INTENTIONALLY LEFT BLANK

AN ABSTRACT OF THE DISSERTATION OF

Andres Cardenas for the degree of Doctor of Philosophy in Public Health presented on September 10, 2015.

Title: *In Utero* Exposure to Arsenic and Mercury: Epigenome-Wide Associations and Newborn Health Outcomes

Abstract approved:

____

Molly L. Kile

Mercury and arsenic are known developmental toxicants and environmental exposures are ubiquitous worldwide from natural and anthropogenic sources. Prenatal exposure to both contaminants are independently associated with adverse perinatal health outcomes and latent disease risk that could be in part mediated by epigenetic reprogramming events. Fetal programming events involving DNA methylation occur early during fetal development and are sensitive to environmental stimuli, potentially influencing disease risk throughout the life course. This research evaluated the epigenetic disruption of DNA methylation associated with exposure to mercury and arsenic *in utero*, and the potential impact on infant birth outcomes.

The first study provides evidence that exposure to mercury during fetal development can contribute to epigenetic variability and immune cell proportion changes in infant cord-blood collected at birth. The data also suggests that exposure to mercury and arsenic, even at low levels, may interact to impact the epigenome in cord-blood. The second study evaluated the epigenome wide association of arsenic exposure *in utero* for placenta, umbilical artery and umbilical vein endothelial cells (HUVEC). Our results show that prenatal arsenic exposure alters DNA methylation of umbilical artery and placenta but there is limited evidence for an association in HUVEC. We also identified DNA methylation disruption of key biological pathways related to adverse health

outcomes previously associated with arsenic exposure in epidemiological studies. Lastly, the third study evaluated the direct association between prenatal arsenic exposure, birth gestational age and birth weight, as well as the indirect effect of arsenic exposure mediated through DNA methylation in cord-blood. We observed that the inverse relationship between arsenic exposure *in utero* and birth gestational age was completely mediated through DNA methylation of selected CpG loci. No association or mediation was observed between prenatal arsenic exposure and infant birth weight.

These results provide evidence that exposure to arsenic and mercury *in utero*, two common environmental contaminants, can influence epigenetic programming of DNA methylation in different biological human tissues. Particularly, the effect of arsenic exposure on gestational age was mediated through CpG methylation of infant cord-blood, which is a critical parameter of infant health and associated with disease later in life. Future studies should evaluate if these epigenetic changes are persistent and associated with disease risk later in life. The design of public health interventions that target prevention of environmental exposures *in utero* that may increase disease risk later in life provide a unique opportunity for early disease prevention. Overall, these results will guide future research and inform regulatory guidelines to help reduce arsenic and mercury exposure, particularly during pregnancy.

©Copyright by Andres Cardenas
September 10, 2015
All Rights Reserved

In Utero Exposure to Arsenic and Mercury: Epigenome-Wide Associations and Newborn Health Outcomes

by
Andres Cardenas

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented September 10, 2015
Commencement June 2016

Doctor of Philosophy dissertation of Andres Cardenas presented on September 10, 2015

APPROVED:

Major Professor, representing Public Health

Co-Director of the School of Biological and Population Health Sciences

---

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Andres Cardenas, Author

ACKNOWLEDGEMENTS

I would like to thank many individuals, organizations and donors that supported my research, education and personal growth during my time at Oregon State University. First, I would like to express my most honest gratitude to Dr. Molly Kile for her continued support, mentoring, guidance and encouragement. I was very fortunate not only to have an outstanding scholar as a mentor but also an approachable, caring, genuine and helpful individual. Thank you for training me in conducting quality research and taking me halfway across the world to learn the trait. Dr. Andy Houseman, you were patience, understanding and instrumental in teaching me everything that I know about handling high dimensional data. I truly appreciate the countless amount of hours that you spent teaching me not only skills but also providing guidance and mentoring in my life. You both have made a large impact in my life and played a pivotal role in my career. I would also like to thank my committee members for their advice, feedback, flexibility and interest in my success. I truly appreciate the encouragement and advice provided by Dr. Anna Harding in pursuing a doctorate degree, your support and guidance were instrumental in my decision to continue in academia. I would also like to thank Dr. Ellen Smit for her continued input in several research projects, your feedback and ideas have always been insightful and useful. To my graduate council representative Dr. Emily Ho for making this process as smooth as possible and ongoing support.

To my family, for encouraging me to continue my education. My parents, Alice and Jose and my brothers Jose and Ruander for your continued support and commitment to my success. You all have taught me the value of hard work and the importance of having goals. Maria, you have been absolutely supportive, encouraging and caring through this whole process. Thank you for your patience, help and continued unconditional support.

Andres Cardenas

CONTRIBUTION OF AUTHORS

Andres Cardenas conceptualized the research questions, analyzed the data, interpreted the results and drafted all three manuscripts.

Dr. Molly L. Kile helped conceptualize the study design and research questions for the second and third manuscripts. Supervised the analysis, helped with the interpretation of findings and provided critical review of all three manuscripts.

Dr. EA Houseman supervised all statistical analyses, provided technical assistance, helped with the interpretation of findings and provided feedback in all three manuscripts.

Dr. Devin C. Koestler supervised the data analysis for the first manuscript.

Dr. Brian P. Jackson helped in the design and exposure analysis for the first manuscript.

Dr. Margaret R. Karagas conceptualized the study design, supervised the execution of all aspects of the birth cohort and provided feedback for the first manuscript.

Dr. Carmen J. Marsit helped conceptualize the research question, supervised the analysis plan and provided key guidance in the interpretation of results for the first manuscript.

Dr. Andrea Baccarelli provided technical assistance and feedback in the design and interpretation of the second and third manuscripts.

Dr. Robert O. Wright provided feedback in the study design for the second manuscript.

Dr. David C. Christiani conceptualized the study design, helped with the interpretation of findings and provided feedback for the second and third manuscripts.

Dr. Quazi Quamruzzaman, Dr. Mahmuder Rahman and Mr. Golam Mostofa organized the sample collection, executed and provided feedback in the design of the birth cohort used for the second and third manuscripts.

TABLE OF CONTENTS

Page

Chapter 1 - INTRODUCTION 1
- Developmental Origins of Health and Disease Hypothesis 1
- Epigenetics: The Role of DNA Methylation in Fetal Programming 3
- Environmental Influence on Epigenetic Regulation and Adult Disease 5
- Inorganic Arsenic and Mercury: Sources of Exposure 7
- Inorganic Arsenic Toxicokinetics and Biomarkers 9
- Methylmercury Toxicokinetics and Biomarkers 10
- Birth Weight and Gestational Age as Outcomes 11
- Arsenic Exposure, Birth Weight and Neurodevelopment 12
- Methylmercury Exposure and Neurodevelopment 14
- Epigenetic effects of Inorganic Arsenic 15
- Epigenetic effects of Methylmercury 16
- Specific Aims 17

Chapter 2 - FIRST MANUSCRIPT 19
- Abstract 20
- Introduction 21
- Materials & Methods 23
  - Study Population 23
  - Toenail Hg Assessment 23
  - Urinary Arsenic Assestment 24
  - DNA methylation assessment and Preprocessing 25
  - Statistical Analysis 25
- Results 27
- Discussion 28
- Conclusions 32
- Acknowledgements 32

Chapter 3 SECOND MANUSCRIPT 37
- Abstract 38

TABLE OF CONTENTS (Continued)

Page

Introduction 39
Materials & Methods 41
Study Population 41
Drinking Water Arsenic 42
Tissue Collection: Umbilical Artery, Placenta & HUVEC 42
DNA Methylation Assessment and Quality Control 43
Statistical Analysis 44
Results 45
Arterial Tissue 46
Placenta Tissue 47
Umbilical Vein Endothelial Cells (HUVEC) 48
Discussion 49
Conclusions 53
Acknowledgements 53
Chapter 4 - THIRD MANUSCRIPT 62
Abstract 63
Introduction 64
Materials & Methods 66
Study Population 66
Maternal Drinking Water Arsenic 66
Cord-Blood DNA Methylation and Quality Control 67
Statistical Analysis 67
Results 69
CpG Methylation and Birth Gestational Age 70
CpG Methylation and Birth Weight 72
Discussion 72
Conclusions 75
Acknowledgements 76

TABLE OF CONTENTS (Continued)

Page

Chapter 5 SUMMARY AND CONCLUSIONS 88

Aim 1 88

Aim 2 89

Aim 3 90

Summary 90

BIBLIOGRAPHY 92

LIST OF FIGURES

Figure	Page
Figure 2.1. Differentially methylated loci based on maternal toenail-Hg.	35
Figure 2.2. Differentially methylated loci based on the interaction between maternal toenail-Hg and urinary-As.	36
Figure 3.1. Locus-by-locus epigenome-wide analysis for umbilical artery	59
Figure 3.2. Locus-by-locus epigenome-wide analysis for placenta	60
Figure 3.3. Locus-by-locus epigenome-wide analysis for HUVEC	61
Figure 4.1. Experimental approach for the discovery and validation of CpG loci and mediated effects	81
Figure 4.2. Conceptual Structural Equation Model (SEM) for the direct and indirect effect of exposure and infant birth outcomes	82
Figure 4.3. Differentially methylated CpG loci in cord-blood relative to prenatal arsenic exposure	83
Figure 4.4. Volcano plots for the association among top CpGs and infant health outcomes	84
Figure 4.5. Correlation among top 10 CpGs Associated with As exposure in utero and gestational age at birth.	85
Figure 4.6. Proportion of variance explained by the first four principal components of DNA methylation	86
Figure 4.7. Structural Equation Model (SEM) conceptualized for the mediated association	87

LIST OF TABLES

Table Page

Table 2.1. Selected maternal and infant characteristics and Hg and As concentrations... 33

Table 2.2. Estimated change in the proportion of leukocyte composition in cord blood in relationship to $\log_2$ toenail Hg concentration.. 34

Table 3.1. Sample characteristics for the 52 mother-infant pairs eligible for the analysis. 54

Table 3.2. Omnibus association of DNA methylation of KEGG biological pathways and prenatal arsenic exposure in artery, placenta and HUVEC............................ 55

Table 4.1. Selected sample characteristics of the study population................................ 77

Table 4.2. Top-10 CpG loci significantly associated with birth gestational age, ... 78

Table 4.3. Structural equation model for the mediated effect of arsenic exposure on birth gestational age.. 79

Table 4.4. Fit indices for the final the structural equation model conceptualized 80

LIST OF APPENDICES

Appendix Page

Appendix A: Supplemental Materials for Chapter 2 115

Appendix B: Supplemental Materials for Chapter 3 127

Appendix C: Supplemental Materials for Chapter 4 132

LIST OF APPENDIX FIGURES

Figure Page

Figure A.1. Density distributions of the DNA methylation by sample 125

Figure A.2. Principal Component Analysis (PCA) of DNA methylation by plate......... 126

Figure B.1. Relative occurrence of the top differentially methylated loci relative to CpG island location in umbilical artery 129

Figure B.2. Relative occurrence of the top differentially methylated loci relative to CpG island location in placenta 130

Figure B.3. Overlap for the unadjusted and cell mixture adjusted results among the top 100 131

Figure C.1. Association between prenatal As exposure and CpG methylation of selected loci 132

Figure C.2. Associations between birth gestational age and CpG methylation of selected loci 133

LIST OF APPENDIX TABLES

Table | Page

Table A.1. Association between the three main principal components and array plate . 115

Table A.2. Estimated change in the proportion of leukocyte composition in cord blood in relationship to the multiplicative interaction between $\log_2$ toenail Hg and urinary As interaction ... 116

Table A.3. Individual CpG lookup of previously reported DNA methylation changes in relationship to Hg exposure in candidate genes studies in relationship to mercury exposure from the New Hampshire Birth Cohort... 117

Table A.4. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal Hg exposure ... 119

Table A.5. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal co-exposure to Hg and As... 122

Table B.1. Top-10 CpGs for unadjusted cell mixture analysis in artery ... 127

Table B.2. Top-10 CpGs for adjusted cell mixture analysis in artery ... 127

Table B.3. Top-10 CpGs for unadjusted cell mixture analysis in placenta ... 127

Table B.4. Top-10 CpGs for adjusted cell mixture analysis in placenta ... 128

Table B.5. Top-10 CpGs for unadjusted cell mixture analysis in HUVEC... 128

Table B.6. Top-10 CpGs for adjusted cell mixture analysis in HUVEC... 128

In Utero Exposure to Arsenic and Mercury: Epigenome-Wide Associations and Newborn Health Outcomes

Chapter 1 - INTRODUCTION

Developmental Origins of Health and Disease Hypothesis

Many common and chronic diseases in the population are mediated through genetic and environmental factors and previous research has shown that the period of fetal development is adaptive to exogenous stimuli such as nutritional factors and environmental exposures (Perera & Herbstman, 2011; Wu, Bazer, Cudd, Meininger, & Spencer, 2004). This fetal developmental adaptive process is referred to as fetal programming, in which the developing fetus has the potential to modulate and adapt the expression of some traits or phenotypes in response to the environment. This process can provide evolutionary advantage for the fetus by regulating gene-expression and phenotypic characteristics to better prepare the offspring in utero in response to environmental conditions experienced during development, particularly nutritional conditions (Gluckman, Hanson, & Low, 2011). However, disruption of this fetal programming process and mismatch of the expected environment during development has been associated with an increase in disease risk and susceptibility later in life (Catalano, Presley, Minium, & Hauguel-de Mouzon, 2009; Entringer, Buss, & Wadhwa, 2010; Godfrey, Lillycrop, Burdge, Gluckman, & Hanson, 2007). The observed association of prenatal exposures, environmental stimuli and nutritional conditions to influence an individual's disease risk or phenotypic traits later in life is referred to as the Developmental Origins of Health and Disease (DOHaD) hypothesis (Barker, 2003). This hypothesis states that intra-uterine nutrition and environmental conditions have the ability to influence gene expression and metabolism causing life-long changes in phenotypes including disease susceptibility and latent disease risk (Barker, 2012).

The main proposed mechanism for fetal programming are epigenetic modifications. These are heritable changes in gene expression that occur without direct changes to the DNA sequence (Bollati & Baccarelli, 2010). Several epigenetic

modifications have been identified that include but are not limited to: DNA methylation, histone modifications and regulation of non-coding RNAs (Jirtle & Skinner, 2007). These epigenetic mechanisms have been shown to regulate gene expression, and also been observed to have some plasticity during development likely due to intrauterine conditions and adaptation (Christensen & Marsit, 2011). Fetal programming events in part mediated through epigenetic mechanisms occurs early during fetal development, a period hypothesized to be the most sensitive to epigenetic disruption. Particularly, early stages of embryonic development have been shown to be sensitive to environmental insults and nutritional conditions (Cortessis et al., 2012). The epigenome has been hypothesized to be the driving mechanism of fetal programming in utero and in part modulate the variability observed in the population for disease onset and progression of many complex and chronic diseases. The process of fetal programming and the DOHaD hypothesis hold the potential to explain the role of environmental exposures at modulating latent disease risk and their identification can generate opportunities for disease interventions and prevention strategies (Foley et al., 2009).

Identifying prenatal environmental exposures that influence the epigenetic makeup of newborns and their associations with children's overall health can provide important insights for preventable exposures that are associated with overall health throughout their life course. To date, Genome-Wide Association studies (GWAs) have only identified a handful of Single Nucleotide Polymorphisms (SNPs) strongly associated with common diseases. At best, they represent a small increment in risk or only explain a relatively small proportion of the variation of disease observed in the population (Manolio et al., 2009). The lack of strong or robust genetic associations with common metabolic diseases in the population has led to the hypothesis that some of these diseases might be in part mediated through epigenetic processes (Gluckman, Hanson, Cooper, & Thornburg, 2008). Furthermore, the discovery of epigenetic changes induced by environmental conditions could help design strategies, interventions and policies targeted at epigenetic mechanisms that could ultimately improve population level health (Burris & Baccarelli, 2014). Theoretically, identifying epigenetic changes can provide an important tool for risk prediction, risk modification, disease biomarkers and exposure biomarkers

(Bakulski & Fallin, 2014). It has also been hypothesized that epigenetic disruption might be the missing link between environmental exposures and common diseases for which a genetic link remains largely absent (Cortessis et al., 2012). Evaluating if ubiquitous environmental toxicants such as arsenic and mercury disrupt the epigenome and their relationship with fetal health outcomes has the potential to provide an important tool for the primary prevention of chronic diseases. Furthermore, numerous studies have shown that most epigenetic changes are reversible, holding promise for the design of treatments and interventions to improve health (Hou, Zhang, Wang, & Baccarelli, 2011).

Epigenetics: The Role of DNA Methylation in Fetal Programming

The hereditary material of nearly all living organisms is comprised of deoxyribonucleic acid (DNA), which contains the genetic blueprint for the development and maintenance of all biological processes. Epigenetically, DNA can be modified by the addition of a methyl group to the 5' position of the pyrimidine ring of a Cytosine nucleotide often followed by a Guanine nucleotide. This methylated dinucleotide sequence is referred to as a CpG (Meaney & Szyf, 2005). DNA methylation has been the most widely studied form of epigenetic alteration. These CpG dinucleotides tend to occur in clusters referred to as CpG islands or in repetitive elements within the genome. An estimated 28 million Cytosines, or between 60-80% of the *CpGs* present in the genome are normally methylated, mainly in repetitive sequences such as transposable elements, including Line (L1) and SINE (Alu) elements. Less than 10% of all CpG dinucleotides in the human genome occur in CpG rich regions, referred to as "CpG islands", outside of repetitive elements (Ehrlich, 2009). The relevance of these CpG islands is that they are located in promoter regions or exons of at least half of the genes across the human genome and are largely resistant to DNA methylation. Furthermore, CpG islands are mostly present at transcription start sites in many housekeeping genes and developmental regulatory genes across the human genome (Deaton & Bird, 2011). The majority of CpGs in transposable elements are often silenced by promoter methylation in normal healthy cells (Z. D. Smith & Meissner, 2013). Conversely, methylation of CpG islands outside of repetitive elements has been reported to halt gene expression and delay replication

(Baylln, Herman, Graff, Vertino, & Issa, 1997). In a oversimplified view of this dynamic, methylation of CpG islands upstream of genes has been shown to repressed gene expression while hypomethylated CpG islands would allow the gene to be normally expressed. However, DNA methylation is a dynamic and variable process with methylation patterns in the genome varying between individuals. For example, age associated epigenetic drift has been observed among identical twins (Martin, 2005). This epigenetic diversity suggests that even among genetically identical organisms, environmental conditions could influence an individual's phenotype such as disease onset or progression. For instance, genetically identically organisms and human monozygotic twins are often discordant in several phenotypes including complex and chronic diseases, suggesting that environmental conditions significantly impact the degree of phenotypic variation thought to occur through an epigenetic mechanisms and fetal programming (Wong, Gottesman, & Petronis, 2005).

In the absence of pathology in normal cell lines, the majority of CpG islands that are outside of repetitive elements remain unmethylated. However, CpG islands that are hypermethylated in promoter regions of genes can lead to gene-silencing. For example, it has been demonstrated that several promoter regions of tumor-suppressor genes are hypermethylated contributing to the carcinogenic process (Esteller, 2002). Hypermethylation of CpG islands located in promoter regions of tumor suppressor genes have been identified as a mechanism for disease progression and it is now well recognized that aberrant DNA methylation is a prevailing signature of many types of cancer (Koestler et al., 2012; Sproul et al., 2012). It has also been shown that genome wide DNA methylation is involved in critical stages of fetal development. During fertilization both parental genomes are quickly demethylated to allow for totipotent and pluripotent cell differentiation followed by the novo methylation shortly after implantation (Z. D. Smith & Meissner, 2013). De novo DNA methylation is established after implantation and this process is critical for cell differentiation and fetal development. This step is hypothesized to be the most critical window for fetal programming in which the fetus can acquire specific phenotypic characteristics through epigenetic regulation and disruption of this process. Therefore, DNA methylation has

been proposed to be the major mechanism implicated in fetal programming in nutritional and environmental studies (Chmurzynska, 2010).

Environmental Influence on Epigenetic Regulation and Adult Disease

The Developmental Origins of Health and Disease (DOHaD) hypothesis originated from the observation of nutritional conditions during development. First in 1986 Barker and colleagues observed high rates of Coronary Heart Disease (CHD) in areas of Wales and England with poor living conditions and high infant mortality, suggesting that nutritional deprivation *in utero* was a risk factor for CHD among this population (Barker & Osmond, 1986). Epidemiological evidence for this hypothesis initially originated from the Dutch Hunger Winter occurring during World War II consequence of the food embargo imposed by Germany on the Netherlands during 1944-45. Individuals born during this period were prenatally exposed to famine and severe caloric deficiencies. Health records of this natural experiment were available for newborns that faced famine prenatally and examined for birth outcomes, mortality and morbidity later in life. The results indicated that nutrient deprived children *in utero* had an increased risk for CHD in adulthood (Painter et al., 2006; T. J. Roseboom et al., 2000; Stanner et al., 1997). Barker and colleagues subsequently described the observed associations between birth weight, length and size at birth with CHD, further supporting that early fetal health outcomes were important predictors of disease morbidity later in life (Barker, 1995). After Barker's seminal work, many other studies also described associations between low birth weight and the incidence of heart disease (Rich-Edwards et al., 1997), hypertension (Law & Shiell, 1996) and type 2 diabetes (Hales et al., 1991) with relationships remaining robust to statistical adjustment for potential confounders. Under the DOHaD hypothesis these epidemiological studies established that birth outcomes such as birth weight, fetal growth and gestational age were important predictors of morbidity and mortality throughout life. Along with low birth weight and infant size, long lasting epigenetic alterations of biologically significant genes have also been observed among individuals exposed to famine early during fetal development (Elmar W Tobi et al., 2015). This suggests that methylation changes might be the link between poor

birth outcomes and disease susceptibility later in life, as suggested by ongoing work (Elmar Wouter Tobi, 2013; Elmar W Tobi et al., 2009).

More recently, in addition to nutritional factors, several environmental exposures have been shown to disrupt the epigenome in both experimental models and epidemiological studies. Some of the toxicants shown to disrupt the epigenome include metals, air pollution, endocrine disrupting chemicals and pesticides. In particular, exposure to several metals have been shown to disrupt DNA methylation including cadmium, arsenic, nickel, chromium and methylmercury of in repetitive elements and candidate gene studies (Martinez-Zamudio & Ha, 2011). The leading hypothesized mechanism for the epigenetic toxicity of metals include the generation of oxidative stress and increased production of reactive oxygen species, partially shown in experimental models (Baccarelli & Bollati, 2009). In human studies, epigenetic alterations due to air pollution exposure have been associated in the development of asthma and the regulation of inflammatory signals involved in disease pathogenesis (Cortessis et al., 2012). For example, tobacco smoke has also been identified as an epigenetic toxicant. Using a genome-wide approach *in utero* exposure to tobacco smoke was associated with differential DNA methylation at 26 *CpG* sites, using cotinine as an objective biomarker of tobacco smoke exposure. This large cohort study identified differentially methylated *CpGs* located in genes associated with the Aryl hydrocarbon Receptor (AhR), important in the detoxification of tobacco smoke exposure (Joubert et al., 2012). This study highlights the utility of human epigenetic studies to uncover potential exposure biomarkers with biological implications. Animal models have also characterized epigenetic alterations associated with varying phenotypes including disease. For example, mice carrying the A^{vy} epiallele have a distinct yellow coat color and varying disease phenotypes like obesity and insulin resistance associated with the degree of DNA methylation at the epiallele. Interestingly, methylation levels of this epiallele can be moderated by exposure to the phytoestrogen genistein, yielding a proof of concept for epigenetic interventions (Cropley, Suter, Beckman, & Martin, 2006).

In summary, several ubiquitous environmental exposures have been associated with disruption of the epigenome and implicated in disease pathogenesis.

Some epigenetic alterations have the potential to be inherited and this process might have lasting effects on population health.

Inorganic Arsenic and Mercury: Sources of Exposure

Arsenic (As) is a natural and abundant element present in the earth's crust and soil. Except for individuals that are occupationally exposed to high levels of arsenic, the main route of human exposure to arsenic is via the ingestion of contaminated drinking water (US Environmental Protection Agency (US EPA), 2001). Arsenic easily dissolves in water from rocks and other minerals present in the soil, entering aquifers in the ground (Hindmarsh, McCurdy, & Savory, 1986). In the Indian subcontinent alone, chronic arsenic poisoning from drinking water affects over 120 million people (Chowdhury et al., 2000). Arsenic has been classified as a known human carcinogen by the International Agency for Research in Cancer (IARC) and by the US EPA. In the United States, a maximum contamination level of 50 µg/L was in effect for public water supplies beginning in 1942. That standard was revised in 2001 and lowered to 10 µg/L by the US Environmental Protection Agency. However, it has been estimated that approximately 13 million Americans currently live in areas where arsenic concentration surpasses this level (A. S. Chen et al., 2004). The maximum contamination level of 10 µg/L was derived using a cost-benefit analysis, with only cancer being the main endpoint evaluated for risk assessment (US EPA, 2001).

Arsenic contaminated drinking water is considered to be the dominant source of exposure and is mostly a concern for communities that rely on groundwater as their source of potable water (US EPA, 2001). Other routes of exposure to arsenic include dietary sources like seafood, juices and grains that take up arsenic from contaminated soils. Grains, particularly rice and cereals, can also take up arsenic from the soil and irrigation water (Cascio et al., 2011). Dietary arsenic is mostly present in its organic form called arsenobetaine, this is relatively less harmful compared to the inorganic species present in water. Dietary organic arsenic that comes from seafood items such as fish and shrimp are not of immediate concern to human health because dietary arsenic is biologically stable and has little toxicity relative to inorganic arsenicals found in water

supplies (Cullen & Reimer, 1989; Navas-Acien, Silbergeld, Pastor-Barriuso, & Guallar, 2009). The inorganic-related species have been demonstrated to be the most toxic to humans and animals, causing skin and body lesions and have also been associated with an increased risk of skin, lung, bladder, kidney and liver cancer among others (Agency for Toxic Substances and Disease Registry (ATSDR), 2000; Yoshida, Yamauchi, & Fan Sun, 2004). Besides being a strong human carcinogen chronic arsenic exposure has recently been identefied as an important immunotoxicant (Burchiel et al., 2009; Dangleben, Skibola, & Smith, 2013; Martin-Chouly et al., 2011; Soto-Peña et al., 2006). Similarly, epidemiological stuies have associated increased arsenic exposure with the occurrence of infectiouses diseases, especially among children exposed *in utero* (Rahman, Vahter, Ekstrom, & Persson, 2010; Raqib et al., 2009). Due to it's potency, widespread occurance and potential for human exposure, arsenic obtained the highest ranking on the 2013 Agency for Toxic Substances and Disease Registry (ATSDR) priority list of 275 hazardous substances (Naujokas et al., 2013).

Mercury (Hg) is also a naturally occurring metallic element present in the earth's crust. Elemental mercury can be found in mineral deposits including coal and release into the environment naturally through volcanic activity. The largest anthropogenic source of mercury in the US comes from coal-burning power plants. Since the industrial revolution, it has been estimated that mercury levels in the ocean have increased at least three-fold (Lamborg et al., 2014). Other sources of exposure include the byproducts of chlorine production, dental fillings using amalgams, thermometers, and batteries. Mercury is converted by some microorganisms into methylmercury (MeHg), a very toxic form that can bio-accumulate in fish and other marine species. MeHg biomagnifies in the food chain with the greater concentration appearing in larger predatory fish, by a factor sometimes greater than a million (Wilson, 2004). Impaired neurological development is the most well document health impact, and mostly a concern during the period of fetal and infant development (Yorifuji, Tsuda, Inoue, Takao, & Harada, 2011). The most common route of Hg exposure *in utero* results from maternal consumption of fish and shellfish during pregnancy, especially large predatory fish like shark and swordfish with greater concentrations of MeHg (Pichichero, Cernichiari, Lopreiato, & Treanor, 2002).

For non-occupational exposed individuals the major route of exposure to mercury is through ingestion in the diet with the consumption of fish and shellfish with high levels of MeHg (Bose-O'Reilly, McCarty, Steckling, & Lettmeier, 2010). Prenatal exposures to MeHg and Hg occurs when maternal exposures levels are also elevated due the fact that MeHg easily crosses the placental and blood-brain barrier. *In utero* exposure can occur even when the mother does not present symptoms of mercury toxicity (Harada, 1995). In the US, the leading source of MeHg exposure is through the dietary sources, mainly from seafood. Non-occupational exposure to MeHg from dietary sources has led the US Food and Drug Administration (FDA) and the US EPA to issue a joint federal advisory that recommends pregnant women to limit their total fish consumption intake to no more than 12 ounces of fish per week, or approximately two average meals (US Food and Drug Administration (FDA), 2014). Based on mercury's toxicity, widespread occurrence and potential for human exposure was ranked third on the 2013 ATSDR priority list of 275 hazardous substances (US EPA, 2014).

Inorganic Arsenic Toxicokinetics and Biomarkers

Arsenic in drinking water is predominantly present in its inorganic form as trivalent (As^{III}) or pentavalent (As^{V}). At least 90% of the ingested inorganic arsenic species are absorbed from the intestine and excreted through the urine (Watanabe & Hirano, 2013). Arsenic metabolism requires reduction and methylation of the trivalent (As^{III}) and pentavalent (As^{V}) species primarily through the arsenic methyltransferase enzyme (AS3MT) which utilizes S-adenosine methionine (SAM) as a methyl donor (Thomas, Waters, & Styblo, 2004). Inorganic forms of arsenic are metabolize via a one-carbon metabolism in the liver resulting in the production of two main metabolites: Monomethylarsonic acid (MMA) and Dimethylarsinic acid (DMA). Subsequently, metabolites are excreted through the urine. Urinary arsenic biomarkers have been shown to be relative stable over time. Total urinary inorganic arsenic has been determined to be highly specific and moderately sensitive at classifying exposures from drinking water in a Bangladeshi cohort (Kile et al., 2009). In the general US population, DMA has been shown to be the predominant metabolite for the detection of inorganic-related arsenic

species (K. L. Caldwell et al., 2008). Total urinary arsenic can be easily influenced by the consumption of seafood that might contain high levels of arsenosugars leading to the excretion of arsenobetaine and arsenocholine, metabolites of organic arsenic thought to be non-toxic (Sabbioni et al., 1991). It has been estimated that approximately 75% of the ingested dose of arsenic is excreted through the urine within 1 to 3 days, mostly as DMA (60-80%) and MMA (10-20%) (Vahter, 2009). However, both methylated arsenic and unmethylated arsenic are capable of crossing the placenta exposing the fetus *in utero* (Hall et al., 2007). Arsenic has also been observed to accumulate in tissue leading to an increase in oxidative stress and disruption of important immunological parameters such as T-cell production (S. Ahmed et al., 2011). Arsenic methylation has been shown to increase in efficiency during the first trimester of pregnancy with an increased excretion of DMA and a decrease in MMA proportions independent of nutritional status (Gardner et al., 2011). Gender differences have also been documented for urinary biomarkers of arsenic exposure. Namely, women have higher fractions of DMA and lower fractions of MMA in the urine compared to men (Hopenhayn-Rich, Biggs, Smith, Kalman, & Moore, 1996). Other potential factors that might also modify the metabolism of arsenic include: folate intake, age, smoking status and genetic polymorphisms (Lindberg et al., 2007).

Methylmercury Toxicokinetics and Biomarkers

Shortly after ingestion, around 95% of methylmercury (MeHg) is absorbed in the gastrointestinal tract and bound to hemoglobin and distributed across the body (Berglund et al., 2005). MeHg can also form low molecular weight complexes with thiol groups that can be easily transported across cell membranes including the placenta. The estimate half-life of MeHg in the human body ranges from 20-70 days (Clarkson & Magos, 2006). MeHg is a strong neurotoxicant, and therefore accurate biomarkers of exposure are essential to investigate dose-response relationships. Previous studies have demonstrated that both hair mercury and total mercury from red blood cells are suitable proxies for MeHg exposure (Berglund et al., 2005). Xenobiotic absorption of metals into the hair and nails occur through the incorporation into the matrix by the formation of keratinized tissue via blood flow during the growth of the nail. The use of toenail as the analytical

matrix is also advantageous, as it is less likely to be influences by external contamination, compared to fingernails or hair (Goullé et al., 2009). Toenail MeHg has been estimated to be an integrated measure of exposure anywhere between 100-140 days before collection, using a growth rate of 0.07 mm/day. Therefore toenail clipping of about 1 mm would reflect blood MeHg levels of about two weeks, 3-5 months prior to collection (Björkman et al., 2007). In a previous pilot study in New Hampshire it was demonstrated that weekly combined finfish and shellfish is moderately correlated with toenail mercury measurements (r=0.48). Therefore, mercury from toenail clippings can capture MeHg exposure from fish and shellfish exposure over an extended period of time retrospectively (Rees, Sturup, Chen, Folt, & Karagas, 2006).

