-86.

48. Kim, G.H., et al., *Translational control of FOG-2 expression in cardiomyocytes by microRNA-130a*. Plos One, 2009. 4(7): p. e6161.

49. Vacchi-Suzzi, C., et al., *Heart structure-specific transcriptomic atlas reveals conserved microRNA-mRNA interactions*. Plos One, 2013. 8(1): p. e52442.

50. Bollati, V., et al., *Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes*. Environ Health Perspect, 2010. 118(6): p. 763-8.

51. Fabbri, M., et al., *Whole genome analysis and microRNAs regulation in HepG2 cells exposed to cadmium*. ALTEX, 2012. 29(2): p. 173-82.

52. Kong, A.P., et al., *Associations between microRNA (miR-21, 126, 155 and 221), albuminuria and heavy metals in Hong Kong Chinese adolescents*. Clin Chim Acta, 2012. 413(13-14): p. 1053-7.

53. Yanez, L., et al., *Arsenic-cadmium interaction in rats: toxic effects in the heart and tissue metal shifts*. Toxicology, 1991. 67(2): p. 227-34.

54. Lu, J., et al., *Metallothionein gene expression in peripheral lymphocytes from cadmium-exposed workers*. Cell Stress Chaperones, 2001. 6(2): p. 97-104.

55. Nordberg, G.F., *Biomarkers of exposure, effects and susceptibility in humans and their application in studies of interactions among metals in China*. Toxicol Lett,

2010. 192(1): p. 45-9.

56. Andrews, G.K., *Cellular zinc sensors: MTF-1 regulation of gene expression*. Biometals, 2001. 14(3-4): p. 223-237.

57. Place, R.F., et al., *MicroRNA-373 induces expression of genes with complementary promoter sequences*. Proc Natl Acad Sci U S A, 2008. 105(5): p. 1608-13.

58. Andrews, G.K., *Regulation of metallothionein gene expression by oxidative stress and metal ions*. Biochemical Pharmacology, 2000. 59(1): p. 95-104.

59. Banerjee, I., et al., *Determination of cell types and numbers during cardiac development in the neonatal and adult rat and mouse*. Am J Physiol Heart Circ Physiol, 2007. 293(3): p. H1883-91.

60. Anson, B.D., K.L. Kolaja, and T.J. Kamp, *Opportunities for use of human iPS cells in predictive toxicology*. Clin Pharmacol Ther, 2011. 89(5): p. 754-8.

61. Chin, M.H., et al., *Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures*. Cell Stem Cell, 2009. 5(1): p. 111-23.

62. Braam, S.R., R. Passier, and C.L. Mummery, *Cardiomyocytes from human pluripotent stem cells in regenerative medicine and drug discovery*. Trends Pharmacol Sci, 2009. 30(10): p. 536-45.

63. Harding, S.E., *Human stem cell-derived cardiomyocytes for pharmacological and toxicological modeling*. Ann N Y Acad Sci, 2011. 1245: p. 48-9.

64. Gunaratne, P.H., *Embryonic stem cell microRNAs: defining factors in induced pluripotent (iPS) and cancer (CSC) stem cells?* Curr Stem Cell Res Ther, 2009. 4(3): p. 168-77.

65. Chaudhry, M.A., et al., *Micro RNA responses to chronic or acute exposures to low dose ionizing radiation*. Mol Biol Rep, 2012. 39(7): p. 7549-58.