Birth Weight and Gestational Age as Outcomes

Weight at birth is an important predictor of not only of infant mortality in the first year of life but also strongly associated with diseases well into adulthood like cardiovascular disease, type 2 diabetes, asthma and all-cause mortality (Risnes et al., 2011; Wilcox, 2001). Some of the potential mechanisms for decrease fetal growth include maternal malnutrition, elevated oxidative stress, altered balance of essential nutrients, reduced nutrient transport to the fetus and toxic environmental exposures (Ishimoto et al., 1997; Ronco, Arguello, Muñoz, Gras, & Llanos, 2005). Barker proposed that environmental factors and nutritional conditions during fetal development could potentially explain the association between birth weight and adult disease in the DOHaD hypothesis. The specific mechanisms in which birth weight influences disease later in life remains an active area of research and extremely important for the design of public health interventions (Wilcox, 2001). Birth weight is often used in epidemiological studies due to its strong association with disease susceptibility later in life and overall mortality. A newborn's low birth weight might be determined by two potential pathways: 1) the newborn was born pre-term or prematurely and 2) the newborn was born at-term but it is Small-for-Gestational-Age (SGA). Birth weight continues to be used as an important indicator of fetal health and a widely used parameter for studying factors that influence life course health (Lynch & Smith, 2005). This outcome is in line with the life course

epidemiological approach to chronic diseases that stipulates that fetal development is a crucial period for determining chronic diseases well into adulthood but could also be potentially modified by lifestyle (Gillman, 2002). Therefore, in many countries including the US interventions to increase birth weight are recommended as strategies to reduced mortality and overall morbidity (Kliegman, Rottman, & Behrman, 1990). In general, birth weight continues to be a widely used metric of fetal health and strongly associated with disease later in life. Its relationship to adult disease provides strong theoretical reasons for its implication in fetal programming as an intermediate phenotype of disease susceptibility hypothesized to occur through an epigenetic mechanism.

Birth gestational age is commonly used in epidemiological studies as it captures relevant neonatal health parameters such as fetal and organ maturity (Leviton, Blair, Dammann, & Allred, 2005). Furthermore, gestational age has also been shown to be a strong predictor of disease risk later in life including cardiovascular disease and neurological impairment (Gagliardi, 2015). The clinical parameter used for gestational age is prematurity, defined as <37 weeks of gestation, and preterm infants have higher rates of neonatal mortality and increase morbidity (Kramer et al., 2000). However, recent findings also suggest a risk gradient for gestational age beyond 37 weeks compared to full term infants (Boyle et al., 2012). For example, recent data indicates that infants born prior to 39 weeks of gestation are at an increased risk of morbidity, particularly respiratory complications (Tita et al., 2009). Reduced gestational age at birth has also been shown to be associated with neonatal DNA methylation patterns in cord-blood and and further hypothesized to mediate disease risk and susceptibility (Platt, 2014; J. W. Schroeder et al., 2011). As an epidemiological outcome, gestational age has remained to be a controversial topic in perinatal epidemiology, as statistical adjustment could potentially lead to bias on the assessment of birth outcomes and newborn's health (Wilcox, Weinberg, & Basso, 2011).

Arsenic Exposure, Birth Weight and Neurodevelopment

A few studies have looked at the relationship between arsenic exposure and birth weight, mostly concluding that arsenic exposure in utero reduces birth weight. In a small

cohort study in Bangladesh prenatal arsenic exposure was associated with a significant decrease in birth weight (Huyck et al., 2007). In an ecological cross-sectional study in Taiwan comparing two different areas that differ in groundwater arsenic concentration a significant decrease in birth weight was observed for the arsenic contaminated area (C.-Y. Yang et al., 2003). In a another large birth cohort study conducted in Bangladesh, arsenic exposure was associated with low birth weight only in the lower range of exposure ($\leq$100 μg/L), and no association was observed for higher exposure levels (Rahman et al., 2009). Another ecological study performed in Chile found a suggestive but non-statistically significant decrease in birth weight with increased levels of arsenic exposure during pregnancy (Hopenhayn et al., 2003). A positive association between arsenic in drinking water and birth weight was observed in an ecological cross-sectional study performed in Mongolia, contradictory to previous findings (S. L. Myers et al., 2010). However, this study lacked personal exposure measurements and thus exposure misclassification could have been present. More recently, a cohort study in New Hampshire identified the expression of the *AQP9* gene and placental expression of the *ENPP2* gene as potential mediators of the association between decrease in birth weight and arsenic exposure (Fei et al., 2013). Overall, most prospective epidemiological studies have documented a decrease in birth weight among arsenic exposed individuals. However, the potential mechanism for this association remains unclear.

Adverse neurological health outcomes have been documented in children relative to prenatal arsenic exposure hypothesized to be in part mediated through epigenetic alterations (Tyler & Allan, 2014). The first evidence of arsenic's neurotoxicity came from clinical arsenic intoxication cases in Japan in 1955 from the use of arsenic contaminated dried milk that caused more than 100 deaths. Survivors have been followed up in epidemiological studies and observed to have an increase risk for epilepsy, severe mental retardation (IQ<50) and hearing disabilities (Dakeishi, Murata, & Grandjean, 2006). The association between arsenic exposure and neurodevelopment has been previously evaluated in arsenic contaminated areas of Bangladesh, where a significant negative association between drinking water arsenic and intellectual function was observed (Wasserman et al., 2007). The association has also been observed cross-

sectionally in the US, with a robust decrease in IQ and other measurements of cognitive performance at low to moderate levels of exposure commonly found in the US (Wasserman et al., 2014). A recent review has summarized most of the current epidemiological studies of arsenic neurotoxicity concluding that only a few well designed studies have been able to correlate exposure with neurodevelopment and therefore firm conclusions can't be made with the current available evidence. It was also suggested that future epidemiological studies capture arsenic exposure in sensitive periods of development *in utero*, suggesting a possible epigenetic component for the observed relationship (Bellinger, 2013). Therefore, it is crucial to have accurate exposure classification prospectively during sensitive periods of development to reduce any possible bias from exposure misclassification.

Methylmercury Exposure and Neurodevelopment

A pollution incident in 1956 documented the dramatic effects of MeHg poisoning from the consumption of contaminated fish from the Minamata Bay, Japan. From this incident a wide range of negative neurological outcomes were observed among exposed individuals, coining the clinical presentation MeHg poisoning as "Minamata disease" (Eto, 2000). A second historic event occurred in 1971 when grain treated with a Hg-based fungicide was distributed to the rural Iraqi population to be used for planting. However, many individuals misused this grain for human consumption leading to more than 6500 individuals being hospitalized. In both Japan and Iraq individuals exposed to Hg suffered from mental retardation, cerebral palsy, deafness and blindness. This relationship was stronger and became more apparent among individuals exposed *in utero* (Counter & Buchanan, 2004). Consequently, the association between MeHg and infant neurodevelopment has been extensively evaluated in epidemiological studies. Two large prospective cohort studies in the Faroe Islands and the Seychelles have continued to evaluate neuropsychological performance in early childhood and *in utero* exposure to MeHg. The Faroe Islands cohort study has documented a positive association between exposure *in utero* and neurobehavioral deficits at levels of exposure commonly found in other populations (Grandjean et al., 1997). The Seychelles cohort has documented a

decrease in fine motor function skills for children exposed *in utero* but no relationship was observed for higher order cognitive function (G. J. Myers et al., 2003). A recent review of the literature concluded that low-level MeHg exposure was consistently associated with a decrease in early childhood neurocognitive function in most well designed epidemiological studies. However, questions about the dose-response relationship, mechanism of action and potential effect modifiers such as gender and timing of exposure remain active areas of research (Karagas et al., 2012). An important remark from most studies is that capturing the timing of exposure during fetal development is critical to assess the impact on neurodevelopment as well as the sensitivity of the neurological assessment (Davidson, Myers, Weiss, Shamlaye, & Cox, 2006; G. J. Myers et al., 2009).

The observed latent adverse neurological effects of both mercury and arsenic exposure, particularly *in utero*, could potentially share a common mode of action through epigenetic disruption during fetal development. However, no study to date has evaluated the joint epigenetic impact of both exposures.

Epigenetic effects of Inorganic Arsenic

Laboratory studies conducted to date indicate that exposure to arsenic can induce DNA methylation changes *in vitro* in different cell models (Davis, Uthus, & Finley, 2000; Mass & Wang, 1997; Reichard, Schnekenburger, & Puga, 2007). The lack of carcinogenic activity of arsenical compounds in animal models and its inability to cause point mutations in the genome has led to the suggestion that epigenetic mechanisms might be driving the observed relationship between arsenic exposure and several types of cancer (Anders et al., 2004). In fact, hypermethylation of *CpGs* of the following genes implicated in carcinogenesis *p53, p16, RASSF1A and PRSS3* have been documented with increasing exposure to arsenic in human subjects (Chanda et al., 2006; Marsit et al., 2005). Two Epigenome-Wide Association Studies (EWAs) have been conducted to date in cord-blood to investigate the epigenetic effect of prenatal arsenic exposure. Koestler *et al.* described that exposure to arsenic *in utero* measured in maternal urine can alter DNA methylation at two *CpG* sites: cg08884395 (*ESR1*) and cg27514608 (*PPARGC1A*). These

two genes have been previously identified in coding estrogen receptors and involved in the development of Bowen's disease (skin carcinoma), respectively (Koestler, Avissar-Whiting, Houseman, Karagas, & Marsit, 2013). However, this relationship did not remained significant after adjusting for multiple comparisons. Kile et al. reported an epigenome-wide association of prenatal arsenic exposure from the mother's drinking water and DNA methylation at several *CpG* sites after controlling for cellular heterogeneity. Specifically, only one *CpG* located in an island but not associated with any nearby gene remained significant after adjusting for multiple comparisons: cg00498691. However, in this study differential DNA methylation was associated with important biological pathways such as the maturity onset of diabetes, the hematopoietic cell lineage and the renin-angiotensin system, previously implicated in arsenic's toxicity (Kile et al., 2014). Both experimental and epidemiological studies suggest that arsenic exposure has the ability to disrupt the epigenome and might also play a role in arsenic induced carcinogenesis. However, the epigenetic effect of prenatal arsenicexposure has not been evaluated in placenta, artery or HUVEC tissues and the potential effect on fetal health remains unclear and warrants further investigation.

Epigenetic effects of Methylmercury

Currently, only a few human studies have evaluated the association between methylmercury exposure and DNA mehtylation. Using a cancer genome-wide methylation panel (Illumina GoldenGate Methylation Cancer Panel I) an increase in methylation was observed for the *GSTM1*gene for individuals exposed to high levels of methylmercury (Hanna et al., 2012). A second study of dental health professionals used a candidate gene approach for DNA methylation measurements and found that a promoter region of the selono-protein P plasma 1 (SEEP1) gene was hypomethylated with respect to mercury exposure (Goodrich, Basu, Franzblau, & Dolinoy, 2013). However, no overall association was observed between total exposure and global DNA methylation levels on both studies. Other types of epigenetic modifications like histone acetylation have also been studied in animal models. For example, in a mouse model exposure to methylmercury was associated with H3K27 trimethylation and H3 acetylation in brain

tissue, a potential pathway for altered neuroperformance (Onishchenko, Karpova, Sabri, Castrén, & Ceccatelli, 2008). The mechanism in which MeHg might disrupt the epigenome or causes neurodevelopmental damage remains to be defined. However, it has been hypothesized that due to the strong affinity of mercury to thiol groups and other selono groups MeHg might induce oxidative stress modulating its epigenetic toxicity at the cellular compartment (Farina, Rocha, & Aschner, 2010). Due to its ability to cross the placenta and the blood-brain barrier, mercury is thought to be the most toxic during fetal development (Vahter et al., 2000).

Specific Aims

The scope of this research is to investigate the impact of prenatal exposure to arsenic and mercury on the epigenome of newborns using two prospective birth cohorts. This goal was completed in three separate studies presented in the next three chapters. The first study evaluated the epigenome wide association in cord-blood of mercury exposure *in utero* and tested the interaction with prenatal arsenic exposure. The second study evaluated the epigenome wide association of prenatal arsenic exposure in three different tissues collected at birth: placenta, umbilical artery and umbilical vein endothelial cells (HUVEC). Finally, the last study evaluated the potential for CpG methylation levels in cord-blood to mediate the association between prenatal arsenic exposure and infant birth outcomes. The specific aims and hypotheses of this research are outlined below:

Specific Aim 1: To investigate the epigenome wide association of cord-blood and prenatal mercury exposure and evaluate its interaction with prenatal arsenic exposure using data collected in a prospective birth cohort recruited in New Hampshire.
Hypothesis 1a: Prenatal mercury exposure measured in maternal toenails will be associated with a dose-dependent differential DNA methylation in cord-blood collected at birth after adjusting for cellular heterogeneity.

Hypothesis 1b: Increasing prenatal levels of both maternal toenail mercury and urinary arsenic will be associated with differential DNA methylation in cord-blood independent of their main epigenetic effects after adjusting for cellular heterogeneity.

Specific Aim 2: To investigate the epigenome wide association of prenatal arsenic exposure in three different tissue types collected at birth: placenta, human umbilical vein endothelial cells (HUVEC) and umbilical artery using data collected in a prospective birth cohort recruited in Bangladesh.
Hypothesis 2: Prenatal arsenic exposure measured during the first trimester of pregnancy will be associated with a dose-dependent differential DNA methylation in three different tissues developed during fetal growth and collected at birth: placenta, HUVEC and umbilical artery.

Specific Aim 3: To evaluate the association between differentially methylated CpG loci in umbilical cord-blood with birth weight and gestational age using data collected in a prospective birth cohort recruited in Bangladesh.
Hypothesis 3a: Differentially methylated CpG loci in cord-blood will be significantly associated with lower birth weight.
Hypothesis 3b: Differentially methylated CpG loci in cord-blood will be significantly associated with decreased gestational age.

Chapter 2 - FIRST MANUSCRIPT

Differential DNA Methylation in Umbilical Cord Blood of Infants Exposed to Mercury and Arsenic *in utero*

Andres Cardenas[1]

Devin C. Koestler[2]

E. Andres Houseman[1]

Brian P. Jackson[3]

Molly L. Kile[1]

Margaret R. Karagas[3]

Carmen J. Marsit[3,4]

[1]School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, Corvallis, OR USA

[2]Department of Biostatistics, University of Kansas Medical Center, Kansas City, KS USA

[3]Department of Epidemiology, Geisel School of Medicine at Dartmouth College, Lebanon, NH USA

[4]Department of Pharmacology and Toxicology, Geisel School of Medicine at Dartmouth College, Hanover, NH USA

Epigenetics

Published in Volume 10, issue 6, pages 508-515

Abstract

Mercury and arsenic are known developmental toxicants. Prenatal exposures are associated with adverse childhood health outcomes that could be in part mediated by epigenetic alterations that may also contribute to altered immune profiles. In this study, we examined the association between prenatal mercury exposure on both DNA methylation and white blood cell composition of cord blood, and evaluated the interaction with prenatal arsenic exposure. A total of 138 mother-infant pairs with postpartum maternal toenail mercury, prenatal urinary arsenic concentrations, and newborn cord blood were assessed using the Illumina Infinium Methylation450K array. White blood cell composition was inferred from DNA methylation measurements. A doubling in toenail mercury concentration was associated with a 2.5% decrease (95% CI: 5.0%, 1.0%) in the estimated monocyte proportion. An increase of 3.5% (95% CI: 1.0, 7.0) in B-cell proportion was observed for females only. Among the top 100 CpGs associated with toenail mercury levels (ranked on *p*-value), there was a significant enrichment of loci located in North shore regions of CpG islands (P=0.049), and the majority of these loci were hypermethylated (85%). Among the top 100 CpGs for the interaction between arsenic and mercury, there was a greater than expected proportion of loci located in CpG islands (P=0.045) and in South shore regions (P=0.009) and all of these loci were hypermethylated. This work supports the hypothesis that mercury may be contributing to epigenetic variability and immune cell proportion changes, and suggests that *in utero* exposure to mercury and arsenic, even at low levels, may interact to impact the epigenome.

Introduction

Mercury (Hg) enters the environment from natural and anthropogenic sources and its concentration has increased at least three-fold in the ocean since the industrial revolution (Lamborg et al., 2014). The largest anthropogenic source of mercury emissions in the US comes from coal-burning power plants; others sources include byproducts of chlorine production, dental fillings using amalgams, thermometers, broken fluorescent bulbs and batteries (Tewalt, Bragg, & Finkelman, 2001). Mercury that is deposited in aquatic systems is biotransformed to methylmercury (MeHg) by anaerobic microbes. MeHgcan then bioaccumulate in species of high trophic level. For the general population, consumption of fish is the primary route of human exposure to MeHg. Since MeHg can cross the placenta, maternal consumption of fish, especially large predatory fish such as shark, tile fish and king mackerel, leads to fetal exposure *in utero* (Oken et al., 2012; Vahter et al., 2000).

Epidemiological studies have found that prenatal exposure to mercury *in utero* is associated with poor cognitive development and behavioral disorders in children (Boucher et al., 2014; Freire et al., 2010; Grandjean et al., 1997; Sagiv, Thurston, Bellinger, Amarasiriwardena, & Korrick, 2012). A recent review of the literature on low level prenatal MeHg exposure found consistent evidence for early childhood neurocognitive dysfunction, limited evidence for cardiovascular effects and possible associations with fetal growth among susceptible groups (Koestler et al., 2012). Furthermore, mercury exposure at low levels was found to be associated with subclinical signs of autoimmunity among women of reproductive age in the US population (Somers et al., 2015). The latency of health effects observed in prospective epidemiological studies of prenatal Hg exposures is suggestive of an epigenetic mode of action and emerging evidence suggests that placental epigenetic disruption can contribute to Hg neurodevelopmental toxicity (Maccani et al., 2015). However, no studies have investigated the effects of prenatal Hg exposure on DNA methylation in cord blood (Basu, Goodrich, & Head, 2014).

Additionally, there is evidence that MeHg exposure is correlated with other metals, such as arsenic, that are also known epigenetic toxicants. For instance,

epidemiological studies in pregnant women report a moderate positive correlation between biomarkers of Hg and As exposure (Miklavčič et al., 2013; Sanders et al., 2012). Few studies, however, have examined the health effects of co-exposures to toxicants A growing body of evidence indicates that prenatal exposure to arsenic may lead to neurotoxic effects and is most severe during brain development and fetal growth (Tolins, Ruchirawat, & Landrigan, 2014; Vahter, 2008). Emerging studies have provided evidence that prenatal exposure to As may disrupt the epigenome of newborns, a potential mechanism that might explain its latent health effects (Argos et al., 2015; Kile et al., 2012; Kile et al., 2014; Koestler et al., 2013).

The ability of some environmental exposures to disrupt epigenetic programming during fetal development provides a mechanism to link *in utero* exposure to toxicants and chronic disease (Marsit, 2015). The epigenome is thought to be most susceptible to environmental toxicants during embryogenesis, a developmental phase characterized by a rapid increase in cell division and epigenetic remodeling of the genomic landscape (Perera & Herbstman, 2011). *In vitro* studies in human neurons have shown that methylmercury is cytotoxic (Sanfeliu, Sebastià, & Kim, 2001). However, the exact mechanism for human toxicity remains ill defined (Farina, Rocha, & Aschner, 2011). Prenatal MeHg exposure has been hypothesized to disrupt calcium homeostasis, alter glutamate homeostasis and generate oxidative stress along with increased reactive oxygen species (Farina et al., 2011). Metal induced DNA methylation changes have been hypothesized to occur through oxidative stress and production of reactive oxygen species as the unifying process for DNA methylation disruption (Baccarelli & Bollati, 2009) Oxidative DNA damage has been shown to interfere with DNA methyltransferase activity resulting in altered methylation patterns of cytosine residues at CpG sites (Turk, Laayoun, Smith, & Weitzman, 1995). Both arsenic and mercury can alter DNA methyltransferase activity suggesting this as a potential mechanism for epigenetic disruption (Arita & Costa, 2009; Basu et al., 2013). Moreover, inorganic arsenic undergoes methylation to facilitate excretion, using S-adenosylmethionine (SAMe) as a methyl donor, the same methyl donor used in all cellular methylation reactions (Howe et al., 2014). Mercury has been shown to have strong inhibitory action in the methionine

synthase enzyme crucial for SAMe regeneration (Waly et al., 2004). Thus, the oxidative stress action of both arsenic and mercury along with the possibility of mercury exposure depleting levels of SAMe raises the question as to whether co-exposure to these two common environmental contaminants can disrupt the epigenome synergistically.

Studies that investigate the association between mercury exposure and DNA methylation are sparse and are limited to using a candidate gene approach (Goodrich et al., 2013; Ray, Yosim, & Fry, 2014). There is also limited data on the potential impact of prenatal Hg exposure on the distribution of leukocytes. Moreover, as even less is known about the effects of co-exposure to Hg and As on DNA methylation, we investigated both the main effect of prenatal exposure to Hg and its interaction with As exposure *in utero* and evaluated their ability to disrupt DNA methylation in cord blood. Furthermore, we estimated white blood cell composition using reference methylomes of isolated leukocyte subtypes to investigate the association with Hg and co-exposure to As.

Materials & Methods

Study Population

The study population consisted of the 138 initial participants enrolled in the ongoing New Hampshire Birth Cohort Study (NHBCS), which is focused on pregnant women from New Hampshire whose primary household drinking water source was a private well (Gilbert-Diamond et al., 2011). Eligibility criteria included English speaking, English literate, and mentally competent pregnant women 18-45 years of age. Subjects who changed their residence since their last menstrual period or whose home water supply was from a source other than from a private well were excluded from the study. Information about the mother and newborn were ascertained through questionnaires and review of the prenatal and delivery medical records. This study was approved by the Committee for the Protection of Human Subjects at Dartmouth College. All study participants provided written informed consent prior to the study.

Toenail Hg Assessment

At two weeks postpartum, an information packet was mailed to study participants requesting maternal toenail clipping samples within 8 weeks of birth; toenails were stored

at room temperature until analysis. After careful cleaning and washing to remove external contaminants, trace elements were quantified at the Trace Element Analysis Core (Dartmouth College, Hanover, New Hampshire, USA), using inductively coupled plasma mass spectrometry (ICP-MS). Toenails were acid digested with Optima nitric acid (Fisher Scientific, St Louis, Missouri, USA) at 105°C followed by the addition of hydrogen peroxide and further heating the dilution with deionized water. All sample preparation steps were recorded gravimetrically. As a quality control, each batch of analyses included six standard reference material (SRM) samples with known trace element content (GBW 07601, powdered human hair) and six analytical blanks, along with the study samples. The majority of participants (89%) had toenail Hg levels above the limit of detection (LOD) of 0.10 ng/g. Samples that were below the detection limit were replaced with the LOD divided by the square root of two.

Urinary Arsenic Assestment

Spot urine samples were collected at approximately 24-28 weeks gestation and analyzed for individual species of urinary arsenic using a high performance liquid chromatography (HPLC) inductively coupled plasma mass spectrometry (ICP-MS) system as previously described (Gilbert-Diamond et al., 2011; Koestler et al., 2013). This method determined five urinary arsenic species: arsenite (As^{III}), arsenate (As^{V}), dimethylarsinic acid (DMA^{V}) and monomethylarsonic acid (MMA^{V}). The detection limits ranged from 0.04 to 0.21 μg/L for the individual arsenic species. Values for samples with measurements below the LOD were imputed to be the metabolites' LOD divided by the square root of two. The proportion of samples below the LOD was 93% for As^{V}, 58% for As^{III}, 18% for MMA^{V} and 0% and DMA^{V}. Subsequently, total urinary arsenic concentration (Urinary-As) was calculated by summing inorganic arsenic (As^{III} and As^{V}) and the metabolic products MMA^{V} and DMA^{V}. We used total urinary-As as a measure of *in utero* exposure to arsenic because urinary arsenic levels have been shown to be a useful indicator of internal dose (Marchiset-Ferlay, Savanovitch, & Sauvant-Rochat, 2012). Urinary creatinine levels were measured using Cayman's creatinine assay kit and protocol to control for urine dilution.

DNA methylation assessment and Preprocessing

DNA was isolated from cord blood samples using DNeasy® blood & tissue kits (Qiagen, Valencia, CA) and bisulfite converted using the EZ DNA Methylation kit (Zymo, Irvine, CA). Samples were randomized across plates and subsequently subjected to epigenome-wide DNA methylation analysis using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA), which simultaneously profiles the methylation status for > 485,000 CpG sites at single nucleotide resolution. Microarrays were processed at the Biomedical Genomics Center at the University of Minnesota, following standard protocols. Data were extracted and processed from the raw methylation image files using functional normalization, a normalization method to correct for variations between arrays found in the *minfi* package of R (Fortin et al., 2014). This normalization method removes unwanted technical variation using the internal control probes as surrogates of batch effects.

The methylation status for each individual CpG locus was calculated as the ratio of fluorescent signals, referred to as a β-value. This is an interval scale quantity between zero and one interpreted as the fraction of DNA molecules whose target CpG is methylated. We further removed any residual plate/BeadChip effects using ComBat (Johnson, Li, & Rabinovic, 2007). All CpG loci on X and Y chromosomes were excluded from the analysis, to avoid gender-specific methylation bias. Furthermore, we excluded non-specific probes, cross-reactive probes and polymorphic CpGs (at $\geq$5% of the minor allele frequency) previously identified in the 450K array to avoid spurious associations (Y.-A. Chen et al., 2013). Finally, all CpG loci with high background signal that had non-significant detection p-values (P>0.01) indicating poor detection were excluded from the analysis, leaving a total of 348,569 autosomal CpG loci measured in 138 cord blood samples.

Statistical Analysis

DNA methylation β-values were logit transformed to M-values as previously recommended for statistical analysis (Du et al., 2010). To deconvolute the most prevalent sources of variability in DNA methylation across the array, we performed a Principal Component Analysis (PCA) on both the (a) normalized only and (b) normalized and

ComBat adjusted DNA methylation data. The top three principle components (PCs) estimated for each data set were then examined in terms of their association with sample plate using a series of linear regression models. The top PCs computed from the normalized only DNA methylation data were significantly associated with plate, however these associations were attenuated in ComBat adjusted DNA methylation data (Appendix A: Figure A.1, Figure A.2 and Table A.1) and indicate that ComBat was effective in reducing plate-associated variation in DNA methylation. Using the normalized and ComBat adjusted plate methylation data, we implemented a statistical deconvolution method to estimate the proportion of white blood cell composition in cord-blood (Houseman et al., 2012). This method estimates the relative proportion of six major cell types: CD8 T-cells, CD4 T-cells, natural killer (NK) cells, B-cells, monocytes and granulocytes. This was achieved by using discriminatory differentially methylated regions in flow-sorted purified leukocytes from six adult samples found in the *minfi* package of R (Reinius et al., 2012). Subsequently, multivariate linear regression models were used to evaluate the association between the estimated leukocyte proportion and log_2-transformed toenail Hg exposure.

We next implemented a locus-by-locus analysis aimed toward identifying differentially methylated CpG loci based on toenail Hg concentration using *limma* models (Smyth, 2005). Briefly, linear models were used to estimate the dose-response relationship between log_2-transformed values of toenail Hg concentration and M-values. A second model that included the main effect of maternal log_2-transformed toenail Hg and log_2-transformed urinary arsenic levels along with their multiplicative interaction was evaluated while also adjusting for urinary creatinine to account for urine dilution. All models were adjusted for maternal age at delivery, infant sex and the imputed white blood cell distribution from the Houseman projection (Houseman et al., 2012). Although our epigenome-wide approach was exploratory in nature, p-values were adjusted for multiple comparisons by comparing results to a Bonferroni corrected threshold for statistical significance. Lastly, using the top 100 differentially methylated CpG loci based on the lowest p-values, we compared their distribution based on CpG island location to

the rest of the array. All analyses were carried out using the R statistical package, version 3.1.1 (www.r-project.org/).

Results

A total of 138 mother-infant pairs were included in the analysis. The majority of mothers were white (98%), had some level of college education (89%), reported not smoking during pregnancy (83%) and were on average 31 years old at delivery. Maternal toenail Hg ranged from 0.001 to 1.44 μg/g with a median level of 0.07 μg/g. Total maternal urinary arsenic ranged from 0.34 to 17.9 μg/L with a median urinary arsenic concentration of 3.19 μg/L. Demographic characteristic are summarized in Table 2.1.

An overall decrease in the proportion of imputed monocytes was observed in relationship to increasing levels of Hg exposure (β=-2.5% ; 95% CI: -5.0, -1.0). After stratifying by gender, the association remained significant for females (β=-2.6% ; 95% CI: -5.0, -1.0) but not males (β=-1.9% ; 95% CI: -8.0, 4.0). An increase in the proportion of B-cells was also observed with increasing levels of Hg exposure in females only (β=3.5% ; 95% CI: 1.0, 7.0). Results for the relationship between $\log_2$-transformed toenail Hg and imputed white blood cells distribution in cord blood are summarized in Table 2.2. The interaction between Hg and As was not statistically significant for any of the estimated cell types (Appendix A, Table A.2).

In the locus-specific analysis, 11,327 (3.2%) of the 348,569 CpG loci were observed to be differentially methylated in relation to toenail Hg with a nominal *p*-value<0.05 and 9 CpG loci were differentially methylated with a nominal *p*-value <0.0001, after controlling for the imputed leukocyte distribution, infant sex and maternal age at enrollment. Results are shown in Figure 2.1 (A), which depicts the $-\log_{10}(p\text{-value})$ on the y-axis along with regression coefficients for $\log_2$-transformed Hg exposure on the x-axis for each individual CpG loci. However, no observed relationships remained significant after controlling for multiple comparisons using a Bonferroni threshold. Among the top 100 differentially methylated CpGs ranked on lowest *p*-value for Hg exposure, there was a greater than expected proportion of loci located in North shore regions of CpG islands (P=0.049), Figure 2.1 (B). Of these loci located in North shore

regions of CpG islands (n=20), the majority was observed to have an increase in methylation (85%).

For the multiplicative interaction between Hg and As in the locus-specific analysis, 15,620 (4.5%) of the 348,569 CpG loci were differentially methylated with a nominal *p*-value<0.05 and 37 CpG loci were differentially methylated with a nominal *p*-value<0.0001, after controlling for the imputed leukocyte distribution, infant sex, maternal age at delivery, urinary creatinine and the main effects of Hg and As. The multiplicative interaction results for all loci are shown in Figure 2.2 (A). The observed associations did not reach statistical significance using a Bonferroni correction. Among the top 100 differentially methylated CpGs ranked on lowest *p*-value for the multiplicative interaction between Hg and As, there was a greater than expected proportion of loci located in CpG islands (P=0.045) and in South shore regions of CpG islands (P=0.009), Figure 2.2 (B). All CpG loci located within CpG islands (n=43) showed increased methylation (100%). Similarly, among the top 100 loci located in South shore regions of CpG islands (n=19), all were observed to be hypermethylated (100%). The top 100 differentially methylated CpG loci based on toenail Hg exposure and interaction with urinary As are provided in Appendix A, Table A.4 and Table A.5, respectively.

Discussion

Our study provides evidence that Hg exposure *in utero* shifted the underlying leukocyte composition in cord blood leading to a decrease in the proportion of monocytes and an increase in the proportion B-cells in female infants. Prenatal Hg exposure, as well as, co-exposure to Hg and As may also have the potential to influence the epigenome of cord blood. The effect of Hg, as well as co-exposure to Hg and As, on the epigenome of cord blood was not statistically significant after accounting for the potential for false positives using a conservative Bonferroni correction. It is important to note, however, that this population was exposed to low Hg levels. For instance, the median concentration of toenail Hg in our study was 0.07 μg/g whereas toenail Hg levels measured in 27 healthy adults within the same study area of New Hampshire had median toenail Hg

concentrations of 0.16 µg/g, ranging from 0.04 to 1.15 µg/g (Rees et al., 2006). In another exposure study of 54 healthy Japanese pregnant women, the geometric mean of toenail Hg was 0.46 µg/g and ranged from 0.36 to 0.62 µg/g (Sakamoto et al., 2015). This could be due in part by dietary counseling pregnant women in the US receive which includes fish consumption advisories by health care providers and federal agencies to reduce prenatal MeHg exposure (Oken et al., 2003). Subsequently, the low levels of exposure found in this study during pregnancy are not surprising given the high level of literacy and education reported among the study population. Future studies may benefit from examining the epigenetic effects of prenatal Hg exposure in other populations with elevated exposure levels.

Several animal models have shown that mercury exposure is an *immunotoxicant* (Bose-O'Reilly et al., 2010). Particularly, studies have observed both monocyte and lymphocyte apoptosis *in vitro* (Shenker, Maserejian, Zhang, & McKinlay, 2008). This is consistent with our results where prenatal toenail-Hg exposure was associated with a decrease in the proportion of monocytes. Interestingly, the observed association was stronger for females where an increase in the estimated proportion of B-cells was also observed. The observed increase in B-cells is consistent with *in vivo* studies of mice exposed to Hg were both B-cell activation and autoantibody production have been documented (Abedi-Valugerdi, 2009). The sex differences observed in response to Hg exposure warrant further investigation, and indicate that Hg should be considered in studies examining outcomes including autoimmune disorders which are consistent with B-cell activation and have sexual dichotomy (S. A. Ahmed et al., 1999). The potential health implications for the observed decrease in monocytes remain to be determined.

While we know of no prior study that investigated the effect of prenatal Hg exposure on the epigenome, one study did examine the relationship between urinary Hg and DNA methylation changes in blood using a cancer-focused array (GoldenGate Cancer Panel I) among 58 women undergoing ovarian stimulation for *in vitro* fertilization. These researchers observed a significant increase in DNA methylation of promoter regions of the *GSTM1/5* genes among women with high urinary Hg levels (Hanna et al., 2012). A second study among dental health professionals using a candidate

gene approach to examine DNA methylation from buccal cells found that increasing levels of Hg exposure was associated with hypomethylation of CpG islands in the promoter region of the *SEPP1* gene (Goodrich et al., 2013). There is also evidence from animal models that epigenetic regulation of the Brain-Derived Neurotrophic Factor (*BDNF*) gene mediates gene suppression linked to behavioral changes in mice exposed to MeHg (Onishchenko et al., 2008).

To evaluate these previous findings in our cohort, we performed an individual CpG lookup based of the *SEPP1*, *GSTM5* and *BDNF* genes using a nominal p-value<0.05 for significance. The GSTM1 gene is polymorphic in the human population and therefore excluded a priori from our epigenome-wide analysis (Lee et al., 2010). Two loci (cg09606766; cg01636003) located in South shore regions of CpG islands of the *BDNF* gene were hypomethylated (P=0.004) while one loci (cg18595174) located in an open sea region was hypermethylated (P=0.015) with increasing toenail mercury concentrations. *For the SEPP1* gene, one loci (cg04502814) located in an open sea region increase in methylation with increasing toenail mercury concentrations (P=0.004). No other CpGs in the *SEPP1* or *GSTM5 gene*s were found to be differentially methylated. Results from the reverse CpG look up are summarized in Appendix A Table A.3.

The observed hypermethylation of CpG islands among the top ranked CpG loci for the interaction between As and Hg is of particular interest given that methylation of CpG islands in promoter regions has been established as a mechanism of gene silencing (Jaenisch & Bird, 2003). Furthermore, the observed hypermethylation of North shore regions with increasing Hg exposure levels and South shore regions in relation to Hg and As co-exposure could be biologically relevant as shore regions of CpG islands have been shown to be highly variable, play an important role in tissue differentiation and potentially drive pathogenesis (Irizarry et al., 2009; Rao et al., 2012).

While our study did not measure functional gene expression, some CpGs were located near genes that have been previously implicated in epigenetic mechanisms for disease. For example, among the top 100 differentially methylated CpGs in relationship to prenatal Hg levels, two loci (cg27458888; cg05881762) located in South shore regions of CpG islands of the Ubiquitin protein ligase E3A gene (*UBE3A*) were observed to be

hypermethylated. This gene is subject to methylation-dependent genomic imprinting and has been previously associated with Angelman syndrome, a neurodevelopmental disorder characterized by severe intellectual disability (Horsthemke & Wagstaff, 2008). Among the top 100 loci that were differentially methylated for the interaction between Hg and As, two loci (cg12419685; cg17250863) located in CpG islands of the Gamma-Glutamyltransferase 7 gene (*GGT7*) were hypermethylated. The *GGT7*gene is involved in the metabolism of glutathione (GSH) and adequate levels of GSH have been shown to protect against MeHg neurotoxicity *in vivo* (Ceccatelli, Daré, & Moors, 2010; Kaur, Aschner, & Syversen, 2006). In neurons, glutathione has been shown to be a physiological reservoir of glutamate, an important neurotransmitter involve in cognitive processes and previously implicated in MeHg-induced neurotoxicity (Farina et al., 2011; Koga et al., 2011).

There are a number of limitations to our current study. First, the study population was mostly white with a high level of education, which could limit the generalizability of our findings. The exposure levels to Hg found are relatively low but might reflect common levels of exposure among pregnant women in the US. Although the exposure assessment relies on toenail Hg concentrations, this is an objective biomarker of exposure previously validated and reflective of general patterns of dietary MeHg intake (Rees et al., 2006). Moreover, Hg concentrations in toenails at birth have been shown to capture MeHg levels approximately 5 months retroactively, reflecting MeHg exposure throughout the third trimester (Sakamoto et al., 2015). The leukocyte composition was projected from DNA methylation measurements taken from adult males and this might not accurately depict the white blood cell composition of newborns. Although DNA methylation changes in blood might serve as a biomarker of exposure and potentially disease, the biological relevance of the observed changes in cord blood is unknown and might not necessarily result in functional changes in gene expression. As demonstrated in previous studies, only a fraction of DNA methylation changes may result in differential gene expression (Argos et al., 2015; Steegenga et al., 2014). However, some consistent epigenetic variation between blood and brain tissue of humans have outlined the utility of whole blood in epidemiological studies (Davies et al., 2012). While many differentially

methylated loci had a relatively small *p*-value none met the Bonferroni correction criteria of statistical significance. Therefore, these results need to be replicated and confirmed in separate studies.

Conclusions

In conclusion, this study provides evidence that *in utero* exposure to mercury can affect leukocyte composition and may disrupt the epigenome even at low levels. Furthermore, exposure to both arsenic and mercury *in utero* may interact jointly to affect the epigenome by hypermethylating relevant CpG regions that have the potential to influence neurodevelopment and other childhood health outcomes. This would suggest that epigenomic alterations should also be considered in order to understand the toxic mechanisms of these exposures and their impact on children's health.

Acknowledgements

This work was supported by the US National Institute of Environmental Health Sciences (NIEHS) grants P42 ES016465(KC Donnelly Fellowship), P01 ES022832, US EPA grant RD83544201, and by a CTSA grant from NCATS awarded to the University of Kansas Medical Center for Frontiers: The Heartland Institute for Clinical and Translational Research # KL2TR000119. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, the US EPA does not endorse the purchase of any commercial products or services mentioned in the presentation.

Table 2.1. Selected maternal and infant characteristics and Hg and As concentrations

Characteristics	**Mean ± SD or n (%)**
No. Mother Infant Pairs	N=138
Maternal age at delivery (years)	31 ± 4.6
Gestational age (weeks)	39.3 ± 1.39
Pre-pregnancy BMI (kg/m^2)	24.3 ± 4.2
Infant Birth weight (grams)	3415 ± 562
Infant Sex	
Male	68 (49%)
Female	70 (51%)
Maternal Smoking During Pregnancy	
No	115 (83%)
Yes	7 (5%)
Unknown	16 (12%)
Maternal Race	
White	135 (98%)
Other	3 (2%)
Education	
Less than 11th grade	2 (2%)
High school graduate or equivalent	13 (9%)
Junior college graduate or some college	26 (19%)
College graduate	52 (38%)
Post-graduate education	28 (20%)
Unknown	17 (12%)
Exposure	**Median (min-max)**
Maternal Toenail Hg (µg/g)	0.07 (0.001, 1.44)
[a]Maternal Urinary Arsenic (µg/L)	3.19 (0.34-17.9)

[a]Maternal Urinary Arsenic=As^{III}+As^{V}+MMA+DMA

Table 2.2. Estimated change in the proportion of leukocyte composition in cord blood in relationship to log_2 toenail Hg concentration.

			Stratified by Sex			
	Overall Association		**Males**		**Females**	
Cell type	**% Change (95% CI)**	***P*-value**	**% Change (95% CI)**	***P*-value**	**% Change (95% CI)**	***P*-value**
CD8T	0.5 (-3.0, 4.0)	0.742	-3.5 (-11.0, 5.0)	0.414	1.4 (-2.0, 5.0)	0.395
CD4T	1.1 (-3.0, 6.0)	0.643	-6.6 (-18.0, 5.0)	0.274	2.3 (-3.0, 7.0)	0.367
NK cells	0.6 (-3.0, 5.0)	0.762	2.7 (-8.0, 1.3)	0.621	0.1 (-4.0, 5.0)	0.991
B-cells	2.3 (-1.0, 5.0)	0.116	-4.9 (-13.0, 3.0)	0.211	3.5 (1.0, 6.7)	0.027
Monocytes	-2.5 (-5.0, -1.0)	0.019	-1.9 (-8.0, 4.0)	0.508	-2.6 (-5.0, -1.0)	0.026
Granulocytes	-2.2 (-9.0, 5.0)	0.543	1.3 (-6.0, 3.3)	0.173	-5.0 (-1.3, 3.0)	0.196

Estimates from a multivariate linear regression model adjusted for infant sex and maternal age at delivery

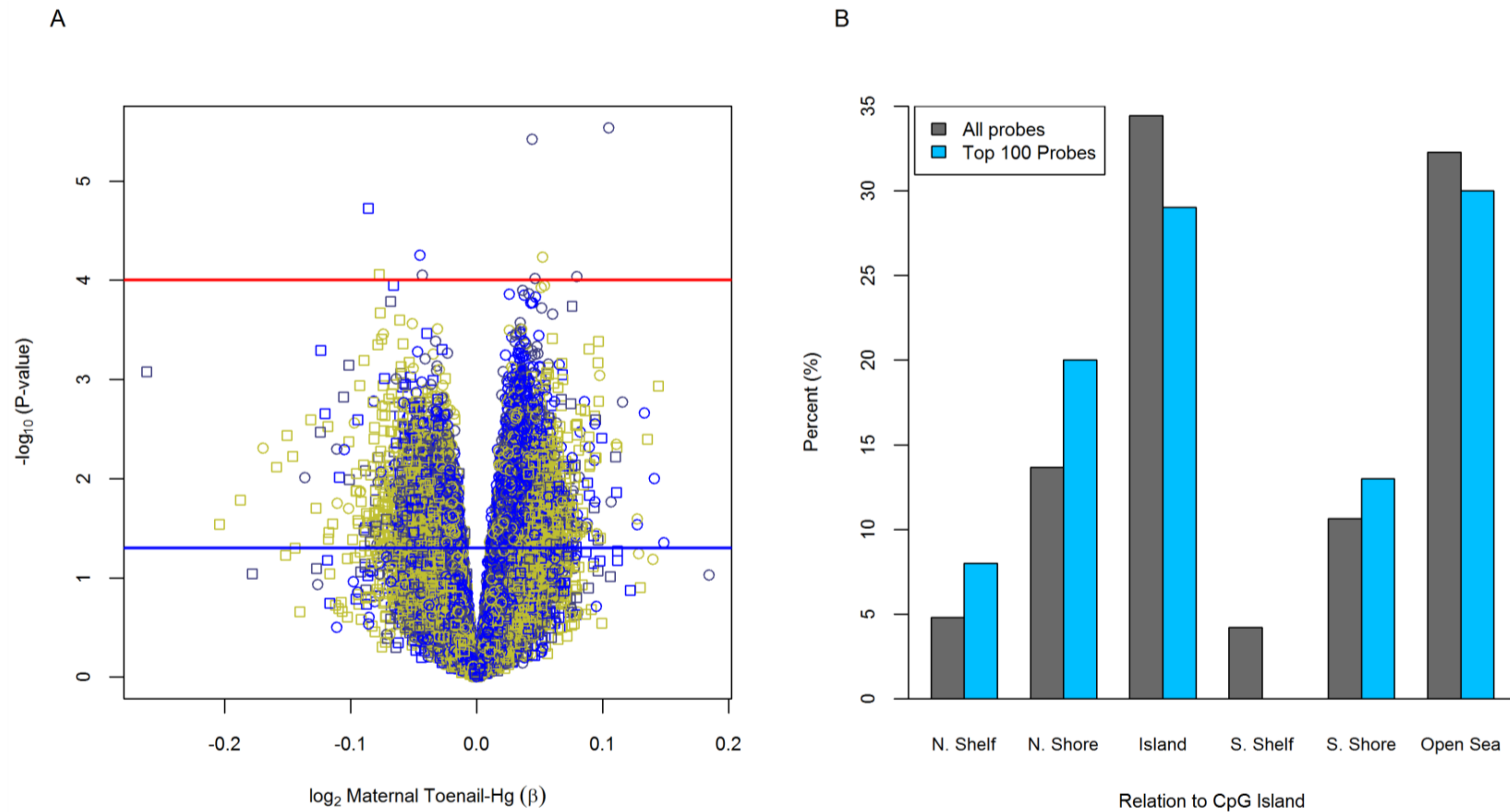

Figure 2.1. Differentially methylated loci based on maternal toenail-Hg.
Legend: **(A)** Volcano plot for the relationship between $\log_2$ toenail-Hg on DNA methylation at all 348,569 CpGs. Red and blue lines indicate $-\log_{10}(1\times10^{-4})$ and $-\log_{10}(0.05)$ p-values, respectively. Colors: yellow=CpG island, black=CpG Shore, blue=Shelfs and Open sea. Symbols: Circle=Infinium Type II, Square= Infinium Type I **(B)** Location of the top 100-CpGs on the basis of p-values compared to all CpGs on the methylation array. Among the top 100 CpGs with lowest p-value, there was a significant enrichment in the N. Shore regions of CpG islands (P=0.049).

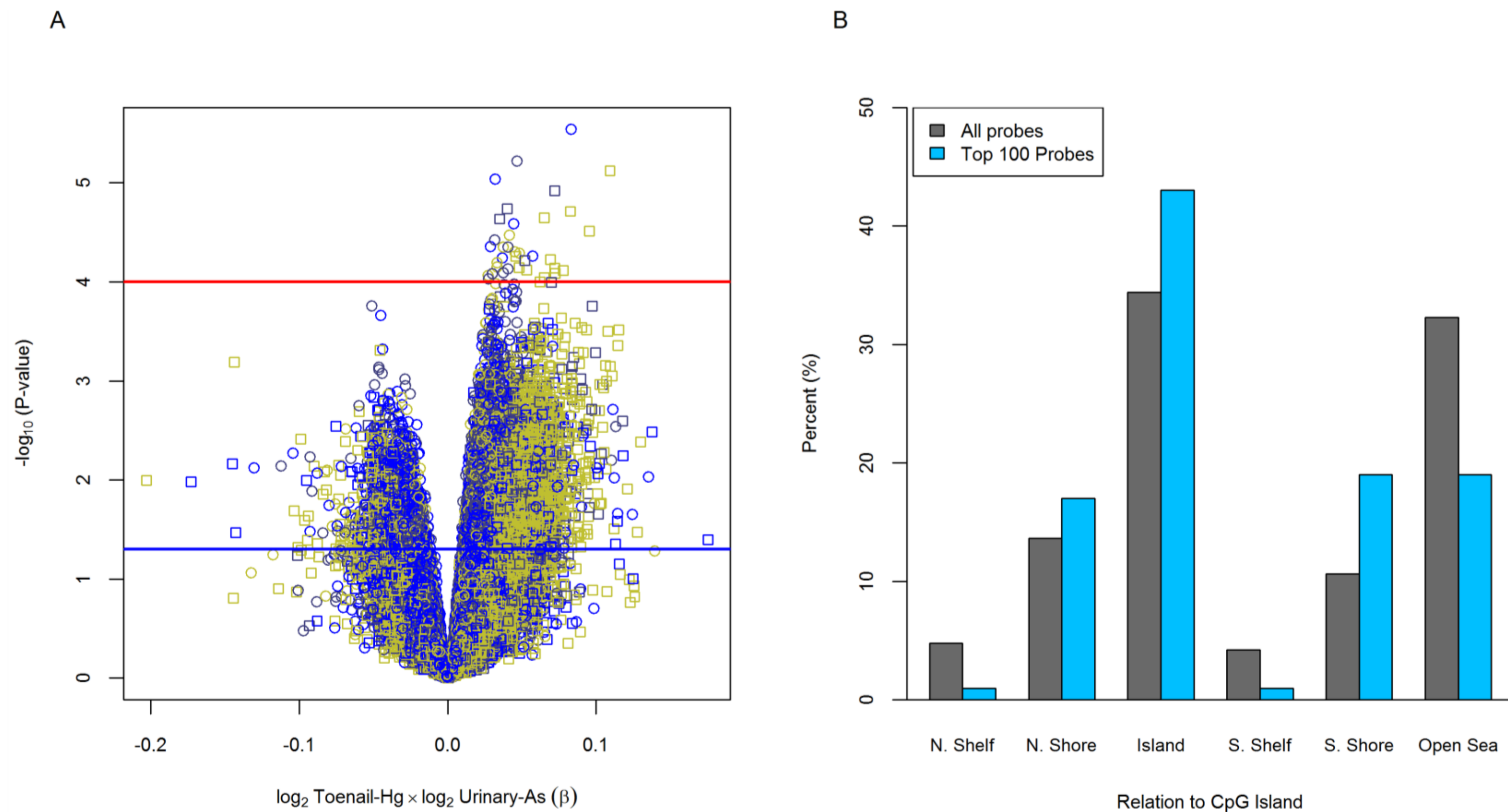

Figure 2.2. Differentially methylated loci based on the interaction between maternal toenail-Hg and urinary-As. Legend: (**A**) Volcano plot for the relationship between the multiplicative interaction of $\log_2$ toenail-Hg and $\log_2$ urinary-As on DNA methylation at all 348,569 CpGs. Red and blue lines indicate $-\log_{10}(1\times10^{-4})$ and $-\log_{10}(0.05)$ p-values, respectively. Colors: yellow=CpG island, black=CpG Shore, blue=Shelfs and Open sea. Symbols: Circle=Infinium Type II, Square= Infinium Type I (**B**) Location of the top 100-CpGs on the basis of p-values compared to all CpGs on the methylation array. Among the top 100 CpGs with lowest p-value, there was a significant enrichment of CpG islands (P=0.045) and S. Shores regions (P=0.009)

.

Chapter 3 SECOND MANUSCRIPT

Arsenic Exposure *in utero*: Epigenome-Wide Associations in Placenta, Artery and Human Umbilical Vein Endothelial Cells

Andres Cardenas[1]

E. Andres Houseman[1]

Andrea A. Baccarelli[2]

Quazi Quamruzzaman[3]

Mahmuder Rahman[3]

Golam Mostofa[3]

Robert O. Wright[4]

David C. Christiani[2]

Molly L. Kile[1]

[1] School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, Corvallis, OR USA

[2] Harvard T.H. Chan School of Public Health; Boston, MA USA

[3] Dhaka Community Hospital; Dhaka, Bangladesh

[4] Preventative Medicine and Pediatrics; Mt Sinai School of Medicine; New York, NY USA

Epigenetics

Under review

Abstract

Exposure to arsenic early in life has been associated with increased risk of several chronic diseases and is believed to alter epigenetic programming *in utero*. In the present study, we evaluate the epigenome-wide association of arsenic exposure in utero and DNA methylation in placenta (n=37), umbilical artery (n=45) and human umbilical vein endothelial cells (HUVEC) (n=52) in a birth cohort using the Infinium HumanMethylation450 BeadChip array. Unadjusted and cell mixture adjusted associations for each tissue were examined along with enrichment analyses relative to CpG island location and omnibus permutation tests of association among biological pathways. One CpG in artery (cg26587014) and four CpGs in placenta (cg12825509; cg20554753; cg23439277; cg21055948) reached a Bonferroni adjusted level of significance. Several CpGs were differentially methylated in artery and placenta when controlling the false discovery rate (q-value<0.05), but none in HUVEC. Enrichment of hypomethylated CpG islands was observed for artery while hypermethylation of open sea regions were present in placenta relative to prenatal arsenic exposure. The melanogenesis pathway was differentially methylated in artery (Max F $P<0.001$), placenta (Max F $P<0.001$) and HUVEC (Max F $P=0.002$). Similarly, the insulin signaling pathway was differentially methylated in artery (Max F $P=0.02$), placenta (Max F $P=0.03$) and HUVEC (Max F $P=0.002$). Our results show that prenatal arsenic exposure can alter DNA methylation in artery and placenta but not in HUVEC. Further studies are needed to determine if these alterations in DNA methylation mediate the effect of prenatal arsenic exposure and health outcomes later in life.

Introduction

Over 200 million individuals worldwide are exposed to elevated levels of inorganic arsenic. This is a public health concern because arsenic is a known human carcinogen and chronic exposure is associated with the development of skin, lung, bladder, kidney, liver and potentially prostate cancer (Argos, 2015). Particularly, early life exposure to arsenic has been associated with the development of many latent health effects including carcinogenesis (Bailey KA, 2015). Human ecological studies from the Antofagasta region of Chile have associated prenatal and early childhood exposure to arsenic from contaminated municipal water with increased risk of lung and bladder cancer later in life (Steinmaus et al., 2014). Increased mortality from acute myocardial infarction and cancers of the bladder, kidney, lung, and liver have also been reported from this population decades after the exposure declined (Yuan et al., 2010; Yuan et al., 2007).

Exposure to arsenic has also been associated with increased prevalence of type 2 diabetes mellitus and cardiovascular disease in epidemiological studies (Moon, Guallar, & Navas-Acien, 2012; Navas-Acien et al., 2006; Pan et al., 2013). Animal models support the involvement of transplacental arsenic exposure in the development and progression of atherosclerosis, consistent with human studies linking early life exposure and cardiovascular disease (Farzan, Karagas, & Chen, 2013; Srivastava, D'Souza, Sen, & States, 2007). Emerging evidence also indicates that exposure to arsenic can disrupt normal immune function and *in utero* exposure can increase the susceptibility and severity of infections later in life (Cardenas, Smit, et al., 2015; Farzan, Korrick, et al., 2013; Rahman et al., 2010). Furthermore, arsenic exposure during fetal development has been associated with growth restrictions and adverse perinatal health outcomes such as low birth weight, still births, infant mortality, and preterm births (Quansah et al., 2015). Lastly, latent adverse neurological health outcomes have also been documented with maternal exposure to arsenic during pregnancy (Hamadani et al., 2011; Tanaka, Tsukuma, & Oshima, 2010).

The exact molecular mechanisms of the toxicological effects attributed to arsenic exposure remains elusive and no single mechanism has been identified in the

development of arsenic associated diseases and the observed latency of health effects (Bailey & Fry, 2014). However, the latency of health effects documented in epidemiological studies and animal models along with the observed susceptibility of prenatal exposures are suggestive of an epigenetic mode of action. Fetal programming events involving DNA methylation occur at critical windows of fetal development in a cell-specific manner shown to be sensitive to environmental exposures (Marsit, 2015). Experimental evidence from animal models demonstrate that transplacental exposure to arsenic leads to epigenetic alterations, changes in gene expression and increased incidence of tumors in the offspring (Waalkes, Liu, & Diwan, 2007; Waalkes, Qu, Tokar, Kissling, & Dixon, 2014). Therefore, it is postulated that epigenomic regulation including, but not limited to, DNA methylation is a potential mechanism of arsenic induced carcinogenesis and latent disease risk (Bailey KA, 2015; Reichard & Puga, 2010; Xie et al., 2007).

Several human studies have evaluated the impact of prenatal arsenic exposure on the cord blood and whole blood epigenome, recently reviewed by M. Argos (Argos, 2015). Among these epidemiological studies evaluating cord blood or whole blood DNA methylation no common loci has been identified to be differentially methylated across studies. However, significant DNA methylation disruption of unique loci along with enrichment of key regulatory CpG regions has been documented across different study populations (Argos et al., 2015; Broberg et al., 2014; Kile et al., 2014; Koestler et al., 2013; X. Liu et al., 2014; Rojas et al., 2015; Seow et al., 2014). Besides studies that examined cord and whole blood epigenome, only one study to date has evaluated the association between arsenic exposure and CpG methylation of a target tissue by evaluating DNA methylation in urothelial carcinoma samples (T.-Y. Yang et al., 2014). This study found significant hypermethylation of loci in urothelial carcinoma samples of patients historically exposed to arsenic from drinking water compared to unexposed cases.

Epigenetic reprogramming during fetal development resulting from transplacental exposure is one of the main hypothesized mechanisms of arsenic’s associated-disease (Bailey KA, 2015). To further our understanding of how prenatal arsenic exposure could

alter epigenetic programming it is important to evaluate its effect on different tissues with diverse cellular compositions. Evaluating if exposure to arsenic in utero alters DNA methylation of different tissues could yield insights into the etiology of toxicant-mediated disease and epigenetic modifications of relevant tissues with specific biological functions. Subsequently, we examined the association between maternal drinking water arsenic as a proxy of transplacental exposure during fetal development and the epigenome of placenta, umbilical artery and Human Umbilical Vein Endothelial Cells (HUVEC) from a birth cohort conducted in arsenic affected regions of Bangladesh.

Materials & Methods

Study Population

This pilot study was nested within an established birth cohort recruited in Bangladesh (2007-2011) and designed to characterize the potential epigenetic disruption associated with arsenic exposure during pregnancy in different tissues collected at birth. A more detailed explanation of the full birth cohort has been published previously (Kile et al., 2014). Briefly, pregnant women with $\leq$ 16 weeks of gestation confirmed by ultrasound were enrolled in a prospective reproductive birth cohort in Bangladesh. Trained health care workers at community health clinics in Sirajdikhan and Birahimpur recruited pregnant women 18 years of age or older that used a tube-well as their primary drinking water source, planned to live at their current residency during the duration of the pregnancy and received prenatal health care at Dhaka Community Hospital (DCH) or affiliated community clinic. Study participants agreed to deliver at DCH or at home with a DCH trained midwife. Informed consent was obtained from all participants prior to enrollment. All participants were provided with prenatal care and prenatal vitamins offered by DCH. This study was approved by the Human Research Committees at the Harvard School of Public Health, Oregon State University and Dhaka Community Hospital Trust.

Three distinct tissues were collected at the time of delivery including: artery from the umbilical cord, placenta, and endothelial cells isolated from the umbilical vein. Since the goal of this pilot study was to examine the potential exposure-response relationship

between arsenic and DNA methylation, specimens were selected based on maternal drinking water arsenic concentrations at study enrollment to cover a wide range of exposures (<1-510 µg/L). A total of 37 placenta samples, 45 artery samples and 52 HUVEC samples were included in the final analysis.

Drinking Water Arsenic

Water samples were collected from the tube-well identified by participants as their main source of drinking water at the time of their enrollment into the study as previously described (Kile et al., 2014). Briefly, water samples were collected in a 50-mL polypropylene tubes (BD Falcon, BD Bioscience, Bedford, MA), preserved with Reagent Grade nitric acid (Merck, Germany) to a pH<2 and stored at room temperature. Arsenic concentrations were measured by inductively coupled plasma-mass spectrometry (ICP-MS) using the US EPA method 200.8 to determined metals in water (Environmental Laboratory Services, North Syracuse, New York) (Creed, Brockhoff, & Martin, 1994). Average percent recovery for Arsenic from plasmaCal multi-element QC standard #1 solution (SCP Science) was 102% ± 7%. The limit of detection (LOD) for arsenic in drinking water was 1 µg/L.

Tissue Collection: Umbilical Artery, Placenta & HUVEC

Trained medical workers present at delivery collected a sample of the umbilical cord and placenta immediately after the delivery was completed. Using sterile techniques, approximately 5-7 cm of umbilical vein was dissected out of fresh umbilical cord and rinsed with phosphate buffered saline solution to remove external contamination. The vein lumen was then bisected and the interior cavity was flushed with approximately 100 mL of phosphate buffered solution to remove blood. The interior lumen wall was gently rubbed using a sterile cytology brush to collected endothelial cells. The cytology brush was then vortexed in 1 mL of cell lysis solution (Qiagen) to transfer the cells. The cell lysis solution was then stored at 4 ^{0}C. Samples were shipped to Harvard School of Public Health where the DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen) following manufacturer's instructions.

Approximately 1 cm of umbilical cord artery was dissected out of fresh umbilical cord, the exterior of the artery was scraped to remove Wharton's Jelly, and rinsed with

phosphate buffered saline solution to remove blood. The arterial cross section was placed in 2 mL of RNAse later and stored at -20 ^{0}C. Samples were shipped to Harvard School of Public Health on dry ice. The artery sample was then minced using a sterile scalpel and added to Maxwell Cell DNA Purification kits (Promega) with an additional 20 μL of Proteinase K (Qiagen). Samples were allowed to sit for 30 minutes before being extracted using the Maxwell 16 Research instrument following manufacturer's instructions.

For placenta samples, a one centimeter tissue plug was excised from fresh placenta. The tissue plug was placed into a sterile vial and covered with Tissue-Tek O.C.T. gel (Electron Microscopy Sciences) and frozen at -20 ^{0}C. Samples were then shipped to Harvard School of Public Health on dry ice. Next, approximately 10 grams of placenta tissue was removed from the plug and minced using a sterile scalpel and added to Maxwell Cell DNA Purification kits (Promega) with an additional 20 μL of Proteinase K (Qiagen). Samples were allowed to sit for 30 minutes before being extracted using the Maxwell 16 Research instrument following manufacturer's instructions.

DNA Methylation Assessment and Quality Control

DNA was shipped to the University of Minnesota's Biomedical Genomic Center that quantified DNA methylation using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA) following standard manufacturer's protocols. The HumanMethylation450 BeadChip measures DNA methylation at > 485,000 CpG sites at single nucleotide resolution, covering 99% of the RefSeq genes.

Tissues were analyzed in separate plates and randomly allocated within different chips. Data were obtained and processed from raw methylation image files and normalized using internal control probes via the functional normalization method with two principal components to account for technical variation between samples using the *minfi* package of R (Fortin et al., 2014). DNA methylation was estimated at each CpG as the fraction of DNA molecules whose target CpG loci is methylated and referred to as β-values. Measurements at CpG loci on X and Y chromosomes were excluded from the analysis to avoid gender-specific methylation bias. Previously identified non-specific and cross-reactive probes within the array along with polymorphic CpG loci ($\geq$ 5% of the minor allele frequency) were removed for the analysis (Y.-A. Chen et al., 2013).

Furthermore, a detection *P-value* was computed for all CpGs and probes with non-significant detection (*P*>0.01) in greater than 10% of the samples were removed from the analysis. After quality control, the total number of autosomal CpGs left in the analysis were 374,320 loci for artery, 365,994 loci for placenta and 347,650 loci in HUVEC samples. Finally, a beta-mixture quantile intra sample normalization procedure (BMIQ) was further applied to the data to reduce the potential bias that can arise from type 2 probes as previously described (Teschendorff et al., 2013). Strip plots and signal intensities of control probes were visually examined for bisulfite conversion, probe hybridization and single base extension. Density plots for the β-values were examined for all samples at each normalization step described above.

Statistical Analysis

Unadjusted and Cell-adjusted Locus-by-Locus Analysis

We first aimed to identify differentially methylated CpG loci in relationship to prenatal arsenic exposure from maternal drinking water. Maternal arsenic concentration in water was right skewed and subsequently $\log_2$-tranformed. In order to evaluate linear associations between prenatal exposure to arsenic and differentially methylated CpG loci, β-values were logit-transformed to M-values previously described to be more appropriate for differential analysis of DNA methylation (Du et al., 2010). In the locus-by-locus approach, two different but complementary methodologies were implemented. First, the linear association between individual CpG methylation on the M-value scale and $\log_2$-transformed arsenic was evaluated adjusting for infant sex using the *limma* function found in the *minfi* package of R. Second, due to the lack of reference methylomes of isolated cell types in placenta, artery or HUVEC, a novel reference-free method of adjusting for cellular heterogeneity was implemented using the *RefFreeEWAS* package of R. The reference-free method is an extension of the original Houseman method that utilizes a deconvolution approach similar to surrogate variable analysis (SVA) that is data driven to identify latent variables or dimensions as surrogates of cellular composition (Houseman, Molitor, & Marsit, 2014). Using this method, the sex adjusted linear association between individual CpG methylation on the β-value scale and $\log_2$-transformed maternal drinking water arsenic was evaluated using 1000 bootstrap samples

for estimating the standard errors of association in placenta, umbilical artery and HUVEC. Results from the unadjusted *limma* models and the reference-free cell mixture adjusted analyses were compared within tissues and across tissues. Enrichment analyses for the distribution of CpGs relative to CpG island location of the top differentially methylated loci based on a q-value<0.05 or a nominal $P<1x10^{-4}$, were compared to the distribution of probes on the rest of the array.

Biological Pathway Analysis

Omnibus permutation based tests and *p*-values were obtained by mapping subsets of CpGs to their associated genes in specific KEGG biological pathways. Gene sets were compiled from the Kyoto Encyclopedia of Genes and Genomes (KEGG) corresponding to specific biological pathways using the Entrez IDs matched to KEGG biological pathways using the Bioconductor library *org.Hs.eg.db*. The permutation distribution was obtained from unadjusted cell mixture models by permuting the exposure with respect to measured DNA methylation over subgroups of CpGs defined by biological pathways (1,000 permutations). Pathway based associations of DNA methylation with prenatal arsenic exposure as a continuous variable were summarized using a maximum nominal F-statistics *p*-value (akin to a minimum *p*-value) and an average nominal F-statistic *p*-value. The maximum and minimum F-statistic *p*-values are better suited for detecting a small number of strong associations and a large number of more variable associations, respectively, as previously described (Kile et al., 2014).

All statistical analyses were performed using the R statistical package version 3.2.0 (http://www.R-project.org).

Results

The sample size varied by tissue type with a maximum of 52 samples present for HUVEC followed by 45 samples in umbilical artery and 37 placenta samples. Arsenic concentration in maternal drinking water at study enrollment ranged from below the detection limit of <1µg/L to 510 µg/L with a mean exposure concentration of 63.7 µg/L. Selected sample characteristics are shown in Table 3.1.

Arterial Tissue

Locus-by-Locus Analysis: In the analysis that was unadjusted for cellular composition, one CpG loci (cg26587014) located in chromosome 19 and not annotated to any gene was differentially methylated in arterial tissue in relation to arsenic exposure using a Bonferroni threshold for statistical significance ($P<1.33\times10^{-7}$). Controlling for the false discovery rate at 5% (q-value<0.05) revealed 2,105 CpGs that were differentially methylated relative to log_2-transformed maternal drinking water arsenic. However, after adjusting for cellular composition using the Houseman reference-free method, no loci reached a Bonferroni corrected level of significance or a q-value<0.05. Unadjusted and adjusted results are shown in Figure 3.1 (A) and Figure 3.1 (B), respectively. The top 10 differentially methylated loci ranked on lowest *P*-value are summarized in Appendix B Table B.1 and Table B.2 for unadjusted and cell mixture adjusted analyses, respectively. In unadjusted analyses, differentially methylated loci with a q-value<0.05 were disproportionately located in CpG islands (54%) compared to the distribution of CpG island probes in the rest of array (33%) ($P<1x10^{-4}$), Appendix B, Figure B.1 (A). The majority of unadjusted hypomethylated loci with a q-value<0.05 were located in CpG islands (83%), Figure 3.1 (C). After adjusting for cellular heterogeneity, a similar enrichment of hypomethylated loci in CpG islands was observed among top loci having a nominal p-value$<1x10^{-4}$, Appendix B Figure B.1 (B).

Biological Pathway Analysis: Omnibus permutation based tests revealed significant associations between *in utero* exposure to arsenic and epigenetic disruption of KEGG biological pathways in arterial tissue (Mean F-statistics P=0.009 and maximum F-statistic P=0.006). Pathways that were observed to have the strongest association based on the lowest mean F-static level of significance (P=0.004) were: maturity onset of diabetes of the young (hsa04950), primary immunodeficiency (hsa05340), ABC transporters (hsa02010), allograft rejection (hsa05330) and vibrio cholerae infection (hsa05110). Differentially methylated pathways observed to have a strong association using a maximum F-statistic level of significance (P<0.001) included: the Hedgehog signaling pathway (hsa04340), Melanogenesis (hsa04916), Wnt signaling pathway (hsa04310), Basal cell carcinoma (hsa05217), DNA replication (hsa03030) and the p53 signaling

pathway (hsa04115). The summary for all associations between maternal drinking water arsenic and epigenetic disruption of KEGG biological pathways are shown in Table 3.2.

Placenta Tissue

Locus-by-Locus Analysis: In the analyses that were unadjusted for cellular composition, no single CpG loci reached Bonferroni adjusted significance in placenta ($P<1.37\times10^{-7}$). However, two CpG loci (cg26390526; cg03857453) annotated to the Epidermal Filaggrin gene (*FLG*) and the nuclear receptor subfamily 3, group C, member 1 glucocorticoid receptor gene (*NR3C1*) were hypermethylated relative to maternal drinking water arsenic after controlling for the false-discovery rate (q-value<0.05). In analyses that adjusted for cell mixture in the placenta, four CpGs reached Bonferroni adjusted significance: cg12825509 (*TRA2B* gene), cg20554753, cg23439277 (*PLCE1* gene) and cg21055948 (*CD36* gene). Moreover, analyses adjusted for cellular heterogeneity revealed 518 CpG loci that were differentially methylated after controlling for the false discovery rate (q-value<0.05). Unadjusted and cell mixture adjusted results are shown in Figure 3.2 (A) and Figure 3.2 (B), respectively. The top 10 differentially methylated loci ranked on lowest *P*-value are summarized in Appendix B Table B.3 and Table B.4 for unadjusted and cell mixture adjusted results, respectively. For the top unadjusted differentially methylated loci with a nominal $P<1\text{x}10^{-4}$ a disproportionate amount of CpGs were located within open sea regions of CpG islands (76%) compared with the distribution of open sea loci in the rest of array (33%) ($P<1\text{x}10^{-4}$), Appendix B: Figure B.2 (A). Among these loci, the great majority of hypermethylated CpGs were located within open sea regions (89%) relative to CpG islands, Figure 3.2 (C). For the cell mixture adjusted analyses a similar enrichment of hypermethylated loci in open sea regions was observed for loci with a q-value<0.05, Appendix B: Figure B.2 (B).

Biological Pathway Analysis: Omnibus permutation based tests indicated that exposure to arsenic *in utero* disrupts methylation of a small number of CpGs within KEGG biological pathways in the placenta tissue (Omnibus maximum F-statistic $P=0.004$). However, a marginal association among KEGG biological pathways and arsenic exposure was observed using an omnibus Mean F-statistic test for association ($P=0.108$). KEGG biological pathways that were differentially methylated in relationship to arsenic

exposure with a maximum F-statistic $P<1\times10^{-3}$ included: Melanogenesis (hsa04916), Neuroactive ligand-receptor interaction (hsa04080), Calcium signaling pathway (hsa04020), GnRH signaling pathway (hsa04912), Dilated cardiomyopathy (hsa05414), Gap junction (hsa04540), Vasopressin-regulated water reabsorption (hsa04962), Vascular smooth muscle contraction (hsa04270), Oocyte meiosis (hsa04114), Vibrio cholerae infection (hsa05110), Progesterone-mediated oocyte maturation (hsa04914) and the Peroxisome pathway (hsa04146). Several other pathways were significantly associated with arsenic exposure using a maximum F-statistic $P<0.05$ and summarized in Table 3.2.

Umbilical Vein Endothelial Cells (HUVEC)

Locus-by-Locus Analysis: In both unadjusted and cell mixture adjusted analyses no single CpG loci was associated with arsenic exposure at a Bonferroni corrected level of significance ($P<1.44\times10^{-7}$) or a q-value<0.05 after controlling for the false discovery rate, Figure 3.3 (A) and Figure 3.3 (B). Among the top 31 CpG loci with a nominal $P<1\times10^{-4}$ no significant differences were present for the occurrence of top loci relative to CpG island location compared to the rest of the array for unadjusted analyses, Figure 3.3 (C). The top 10 differentially methylated loci ranked by lowest *p*-value are summarized in Appendix B, Table B.5 and Table B.6, for unadjusted and cell mixture adjusted results, respectively.

Biological Pathway Analysis: Omnibus permutation tests for association among KEGG biological pathways indicated that arsenic exposure was not significantly associated with a large number of changes in DNA methylation across pathways in HUVEC (Mean F-statistic $P=0.129$) and the presence of a small number of strong associations was borderline significant (Max F-statistic $P=0.06$). Few individual biological pathways reached statistical significance using a maximum F-statistic level of significance. The top differentially methylated biological pathways (maximum F-statistic $P=0.002$) in HUVEC included: Melanogenesis (hsa04916), Wnt signaling pathway (hsa04310), Basal cell carcinoma (hsa05217) and the Insulin signaling pathway (hsa04910), all KEGG biological pathway based associations are summarized in supplementary Table 3.2.

The overlap among CpGs within each tissue for unadjusted and cell mixture adjusted analyses using the top 100 differentially methylated CpGs was 26 loci in artery,

21 loci in placenta and 33 loci for HUVEC, Appendix B, Figure B.3. Among the top 100 differentially methylated loci, only one CpG (cg21002651) located within the body of the *CASP1* gene was differentially methylated across two tissues in unadjusted analyses. This loci was hypomethylated in placenta (β=-0.20, P=5.73x10^{-6}) but hypermethylated in HUVEC (β=0.20, P=1.29x10^{-4}) in relationship to maternal drinking water arsenic. No other CpGs overlapped in unadjusted or adjusted analyses.

Discussion

Our study provides evidence that *in utero* exposure to arsenic can disrupt DNA methylation of artery and placenta tissues but the association with umbilical vein endothelial vein cells was marginal. However, the association of prenatal arsenic exposure on the epigenome on artery and placenta depended on the cell mixture adjustment. For instance, the association in artery was attenuated after controlling for cellular heterogeneity but strengthened in placenta. *In utero* exposure to arsenic was also associated with DNA methylation levels of key biological pathways across tissues providing new insights into the potential etiology of arsenic-mediated diseases with a plausible epigenetic reprogramming component.

In normal tissue, the majority of CpG islands remain unmethylated and methylation of CpG islands located within promoter regions of genes is an established mechanism of gene silencing (Jones, 2012). Conversely, we observed an enrichment of hypomethylated loci in CpG islands relative to prenatal arsenic exposure. This is of particular interest because both animal and human studies have demonstrated that DNA hypomethylation occurs in atherosclerotic lesions and that hypomethylation of CpG islands is observed broadly in human atherosclerotic arteries (Castillo-Díaz, Garay-Sevilla, Hernández-González, Solís-Martínez, & Zaina, 2010; Hiltunen & Ylä-Herttuala, 2003). Similarly, in a case-control study of patients suffering from giant cell arteritis CpG hypomethylation of arterial disease tissue was observed compared to controls further supporting the involvement of epigenetic hypomethylation in arterial disease pathogenesis (Coit, De Lott, Nan, Elner, & Sawalha, 2015). In animal models, *in utero* arsenic exposure has been shown to induce the early onset of atherosclerosis along with

epidemiological studies linking early life exposure with cardiovascular disease (Farzan, Karagas, et al., 2013; Srivastava et al., 2007). Therefore, it is possible that the observed hypomethylation of influential genomic regions such as CpG islands could play a role in the development of arsenic-associated cardiovascular disease, particularly atherosclerosis of arterial tissue. Another early observation from epigenetic cancer studies was the global hypomethylation of tumor samples compared to normal tissue mainly at repetitive genomic elements and that hypomethylation of these regions can lead to hypermethylation of tumor suppressor genes (Ehrlich, 2009). Along with this observation, previous studies of arsenic exposure have characterized hypermethylation of the promoter region of the p53 gene a mechanisms hypothesized to contribute to the carcinogenesis of arsenical compounds (Intarasunanont et al., 2012; Mass & Wang, 1997). Consistent with these reports, our gene set analysis shows that CpG methylation within the p53 signaling pathway is associated with arsenic exposure during pregnancy suggesting that artery could be a target tissue for the epigenetic toxicity of arsenic, and potentially involved in carcinogenesis.

The placenta is an important regulator of fetal development and intrauterine growth that plays a crucial role mediating the maternal and fetal environment. Furthermore, the placenta is a unique epigenetic target organ as the majority of imprinted genes in animal models are both expressed and imprinted in the placenta and hypothesized to contribute to fetal neurodevelopment (Lesseur, Paquette, & Marsit, 2014; Tunster, Jensen, & John, 2013). Interestingly, in unadjusted analyses a CpG located in the body of the glucocorticoid receptor gene (*NR3C1*) was significantly hypermethylated in the placenta relative to prenatal arsenic exposure. Previous studies have shown that hypermethylation of the *NR3C1* gene influences cortisol response, infant behavior and self-regulation (Conradt et al., 2015; Oberlander et al., 2008). In line our finding a recent experimental study demonstrated that exposure to arsenic *in utero* lowers the activity of the glucocorticoid receptor pathway and these changes were maintained into adolescence of the mouse model (K. E. Caldwell, Labrecque, Solomon, Ali, & Allan, 2015). The placenta has also been characterized as one of the hypomethylated tissues as LINE-1 elements have lower levels of methylation when compared to other tissues. Furthermore,

it has been shown that normal human placenta contains partially methylated domains (37%) with the ability to suppress genes and impact tissue-specific functions independent of the tissue of origin (D. I. Schroeder et al., 2013). The observed hypermethylation of open sea regions relative to CpG island location could have implications for normal methylation of LINE-1 elements and partially methylated domains, potentially affecting normal biological function and development of the placenta.

A few KEGG biological pathways were differentially methylated in relation to maternal drinking water arsenic in all three tissues. Namely, DNA methylation of the melanogenesis pathway was strongly associated with exposure to arsenic in artery, placenta and HUVEC. An early clinical symptom of arsenicosis include the appearance of hyperpigmentation changes of the skin in the trunk, neck and chest regions of the body eventually progressing to the palmar and plantar regions and eventually leading to hyperkeratosis (Wahed, Rahman, & Vahter, 2006). Consistent with the differential methylation of this pathway in our data, arsenic-associated alterations in DNA methylation of leukocytes has been previously associated with increased risk of developing skin lesions (Pilsner et al., 2009). The most common arsenic induced skin cancers are Bowen's disease, Basal cell carcinoma (BCC) and Squamous cell carcinoma (SCC) (H.-S. Yu, Liao, & Chai, 2006). Our biological pathway analysis revealed significant association between maternal drinking water arsenic and the BCC biological pathway in artery and HUVEC but not placenta. Therefore, as previously suggested future studies should consider epigenomic regulation as a potential mechanism of arsenic mediated skin cancer and lesions (Seow et al., 2014). Lastly, the insulin signaling pathway was observed to be differentially methylated across all three tissues with respect to arsenic exposure. Exposure to arsenic has been consistently associated with Type 2 diabetes and insulin resistance in both animal models and epidemiological studies (Pan et al., 2013). Previous studies have documented the epigenetic disruption of several genes involved in the development of diabetes and insulin resistance for individuals chronically exposed to arsenic (Bailey et al., 2013).

It is crucial to highlight that HUVEC is a homogenous tissue in terms of cellular composition and was not significantly disrupted in the locus-by-locus analysis and

marginally associated among some biological pathways. However, artery and placenta, both representing a diverse mixture of cell types, were observed to be differentially methylated relative to prenatal arsenic exposure. DNA methylation is cell specific playing a key role in tissue differentiation and lineage commitment making this process particularly vulnerable to environmental stimuli and exposures during fetal development. Epidemiologic studies often rely on preserved samples and have limited fresh tissue availability making the sorting or isolation of target cell types not feasible. Therefore, future experimental studies should evaluate the development of cancer stem cells (CSCs) and alterations to the immune function as factors or intermediary mechanisms of the observed epigenetic perturbations, as others have also suggested (Bailey KA, 2015). Moreover, the interaction between prenatal arsenic exposure and other transplacental contaminants should also be considered, as prenatal exposure to arsenic has been previously shown to interact with other prenatal exposures such as mercury (Cardenas, Koestler, et al., 2015).

One of the major strengths for the present study is the epigenome-wide analysis of three different tissues yielding insights for the potential biological impact of arsenic exposure during fetal development. Also, the prospective design of this birth cohort along with the exposure assessment early during pregnancy are important qualities that strengthens the temporality of the epigenetic perturbations reported. Although the present study relies on a single water sample during early pregnancy and exposure misclassification cannot be ruled out, previous studies in rural Bangladesh have demonstrated that drinking water arsenic exposure is relatively constant and correlated with biomarkers of internal doses, such as urine and toenails (Kile et al., 2009; Kile et al., 2005) and that arsenic readily crosses the placenta (Concha, Vogler, Lezcano, Nermell, & Vahter, 1998). Additionally, the availability of umbilical samples at birth provides one of the few opportunities for examining epigenetic programming in cardiovascular target tissue in a non-invasive and feasible manner. There are a number of important limitations to our current study including the relatively small sample size and the lack of validation using a complementary DNA methylation method due to sample availability. The lack of reference methylomes for placenta, artery and HUVEC also raise an important challenge

when interpreting the observed epigenetic perturbations in tissues that might represent a mixture of cell types such as artery or placenta. However, we implemented a complementary bioinformatics method to adjust for cellular heterogeneity to identify potential perturbations in loci hypothesized to be associated with methylation levels independent of cellular heterogeneity. Finally, gene expression was not measured and the observed changes in DNA methylation need to be further confirmed and evaluated. Particularly significant association between DNA methylation among KEGG biological pathways might not result in functional gene expression alterations.

Conclusions

In conclusion, we show that prenatal arsenic exposure is associated with altered DNA methylation of umbilical artery and placenta tissue but evidence of an association for HUVEC is limited. Furthermore, we present evidence of DNA methylation disruption of key biological pathways across different tissues holding the potential to mediate arsenic-associated diseases previously described from exposures *in utero*.

Acknowledgements

This work was supported by the US National Institute of Environmental Health Sciences (NIEHS) grants R01 ES015533, K01 ES017800, R01 ES016454, P30 ES000210, and P30 ES000002.

Table 3.1. Sample characteristics for the 52 mother-infant pairs eligible for the analysis.

Sample characteristics	**Mean±SD**	**Range**
Drinking water arsenic at recruitment (μg/L)	63.7±116.5	<1 - 510
Gestational age at recruitment (weeks)	12.2±2.5	6 - 16
Gestational age at delivery (weeks)	37.6±2.1	33 - 41
Birth weight (grams)	2923±372	2080 - 4050
Gender	**N (%)**	
Male	33 (63.5 %)	
Female	19 (36.5%)	
Number of samples available by tissue	**n**	**CpG loci analyzed**
HUVEC	52	347,650
Artery	45	374,320
Placenta	37	365,994

Table 3.2. Omnibus association of DNA methylation of KEGG biological pathways and prenatal arsenic exposure in artery, placenta and HUVEC.

		Artery Tissue		**Placenta Tissue**		**HUVEC Tissue**	
KEGG Entry ID	**KEGG Biological Pathway Name**	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***
Any KEGG Pathway	---	**0.009**	**0.006**	**0.108**	**0.004**	**0.129**	**0.06**
Any Gene within Pathway	---	**0.009**	**0.002**	**0.111**	**0.008**	**0.126**	**0.138**
hsa02010	ABC transporters	0.004	0.021	0.109	0.008	0.135	0.321
hsa03010	Ribosome	0.017	0.019	0.042	0.116	0.096	0.203
hsa03018	RNA degradation	0.023	0.031	0.061	0.062	0.093	0.261
hsa03020	RNA polymerase	0.007	0.002	0.117	0.143	0.14	0.499
hsa03022	Basal transcription factors	0.02	0.089	0.094	0.255	0.081	0.069
hsa03030	DNA replication	0.01	<0.001	0.069	0.039	0.11	0.538
hsa03040	Spliceosome	0.023	0.017	0.033	0.082	0.086	0.005
hsa03050	Proteasome	0.016	0.08	0.078	0.114	0.086	0.806
hsa03060	Protein export	0.027	0.047	0.107	0.197	0.07	0.143
hsa03320	PPAR signaling pathway	0.007	0.054	0.188	0.228	0.141	0.188
hsa03410	Base excision repair	0.008	0.047	0.131	0.078	0.126	0.729
hsa03420	Nucleotide excision repair	0.013	0.002	0.081	0.076	0.132	0.713
hsa03430	Mismatch repair	0.018	0.045	0.145	0.277	0.101	0.629
hsa03440	Homologous recombination	0.02	0.046	0.086	0.314	0.124	0.449
hsa03450	Non-homologous end-joining	0.027	0.072	0.114	0.024	0.048	0.14
hsa04010	MAPK signaling pathway	0.008	0.029	0.123	0.066	0.148	0.834
hsa04012	ErbB signaling pathway	0.01	0.025	0.113	0.023	0.136	0.76
hsa04020	Calcium signaling pathway	0.005	0.012	0.168	<0.001	0.178	0.801
hsa04060	Cytokine-cytokine receptor interaction	0.012	0.053	0.179	0.365	0.138	0.387
hsa04062	Chemokine signaling pathway	0.01	0.064	0.126	0.001	0.171	0.121
hsa04080	Neuroactive ligand-receptor interaction	0.006	0.002	0.182	<0.001	0.147	0.689
hsa04110	Cell cycle	0.015	0.024	0.044	0.239	0.136	0.455
hsa04114	Oocyte meiosis	0.015	0.101	0.116	<0.001	0.092	0.444
hsa04115	p53 signaling pathway	0.01	<0.001	0.13	0.196	0.159	0.288
hsa04120	Ubiquitin mediated proteolysis	0.015	0.059	0.061	0.015	0.097	0.163
hsa04130	SNARE interactions in vesicular transport	0.01	0.009	0.149	0.095	0.152	0.326
hsa04140	Regulation of autophagy	0.011	0.06	0.043	0.01	0.053	0.473

Table 3.2. Omnibus association of DNA methylation of KEGG biological pathways and prenatal arsenic exposure in artery, placenta and HUVEC (Continued)

		Artery Tissue		**Placenta Tissue**		**HUVEC Tissue**	
KEGG Entry ID	**KEGG Biological Pathway Name**	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***
hsa04142	Lysosome	0.007	0.026	0.151	0.011	0.128	0.223
hsa04144	Endocytosis	0.008	0.046	0.086	0.019	0.135	0.127
hsa04146	Peroxisome	0.01	0.074	0.064	<0.001	0.083	0.292
hsa04150	mTOR signaling pathway	0.006	0.121	0.054	0.012	0.129	0.262
hsa04210	Apoptosis	0.011	0.007	0.161	0.226	0.108	0.655
hsa04260	Cardiac muscle contraction	0.011	0.002	0.091	0.016	0.107	0.484
hsa04270	Vascular smooth muscle contraction	0.008	0.005	0.124	<0.001	0.148	0.307
hsa04310	Wnt signaling pathway	0.009	<0.001	0.148	0.156	0.154	0.002
hsa04320	Dorso-ventral axis formation	0.008	0.009	0.203	0.006	0.183	0.599
hsa04330	Notch signaling pathway	0.005	0.01	0.093	0.052	0.196	0.521
hsa04340	Hedgehog signaling pathway	0.008	<0.001	0.144	0.197	0.114	0.004
hsa04350	TGF-beta signaling pathway	0.015	0.097	0.122	0.016	0.101	0.081
hsa04360	Axon guidance	0.009	0.021	0.181	0.05	0.149	0.786
hsa04370	VEGF signaling pathway	0.009	0.058	0.157	0.018	0.161	0.697
hsa04510	Focal adhesion	0.008	0.006	0.151	0.057	0.182	0.267
hsa04512	ECM-receptor interaction	0.006	0.096	0.166	0.208	0.228	0.582
hsa04514	Cell adhesion molecules (CAMs)	0.012	0.031	0.348	0.019	0.168	0.518
hsa04520	Adherens junction	0.016	0.031	0.094	0.012	0.157	0.123
hsa04530	Tight junction	0.012	0.004	0.089	0.052	0.152	0.18
hsa04540	Gap junction	0.011	0.068	0.13	<0.001	0.118	0.869
hsa04610	Complement and coagulation cascades	0.016	0.164	0.383	0.312	0.2	0.763
hsa04612	Antigen processing and presentation	0.005	0.042	0.139	0.036	0.126	0.425
hsa04614	Renin-angiotensin system	0.007	0.141	0.308	0.037	0.186	0.51
hsa04620	Toll-like receptor signaling pathway	0.016	0.037	0.09	0.012	0.101	0.634
hsa04621	NOD-like receptor signaling pathway	0.02	0.146	0.118	0.007	0.136	0.067
hsa04622	RIG-I-like receptor signaling pathway	0.013	0.001	0.065	0.063	0.1	0.658
hsa04623	Cytosolic DNA-sensing pathway	0.012	0.005	0.092	0.085	0.122	0.036
hsa04630	Jak-STAT signaling pathway	0.013	0.061	0.133	0.243	0.124	0.153
hsa04640	Hematopoietic cell lineage	0.011	0.062	0.332	0.04	0.194	0.586
hsa04650	Natural killer cell mediated cytotoxicity	0.007	0.089	0.257	0.022	0.15	0.454
hsa04660	T cell receptor signaling pathway	0.009	0.06	0.182	0.029	0.16	0.451

Table 3.2. Omnibus association of DNA methylation of KEGG biological pathways and prenatal arsenic exposure in artery, placenta and HUVEC (Continued)

		Artery Tissue		Placenta Tissue		HUVEC Tissue	
KEGG Entry ID	**KEGG Biological Pathway Name**	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***
hsa04662	B cell receptor signaling pathway	0.012	0.054	0.195	0.022	0.152	0.687
hsa04664	Fc epsilon RI signaling pathway	0.011	0.05	0.223	0.017	0.176	0.643
hsa04666	Fc gamma R-mediated phagocytosis	0.012	0.065	0.187	0.024	0.145	0.215
hsa04670	Leukocyte transendothelial migration	0.008	0.003	0.155	0.051	0.181	0.138
hsa04672	Intestinal immune network for IgA production	0.006	0.146	0.301	0.103	0.123	0.531
hsa04710	Circadian rhythm - mammal	0.01	0.076	0.094	0.051	0.21	0.69
hsa04720	Long-term potentiation	0.009	0.029	0.123	0.024	0.112	0.505
hsa04722	Neurotrophin signaling pathway	0.009	0.039	0.131	0.028	0.132	0.486
hsa04730	Long-term depression	0.013	0.051	0.243	0.017	0.132	0.502
hsa04740	Olfactory transduction	0.033	0.077	0.15	0.129	0.123	0.055
hsa04742	Taste transduction	0.007	0.045	0.234	0.054	0.299	0.949
hsa04810	Regulation of actin cytoskeleton	0.008	0.005	0.173	0.038	0.162	0.779
hsa04910	Insulin signaling pathway	0.005	0.02	0.09	0.03	0.189	0.002
hsa04912	GnRH signaling pathway	0.008	0.061	0.147	<0.001	0.141	0.618
hsa04914	Progesterone-mediated oocyte maturation	0.01	0.021	0.08	<0.001	0.15	0.826
hsa04916	Melanogenesis	0.008	<0.001	0.219	<0.001	0.17	0.002
hsa04920	Adipocytokine signaling pathway	0.009	0.003	0.181	0.54	0.162	0.114
hsa04930	Type II diabetes mellitus	0.006	0.038	0.045	0.024	0.252	0.383
hsa04940	Type I diabetes mellitus	0.005	0.023	0.167	0.047	0.105	0.416
hsa04950	Maturity onset diabetes of the young	0.004	0.007	0.084	0.025	0.237	0.096
hsa04960	Aldosterone-regulated sodium reabsorption	0.008	0.087	0.1	0.014	0.126	0.462
hsa04962	Vasopressin-regulated water reabsorption	0.017	0.089	0.124	<0.001	0.133	0.449
hsa04964	Proximal tubule bicarbonate reclamation	0.005	0.054	0.167	0.274	0.11	0.561
hsa04966	Collecting duct acid secretion	0.005	0.01	0.123	0.017	0.186	0.582
hsa05010	Alzheimer's disease	0.012	0.004	0.065	0.019	0.096	0.32
hsa05012	Parkinson's disease	0.015	0.003	0.089	0.01	0.086	0.209
hsa05014	Amyotrophic lateral sclerosis (ALS)	0.01	0.002	0.052	0.204	0.09	0.09
hsa05016	Huntington's disease	0.013	0.036	0.055	0.013	0.098	0.33
hsa05020	Prion diseases	0.01	0.006	0.235	0.005	0.11	0.168

Table 3.2. Omnibus association of DNA methylation for KEGG biological pathways and prenatal arsenic exposure in artery, placenta and HUVEC (Continued)

		Artery Tissue		Placenta Tissue		HUVEC Tissue	
KEGG Entry ID	**KEGG Biological Pathway Name**	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***	**Mean F-statistic *P***	**Max F-statistic *P***
hsa05110	Vibrio cholerae infection	0.004	0.075	0.106	<0.001	0.179	0.773
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	0.015	0.033	0.063	0.03	0.103	0.753
hsa05130	Pathogenic Escherichia coli infection	0.01	0.191	0.063	0.068	0.105	0.09
hsa05140	Leishmania infection	0.011	0.307	0.31	0.014	0.107	0.648
hsa05200	Pathways in cancer	0.01	0.001	0.113	0.005	0.13	0.011
hsa05210	Colorectal cancer	0.017	0.012	0.069	0.012	0.121	0.091
hsa05211	Renal cell carcinoma	0.01	0.036	0.056	0.002	0.098	0.645
hsa05212	Pancreatic cancer	0.017	0.009	0.091	0.01	0.144	0.113
hsa05213	Endometrial cancer	0.017	0.046	0.081	0.013	0.12	0.082
hsa05214	Glioma	0.01	0.051	0.126	0.014	0.149	0.334
hsa05215	Prostate cancer	0.017	0.065	0.131	0.012	0.127	0.12
hsa05216	Thyroid cancer	0.008	0.016	0.092	0.009	0.135	0.045
hsa05217	Basal cell carcinoma	0.009	<0.001	0.119	0.084	0.156	0.002
hsa05218	Melanoma	0.016	0.017	0.154	0.016	0.12	0.332
hsa05219	Bladder cancer	0.009	0.028	0.079	0.004	0.226	0.692
hsa05220	Chronic myeloid leukemia	0.01	0.03	0.071	0.01	0.156	0.729
hsa05221	Acute myeloid leukemia	0.01	0.041	0.175	0.013	0.126	0.115
hsa05222	Small cell lung cancer	0.019	0.008	0.095	0.105	0.112	0.34
hsa05223	Non-small cell lung cancer	0.014	0.048	0.094	0.01	0.15	0.468
hsa05310	Asthma	0.006	0.247	0.804	0.372	0.118	0.36
hsa05320	Autoimmune thyroid disease	0.005	0.045	0.575	0.492	0.123	0.522
hsa05322	Systemic lupus erythematosus	0.022	0.01	0.134	0.051	0.099	0.29
hsa05330	Allograft rejection	0.004	0.042	0.673	0.422	0.117	0.482
hsa05332	Graft-versus-host disease	0.006	0.042	0.528	0.43	0.118	0.475
hsa05340	Primary immunodeficiency	0.004	0.015	0.089	0.02	0.197	0.18
hsa05410	Hypertrophic cardiomyopathy (HCM)	0.009	0.001	0.125	0.253	0.158	0.566
hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	0.019	0.151	0.113	0.274	0.139	0.134
hsa05414	Dilated cardiomyopathy	0.008	0.003	0.137	<0.001	0.154	0.609
hsa05416	Viral myocarditis	0.009	0.026	0.255	0.044	0.147	0.78

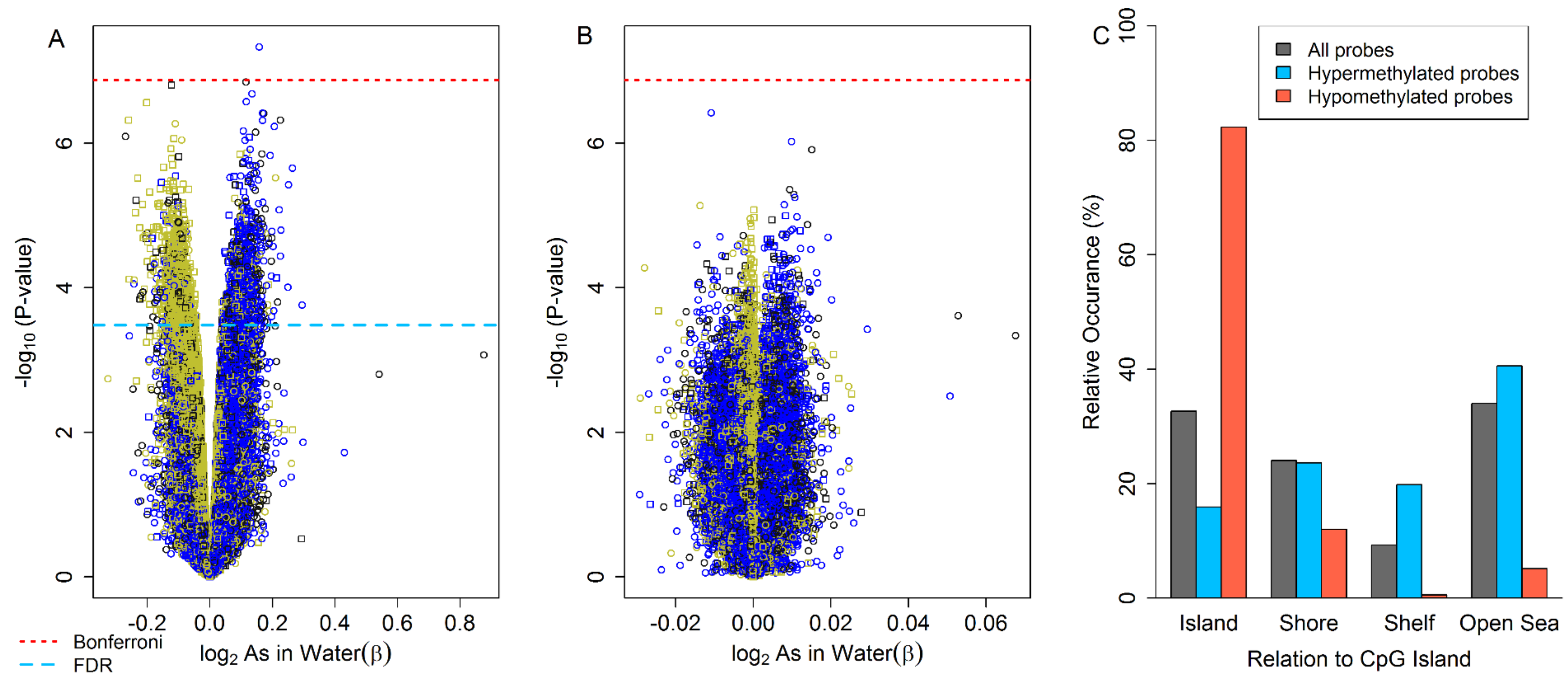

Figure 3.1. Locus-by-locus epigenome-wide analysis for umbilical artery
Legend: Volcano plots for the association between $\log_2$-transformed maternal drinking water arsenic (**A**) unadjusted for cellular heterogeneity and (**B**) adjusting for cellular heterogeneity using the Houseman reference-free method. (**C**) Distribution of differentially methylated loci (q-value<0.05) relative to CpG islands for the unadjusted cell mixture analysis.

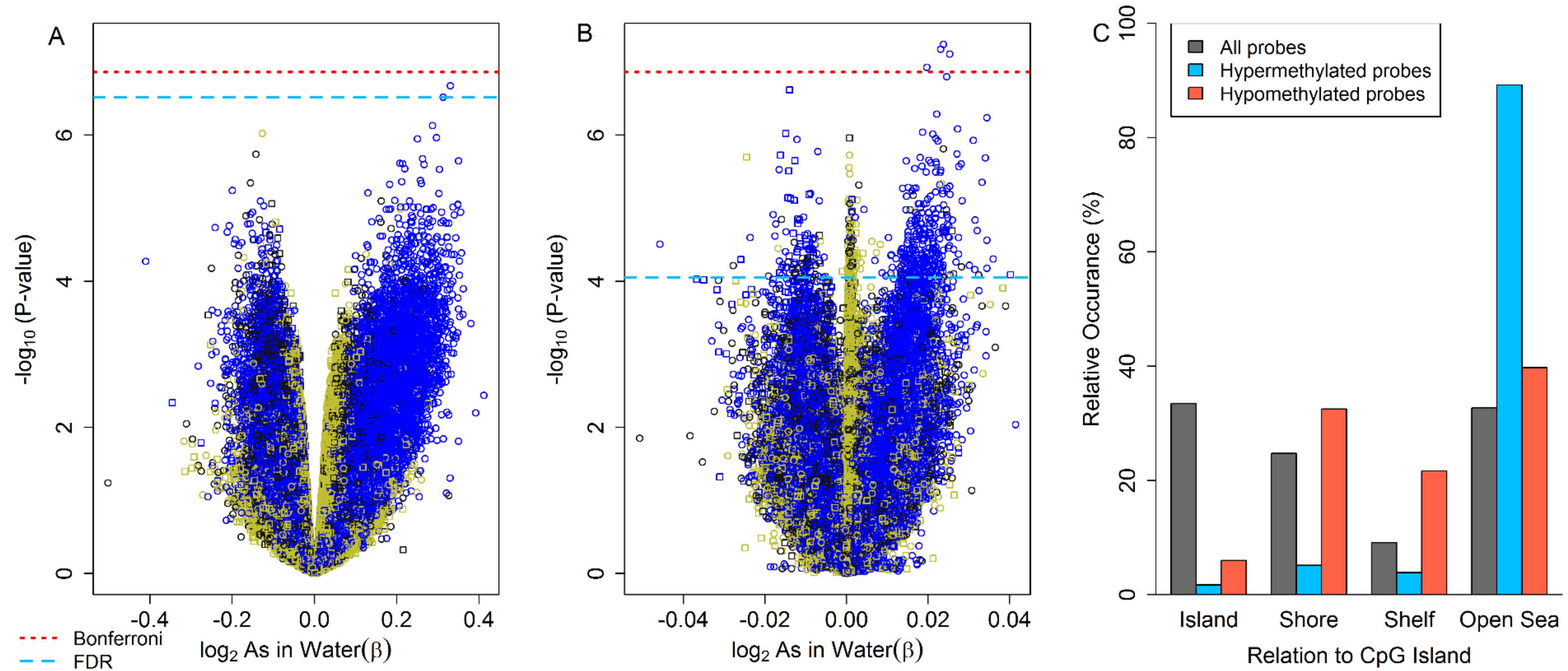

Figure 3.2. Locus-by-locus epigenome-wide analysis for placenta
Legend: Volcano plots for the association between $\log_2$-transformed maternal drinking water arsenic (**A**) unadjusted for cellular heterogeneity and (**B**) adjusting for cellular heterogeneity using the Houseman reference-free method. (**C**) Distribution of differentially methylated loci (nominal $P<1x10^{-4}$) relative to CpG islands for the unadjusted cell mixture analysis.

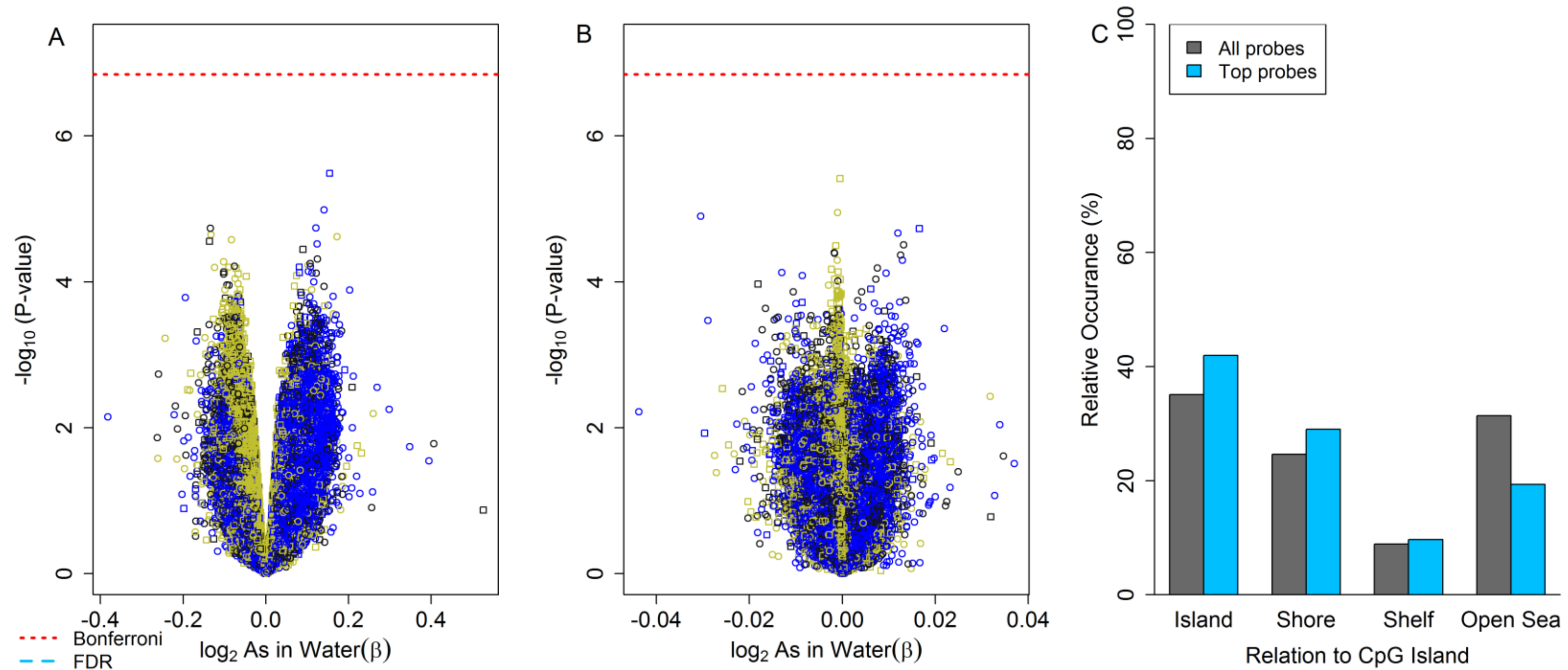

Figure 3.3. Locus-by-locus epigenome-wide analysis for HUVEC
Legend: Volcano plots for the association between $\log_2$-transformed maternal drinking water arsenic: (**A**) unadjusted for cellular heterogeneity and (**B**) adjusting for cellular heterogeneity using the Houseman reference-free method. (**C**) Distribution of differentially methylated loci (nominal *p*-value<$1\text{x}10^{-4}$) relative to CpG islands for the unadjusted cell mixture analysis.

Chapter 4 - THIRD MANUSCRIPT

DNA Methylation in Cord-blood as Mediator of the Association between Arsenic Exposure *in utero* and Infant Birth Outcomes

Andres Cardenas[1]

E. Andres Houseman[1]

Andrea A. Baccarelli[2]

Quazi Quamruzzaman[3]

Mahmuder Rahman[3]

Golam Mostofa[3]

David C. Christiani[2]

Molly L. Kile[1]

[1]School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, Corvallis, OR USA

[2]Harvard T.H. Chan School of Public Health; Boston, MA USA

[3]Dhaka Community Hospital; Dhaka, Bangladesh

Epigenetics

Submission Pending

Abstract

Arsenic is a known developmental toxicant and exposure *in utero* has been associated with poor neonatal health outcomes as well as with increased disease risk later in life that could be in part mediated through epigenetic programming events. Therefore, we evaluated the direct association between maternal drinking water arsenic, birth gestational age and birth weight, as well as the indirect effect of exposure mediated through DNA methylation. Personal maternal drinking water arsenic was measured using inductively-coupled plasma mass spectrometry (ICP-MS) at enrollment ($\leq$16 weeks of gestation). DNA methylation was measured using the Infinium HumanMethylation 450BeadChip array at a single nucleotide resolution from umbilical cord-blood (n=44). A total of 396 loci were differentially methylated by arsenic exposure based on a priori criteria ($P<1.10 \times 10^{-6}$ and |β regression|>0.10). From these, 139 CpGs were associated with birth gestational age and one with birth weight ($P<0.05$). Among loci associated with gestational age, the top-ten (ranked on lowest *p*-value) were selected for mediation analysis. In a structural equation model, a doubling in maternal drinking water arsenic during pregnancy decreased birth gestational age by two days and this was fully mediated through the main principal component of the top-ten CpGs associated with gestational age (β=-0.29, 95% CI:-0.42, -0.15; $P<0.001$). No significant mediation was observed for the single CpG associated with birth weight. We show that DNA methylation has the potential to mediate the effect of prenatal arsenic exposure on gestational age. Future research should validate these findings and evaluate if these epigenetic changes are persistent and associated with disease risk.

Introduction

Arsenic is a global public health problem with >200 million people estimated to be chronically exposed to concentrations that surpass the World Health Organization (WHO) and the U.S. Environmental Protection Agency standard of 10 µg/L in drinking water. In the country of Bangladesh alone, an estimated 20 million and 40 million people are exposed to levels that surpass the national 50 µg/L and the WHO 10 µg/L standards, respectively (World Health Organization (WHO), 2008; Naujokas et al., 2013). Arsenic is classified by the International Agency for Research on Cancer (IARC) as a known human carcinogen (Group 1) (International Agency for Research on Cancer, 2012). In addition to being a human carcinogen, exposure to arsenic has been associated with a wide range of health effects that include immune system disruption, cardiovascular disease, atherosclerosis, type 2 diabetes, hypertension and skin lesions (Abhyankar, Jones, Guallar, & Navas-Acien, 2011; Cardenas, Smit, et al., 2015; Dangleben et al., 2013; Quansah et al., 2015; Tyler & Allan, 2014). Arsenic readily crosses the placenta and maternal exposures are correlated with fetal concentrations affecting newborn's health and normal development (Concha et al., 1998). Exposure to arsenic during fetal development has been consistently associated with adverse perinatal health outcomes including increased risk of spontaneous abortions, still birth, reduced birth weight and both neonatal and infant mortality (Quansah et al., 2015). Furthermore, latent health effects have been associated with exposures occurring prenatally that include increased susceptibility to infections, increased cancer mortality and adverse neurodevelopmental health outcomes (Farzan, Karagas, et al., 2013; Farzan, Korrick, et al., 2013; A. H. Smith et al., 2006)

Of public health importance is the observed ability of early life arsenic exposure, particularly exposures occurring *in utero*, to increase disease risk and susceptibility to adverse health conditions later in life (Bailey et al., 2015; Farzan, Karagas, et al., 2013). The disruption of fetal programming events through an epigenetic mechanism has been postulated as the leading hypothesis for the documented latency of adverse health effects associated with arsenic's toxicity (Bailey & Fry, 2014; Thomas, 2013). Epigenetic remodeling of DNA methylation is crucial for fetal reprogramming events that occur

rapidly following fertilization and could hold the potential to mediate latent disease phenotypes observed in epidemiological studies and experimental models (Marsit, 2015). Although there is some evidence of epigenetic disruption acting as a mediator of disease and adverse phenotypes in experimental animal models, epidemiological data for this is still short-coming, particularly for environmental exposures other that maternal nutrition (Saffery & Novakovic, 2014). For example, in a seminal study manipulating methyl donor bioavailability during pregnancy for the Agouti viable yellow (A^{vy}) mouse induced a change in DNA methylation of a gene promoter that is associated with offspring coat color and disease phenotypes (Dolinoy, 2008). Epigenetic changes in DNA methylation predominantly during development have been proposed to be the main mechanism driving the Developmental Origins of Health and Disease Hypothesis (DOHaD) that states that early nutritional and environmental exposures during fetal development have the ability to influence metabolism and chronic diseases later in life. Epidemiological evidence for this hypothesis initially originated from the Dutch Hunger Winter occurring during World War II consequence of the food embargo imposed by Germany on the Netherlands during 1944-45. Individuals born during this period were prenatally exposed to famine and severe caloric deficiencies. From this natural experiment it was observed that *in utero* famine exposed individuals had low birth weight that correlated with increased risk of coronary heart disease, mental health disorders, obesity, glucose intolerance and breast cancer (T. Roseboom, de Rooij, & Painter, 2006). Persistent epigenetic changes in DNA methylation of famine exposed individuals have been documented decades later and hypothesized to mediate the latency of health effects (Heijmans et al., 2008). In addition, recent epidemiological studies support the ongoing hypothesis that low birth weight or pathways leading to low birth weight associated with latent disease risk are mediated through DNA methylation changes (Engel et al., 2014; Simpkin et al., 2015).

Several studies have examined the epigenome-wide association of DNA methylation in cord or adult whole blood and arsenic exposure. However, most of these studies were restricted to the discovery of CpG loci or differentially methylated regions in relationship to exposure (Argos, 2015). In this study we aimed to test for mediation of

CpG methylation in the relationship between prenatal arsenic exposure birth weight and gestational age using data from our previously reported EWAS in cord-blood (Kile et al., 2014). Furthermore, we introduced an experimental framework for the discovery, evaluation and validation of candidate CpG loci as mediators of prenatal conditions that can be extended to other prenatal exposures.

Materials & Methods

Study Population

The study sample has been previously described in detailed elsewhere (Kile et al., 2014). Briefly, pregnant women of ≤16 weeks of gestational age were recruited into a prospective birth cohort by Dhaka Community Hospital (DCH) in Bangladesh. Inclusion criteria included having a single pregnancy, using a tube well as the main source of drinking water and planning to live in their current residence for the duration of the pregnancy. Women received monthly prenatal vitamins and gave birth at a local clinic or at home with DCH trained medical personnel. This study was approved by the Human Research Committees at the Harvard School of Public Health, Oregon State University and Dhaka Community Hospital Trust.

Maternal Drinking Water Arsenic

Arsenic exposure was measured from tube-wells identified by participants as the main source of drinking water at the time of enrollment as previously described (Kile et al., 2014). Briefly, water samples were collected in a 50-mL polypropylene tubes (BD Falcon, BD Bioscience, Bedford, MA), preserved with reagent grade nitric acid (Merck, Germany) to a pH<2 and stored at room temperature. Drinking water arsenic concentrations were measured by inductively coupled plasma-mass spectrometry (ICP-MS) using the US EPA method 200.8 to determined metals in water (Environmental Laboratory Services, North Syracuse, New York) (Creed et al., 1994). Average percent recovery for arsenic from plasmaCal multi-element QC standard #1 solution (SCP Science) was 102% ± 7%. The limit of detection (LOD) for arsenic was 1 μg/L.

Cord-Blood DNA Methylation and Quality Control

A sample of umbilical cord-blood was collected after delivery into an EDTA-coated vacutainer tube (B.D. Scientific). Subsequently, DNA was extracted from whole blood using the Purgene DNA isolation solutions (Qiagen/Gentra Systems) following manufacturer's instructions. DNA samples were analyzed for DNA methylation at the University of Minnesota's Biomedical Genomic Center using the Illumina Infinium HumanMethylation 450 BeadChip (Illumina, San Diego, CA) which simultaneously profiles the methylation status for >485,000 CpG sites at a single nucleotide resolution covering 99% of the RefSeq genes.

Samples were analyzed in one plate and randomly allocated to 16 chips. Methylation image files were normalized using internal control probes via the functional normalization method with two principal components to account for technical variation between samples using the *minfi* package of R (Fortin et al., 2014). Methylation measurements at CpG loci on X and Y chromosomes were excluded from the analysis to avoid gender-specific methylation bias. Previously identified non-specific and cross-reactive probes within the array along with polymorphic CpG loci ($\geq$ 5% of the minor allele frequency) were excluded for the analysis (Y.-A. Chen et al., 2013). Additionally, detection *p-values* were computed for all CpGs and probes with non-significant detection (P>0.01) in greater than 10% of the samples were also excluded from the analysis. Lastly, a beta-mixture quantile intra sample normalization procedure (BMIQ) was further applied to reduce the potential bias that can arise from type 2 probes as previously described (Teschendorff et al., 2013). The total number of autosomal loci left for analysis after quality control procedures were 383,940 CpGs. Methylation values were logit-transformed to M-values to evaluate the sex adjusted linear association between CpG methylation and prenatal arsenic exposure.

Statistical Analysis

Epigenome-Wide Association (Step 1): A three stage filtering method was implemented to identify candidate loci that could mediate the association between prenatal arsenic exposure and infant birth outcomes. Namely, the first step was to conduct an unadjusted cell composition Epigenome-Wide Association Study (EWAS) in

cord-blood (step 1, Figure 1). The EWAS in cord-blood has been previously published for this sample but we re-analyzed the data to implement the latest technical processing steps for 450K data described in the quality control section (Kile et al., 2014). After quality control, the sex adjusted linear association between maternal drinking water arsenic and individual CpG methylation levels were evaluated using the *limma* function (linear models for microarray analysis) from the minfi package of R. A selection criteria was set a priori for both significance ($P<1x10^{-6}$) and effect size ($|\beta|>0.10$) to identify differentially methylated CpGs associated with arsenic exposure *in utero* on the M-value scale. The unadjusted EWAS was chosen due to the fact that adjusting for white blood cell composition in the rest of the archived samples is not feasible when performing pyrosequencing of the selected CpGs for subsequent validation. To ensure that selected loci were not influenced by white blood changes, the top 600 probes (Houseman-probes) used to differentiate white blood cell composition from reference methylomes were cross-referenced with selected CpGs (Houseman et al., 2012).

Phenotype Association (Step 2): All loci selected from the EWAS using the a priori criteria were then evaluated for their sex adjusted association with both birth gestational age and birth weight using multivariate linear regression models on a CpG-by-CpG basis (step 2, Figure 1). Two selection criteria for association between individual CpG methylation levels and birth outcomes was used. Specifically, CpGs were considered to be significantly associated with birth gestational age and birth weight if: 1) they reached an uncorrected level of significance of $P<0.05$ and 2) if multiple loci were associated only the top-ten loci ranked on lowest *P*-value would be selected for subsequent validation using pyrosequencing and the mediation approach proposed.

Mediation Analysis (Step 3): Finally, CpGs identified to be differentially methylated relative to prenatal arsenic exposure in step 1 and also found to be significantly associated with either birth gestational age or birth weight in step 2 were evaluated for their potential to mediate the effect of exposure on these birth outcomes using structural equation models (SEMs) (step 3, Figure1). First, we check conditions previously postulated for a variable to be considered a potential mediator (Baron & Kenny, 1986). Specifically, mediation requires a significant association between the

exposure and the outcome, a significant association between the mediator and the exposure and a statistically significant association of the mediator to the outcome while controlling for the exposure.

We proposed a conceptual model for mediation based on the a priori assumption that a mediated effect through CpG DNA methylation is biologically plausible (Figure 2). First, we tested the independent effects of exposure on birth outcomes, exposure on CpG methylation and CpG methylation on birth outcomes while adjusting for the exposure. We then conceptualized a model in which the direct effects between prenatal arsenic exposure, CpG methylation and birth outcomes (a, b, c) was evaluated while also testing the indirect effect of exposure on birth outcomes mediated through CpG methylation levels (a•b) while adjusting for infant sex (Figure 2). Two SEMs were used to evaluate the direct effect of log_2-transformed maternal drinking water arsenic on both birth gestational age and birth weight. Bias corrected standard errors and 95% bootstrap Confidence Intervals (CIs) were calculated from 10,000 replicates as the sample size available was relatively small.

Histograms and scatter plots along with regression lines and locally weighted smoothing lines were plotted for bivariate association between exposure, methylation and birth outcomes. All pair wise Pearson correlations coefficients were evaluated among the top candidate CpGs considered for mediation analyses. Due to high correlation among all top-ten CpGs found to be associated with birth gestational age, a post-hoc Principal Component Analysis (PCA) was implemented to deconvolute the major source of variability into a single factor. The scores from the first principal component that accounted for the maximum amount of variability of the methylation levels of all top-ten CpGs were then evaluated as a mediator of the exposure and gestational age relationship into the conceptualized SEM model. All analyses were carried out using the R statistical package, version 3.1.3 (www.r-project.org/).

Results

A total of 29 male (65.9%) and 15 female (34.1%) infants had cord-blood DNA methylation measurements along with maternal drinking water arsenic concentrations

available for analysis. The mean maternal drinking water arsenic concentration at ≤16 weeks of gestation was 63.7 μg/L (range: <1-510 μg/L). The mean gestational age at delivery was 37.6 weeks (range: 33-41 weeks) and the average birth weight was 2923 grams (range: 2080-4050 grams). Selected sample characteristics are summarized in Table 1.

A total of 87,849 CpG loci were observed to be differentially methylated relative to maternal drinking water arsenic after controlling for the false discovery rate (q-value<0.05) and 82 loci reached a Bonferroni adjusted level of significance ($P<1.30\text{x}10^{-7}$) in unadjusted cell mixture analyses. Using our a priori selection criteria for significance of $P<1.10\text{X}10^{-6}$ and effect size of |β regression|>0.10, a total of 396 loci were selected to evaluate their association with birth weight and gestational age, Figure 3. Among the selected 396 CpG loci identified to be differentially methylated relative to prenatal arsenic exposure, none overlapped with the top 600 Houseman-probes used in differentiating white blood cell composition of whole blood samples when compared to the reference methylome set (Reinius et al., 2012).

CpG Methylation and Birth Gestational Age

Multivariate linear regression models adjusted for sex revealed that methylation levels of 139 CpGs (35.1%) from the 396 candidate loci were significantly associated with birth gestational age ($P<0.05$), Figure 4A. Among these loci, the top 10 CpGs ranked on lowest *p*-value were selected to be evaluated as mediators of the exposure and birth outcome relationship. Six of the top ten loci were located in CpG islands and the other four in shore regions of CpG islands among unique genes and chromosomes, Table 2. Nine of the top-ten CpGs were hypermethylated relative to prenatal arsenic exposure and only one was observed to be hypomethylated (Appendix C: Figure C.1). The nine CpGs observed to be positively associated with $\log_2$-transformed arsenic exposure (hypermethylated) were inversely associated with gestational age at birth while the single hypomethylated CpG loci was positively associated with birth gestational age (Appendix C: Figure C.2).

Nine of the top-ten CpG loci observed to be hypermethylated relative to prenatal arsenic exposure were positively correlated (Pearson's ρ range: 0.61 to .90) while the

single hypomethylated loci was negatively correlated with the other nine (Pearson's *ρ* range: -0.85 to -0.68), Figure 7. The Principal Component Analysis (PCA) of the top-ten loci selected demonstrated that 80% of the variance was accounted in the first and main principal component (Figure 8A) and that the cumulative variance explained by four principal components was 93% (Figure 8B).

Before implementing the SEM we evaluated the assumptions for mediation analysis described in the methods section. Namely, log_2-transformed arsenic was significantly associated with birth gestational age (β=-0.25, 95% CI: -0.47, -0.05; $P<0.017$) and with the scores for the first principal component capturing the maximum amount of variation for the methylation levels of the top-ten loci (β=0.70, 95% CI: 0.49, 0.88; $P<0.001$). In turn the scores for the first principal component were significantly associated with birth gestational age while also including log_2-transformed arsenic exposure in the model (β=-0.47, 95% CI: -0.77, -0.17; $P=0.003$), meeting the postulated conditions for mediation.

In the sex adjusted conceptual SEM log_2-tranformed maternal drinking water arsenic was positively associated with the first principal component (PC_1) scores of methylation levels for the top-ten CpGs (β=0.69, 95% CI: 0.50, 0.87; $P<0.001$). The principal component scores of PC_1 were negatively associated with birth gestational age (β=-0.42, 95% CI: -0.59, -0.25; $P<0.001$), Figure 9. The effect of prenatal arsenic exposure on birth gestational age was completely mediated through the first principal component scores. Specifically, a doubling in maternal drinking water arsenic during pregnancy decreased birth gestational age by 0.29 weeks or approximately two days and this was fully mediated through the PC_1 scores for the methylation levels of the selected top-ten CpGs (β=-0.29, 95% CI:-0.42, -0.15; $P<0.001$). The direct effect of maternal drinking water arsenic on birth gestational age was non-significant after accounting for the mediation pathway and therefore not included in the final mediation model (β=0.06, 95% CI:-0.18, 0.30; $P=0.617$). The direct and indirect results for the conceptual model are summarized in Table 3. This final SEM conformed to all model fit indices for good fit summarized in Table 4.

CpG Methylation and Birth Weight

From the 396 candidate CpGs only one loci (cg24484905) located in an open sea region of the *DAB1* gene and observed to be hypomethylated relative to prenatal arsenic exposure was also associated with birth weight (P=0.035), Figure 4B. No direct significant association was observed between maternal drinking water arsenic and birth weight (β=0.977, 95% CI:-27.40, 46.93; P=0.587) and the direct effect of methylation levels of the *DAB1* loci (cg24484905) on birth weight was significant after controlling for arsenic exposure (β=-2717, 95% CI:-5186, -248; P=0.032). However, no significant mediation for the effect of prenatal arsenic exposure on birth weight was observed through methylation levels of this single loci (β=0.02, 95% CI:-0.01, 0.05; P=0.147).

Discussion

In the current study we show that maternal drinking water arsenic during pregnancy decreases birth gestational age and the association is fully mediated through DNA methylation levels of selected CpG loci. However, no significant mediation or direct association was observed between maternal drinking water arsenic and birth weight. We also introduced an experimental approach for the discovery, evaluation and validation of candidate CpG loci as mediators of prenatal conditions or exposures that can be implemented in prospective epidemiological studies and birth cohorts.

To our knowledge, only one previous study conducted by Rojas et al. has evaluated the epigenome-wide association of prenatal arsenic exposure in association with birth outcomes (Rojas et al., 2015). Namely, the authors observed 7 unique loci significantly associated with prenatal arsenic exposure that also correlated with birth gestational age, head circumference or placental weight. None of the CpGs found by this group were within our top-ten differentially methylated loci found to be correlated with gestational age or the loci associated with birth weight. This could be potentially attributed to differences in the timing of the exposure assessment and biomarker as Rojas and colleagues used urinary arsenic measurements at the time of delivery. Differences in timing and biomarker of exposures could yield different epigenetic toxicity profiles and also diverge in the evaluation of chronic arsenic exposure during pregnancy. However,

most loci found in their study were associated with gestational age and none with birth weight similar to our results. This is also consistent with our previously reported finding in which we show that the effect of arsenic on birth weight is mediated through birth gestational age and to a lesser extend with maternal weight gain during pregnancy within this birth cohort (Kile et al., 2015). Another study of arsenic exposure *in utero* using a candidate gene approach found that the expression of *AQP9*, which is a cell membrane channel, has the ability to mediate the effect of arsenic on birth weight. However, it is unknown if the *AQP9* has an epigenomic control mechanism but raise the possibility that prenatal exposure can influence size at birth (Fei et al., 2013).

Interestingly, two CpGs (cg01163597; cg04874129) located in two genes of the solute carrier (SLC) superfamily were observed to mediate the association between gestational age at birth and prenatal arsenic exposure (*SLC22A23; SLC6A2*). The *SLC6A2* is an neurotransmitter transporter across the cell membrane and has been shown to be up-regulated by exposure to arsenic in animal models (P. Liu, Piao, Wang, & Hong, 2008). Furthermore, hypermethylation of this gene has been associated with esophageal carcinogenesis and non-small cell lung cancer (Carvalho et al., 2012; Xu et al., 2013). The *SLC22A23* gene is a novel solute carrier protein transporter and its function has not been well characterized but abundant expression in the brain and liver has been observed (Bennett, Liu, Hoelting, & Stoll, 2011). It has been proposed that even though these transporters exist for endogenous substances, drugs, non-essential metals and environmental toxins could cross the cell membrane through these transporters. However, the physiological purpose in more than half of these transporters remain to be characterized (He, Vasiliou, & Nebert, 2009).

In this study, the only CpG observed to be hypermethylated relative to prenatal arsenic exposure and positively associated with gestational was located in a north shore region of a CpG island in the body of the *S100A6* gene, involved in a Ca^{+2} dependent insulin release. Down regulation of this specific gene has been associated with intrauterine growth restrictions (Sitras, Paulssen, Leirvik, Vårtun, & Acharya, 2009). In addition, high expression levels of this protein has been observed in the human heart and in experimental models increased cardiac expression has been shown to be anti-

hypertrophic (Tsoporis et al., 2005). Another top loci on the *MIR124-3* gene, producing a microRNA predominantly expressed in the nervous system, was observed to be hypomethylated by prenatal arsenic exposure and mediate the association with gestational age. This specific micro-RNA has been correlated with tumor size and disease recurrence of non-small cell lung cancer and renal cell carcinoma (Gebauer et al., 2013; Kitano et al., 2011). It has been also shown to affect neuron growth and differentiation *in vitro* (J.-Y. Yu, Chung, Deo, Thompson, & Turner, 2008).

Gestational age as an outcome is a biologically significant parameter. The clinical phenotype for early gestational age is prematurity defined as <37 weeks of gestations and preterm infants have higher rates of mortality and increase neonatal morbidity. However, recent findings also suggest a risk gradient for gestational age beyond 37 weeks compared to full term infants (Boyle et al., 2012). Reduced gestational age at birth is associated with many adverse long-term health outcomes hypothesized to be mediated by DNA methylation measurements (Platt, 2014; J. W. Schroeder et al., 2011). Our mediation approach showed that selected CpGs had the potential to mediate newborn's gestational age. Recently few studies have started to use mediation approaches to understand the effect of environmental exposures on relevant phenotypes. For example, the effect of smoking on birth weight has been shown to be mediated by individual CpG methylation as well as the association between air pollution and blood pressure (Bellavia et al., 2013; Küpers et al., 2015). We further proposed that gestational age is an intermediate phenotype of disease risk later in life. Future prospective studies should evaluate if these epigenetic perturbations are persistent as the individual ages and also test if certain birth outcomes are an intermediate phenotype to for a clinical disease stage.

Although no direct or mediated association was observed between prenatal arsenic exposure and birth weight, the CpG in the *DAB1* gene was associated with both prenatal arsenic exposure and birth weight. *DAB1* expression has been shown to play an important role in brain ontogenesis and shown to be highly methylated in placentas of different species (D. I. Schroeder et al., 2015).

The present study has many strengths. First, the prospective measurements of the exposure, methylation and subsequent birth outcome present the possibility of testing for

mediation that is chronologically possible. We also used an objective personal exposure measure early during pregnancy in which many of the fetal programming events take place. However, we do rely on a single personal water sample so exposure misclassification cannot be ruled out but other studies in Bangladesh have shown that drinking water arsenic exposures are relative constant and correlate with biomarkers of internal dose (Kile et al., 2009; Kile et al., 2005). There are also some important limitations to be considered. Namely, functional gene expression was not evaluated and the observed epigenetic disruption might not lead to physiological changes in gene expression. Although the unadjusted EWAS was used for the identification of differentially methylated loci the potential of confounding by shifts in white blood cell composition was minimize by ensuring that the probes selected did not differentiate cell types when using the Houseman method. Finally, our relatively small sample size is an important limitation as it does not allow us to adjust for other potential confounders. However, validation of the top CpGs found to mediate the association between gestational age and prenatal arsenic exposure is currently underway using pyrosequencing in a larger independent set of the birth cohort.

Conclusions

In summary, we show that prenatal arsenic exposure from maternal drinking water during pregnancy is inversely associated with birth gestational age and the association is completely mediated by methylation levels of several candidate CpG loci. However, no direct or mediated association was observed for birth weight. Furthermore, we introduce a framework for the discovery, evaluation and validation of candidate CpG loci as mediators of adverse health outcomes to be used for prospective studies. Our results support the hypothesis that arsenic exposure *in utero* can disrupt fetal programming leading to phenotypic consequences that play a role in the developmental origins of health and disease.

Acknowledgements

This work was supported by the US National Institute of Environmental Health Sciences (NIEHS) grants R01 ES015533 and K01 ES017800

Table 4.1. Selected sample characteristics of the study population

Sample characteristics	**Mean±SD**	**Range**
Drinking water As at recruitment (µg/L)	63.7±116.5	<1 - 510
Gestational age at recruitment (weeks)	12.2±2.5	6 - 16
Gestational age at delivery (weeks)	37.6±2.1	33 - 41
Birth weight (grams)	2923±372	2080 - 4050
Gender	**N (%)**	
Male	29 (65.9 %)	---
Female	15 (34.1%)	

Table 4.2. Top-10 CpG loci significantly associated with birth gestational age,

CpG ID	P-value	β Coeff.	Relation to CpG Island	Chromosome	Gene	Gene Group
cg01163597	0.00029	-62.25	N_Shore	Chr6	SLC22A23	Body
cg16081457	0.00031	-9.0	S_Shore	Chr12		
cg06522054	0.00038	-68.1	Island	Chr18	GNAL;GNAL;GNAL	1stExon;Body;Body
cg20382695	0.00055	-52.8	Island	Chr10	ATRNL1	Body
cg24937280	0.00110	-56.7	Island	Chr5	MCC	Body
cg01910639	0.00127	14.7	N_Shore	Chr1	S100A6	Body
cg18115406	0.00135	-35.9	Island	Chr9	LMX1B	TSS200
cg04874129	0.00149	-31.1	Island	Chr16	SLC6A2	1stExon
cg20277905	0.00168	-39.9	Island	Chr20	MIR124-3	TSS200
cg00398764	0.00171	-28.0	N_Shore	Chr15		

Table 4.3. Structural equation model for the mediated effect of arsenic exposure on birth gestational age.

Pathway	**Effect**	**β Coefficient (95% CIs)**	**P-value**
$\log_2$ (As Water) → PC_1	Direct	0.69 (0.50, 0.87)	<0.001
PC_1 → Birth gestational age	Direct	-0.42 (-0.59, -0.25)	<0.001
Sex → Birth gestational age	Direct	-0.05 (-1.17, 1.07)	0.934
$\log_2$ (As Water) → Birth gestational age	Direct	*NS*	
$\log_2$ (As Water) → PC_1 → Birth gestational age	Indirect	-0.29 (-0.42, -0.15)	<0.001

Table 4.4. Fit indices for the final the structural equation model conceptualized

Index	**Criterion for Good Fit**	**Model Fit**
χ^2 p-value	>0.05	0.805
Root Mean Square Error of Approximation (RMSEA)	<0.05	<0.001
Comparative Fit Index	>0.95	1
Tucker-Lewisnon-normed Fit Index	>0.90	1
Standardized Root Mean Squared Residual	>0.05	0.018
Coefficient of Determination	*NA*	0.54

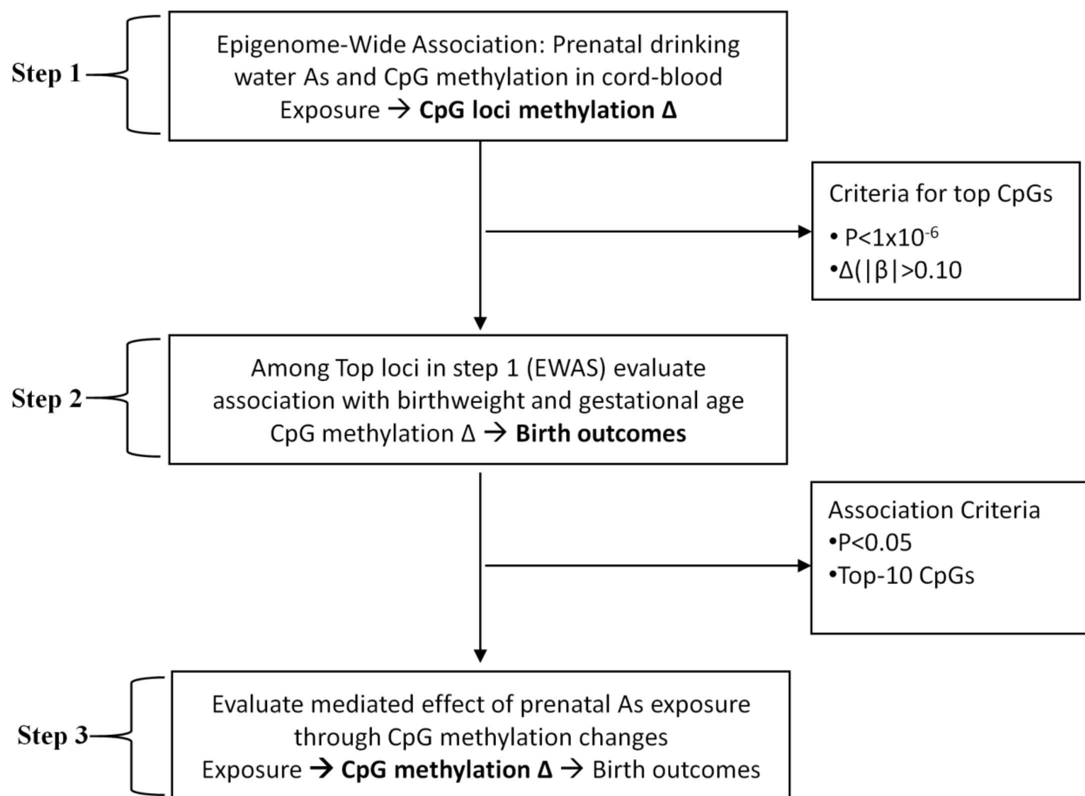

Figure 4.1. Experimental approach for the discovery and validation of CpG loci and mediated effects

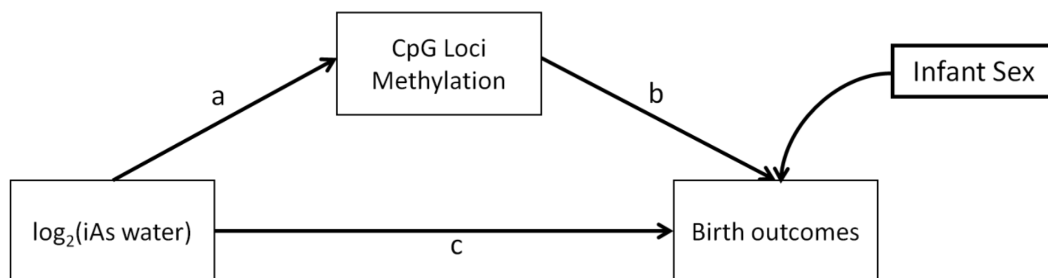

Figure 4.2. Conceptual Structural Equation Model (SEM) for the direct and indirect effect of exposure and infant birth outcomes

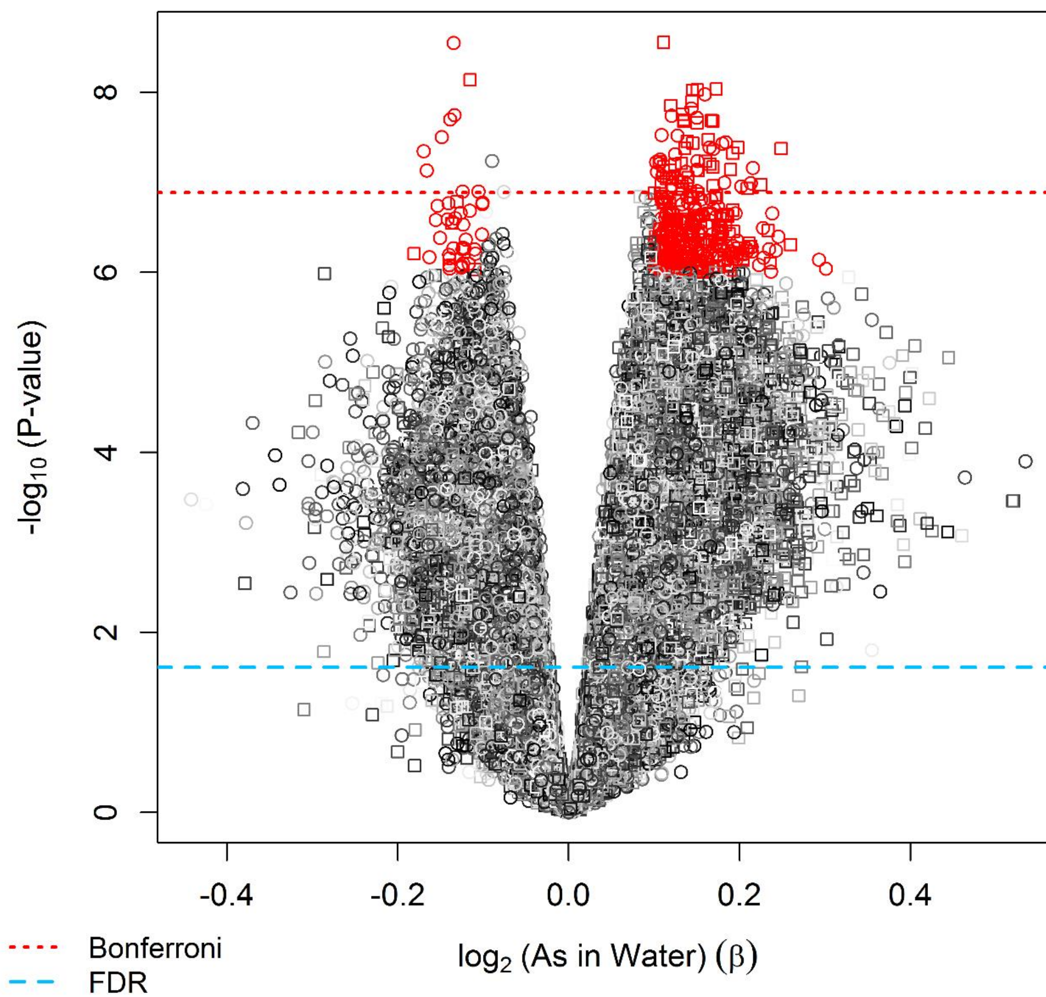

Figure 4.3. Differentially methylated CpG loci in cord-blood relative to prenatal arsenic exposure

Legend: red; loci selected using a priori selection criteria of $|\beta|>0.10$ and nominal $P<1x10^{-6}$ for association (Step 1)

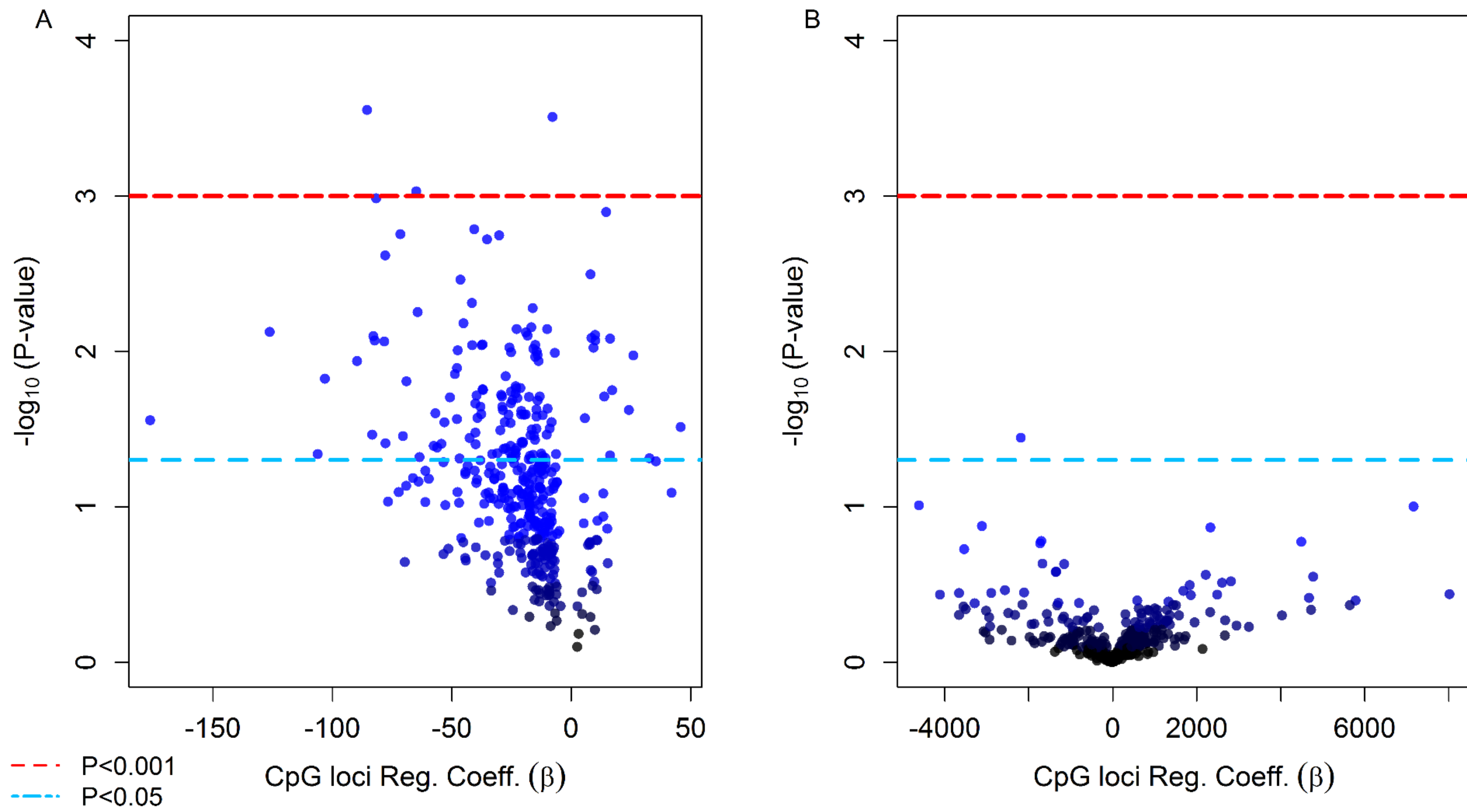

Figure 4.4. Volcano plots for the association among top CpGs and infant health outcomes
Legend: (**A**) birth gestational age and (**B**) birth weight

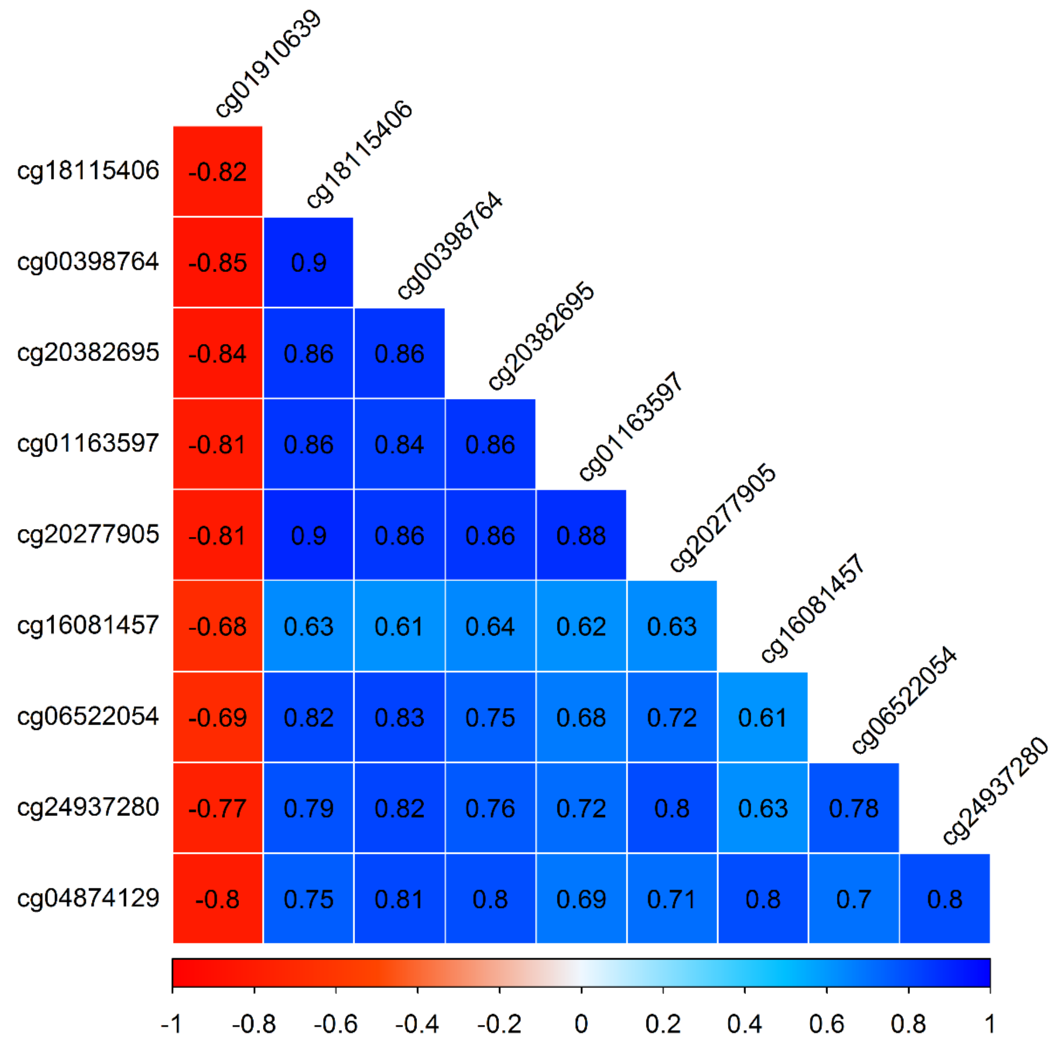

Figure 4.5. Correlation among top 10 CpGs Associated with As exposure in utero and gestational age at birth.

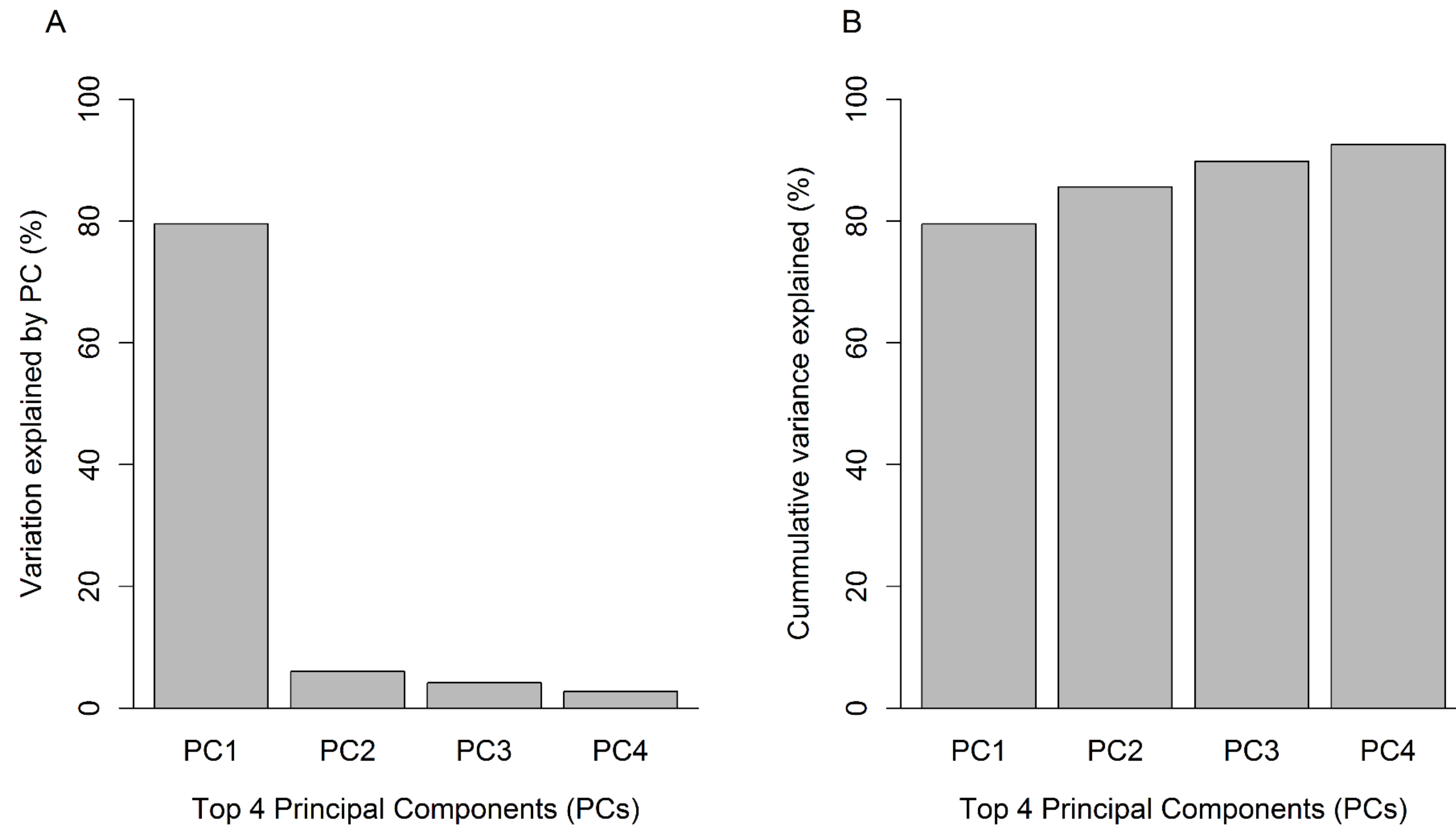

Figure 4.6. Proportion of variance explained by the first four principal components of DNA methylation
Legend: (**A**) Proportion of variance explained by each PC and (**B**) cumulative proportion of variance explained by all 4 PCs

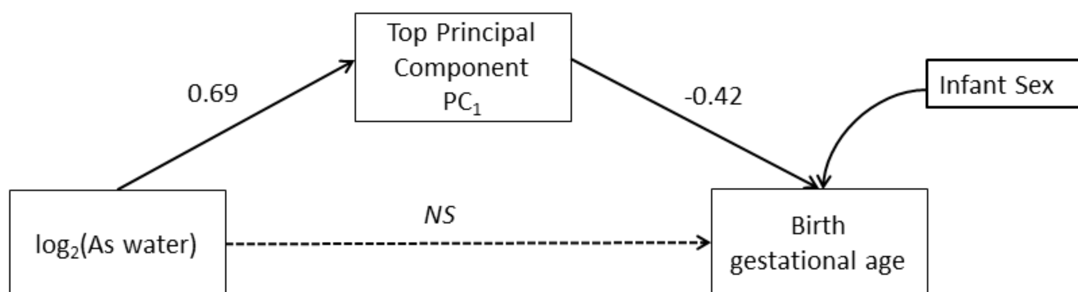

Figure 4.7. Structural Equation Model (SEM) conceptualized for the mediated association
Legend: Mediated association of the main principal component (PC_1) that explained 80% of the variance for the top-10 CpG loci and birth gestational age.

Chapter 5 SUMMARY AND CONCLUSIONS

Arsenic and mercury are common environmental contaminants worldwide that have prevalent dietary exposure sources, posing a public health threat to several populations including pregnant women. Prenatal environmental exposures are of public health interest as they can significantly impact immediate and latent health trajectories of children over the life course. The leading hypothesized mechanism on how the prenatal environment influences disease risk and susceptibility later in life are early epigenetic programming events that occur during embryogenesis and fetal development. This research evaluated the association between epigenetic disruption of DNA methylation and exposure to mercury and arsenic *in utero* for different tissues collected at birth. The conclusions from each proposed specific aim are summarized below:

Aim 1

The first specific aim was to evaluate if prenatal exposure to mercury influenced the cord-blood epigenome of newborns at birth and to test whether prenatal mercury exposure influenced white blood cell composition. We also tested the hypothesis that both mercury and arsenic interact to influence the cord-blood epigenome. This aim was addressed in the first study using a birth cohort conducted in New Hampshire, USA and presented in greater detailed in chapter 2.

Our findings support the hypothesis that *in utero* exposure to mercury can impact important regulatory epigenomic regions at birth. We observed that prenatal exposure to mercury is associated with shifts in the estimated white blood cell composition of cord-blood of newborns potentially altering immune response and regulation. Specifically, we observed a decrease in the estimated proportion of monocytes for the entire sample and an increase in B-cell proportion among females only. In addition, prenatal exposure to both arsenic and mercury were associated with hypermethylation of individual CpG loci and CpG islands. Loci observed to be differentially methylated were found within genes previously implicated in neurodevelopment, a known developmental effect of mercury

and arsenic exposure. Although not a single loci reached the conservative Bonferroni criteria for statistical significance, our results suggest that exposure to mercury may contribute to epigenetic variability at regulatory epigenomic regions. Future studies should considered epigenomic alterations of these two exposures to understand their mechanisms and implications for children's health along with transient and long-term leukocyte shifts.

Aim 2

The second specific aim evaluated the potential for prenatal arsenic exposure to disrupt the epigenome of placenta, umbilical vein endothelial cells (HUVEC) and umbilical artery tissues collected at birth. This specific aim was addressed in the second study using a birth cohort recruited in Bangladesh and presented in detail in chapter 3.

Our results show that prenatal arsenic exposure alters DNA methylation of placenta and umbilical artery but not HUVEC of individual CpG loci. Enrichment of hypomethylated loci within CpG islands of umbilical artery was observed relative to prenatal arsenic exposure and hypermethylation of open sea regions of CpG islands was documented for placenta tissue. We also observed significant epigenetic disruption of key biological pathways previously identified to be involved in arsenic-mediated diseases across the three different tissues. Namely, significant DNA methylation disruption of the melanogesis and insulin signaling pathways was observed along with differential methylation of several biological pathways involved in carcinogenesis relative to arsenic exposure *in utero*. Supporting the hypothesis that arsenic's toxicity and documented adverse health effects may be in part explained by early epigenetic disruption and have fetal origins. Our data also highlights the issue of controlling for cellular heterogeneity when analyzing DNA methylation data from tissues consisting of different cell mixtures as the results and interpretation of findings may dependent on the analytical approach. Furthermore, DNA methylation of the most homogenous tissue (HUVEC) was relatively unaffected by prenatal arsenic exposure in comparison to placenta and umbilical artery that represent a diverse mixture of cells. Future experimental studies should evaluate arsenic's ability to alter cellular differentiation and composition.

These results suggest that prenatal arsenic exposure has the potential to disrupt fetal programming in umbilical artery and placenta tissue but not HUVEC and provides insight into the potential epigenetic pathways that might mediate arsenic-associated diseases.

Aim 3

The third and final specific aim evaluated the potential for DNA methylation in cord-blood to mediate the association between prenatal arsenic exposure and birth outcomes. Namely, the direct and indirect effect of maternal drinking water arsenic during pregnancy on both birth gestational age and birth weight were estimated. This aim was accomplished in the third study using a birth cohort recruited in Bangladesh as outlined in chapter 4.

Our data supports the hypothesis that exposure to arsenic *in utero* is inversely associated with birth gestational age and that this association is fully mediated through DNA methylation of selected CpG loci using an epigenome wide approach. The direct association between arsenic exposure *in utero* and birth weight was marginal and the data did not support a mediated effect through CpG methylation levels. An experimental approach for the discovery, evaluation and validation of candidate loci hypothesized to mediate phenotypes associated with prenatal conditions was also proposed and implemented for this aim. These results support the hypothesis that prenatal exposure to arsenic can disrupt DNA methylation in cord-blood potentially mediating an important neonatal health indicator that has been previously associated with disease susceptibility later in life.

Summary

Our results show that exposure to both arsenic and mercury have the potential to disrupt fetal programming through DNA methylation changes of different tissues. We show in the case of arsenic exposure an inverse association with gestational age mediated through CpG methylation, an intermediate phenotype of latent disease risk. This research provides evidence that prenatal environmental exposures can affect fetal programming playing a role in the developmental origins of health and disease hypothesis. Future research should evaluate the persistence of these epigenetic alterations and test if the

observed disruption has disease associated implications later in life. Reducing environmental exposures to both arsenic and mercury during pregnancy remain critical public health interventions that should be prioritized. Public health interventions that target early prevention of environmental exposures *in utero* that may increase disease risk later in life represent the quintessence of primary prevention. Therefore, reducing exposure to arsenic and mercury during pregnancy has the potential to be a valuable public health intervention.

BIBLIOGRAPHY

Abedi-Valugerdi, M. (2009). Mercury and silver induce B cell activation and anti-nucleolar autoantibody production in outbred mouse stocks: are environmental factors more important than the susceptibility genes in connection with autoimmunity? *Clinical & Experimental Immunology, 155*(1), 117-124.

Abhyankar, L. N., Jones, M. R., Guallar, E., & Navas-Acien, A. (2011). Arsenic exposure and hypertension: a systematic review. *Environmental health perspectives, 120*(4), 494-500.

Agency for Toxic Substances and Disease Registry (ATSDR). (2000). Toxicological Profile for Arsenic: US Department of Public Health and Human Services Atlanta, GA.

Ahmed, S., Khoda, S. M.-e., Rekha, R. S., Gardner, R. M., Ameer, S. S., Moore, S., . . . Raqib, R. (2011). Arsenic-associated oxidative stress, inflammation, and immune disruption in human placenta and cord blood. *Environmental health perspectives, 119*(2), 258-264.

Ahmed, S. A., Hissong, B., Verthelyi, D., Donner, K., Becker, K., & Karpuzoglu-Sahin, E. (1999). Gender and risk of autoimmune diseases: possible role of estrogenic compounds. *Environmental health perspectives, 107*(Suppl 5), 681.

Anders, W., Bull, R., Cantor, K., Chakraborti, D., Chen, C., DeAngelo, A., . . . Gebel, T. (2004). *Some drinking-water disinfectants and contaminants, including arsenic* (Vol. 84): World Health Organization.

Argos, M. (2015). Arsenic Exposure and Epigenetic Alterations: Recent Findings Based on the Illumina 450K DNA Methylation Array. *Current Environmental Health Reports, 2*(2), 137-144.

Argos, M., Chen, L., Jasmine, F., Tong, L., Pierce, B. L., Roy, S., . . . Parvez, F. (2015). Gene-specific differential DNA methylation and chronic arsenic exposure in an epigenome-wide association study of adults in Bangladesh. *Environ Health Perspect, 123*(1), 64-71.

Arita, A., & Costa, M. (2009). Epigenetics in metal carcinogenesis: nickel, arsenic, chromium and cadmium. *Metallomics, 1*(3), 222-228.

Baccarelli, A., & Bollati, V. (2009). Epigenetics and environmental chemicals. *Current opinion in pediatrics, 21*(2), 243.

Bailey, K. A., & Fry, R. C. (2014). Arsenic-associated changes to the epigenome: what are the functional consequences? *Current Environmental Health Reports, 1*(1), 22-34.

Bailey KA, S. A., Tokar EJ, Graziano JH, Kim KW, Navasumrit P, Ruchirawat M, Thiantanawat A, Suk WA, Fry RC. (2015). Mechanisms Underlying Latent Disease Risk Associated with Early-Life Arsenic Exposure: Current Research Trends and Scientific Gaps. *Environ Health Perspect, [Epub] DOI:10.1289/ehp.1409360*.

Bailey, K. A., Smith, A. H., Tokar, E. J., Graziano, J. H., Kim, K.-W., Navasumrit, P., . . . Fry, R. C. (2015). Mechanisms Underlying Latent Disease Risk Associated with Early-Life Arsenic Exposure: Current Research Trends and Scientific Gaps. *Environmental health perspectives*.

Bailey, K. A., Wu, M. C., Ward, W. O., Smeester, L., Rager, J. E., García-Vargas, G., . . . Fry, R. C. (2013). Arsenic and the epigenome: interindividual differences in arsenic metabolism related to distinct patterns of DNA methylation. *Journal of biochemical and molecular toxicology, 27*(2), 106-115.

Bakulski, K. M., & Fallin, M. D. (2014). Epigenetic epidemiology: Promises for public health research. *Environmental and molecular mutagenesis, 55*(3), 171-183.

Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ, 311*(6998), 171-174.

Barker, D. J. (2003). EDITORIAL: The developmental origins of adult disease. *European journal of epidemiology, 18*(8), 733-736.

Barker, D. J. (2012). Developmental origins of chronic disease. *Public health, 126*(3), 185-189.

Barker, D. J., & Osmond, C. (1986). Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *The Lancet, 327*(8489), 1077-1081.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology, 51*(6), 1173.

Basu, N., Goodrich, J. M., & Head, J. (2014). Ecogenetics of mercury: From genetic polymorphisms and epigenetics to risk assessment and decision-making. *Environmental Toxicology and Chemistry, 33*(6), 1248-1258.

Basu, N., Head, J., Nam, D.-H., Pilsner, J. R., Carvan, M. J., Chan, H. M., . . . Scheuhammer, A. M. (2013). Effects of methylmercury on epigenetic markers in three model species: mink, chicken and yellow perch. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology, 157*(3), 322-327.

Baylln, S. B., Herman, J. G., Graff, J. R., Vertino, P. M., & Issa, J.-P. (1997). Alterations in DNA methylation: a fundamental aspect of neoplasia. *Advances in cancer research, 72*, 141-196.

Bellavia, A., Urch, B., Speck, M., Brook, R. D., Scott, J. A., Albetti, B., . . . Bertazzi, P. A. (2013). DNA hypomethylation, ambient particulate matter, and increased blood pressure: findings from controlled human exposure experiments. *Journal of the American Heart Association, 2*(3), e000212.

Bellinger, D. C. (2013). Inorganic arsenic exposure and children's neurodevelopment: A review of the evidence. *Toxics, 1*(1), 2-17.

Bennett, K. M., Liu, J., Hoelting, C., & Stoll, J. (2011). Expression and analysis of two novel rat organic cation transporter homologs, SLC22A17 and SLC22A23. *Molecular and cellular biochemistry, 352*(1-2), 143-154.

Berglund, M., Lind, B., Björnberg, K. A., Palm, B., Einarsson, Ö., & Vahter, M. (2005). Inter-individual variations of human mercury exposure biomarkers: a cross-sectional assessment. *Environmental Health, 4*(1), 20.

Björkman, L., Lundekvam, B. F., Lægreid, T., Bertelsen, B. I., Morild, I., Lilleng, P., . . . Vahter, M. (2007). Mercury in human brain, blood, muscle and toenails in relation to exposure: an autopsy study. *Environ Health, 6*(30), 13.

Bollati, V., & Baccarelli, A. (2010). Environmental epigenetics. *Heredity, 105*(1), 105-112.

Bose-O'Reilly, S., McCarty, K. M., Steckling, N., & Lettmeier, B. (2010). Mercury exposure and children's health. *Current problems in pediatric and adolescent health care, 40*(8), 186-215.

Boucher, O., Muckle, G., Jacobson, J. L., Carter, R. C., Kaplan-Estrin, M., Ayotte, P., . . . Jacobson, S. W. (2014). Domain-specific effects of prenatal exposure to PCBs, mercury, and lead on infant cognition: results from the Environmental Contaminants and Child Development Study in Nunavik. *Environmental health perspectives, 122*(3), 310.

Boyle, E. M., Poulsen, G., Field, D. J., Kurinczuk, J. J., Wolke, D., Alfirevic, Z., & Quigley, M. A. (2012). Effects of gestational age at birth on health outcomes at 3 and 5 years of age: population based cohort study. *BMJ, 344*, e896.

Broberg, K., Ahmed, S., Engström, K., Hossain, M., Jurkovic Mlakar, S., Bottai, M., . . . Vahter, M. (2014). Arsenic exposure in early pregnancy alters genome-wide DNA methylation in cord blood, particularly in boys. *Journal of developmental origins of health and disease, 5*(04), 288-298.

Burchiel, S. W., Mitchell, L. A., Lauer, F. T., Sun, X., McDonald, J. D., Hudson, L. G., & Liu, K. J. (2009). Immunotoxicity and biodistribution analysis of arsenic trioxide in C57Bl/6 mice following a 2-week inhalation exposure. *Toxicology and applied pharmacology, 241*(3), 253-259.

Burris, H. H., & Baccarelli, A. A. (2014). Environmental epigenetics: from novelty to scientific discipline. *Journal of Applied Toxicology, 34*(2), 113-116.

Caldwell, K. E., Labrecque, M. T., Solomon, B. R., Ali, A., & Allan, A. M. (2015). Prenatal arsenic exposure alters the programming of the glucocorticoid signaling system during embryonic development. *Neurotoxicology and teratology, 47*, 66-79.

Caldwell, K. L., Jones, R. L., Verdon, C. P., Jarrett, J. M., Caudill, S. P., & Osterloh, J. D. (2008). Levels of urinary total and speciated arsenic in the US population: National Health and Nutrition Examination Survey 2003–2004. *Journal of Exposure Science and Environmental Epidemiology, 19*(1), 59-68.

Cardenas, A., Koestler, D. C., Houseman, E. A., Jackson, B. P., Kile, M. L., Karagas, M. R., & Marsit, C. J. (2015). Differential DNA Methylation in Umbilical Cord Blood of Infants Exposed to Mercury and Arsenic in utero. *Epigenetics, (10)6*, 508-515.

Cardenas, A., Smit, E., Houseman, E. A., Kerkvliet, N. I., Bethel, J. W., & Kile, M. L. (2015). Arsenic exposure and prevalence of the varicella zoster virus in the United States: NHANES (2003-2004 and 2009-2010). *Environ Health Perspect, 123*(6), 590-596. doi: 10.1289/ehp.1408731

Carvalho, R. H., Haberle, V., Hou, J., van Gent, T., Thongjuea, S., van IJcken, W., . . . Sieuwerts, A. (2012). Genome-wide DNA methylation profiling of non-small cell lung carcinomas. *Epigenetics Chromatin, 5*(1), 9.

Cascio, C., Raab, A., Jenkins, R. O., Feldmann, J., Meharg, A. A., & Haris, P. I. (2011). The impact of a rice based diet on urinary arsenic. [Article]. *Journal of Environmental Monitoring, 13*(2), 257-265. doi: 10.1039/c0em00482k

Castillo-Díaz, S. A., Garay-Sevilla, M. E., Hernández-González, M. A., Solís-Martínez, M. O., & Zaina, S. (2010). Extensive demethylation of normally hypermethylated CpG islands occurs in human atherosclerotic arteries. *International journal of molecular medicine, 26*(5), 691-700.

Catalano, P. M., Presley, L., Minium, J., & Hauguel-de Mouzon, S. (2009). Fetuses of obese mothers develop insulin resistance in utero. *Diabetes care, 32*(6), 1076-1080.

Ceccatelli, S., Daré, E., & Moors, M. (2010). Methylmercury-induced neurotoxicity and apoptosis. *Chemico-biological interactions, 188*(2), 301-308.

Chanda, S., Dasgupta, U. B., GuhaMazumder, D., Gupta, M., Chaudhuri, U., Lahiri, S., . . . Chatterjee, D. (2006). DNA hypermethylation of promoter of gene p53 and p16 in arsenic-exposed people with and without malignancy. *Toxicological Sciences, 89*(2), 431-437.

Chen, A. S., Wang, L., Oxenham, J. L., Condit, W. E., Sorg, T. J., & Supply, W. (2004). *Capital Costs of Arsenic Removal Technologies, US EPA Arsenic Removal Technology Demonstration Program, Round 1*: United States Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory.

Chen, Y.-A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., . . . Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics, 8*(2), 203-209.

Chmurzynska, A. (2010). Fetal programming: link between early nutrition, DNA methylation, and complex diseases. *Nutrition reviews, 68*(2), 87-98.

Chowdhury, U. K., Biswas, B. K., Chowdhury, T. R., Samanta, G., Mandal, B. K., Basu, G. C., . . . Mukherjee, S. K. (2000). Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environmental health perspectives, 108*(5), 393.

Christensen, B. C., & Marsit, C. J. (2011). Epigenomics in environmental health. *Frontiers in genetics, 2*.

Clarkson, T. W., & Magos, L. (2006). The toxicology of mercury and its chemical compounds. *CRC Critical Reviews in Toxicology, 36*(8), 609-662.

Coit, P., De Lott, L. B., Nan, B., Elner, V. M., & Sawalha, A. H. (2015). DNA methylation analysis of the temporal artery microenvironment in giant cell arteritis. *Annals of the rheumatic diseases*, annrheumdis-2014-207116.

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

Conradt, E., Fei, M., LaGasse, L., Tronick, E. Z., Guerin, D., Gorman, D., . . . Lester, B. (2015). Prenatal predictors of infant self-regulation: The contributions of placental DNA methylation of NR3C1 and neuroendocrine activity. *Name: Frontiers in Behavioral Neuroscience, 9*, 130.

Cortessis, V. K., Thomas, D. C., Levine, A. J., Breton, C. V., Mack, T. M., Siegmund, K. D., . . . Laird, P. W. (2012). Environmental epigenetics: prospects for studying epigenetic mediation of exposure–response relationships. *Human genetics, 131*(10), 1565-1589.

Counter, S. A., & Buchanan, L. H. (2004). Mercury exposure in children: a review. *Toxicology and applied pharmacology, 198*(2), 209-230.

Creed, J., Brockhoff, C., & Martin, T. (1994). Method 200.8: Determination of trace elements in waters and wastes by inductively-coupled plasma-mass spectrometry. *Environmental Monitoring Systems Laboratory, Office of Research and Development, US Environmental Protection Agency, Cincinnati, OH, Rev, 5*.

Cropley, J. E., Suter, C. M., Beckman, K. B., & Martin, D. I. (2006). Germ-line epigenetic modification of the murine Avy allele by nutritional supplementation. *Proceedings of the National Academy of Sciences, 103*(46), 17308-17312.

Cullen, W. R., & Reimer, K. J. (1989). Arsenic speciation in the environment. *Chemical Reviews, 89*(4), 713-764. doi: 10.1021/cr00094a002

Dakeishi, M., Murata, K., & Grandjean, P. (2006). Long-term consequences of arsenic poisoning during infancy due to contaminated milk powder. *Environmental Health, 5*(1), 31.

Dangleben, N. L., Skibola, C. F., & Smith, M. T. (2013). Arsenic immunotoxicity: a review. *Environmental health, 12*(1), 73.

Davidson, P. W., Myers, G. J., Weiss, B., Shamlaye, C. F., & Cox, C. (2006). Prenatal methyl mercury exposure from fish consumption and child development: a review of evidence and perspectives from the Seychelles Child Development Study. *Neurotoxicology, 27*(6), 1106-1109.

Davies, M. N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., . . . Troakes, C. (2012). Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol, 13*(6), R43.

Davis, C. D., Uthus, E. O., & Finley, J. W. (2000). Dietary selenium and arsenic affect DNA methylation in vitro in Caco-2 cells and in vivo in rat liver and colon. *The Journal of nutrition, 130*(12), 2903-2909.

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & development, 25*(10), 1010-1022.

Desaulniers, D., Xiao, G.-h., Lian, H., Feng, Y.-L., Zhu, J., Nakai, J., & Bowers, W. J. (2009). Effects of mixtures of polychlorinated biphenyls, methylmercury, and

organochlorine pesticides on hepatic DNA methylation in prepubertal female Sprague-Dawley rats. *International journal of toxicology, 28*(4), 294-307.

Dolinoy, D. C. (2008). The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. *Nutrition reviews, 66*(suppl 1), S7-S11.

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC bioinformatics, 11*(1), 587.

Ehrlich, M. (2009). DNA hypomethylation in cancer cells. *Epigenomics, 1*(2), 239-259.

Engel, S. M., Joubert, B. R., Wu, M. C., Olshan, A. F., Håberg, S. E., Ueland, P. M., . . . Peddada, S. D. (2014). Neonatal genome-wide methylation patterns in relation to birth weight in the Norwegian Mother and Child Cohort. *American journal of epidemiology, 179*(7), 834-842.

Entringer, S., Buss, C., & Wadhwa, P. D. (2010). Prenatal stress and developmental programming of human health and disease risk: concepts and integration of empirical findings. *Current opinion in endocrinology, diabetes, and obesity, 17*(6), 507.

Esteller, M. (2002). CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene, 21*(35), 5427-5440.

Eto, K. (2000). Minamata disease. *Neuropathology, 20*(s1), 14-19.

Farina, M., Rocha, J. B., & Aschner, M. (2011). Mechanisms of methylmercury-induced neurotoxicity: evidence from experimental studies. *Life sciences, 89*(15), 555-563.

Farina, M., Rocha, J. B. T., & Aschner, M. (2010). Oxidative stress and methylmercury-induced neurotoxicity. *Developmental Neurotoxicology Research: Principles, Models, Techniques, Strategies, and Mechanisms*, 357-385.

Farzan, S. F., Karagas, M. R., & Chen, Y. (2013). In utero and early life arsenic exposure in relation to long-term health and disease. *Toxicology and applied pharmacology, 272*(2), 384-390.

Farzan, S. F., Korrick, S., Li, Z., Enelow, R., Gandolfi, A. J., Madan, J., . . . Karagas, M. R. (2013). In utero arsenic exposure and infant infection in a United States cohort: A prospective study. *Environmental research, 126*, 24-30.

Fei, D. L., Koestler, D. C., Li, Z., Giambelli, C., Sanchez-Mejias, A., Gosse, J. A., . . . Robbins, D. J. (2013). Association between In Utero arsenic exposure, placental

gene expression, and infant birth weight: a US birth cohort study. *Environ Health, 12*, 58.

Foley, D. L., Craig, J. M., Morley, R., Olsson, C. J., Dwyer, T., Smith, K., & Saffery, R. (2009). Prospects for epigenetic epidemiology. *American journal of epidemiology, 169*(4), 389-400.

Fortin, J.-P., Labbe, A., Lemire, M., Zanke, B. W., Hudson, T. J., Fertig, E. J., . . . Hansen, K. D. (2014). Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol, 15*(11), 503.

Freire, C., Ramos, R., Lopez-Espinosa, M.-J., Díez, S., Vioque, J., Ballester, F., & Fernández, M.-F. (2010). Hair mercury levels, fish consumption, and cognitive development in preschool children from Granada, Spain. *Environmental research, 110*(1), 96-104.

Gagliardi, L. (2015). On the importance–and unimportance–of gestational age. *Acta Paediatrica, 104*(6), 544-546.

Gardner, R. M., Nermell, B., Kippler, M., Grandér, M., Li, L., Ekström, E.-C., . . . Vahter, M. (2011). Arsenic methylation efficiency increases during the first trimester of pregnancy independent of folate status. *Reproductive toxicology, 31*(2), 210-218.

Gebauer, K., Peters, I., Dubrowinskaja, N., Hennenlotter, J., Abbas, M., Scherer, R., . . . Kuczyk, M. (2013). Hsa-mir-124-3 CpG island methylation is associated with advanced tumours and disease recurrence of patients with clear cell renal cell carcinoma. *British journal of cancer, 108*(1), 131-138.

Gilbert-Diamond, D., Cottingham, K. L., Gruber, J. F., Punshon, T., Sayarath, V., Gandolfi, A. J., . . . Karagas, M. R. (2011). Rice consumption contributes to arsenic exposure in US women. *Proc Natl Acad Sci USA, 108*(51), 20656-20660. doi: 10.1073/pnas.1109127108

Gillman, M. W. (2002). Epidemiological challenges in studying the fetal origins of adult chronic disease. *International journal of epidemiology, 31*(2), 294-299.

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *New England Journal of Medicine, 359*(1), 61-73.

Gluckman, P. D., Hanson, M. A., & Low, F. M. (2011). The role of developmental plasticity and epigenetics in human health. *Birth Defects Research Part C: Embryo Today: Reviews, 93*(1), 12-18.

Godfrey, K. M., Lillycrop, K. A., Burdge, G. C., Gluckman, P. D., & Hanson, M. A. (2007). Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease. *Pediatric research, 61*, 5R-10R.

Goodrich, J. M., Basu, N., Franzblau, A., & Dolinoy, D. C. (2013). Mercury biomarkers and DNA methylation among Michigan dental professionals. *Environmental and molecular mutagenesis, 54*(3), 195-203.

Goullé, J., Saussereau, E., Mahieu, L., Bouige, D., Groenwont, S., Guerbet, M., & Lacroix, C. (2009). Application of inductively coupled plasma mass spectrometry multielement analysis in fingernail and toenail as a biomarker of metal exposure. *Journal of analytical toxicology, 33*(2), 92-98.

Grandjean, P., Weihe, P., White, R. F., Debes, F., Araki, S., Yokoyama, K., . . . Jørgensen, P. J. (1997). Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicology and teratology, 19*(6), 417-428.

Hales, C., Barker, D., Clark, P., Cox, L., Fall, C., Osmond, C., & Winter, P. (1991). Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ: British Medical Journal, 303*(6809), 1019.

Hall, M., Gamble, M., Slavkovich, V., Liu, X., Levy, D., Cheng, Z., . . . Pilsner, J. R. (2007). Determinants of arsenic metabolism: blood arsenic metabolites, plasma folate, cobalamin, and homocysteine concentrations in maternal-newborn pairs. *Environmental health perspectives*, 1503-1509.

Hamadani, J., Tofail, F., Nermell, B., Gardner, R., Shiraji, S., Bottai, M., . . . Vahter, M. (2011). Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-school girls and boys: a population-based cohort study. *International journal of epidemiology, 40*(6), 1593-1604.

Hanna, C. W., Bloom, M. S., Robinson, W. P., Kim, D., Parsons, P. J., vom Saal, F. S., . . . Fujimoto, V. Y. (2012). DNA methylation changes in whole blood is associated with exposure to the environmental contaminants, mercury, lead, cadmium and bisphenol A, in women undergoing ovarian stimulation for IVF. *Human reproduction*, des038.

Harada, M. (1995). Minamata disease: methylmercury poisoning in Japan caused by environmental pollution. *CRC Critical Reviews in Toxicology, 25*(1), 1-24.

He, L., Vasiliou, K., & Nebert, D. W. (2009). Analysis and update of the human solute carrier (SLC) gene superfamily. *Hum Genomics, 3*(2), 195-206.

Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., . . . Lumey, L. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA, 105*(44), 17046-17049.

Hiltunen, M. O., & Ylä-Herttuala, S. (2003). DNA methylation, smooth muscle cells, and atherogenesis. *Arteriosclerosis, thrombosis, and vascular biology, 23*(10), 1750-1753.

Hindmarsh, J. T., McCurdy, R. F., & Savory, J. (1986). Clinical and environmental aspects of arsenic toxicity. *Critical reviews in clinical laboratory sciences, 23*(4), 315-347.

Hopenhayn-Rich, C., Biggs, M. L., Smith, A. H., Kalman, D. A., & Moore, L. E. (1996). Methylation study of a population environmentally exposed to arsenic in drinking water. *Environmental health perspectives, 104*(6), 620.

Hopenhayn, C., Ferreccio, C., Browning, S. R., Huang, B., Peralta, C., Gibb, H., & Hertz-Picciotto, I. (2003). Arsenic exposure from drinking water and birth weight. *Epidemiology, 14*(5), 593-602.

Horsthemke, B., & Wagstaff, J. (2008). Mechanisms of imprinting of the Prader–Willi/Angelman region. *American journal of medical genetics Part A, 146*(16), 2041-2052.

Hou, L., Zhang, X., Wang, D., & Baccarelli, A. (2011). Environmental chemical exposures and human epigenetics. *International journal of epidemiology*, dyr154.

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., . . . Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics, 13*(1), 86.

Houseman, E. A., Molitor, J., & Marsit, C. J. (2014). Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics, 30*(10), 1431-1439.

Howe, C. G., Niedzwiecki, M. M., Hall, M. N., Liu, X., Ilievski, V., Slavkovich, V., . . . Gamble, M. V. (2014). Folate and cobalamin modify associations between S-adenosylmethionine and methylated arsenic metabolites in arsenic-exposed Bangladeshi adults. *The Journal of nutrition, 144*(5), 690-697.

Huyck, K. L., Kile, M. L., Mahiuddin, G., Quamruzzaman, Q., Rahman, M., Breton, C. V., . . . Yousuf, J. (2007). Maternal arsenic exposure associated with low birth weight in Bangladesh. *Journal of Occupational and Environmental Medicine, 49*(10), 1097-1104.

Intarasunanont, P., Navasumrit, P., Waraprasit, S., Chaisatra, K., Suk, W. A., Mahidol, C., & Ruchirawat, M. (2012). Effects of arsenic exposure on DNA methylation in cord blood samples from newborn babies and in a human lymphoblast cell line. *Environ Health, 11*(1), 31.

International Agency for Research on Cancer (IARC). (2012). *A Review of Human Carcinogens: Arsenic, Metals, Fibres, and Dusts. Lyon, France: International Agency for Research on Cancer; 2012*. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, (100).

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., . . . Webster, M. (2009). Genome-wide methylation analysis of human colon cancer reveals similar hypo-and hypermethylation at conserved tissue-specific CpG island shores. *Nature genetics, 41*(2), 178.

Ishimoto, H., Natori, M., Tanaka, M., Miyazaki, T., Kobayashi, T., & Yoshimura, Y. (1997). Role of oxygen-derived free radicals in fetal growth retardation induced by ischemia-reperfusion in rats. *American Journal of Physiology-Heart and Circulatory Physiology, 41*(2), H701.

Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics, 33*, 245-254.

Jirtle, R. L., & Skinner, M. K. (2007). Environmental epigenomics and disease susceptibility. *Nature reviews genetics, 8*(4), 253-262.

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics, 8*(1), 118-127.

Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics, 13*(7), 484-492.

Joubert, B. R., Håberg, S. E., Nilsen, R. M., Wang, X., Vollset, S. E., Murphy, S. K., . . . Cupul-Uicab, L. A. (2012). 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environmental health perspectives, 120*(10), 1425.

Karagas, M., Choi, A. L., Oken, E., Horvat, M., Schoeny, R., Kamai, E., . . . Korrick, S. (2012). Evidence on the human health effects of low level methylmercury exposure. *Environmental health perspectives, 120*(6), 799-806.

Kaur, P., Aschner, M., & Syversen, T. (2006). Glutathione modulation influences methyl mercury induced neurotoxicity in primary cell cultures of neurons and astrocytes. *Neurotoxicology, 27*(4), 492-500.

Kile, M. L., Baccarelli, A., Hoffman, E., Tarantini, L., Quamruzzaman, Q., Rahman, M., . . . Wright, R. O. (2012). Prenatal arsenic exposure and DNA methylation in maternal and umbilical cord blood leukocytes. *Environmental health perspectives, 120*(7), 1061-1066.

Kile, M. L., Cardenas, A., Rodrigues, E., Maitrey, M., Dobson, C., Mostofa, G., . . . Christiani, D. C. (2015). Estimating direct and indirect effects of arsenic exposure during pregnancy on maternal weight gain, gestational age, and birth weight in a Bangladeshi birth cohort. *Epidemiology (Accepted in press).*

Kile, M. L., Hoffman, E., Hsueh, Y.-M., Afroz, S., Quamruzzaman, Q., Rahman, M., . . . Christiani, D. C. (2009). Variability in biomarkers of arsenic exposure and metabolism in adults over time. *Environ Health Perspect, 117*(3), 455-460.

Kile, M. L., Houseman, E. A., Baccarelli, A. A., Quamruzzaman, Q., Rahman, M., Mostofa, G., . . . Christiani, D. C. (2014). Effect of prenatal arsenic exposure on DNA methylation and leukocyte subpopulations in cord blood. *Epigenetics, 9*(5), 774-782.

Kile, M. L., Houseman, E. A., Rodrigues, E., Smith, T. J., Quamruzzaman, Q., Rahman, M., . . . Christiani, D. C. (2005). Toenail arsenic concentrations, GSTT1 gene polymorphisms, and arsenic exposure from drinking water. *Cancer Epidemiology Biomarkers & Prevention, 14*(10), 2419-2426.

Kitano, K., Watanabe, K., Emoto, N., Kage, H., Hamano, E., Nagase, T., . . . Goto, A. (2011). CpG island methylation of microRNAs is associated with tumor size and recurrence of non-small-cell lung cancer. *Cancer science, 102*(12), 2126-2131.

Kliegman, R., Rottman, C., & Behrman, R. (1990). Strategies for the prevention of low birth weight. *American journal of obstetrics and gynecology, 162*(4), 1073-1083.

Koestler, D. C., Avissar-Whiting, M., Houseman, E. A., Karagas, M. R., & Marsit, C. J. (2013). Differential DNA Methylation in Umbilical Cord Blood of Infants Exposed to Low Levels of Arsenic in Utero. *Environmental health perspectives, 121*(8), 971.

Koestler, D. C., Marsit, C. J., Christensen, B. C., Accomando, W., Langevin, S. M., Houseman, E. A., . . . Kelsey, K. T. (2012). Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers. *Cancer Epidemiology Biomarkers & Prevention, 21*(8), 1293-1302.

Koga, M., Serritella, A. V., Messmer, M. M., Hayashi-Takagi, A., Hester, L. D., Snyder, S. H., . . . Sedlak, T. W. (2011). Glutathione is a physiologic reservoir of neuronal glutamate. *Biochemical and biophysical research communications, 409*(4), 596-602.

Kramer, M. S., Demissie, K., Yang, H., Platt, R. W., Sauvé, R., & Liston, R. (2000). The contribution of mild and moderate preterm birth to infant mortality. *JAMA, 284*(7), 843-849.

Küpers, L. K., Xu, X., Jankipersadsing, S. A., Vaez, A., la Bastide-van Gemert, S., Scholtens, S., . . . Felix, J. F. (2015). DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *International journal of epidemiology*, [Epub Ahead of Pring] April 10, 2015. doi: 10.1093/ije/dyv048

Lamborg, C. H., Hammerschmidt, C. R., Bowman, K. L., Swarr, G. J., Munson, K. M., Ohnemus, D. C., . . . Saito, M. A. (2014). A global ocean inventory of anthropogenic mercury based on water column measurements. *Nature, 512*(7512), 65-68.

Law, C. M., & Shiell, A. W. (1996). Is blood pressure inversely related to birth weight? The strength of evidence from a systematic review of the literature. *Journal of hypertension, 14*(8), 935-942.

Lee, B.-E., Hong, Y.-C., Park, H., Ha, M., Koo, B. S., Chang, N., . . . Kim, B.-M. (2010). Interaction between GSTM1/GSTT1 Polymorphism and Blood Mercury on Birth Weight. *Environmental health perspectives, 118*(3), 437.

Lesseur, C., Paquette, A. G., & Marsit, C. J. (2014). Epigenetic regulation of infant neurobehavioral outcomes. *Medical epigenetics, 2*(2), 71-79.

Leviton, A., Blair, E., Dammann, O., & Allred, E. (2005). The wealth of information conveyed by gestational age. *The Journal of pediatrics, 146*(1), 123-127.

Lindberg, A.-L., Kumar, R., Goessler, W., Thirumaran, R., Gurzau, E., Koppova, K., . . . Vahter, M. (2007). Metabolism of low-dose inorganic arsenic in a central European population: influence of sex and genetic polymorphisms. *Environmental health perspectives*, 1081-1086.

Liu, P., Piao, F.-y., Wang, Y.-y., & Hong, Y. (2008). Study on expression of neurotransmitter transmission-related genes in cerebellum of mice exposed to arsenic by gene chip technology. *Journal of Dalian Medical University, 5*, 002.

Liu, X., Zheng, Y., Zhang, W., Zhang, X., LIoyd-Jones, D. M., Baccarelli, A. A., . . . Liu, K. (2014). Blood methylomics in response to arsenic exposure in a low-exposed US population. *Journal of Exposure Science and Environmental Epidemiology, 24*(2), 145-149.

Lynch, J., & Smith, G. D. (2005). A life course approach to chronic disease epidemiology. *Annu. Rev. Public Health, 26*, 1-35.

Maccani, J., Koestler, D. C., Lester, B., Houseman, E. A., Armstrong, D. A., Kelsey, K. T., & Marsit, C. J. (2015). Placental DNA Methylation Related to Both Infant Toenail Mercury and Adverse Neurobehavioral Outcomes. *Environmental health perspectives, Advanced Publication: DOI:10.1289/ehp.1408561*.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., . . . Chakravarti, A. (2009). Finding the missing heritability of complex diseases. *Nature, 461*(7265), 747-753.

Marchiset-Ferlay, N., Savanovitch, C., & Sauvant-Rochat, M.-P. (2012). What is the best biomarker to assess arsenic exposure via drinking water? *Environment international, 39*(1), 150-171.

Marsit, C. J. (2015). Influence of environmental exposure on human epigenetic regulation. *The Journal of experimental biology, 218*(1), 71-79.

Marsit, C. J., Karagas, M. R., Danaee, H., Liu, M., Andrew, A., Schned, A., . . . Kelsey, K. T. (2005). Carcinogen exposure and gene promoter hypermethylation in bladder cancer. *Carcinogenesis, 27*(1), 112-116.

Martin-Chouly, C., Morzadec, C., Bonvalet, M., Galibert, M.-D., Fardel, O., & Vernhet, L. (2011). Inorganic arsenic alters expression of immune and stress response genes in activated primary human T lymphocytes. *Molecular immunology, 48*(6), 956-965.

Martin, G. M. (2005). Epigenetic drift in aging identical twins. *Proc Natl Acad Sci USA, 102*(30), 10413-10414.

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

Mass, M. J., & Wang, L. (1997). Arsenic alters cytosine methylation patterns of the promoter of the tumor suppressor gene p53 in human lung cells: a model for a mechanism of carcinogenesis. *Mutation Research/Reviews in Mutation Research, 386*(3), 263-277.

Meaney, M. J., & Szyf, M. (2005). Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues in clinical neuroscience, 7*(2), 103.

Miklavčič, A., Casetta, A., Tratnik, J. S., Mazej, D., Krsnik, M., Mariuz, M., . . . Horvat, M. (2013). Mercury, arsenic and selenium exposure levels in relation to fish consumption in the Mediterranean area. *Environmental research, 120*, 7-17.

Moon, K., Guallar, E., & Navas-Acien, A. (2012). Arsenic exposure and cardiovascular disease: an updated systematic review. *Current atherosclerosis reports, 14*(6), 542-555.

Myers, G. J., Davidson, P. W., Cox, C., Shamlaye, C. F., Palumbo, D., Cernichiari, E., . . . Huang, L.-S. (2003). Prenatal methylmercury exposure from ocean fish

consumption in the Seychelles child development study. *The Lancet, 361*(9370), 1686-1692.

Myers, G. J., Thurston, S. W., Pearson, A. T., Davidson, P. W., Cox, C., Shamlaye, C. F., . . . Clarkson, T. W. (2009). Postnatal exposure to methyl mercury from fish consumption: a review and new data from the Seychelles Child Development Study. *Neurotoxicology, 30*(3), 338-349.

Myers, S. L., Lobdell, D. T., Liu, Z., Xia, Y., Ren, H., Li, Y., . . . Mendola, P. (2010). Maternal drinking water arsenic exposure and perinatal outcomes in Inner Mongolia, China. *Journal of epidemiology and community health, 64*(4), 325-329.

Naujokas, M. F., Anderson, B., Ahsan, H., Aposhian, H., Graziano, J. H., Thompson, C., & Suk, W. A. (2013). The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environmental health perspectives, 121*(3), 295-302.

Navas-Acien, A., Silbergeld, E. K., Pastor-Barriuso, R., & Guallar, E. (2009). Arsenic exposure and Prevalence of Type 2 Diabetes: Updated Findings from the National Health Nutrition and Examination Survey, 2003-2006. *Epidemiology (Cambridge, Mass.), 20*(6), 816.

Navas-Acien, A., Silbergeld, E. K., Streeter, R. A., Clark, J. M., Burke, T. A., & Guallar, E. (2006). Arsenic exposure and type 2 diabetes: a systematic review of the experimental and epidemiologic evidence. *Environmental health perspectives*, 641-648.

Oberlander, T. F., Weinberg, J., Papsdorf, M., Grunau, R., Misri, S., & Devlin, A. M. (2008). Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics, 3*(2), 97-106.

Oken, E., Choi, A. L., Karagas, M. R., Mariën, K., Rheinberger, C. M., Schoeny, R., . . . Korrick, S. (2012). Which fish should I eat? Perspectives influencing fish consumption choices. *Environmental health perspectives, 120*(6), 790.

Oken, E., Kleinman, K. P., Berland, W. E., Simon, S. R., Rich-Edwards, J. W., & Gillman, M. W. (2003). Decline in fish consumption among pregnant women after a national mercury advisory. *Obstetrics and gynecology, 102*(2), 346.

Onishchenko, N., Karpova, N., Sabri, F., Castrén, E., & Ceccatelli, S. (2008). Long-lasting depression-like behavior and epigenetic changes of BDNF gene expression induced by perinatal exposure to methylmercury. *Journal of neurochemistry, 106*(3), 1378-1387.

Painter, R. C., de Rooij, S. R., Bossuyt, P. M., Simmers, T. A., Osmond, C., Barker, D. J., . . . Roseboom, T. J. (2006). Early onset of coronary artery disease after prenatal exposure to the Dutch famine. *The American journal of clinical nutrition, 84*(2), 322-327.

Pan, W.-C., Seow, W. J., Kile, M. L., Hoffman, E. B., Quamruzzaman, Q., Rahman, M., . . . Christiani, D. C. (2013). Association of low to moderate levels of arsenic exposure with risk of type 2 diabetes in Bangladesh. *American journal of epidemiology, 178*(10), 1563-1570.

Perera, F., & Herbstman, J. (2011). Prenatal environmental exposures, epigenetics, and disease. *Reproductive toxicology, 31*(3), 363-373.

Pichichero, M. E., Cernichiari, E., Lopreiato, J., & Treanor, J. (2002). Mercury concentrations and metabolism in infants receiving vaccines containing thiomersal: a descriptive study. *The Lancet, 360*(9347), 1737-1741.

Pilsner, R. J., Liu, X., Ahsan, H., Ilievski, V., Slavkovich, V., Levy, D., . . . Gamble, M. V. (2009). Folate deficiency, hyperhomocysteinemia, low urinary creatinine, and hypomethylation of leukocyte DNA are risk factors for arsenic-induced skin lesions. *Environmental health perspectives, 117*(2), 254-260.

Platt, M. (2014). Outcomes in preterm infants. *Public health, 128*(5), 399-403.

Quansah, R., Armah, F. A., Essumang, D. K., Luginaah, I., Clarke, E., Marfoh, K., . . . Obiri, S. (2015). Association of Arsenic with Adverse Pregnancy Outcomes/Infant Mortality: A Systematic Review and Meta-Analysis. *Environmental health perspectives, 123*(5), 412.

Rahman, A., Vahter, M., Ekstrom, E.-C., & Persson, L.-Å. (2010). Arsenic exposure in pregnancy increases the risk of lower respiratory tract infection and diarrhea during infancy in Bangladesh. *Environmental health perspectives, 119*(5), 719-724.

Rahman, A., Vahter, M., Smith, A. H., Nermell, B., Yunus, M., El Arifeen, S., . . . Ekström, E.-C. (2009). Arsenic exposure during pregnancy and size at birth: a prospective cohort study in Bangladesh. *American journal of epidemiology, 169*(3), 304-312.

Rao, X., Evans, J., Chae, H., Pilrose, J., Kim, S., Yan, P., . . . Liu, Y. (2012). CpG island shore methylation regulates caveolin-1 expression in breast cancer. *Oncogene, 32*(38), 4519-4528.

Raqib, R., Ahmed, S., Sultana, R., Wagatsuma, Y., Mondal, D., Hoque, A., . . . Persson, L. A. (2009). Effects of in in utero arsenic exposure on child immunity and morbidity in rural Bangladesh. *Toxicology letters, 185*(3), 197-202.

Ray, P. D., Yosim, A., & Fry, R. C. (2014). Incorporating epigenetic data into the risk assessment process for the toxic metals arsenic, cadmium, chromium, lead, and mercury: strategies and challenges. *Frontiers in genetics, 5*.

Rees, J. R., Sturup, S., Chen, C., Folt, C., & Karagas, M. R. (2006). Toenail mercury and dietary fish consumption. *Journal of Exposure Science and Environmental Epidemiology, 17*(1), 25-30.

Reichard, J. F., & Puga, A. (2010). Effects of arsenic exposure on DNA methylation and epigenetic gene regulation. *Epigenomics, 2*(1), 87-104.

Reichard, J. F., Schnekenburger, M., & Puga, A. (2007). Long term low-dose arsenic exposure induces loss of DNA methylation. *Biochemical and biophysical research communications, 352*(1), 188-192.

Reinius, L. E., Acevedo, N., Joerink, M., Pershagen, G., Dahlén, S.-E., Greco, D., . . . Kere, J. (2012). Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One, 7*(7), e41361.

Rich-Edwards, J. W., Stampfer, M. J., Manson, J. E., Rosner, B., Hankinson, S. E., Colditz, G. A., . . . Willet, W. C. (1997). Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. *BMJ, 315*(7105), 396-400.

Risnes, K. R., Vatten, L. J., Baker, J. L., Jameson, K., Sovio, U., Kajantie, E., . . . Painter, R. C. (2011). Birthweight and mortality in adulthood: a systematic review and meta-analysis. *International journal of epidemiology*, dyq267.

Rojas, D., Rager, J. E., Smeester, L., Bailey, K. A., Drobná, Z., Rubio-Andrade, M., . . . Fry, R. C. (2015). Prenatal arsenic exposure and the epigenome: identifying sites of 5-methylcytosine alterations that predict functional changes in gene expression in newborn cord blood and subsequent birth outcomes. *Toxicological Sciences, 143*(1), 97-106.

Ronco, A. M., Arguello, G., Muñoz, L., Gras, N., & Llanos, M. (2005). Metals content in placentas from moderate cigarette consumers: correlation with newborn birth weight. *Biometals, 18*(3), 233-241.

Roseboom, T., de Rooij, S., & Painter, R. (2006). The Dutch famine and its long-term consequences for adult health. *Early human development, 82*(8), 485-491.

Roseboom, T. J., van der Meulen, J. H., Osmond, C., Barker, D. J., Ravelli, A. C., Schroeder-Tanka, J. M., . . . Bleker, O. P. (2000). Coronary heart disease after prenatal exposure to the Dutch famine, 1944–45. *Heart, 84*(6), 595-598.

Sabbioni, E., Fischbach, M., Pozzi, G., Pietra, R., Gallorini, M., & Piette, J. (1991). Cellular retention, toxicity and carcinogenic potential of seafood arsenic. I. Lack of cytotoxicity and transforming activity of arsenobetaine in the BALB/3T3 cell line. *Carcinogenesis, 12*(7), 1287-1291.

Saffery, R., & Novakovic, B. (2014). Epigenetics as the mediator of fetal programming of adult onset disease: what is the evidence? *Acta obstetricia et gynecologica Scandinavica, 93*(11), 1090-1098.

Sagiv, S. K., Thurston, S. W., Bellinger, D. C., Amarasiriwardena, C., & Korrick, S. A. (2012). Prenatal exposure to mercury and fish consumption during pregnancy and attention-deficit/hyperactivity disorder–related behavior in children. *Archives of pediatrics & adolescent medicine, 166*(12), 1123-1131.

Sakamoto, M., Chan, H. M., Domingo, J. L., Oliveira, R. B., Kawakami, S., & Murata, K. (2015). Significance of fingernail and toenail mercury concentrations as biomarkers for prenatal methylmercury exposure in relation to segmental hair mercury concentrations. *Environmental research, 136*, 289-294.

Sanders, A. P., Flood, K., Chiang, S., Herring, A. H., Wolf, L., & Fry, R. C. (2012). Towards prenatal biomonitoring in North Carolina: assessing arsenic, cadmium, mercury, and lead levels in pregnant women. *PLoS One, 7*(3), e31354.

Sanfeliu, C., Sebastià, J., & Kim, S. U. (2001). Methylmercury neurotoxicity in cultures of human neurons, astrocytes, neuroblastoma cells. *Neurotoxicology, 22*(3), 317-327.

Schroeder, D. I., Blair, J. D., Lott, P., Yu, H. O. K., Hong, D., Crary, F., . . . Robinson, W. P. (2013). The human placenta methylome. *Proc Natl Acad Sci USA, 110*(15), 6037-6042.

Schroeder, D. I., Jayashankar, K., Douglas, K. C., Thirkill, T. L., York, D., Dickinson, P. J., . . . Bannasch, D. L. (2015). Early Developmental and Evolutionary Origins of Gene Body DNA Methylation Patterns in Mammalian Placentas. *PLoS Genet, 11*(8), e1005442.

Schroeder, J. W., Conneely, K. N., Cubells, J. F., Kilaru, V., Newport, D. J., Knight, B. T., . . . Tylavsky, F. A. (2011). Neonatal DNA methylation patterns associate with gestational age. *Epigenetics, 6*(12), 1498-1504.

Seow, W. J., Kile, M. L., Baccarelli, A. A., Pan, W. C., Byun, H. M., Mostofa, G., . . . Christiani, D. C. (2014). Epigenome-wide DNA methylation changes with development of arsenic-induced skin lesions in Bangladesh: A case–control follow-up study. *Environmental and molecular mutagenesis, 55*(6), 449-456.

Shenker, B. J., Maserejian, N. N., Zhang, A., & McKinlay, S. (2008). Immune function effects of dental amalgam in children: a randomized clinical trial. *Journal of the American Dental Association (1939), 139*(11), 1496.

Simpkin, A. J., Suderman, M., Gaunt, T. R., Lyttleton, O., McArdle, W. L., Ring, S. M., . . . Relton, C. L. (2015). Longitudinal analysis of DNA methylation associated with birth weight and gestational age. *Human molecular genetics*, ddv119.

Sitras, V., Paulssen, R., Leirvik, J., Vårtun, Å., & Acharya, G. (2009). Placental gene expression profile in intrauterine growth restriction due to placental insufficiency. *Reproductive Sciences, 16*(7), 701-711.

Smith, A. H., Marshall, G., Yuan, Y., Ferreccio, C., Liaw, J., von Ehrenstein, O., . . . Selvin, S. (2006). Increased mortality from lung cancer and bronchiectasis in young adults after exposure to arsenic in utero and in early childhood. *Environmental health perspectives*, 1293-1296.

Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nature reviews genetics, 14*(3), 204-220.

Smyth, G. K. (2005). Limma: linear models for microarray data *Bioinformatics and computational biology solutions using R and Bioconductor* (pp. 397-420): Springer.

Somers, E. C., Ganser, M. A., Warren, J. S., Basu, N., Wang, L., Zick, S. M., & Park, S. K. (2015). Mercury Exposure and Antinuclear Antibodies among Females of Reproductive Age in the United States: NHANES. *Environ Health Perspect* (123), 792–798. doi: 10.1289/ehp.1408751

Soto-Peña, G. A., Luna, A. L., Acosta-Saavedra, L., Conde, P., López-Carrillo, L., Cebrián, M. E., . . . Vega, L. (2006). Assessment of lymphocyte subpopulations and cytokine secretion in children exposed to arsenic. *The FASEB journal, 20*(6), 779-781.

Sproul, D., Kitchen, R. R., Nestor, C. E., Dixon, J. M., Sims, A. H., Harrison, D. J., . . . Meehan, R. R. (2012). Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol, 13*(10), R84.

Srivastava, S., D'Souza, S. E., Sen, U., & States, J. C. (2007). In utero arsenic exposure induces early onset of atherosclerosis in ApoE−/− mice. *Reproductive Toxicology, 23*(3), 449-456.

Stanner, S. A., Bulmer, K., Andres, C., Lantseva, O. E., Borodina, V., Poteen, V., & Yudkin, J. S. (1997). Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad siege study, a cross sectional study. *BMJ, 315*(7119), 1342-1348.

Steegenga, W. T., Boekschoten, M. V., Lute, C., Hooiveld, G. J., de Groot, P. J., Morris, T. J., . . . Müller, M. (2014). Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs. *Age, 36*(3), 1523-1540.

Steinmaus, C., Ferreccio, C., Acevedo, J., Yuan, Y., Liaw, J., Durán, V., . . . Valdés, R. (2014). Increased lung and bladder cancer incidence in adults after in utero and early-life arsenic exposure. *Cancer Epidemiology Biomarkers & Prevention, 23*(8), 1529-1538.

Tanaka, H., Tsukuma, H., & Oshima, A. (2010). Long-term prospective study of 6104 survivors of arsenic poisoning during infancy due to contaminated milk powder in 1955. *Journal of Epidemiology, 20*(6), 439.

Teschendorff, A. E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., & Beck, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics, 29*(2), 189-196.

Tewalt, S. J., Bragg, L. J., & Finkelman, R. B. (2001). *Mercury in US Coal: Abundance, Distribution, and Modes of Occurrence*: US Department of the Interior, US Geological Survey.

Thomas, D. J. (2013). The die is cast: arsenic exposure in early life and disease susceptibility. *Chemical research in toxicology, 26*(12), 1778-1781.

Thomas, D. J., Waters, S. B., & Styblo, M. (2004). Elucidating the pathway for arsenic methylation. *Toxicology and applied pharmacology, 198*(3), 319-326.

Tita, A. T., Landon, M. B., Spong, C. Y., Lai, Y., Leveno, K. J., Varner, M. W., . . . Wapner, R. J. (2009). Timing of elective repeat cesarean delivery at term and neonatal outcomes. *New England Journal of Medicine, 360*(2), 111-120.

Tobi, E. W. (2013). *Epigenetic differences after prenatal adversity: the Dutch hunger winter.* Molecular Epidemiology, Faculty of Medicine/Leiden University Medical Center (LUMC), Leiden University. Retrieved from https://openaccess.leidenuniv.nl/handle/1887/22065

Tobi, E. W., Lumey, L., Talens, R. P., Kremer, D., Putter, H., Stein, A. D., . . . Heijmans, B. T. (2009). DNA methylation differences after exposure to prenatal famine are common and timing-and sex-specific. *Human molecular genetics, 18*(21), 4046-4053.

Tobi, E. W., Slieker, R. C., Stein, A. D., Suchiman, H. E. D., Slagboom, P. E., van Zwet, E. W., . . . Lumey, L. (2015). Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *International journal of epidemiology*, dyv043.

Tolins, M., Ruchirawat, M., & Landrigan, P. (2014). The Developmental Neurotoxicity of Arsenic: Cognitive and Behavioral Consequences of Early Life Exposure. *Annals of global health, 80*(4), 303-314.

Tsoporis, J., Marks, A., Haddad, A., O'Hanlon, D., Jolly, S., & Parker, T. (2005). S100A6 is a negative regulator of the induction of cardiac genes by trophic stimuli in cultured rat myocytes. *Experimental cell research, 303*(2), 471-481.

Tunster, S. J., Jensen, A. B., & John, R. M. (2013). Imprinted genes in mouse placental development and the regulation of fetal energy stores. *Reproduction, 145*(5), R117-R137.

Turk, P. W., Laayoun, A., Smith, S. S., & Weitzman, S. A. (1995). DNA adduct 8-hydroxyl-2′-deoxyguanosine (8-hydroxyguanine) affects function of human DNA methyltransferase. *Carcinogenesis, 16*(5), 1253-1255.

Tyler, C. R., & Allan, A. M. (2014). The effects of arsenic exposure on neurological and cognitive dysfunction in human and rodent studies: a review. *Current Environmental Health Reports, 1*(2), 132-147.

US Environmental Protection Agency (US EPA). (2001). National Primary Drinking Water Regulations: Arsenic and Clarifications to Compliance and New Source Contaminants Monitoring. *Federal Register, 66*, 6975-7066.

US Environmental Protection Agency (US EPA). (2014, May 7, 2014). Priority List of Hazardous Substances Retrieved August 1st, from http://www.atsdr.cdc.gov/spl/

US Food Drug Administration (US FDA). (2014). *What you need to know about mercury in fish and shellfish: 2004 EPA and FDA advice for women who might become pregnant, women who are pregnant, nursing mothers, young children*. Retrieved from http://www.fda.gov/food/resourcesforyou/consumers/ucm110591.htm.

Vahter, M. (2008). Health effects of early life exposure to arsenic. *Basic & Clinical Pharmacology & Toxicology, 102*(2), 204-211.

Vahter, M. (2009). Effects of arsenic on maternal and fetal health. *Annual review of nutrition, 29*, 381-399.

Vahter, M., Åkesson, A., Lind, B., Björs, U., Schütz, A., & Berglund, M. (2000). Longitudinal study of methylmercury and inorganic mercury in blood and urine of pregnant and lactating women, as well as in umbilical cord blood. *Environmental research, 84*(2), 186-194.

Waalkes, M. P., Liu, J., & Diwan, B. A. (2007). Transplacental arsenic carcinogenesis in mice. *Toxicology and applied pharmacology, 222*(3), 271-280.

Waalkes, M. P., Qu, W., Tokar, E. J., Kissling, G. E., & Dixon, D. (2014). Lung tumors in mice induced by "whole-life" inorganic arsenic exposure at human-relevant doses. *Archives of toxicology, 88*(8), 1619-1629.

Wahed, M., Rahman, M., & Vahter, M. (2006). Prevalence of arsenic exposure and skin lesions. A population based survey in Matlab, Bangladesh. *Journal of Epidemiology and Community health*(3), 242-248.

Waly, M., Olteanu, H., Banerjee, R., Choi, S., Mason, J., Parker, B., . . . Benzecry, J. (2004). Activation of methionine synthase by insulin-like growth factor-1 and dopamine: a target for neurodevelopmental toxins and thimerosal. *Molecular psychiatry, 9*(4), 358-370.

Wasserman, G. A., Liu, X., LoIacono, N. J., Kline, J., Factor-Litvak, P., van Geen, A., . . . Schwartz, A. (2014). A cross-sectional study of well water arsenic and child IQ in Maine school children. *Environmental Health, 13*(1), 23.

Wasserman, G. A., Liu, X., Parvez, F., Ahsan, H., Factor-Litvak, P., Kline, J., . . . Levy, D. (2007). Water arsenic exposure and intellectual function in 6-year-old children in Araihazar, Bangladesh. *Environmental health perspectives*, 285-289.

Watanabe, T., & Hirano, S. (2013). Metabolism of arsenic and its toxicological relevance. *Archives of toxicology, 87*(6), 969-979.

Wilcox, A. J. (2001). On the importance—and the unimportance—of birthweight. *International journal of epidemiology, 30*(6), 1233-1241.

Wilcox, A. J., Weinberg, C. R., & Basso, O. (2011). On the pitfalls of adjusting for gestational age at birth. *American journal of epidemiology*, kwr230.

Wilson, J. F. (2004). Balancing the risks and benefits of fish consumption. *Annals of internal medicine, 141*(12), 977-980.

Wong, A. H., Gottesman, I. I., & Petronis, A. (2005). Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Human molecular genetics, 14*(suppl 1), R11-R18.

World Health Organization (WHO). (2008). Guidelines for drinking-water quality: incorporating first and second addenda to third edition, Vol. 1, Recommendations: Geneva: WHO Press.

Wu, G., Bazer, F. W., Cudd, T. A., Meininger, C. J., & Spencer, T. E. (2004). Maternal nutrition and fetal development. *The Journal of nutrition, 134*(9), 2169-2172.

Xie, Y., Liu, J., Benbrahim-Tallaa, L., Ward, J. M., Logsdon, D., Diwan, B. A., & Waalkes, M. P. (2007). Aberrant DNA methylation and gene expression in livers

of newborn mice transplacentally exposed to a hepatocarcinogenic dose of inorganic arsenic. *Toxicology, 236*(1), 7-15.

Xu, E., Gu, J., Hawk, E. T., Wang, K. K., Lai, M., Huang, M., . . . Wu, X. (2013). Genome-wide methylation analysis shows similar patterns in Barrett's esophagus and esophageal adenocarcinoma. *Carcinogenesis, 34*(12), 2750-2756.

Yang, C.-Y., Chang, C.-C., Tsai, S.-S., Chuang, H.-Y., Ho, C.-K., & Wu, T.-N. (2003). Arsenic in drinking water and adverse pregnancy outcome in an arseniasis-endemic area in northeastern Taiwan. *Environmental Research, 91*(1), 29-34.

Yang, T.-Y., Hsu, L.-I., Chiu, A. W., Pu, Y.-S., Wang, S.-H., Liao, Y.-T., . . . Lee, T.-C. (2014). Comparison of genome-wide DNA methylation in urothelial carcinomas of patients with and without arsenic exposure. *Environmental research, 128*, 57-63.

Yorifuji, T., Tsuda, T., Inoue, S., Takao, S., & Harada, M. (2011). Long-term exposure to methylmercury and psychiatric symptoms in residents of Minamata, Japan. *Environment international, 37*(5), 907-913.

Yoshida, T., Yamauchi, H., & Fan Sun, G. (2004). Chronic health effects in people exposed to arsenic via the drinking water: dose–response relationships in review. *Toxicology and applied pharmacology, 198*(3), 243-252.

Yu, H.-S., Liao, W.-T., & Chai, C.-Y. (2006). Arsenic carcinogenesis in the skin. *Journal of biomedical science, 13*(5), 657-666.

Yu, J.-Y., Chung, K.-H., Deo, M., Thompson, R. C., & Turner, D. L. (2008). MicroRNA miR-124 regulates neurite outgrowth during neuronal differentiation. *Experimental cell research, 314*(14), 2618-2633.

Yuan, Y., Marshall, G., Ferreccio, C., Steinmaus, C., Liaw, J., Bates, M., & Smith, A. H. (2010). Kidney cancer mortality: fifty-year latency patterns related to arsenic exposure. *Epidemiology, 21*(1), 103-108.

Yuan, Y., Marshall, G., Ferreccio, C., Steinmaus, C., Selvin, S., Liaw, J., . . . Smith, A. H. (2007). Acute myocardial infarction mortality in comparison with lung and bladder cancer mortality in arsenic-exposed region II of Chile from 1950 to 2000. *American journal of epidemiology, 166*(12), 1381-1391.

Appendix A: Supplemental Materials for Chapter 2

Table A.1. Association between the three main principal components and array plate

	Normalized		**Normalized and ComBat Adjusted**	
Principal Component	Variance explained	[a]*P*	Variance explained	[a]*P*
PC_1	24.1%	<0.001	19.8%	0.994
PC_2	9.4%	0.001	8.7%	0.951
PC_3	6%	<0.001	4.5%	0.996

[a]Omnibus *p*-values for the relationship among the top three principal components (PCs) and plate effect for the normalized data and the normalized data further adjusted for plate effect using ComBat

Table A.2. Estimated change in the proportion of leukocyte composition in cord blood in relationship to the multiplicative interaction between log_2 toenail Hg and urinary As interaction

	Overall Association		Stratified by Sex			
			Males		**Females**	
Cell type	**% Change (95% CI)**	***P*-value**	**% Change (95% CI)**	***P*-value**	**% Change (95% CI)**	***P*-value**
CD8T	-0.6 (-2.0, 0.8)	0.411	0.01 (-2.4, 2.6)	0.947	0.6 (-3.1, 1.9)	0.652
CD4T	-0.2 (-4.6, 7.8)	0.819	2.2 (-1.3, 5.7)	0.216	-0.9 (-4.8, 2.9)	0.620
NK cells	0.1 (-1.8, 2.1)	0.892	0.3 (-2.8, 3.5)	0.841	0.7 (-4.3, 2.7)	0.659
B-cells	0.6 (-1.0, 5.0)	0.407	0.7 (-1.6, 2.9)	0.552	-0.1 (-2.7, 2.3)	0.879
Monocytes	0.1 (-0.8, 1.1)	0.721	0.1 (-1.6, 1.8)	0.896	0.3 (-1.5, 2.1)	0.745
Granulocytes	1.2 (-9.0, 5.0)	0.478	-3.2 (-8.8, 2.4)	0.254	2.2 (-3.8, 8.3)	0.463

[a]Estimates from a multivariate linear regression model adjusted for infant sex, maternal age at delivery and urinary creatinine

Table A.3. Individual CpG lookup of previously reported DNA methylation changes in relationship to Hg exposure in candidate genes studies in relationship to mercury exposure from the New Hampshire Birth Cohort

Gene: Methylation Probe	**Coefficient**	**Relation to Island**	**Nominal *p*-value**
SEPP1 (Goodrich et al., 2013)			
cg04502814	0.025	Open Sea	0.004
cg00886598	-0.002	Open Sea	0.838
GSTM5 (Hanna et al., 2012)			
cg12858902	-0.027	Open Sea	0.449
cg03542222	0.004	Open Sea	0.755
cg14377951	0.007	Open Sea	0.839
cg24467349	-0.010	Open Sea	0.854
cg25210835	0.010	Open Sea	0.857
BDNF (Onishchenko et al., 2008)			
cg09606766	-0.034	S. Shore	0.004
cg01636003	-0.026	S. Shore	0.004
cg18595174	0.032	Open sea	0.015
cg27193031	-0.037	N. Shore	0.072
cg25412831	0.039	Island	0.078
cg24249411	0.016	S. Shore	0.114
cg23497217	0.023	S. Shore	0.119
cg07159484	0.035	Island	0.202
cg14589148	-0.032	Island	0.223
cg11718030	-0.010	Island	0.228
cg20340655	0.011	S. Shore	0.235
cg25381667	-0.031	Island	0.245
cg05818894	0.015	N. Shore	0.249
cg27351358	0.015	N. Shore	0.267
cg15313332	-0.018	N. Shore	0.286
cg23426002	0.010	Open sea	0.313
cg04481212	0.013	N. Shore	0.324
cg23947039	-0.021	Island	0.329
cg26057780	0.018	N. Shore	0.344
cg25962210	0.007	N. Shore	0.364
cg08388004	0.010	Open sea	0.375
cg20954537	-0.009	Island	0.389
cg07238832	0.011	Open sea	0.399
cg09492354	-0.011	N. Shore	0.411
cg03167496	-0.009	Island	0.414
cg14291693	0.010	Open sea	0.440
cg23619332	-0.007	Island	0.452
cg20108357	0.008	N. Shelf	0.454
cg24650785	-0.010	N. Shore	0.466
cg15688670	-0.012	S. Shore	0.514
cg10022526	0.006	Island	0.532
cg06991510	-0.009	S. Shore	0.556
cg05733135	0.008	Island	0.558
cg06979684	0.008	Open sea	0.563
cg02527472	0.006	N. Shore	0.579
cg15462887	-0.005	Island	0.618

Table A.3. Individual CpG lookup of previously reported DNA methylation changes in relationship to Hg exposure in candidate genes studies in relationship to mercury exposure from the New Hampshire Birth Cohort (Continued)

Gene: Methylation Probe	**Coefficient**	**Relation to Island**	**Nominal *p*-value**
cg06684850	-0.005	N. Shore	0.656
cg18117895	0.009	Island	0.709
cg06046431	0.012	Island	0.709
cg26840770	0.004	S. Shore	0.713
cg06816235	0.004	Island	0.748
cg01583131	-0.004	S. Shore	0.756
cg25457956	0.003	Island	0.775
cg06260077	-0.004	N. Shore	0.803
cg22043168	0.004	Island	0.809
cg17413943	-0.002	N. Shore	0.822
cg01418645	0.002	Open sea	0.828
cg04672351	-0.002	S. Shore	0.836
cg05189570	0.003	Open sea	0.862
cg00298481	0.002	Island	0.877
cg10558494	0.002	N. Shore	0.883
cg12448003	0.002	N. Shore	0.883
cg03984780	-0.001	Island	0.883
cg08362738	0.001	Island	0.892
cg26949694	-0.001	Island	0.896
cg06025631	-0.002	Island	0.906
cg25328597	0.001	Island	0.931
cg11241206	0.001	S. Shore	0.936
cg07704699	0.001	N. Shore	0.943
cg10635145	0.001	N. Shore	0.945
cg21010859	0.001	N. Shore	0.946
CDKN2A (Desaulniers et al., 2009)			
cg14430974	0.030	Island	0.082
cg10848754	0.015	Island	0.206
cg04026675	0.010	N. Shore	0.356
cg00718440	-0.003	Island	0.650
cg13601799	0.006	Island	0.738
cg07562918	-0.008	Island	0.373
cg03079681	-0.002	Island	0.847

Table A.4. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal Hg exposure

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg00588920	0.10	2.91E-06	N_Shore	
cg23279152	0.04	3.80E-06	N_Shore	PARM1
cg08994060	-0.09	1.88E-05	OpenSea	PFKFB3
cg23168000	-0.05	5.63E-05	OpenSea	LGMN;LGMN
cg03778029	0.05	5.87E-05	Island	CCDC68;CCDC68
cg03729251	-0.08	8.73E-05	Island	LRBA
cg26771582	-0.04	8.88E-05	N_Shore	FBXO31;FBXO31
cg10791930	0.08	9.23E-05	N_Shore	CEP97
cg21693033	0.05	9.59E-05	S_Shore	KLHDC7A
cg24852565	-0.07	1.13E-04	OpenSea	
cg09339219	0.05	1.13E-04	Island	FZD2
cg10624665	0.05	1.18E-04	Island	FZD2
cg26470696	0.04	1.28E-04	N_Shore	TM6SF2
cg27458888	0.04	1.37E-04	S_Shore	UBE3A
cg25755892	0.03	1.39E-04	OpenSea	PLA2G2A
cg03605454	0.04	1.42E-04	OpenSea	
cg18942298	0.05	1.47E-04	OpenSea	JPH2;JPH2
cg02419835	-0.07	1.64E-04	N_Shore	
cg25402228	0.04	1.69E-04	N_Shelf	
cg23543432	0.04	1.70E-04	OpenSea	
cg18403792	0.04	1.70E-04	OpenSea	LOC645323
cg09972618	0.08	1.83E-04	N_Shore	ISPD;ISPD
cg10178628	0.05	1.91E-04	S_Shore	PLVAP
cg20227255	-0.08	2.14E-04	Island	SRF
cg13149736	0.06	2.21E-04	N_Shore	LYPD6B
cg18009321	-0.06	2.52E-04	Island	LOC284798
cg09997271	0.03	2.69E-04	N_Shore	PDE4A;PDE4A;PDE4A
cg25547332	-0.05	2.74E-04	Island	
cg13854341	0.03	3.10E-04	N_Shore	
cg04374102	-0.03	3.11E-04	Island	TMEM102
cg17105912	0.04	3.14E-04	Island	
cg23781495	0.03	3.19E-04	Island	PAPPA
cg07816809	0.03	3.33E-04	N_Shelf	
cg22983885	0.03	3.42E-04	N_Shore	C1orf128
cg01888869	-0.04	3.43E-04	OpenSea	ITGAX
cg03654304	-0.07	3.50E-04	Island	ZNF598
cg16596440	0.03	3.51E-04	OpenSea	MAML3
cg24765602	0.03	3.59E-04	N_Shore	HSPA2
cg13509756	0.05	3.60E-04	OpenSea	
cg17228105	0.03	3.73E-04	OpenSea	ERG;ERG;ERG;ERG
cg00012701	0.06	3.87E-04	Island	
cg17252808	-0.08	3.91E-04	Island	FAM20B
cg27578381	0.04	4.13E-04	OpenSea	USP36
cg24217567	-0.03	4.16E-04	S_Shore	
cg01454215	0.10	4.16E-04	Island	UCN
cg25600902	0.03	4.16E-04	N_Shore	RHBDD2;RHBDD2
cg15333318	0.04	4.30E-04	S_Shore	NECAB1
cg03618215	-0.06	4.38E-04	Island	GOLGA3

Table A.4. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal Hg exposure (Continued)

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg02537108	0.04	4.45E-04	Island	USP35
cg25653141	-0.08	4.49E-04	Island	BLM
cg08822136	0.03	4.65E-04	N_Shelf	
cg01877814	0.05	4.66E-04	S_Shore	POLG;POLG
cg12498887	0.04	4.87E-04	N_Shore	SYN3;TIMP3;SYN3;SYN3
cg18764516	0.09	4.95E-04	Island	C19orf22
cg00450275	0.03	4.98E-04	OpenSea	TLN2
cg22104744	-0.03	5.00E-04	OpenSea	RASA3
cg03671052	-0.12	5.11E-04	OpenSea	
cg00702593	0.04	5.17E-04	S_Shore	DSCAM
cg08320303	-0.05	5.29E-04	OpenSea	
cg18203974	0.03	5.37E-04	OpenSea	
cg01500402	-0.02	5.43E-04	S_Shore	MLST8
cg12483947	0.04	5.50E-04	OpenSea	SGPL1
cg14275779	0.05	5.53E-04	N_Shore	PLEKHH3
cg02193650	-0.03	5.57E-04	Island	WNT1
cg04729592	0.03	5.58E-04	OpenSea	LOC344595;LOC344595
cg13551894	0.05	5.64E-04	N_Shore	USO1
cg13626585	0.02	5.70E-04	OpenSea	
cg07648698	0.04	5.80E-04	OpenSea	AMPH;AMPH
cg02884053	0.04	5.82E-04	N_Shelf	VIPR2
cg03699385	0.04	5.85E-04	N_Shore	
cg09051342	0.04	5.93E-04	OpenSea	
cg23156226	0.03	6.13E-04	N_Shelf	NCALD
cg15810996	-0.04	6.19E-04	S_Shore	ZBTB7A
cg20750843	-0.09	6.41E-04	Island	CRKL
cg10503827	0.03	6.41E-04	N_Shelf	GPR37
cg08486507	0.03	6.64E-04	N_Shelf	
cg23999973	-0.05	6.70E-04	Island	SUDS3
cg05779081	0.10	6.76E-04	Island	BTAF1
cg14213543	0.07	6.88E-04	Island	NDUFV2
cg05881762	0.06	7.04E-04	S_Shore	UBE3A
cg00753039	0.07	7.08E-04	N_Shelf	
cg22880757	-0.10	7.20E-04	N_Shore	RNF182
cg04300873	-0.03	7.30E-04	S_Shore	ASB16;C17orf65
cg01781374	0.05	7.41E-04	OpenSea	CAMK4
cg23772395	-0.06	7.43E-04	Island	
cg11376706	0.03	7.47E-04	OpenSea	PTPRN2;PTPRN2;PTPRN2
cg02585329	0.04	7.52E-04	OpenSea	
cg24842733	0.03	7.52E-04	Island	SNAP91;SNAP91
cg09088448	0.03	7.54E-04	OpenSea	
cg10579631	0.04	7.59E-04	N_Shore	CHST8;CHST8;CHST8
cg17888090	0.04	7.59E-04	OpenSea	PTGER3
cg04863968	0.05	7.60E-04	OpenSea	CLPS
cg00103984	0.06	7.60E-04	Island	DHX29;SKIV2L2;SKIV2L2
cg24936179	-0.05	7.75E-04	Island	ZBTB46
cg17865033	0.04	7.90E-04	N_Shore	ARNT2
cg09015774	0.03	7.98E-04	Island	CAMTA1

Table A.4. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal Hg exposure (Continued)

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg07892422	-0.03	8.04E-04	Island	HOXC13
cg00784263	0.04	8.07E-04	S_Shore	ZNF480
cg01791798	0.04	8.10E-04	OpenSea	MAST4;MAST4
cg05034363	0.04	8.14E-04	OpenSea	CCL24

Table A.5. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal co-exposure to Hg and As

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg20325200	0.08	2.88E-06	OpenSea	GRHL3;GRHL3;GRHL3
cg10587449	0.05	6.07E-06	S_Shore	
cg05351940	0.11	7.58E-06	Island	ARPC2;ARPC2;ARPC2
cg13422817	0.03	9.24E-06	N_Shelf	FGF6
cg11539055	0.07	1.20E-05	S_Shore	ANTXR2
cg17816637	0.04	1.82E-05	S_Shore	SLC7A4
cg12419685	0.08	1.94E-05	Island	GGT7
cg27544191	0.07	2.25E-05	Island	SBF1
cg05063952	0.04	2.31E-05	N_Shore	STAC2
cg14499274	0.04	2.59E-05	OpenSea	VAV2;VAV2
cg27062369	0.10	3.07E-05	Island	DDIT4L
cg00523450	0.04	3.37E-05	Island	
cg24119717	0.03	3.76E-05	S_Shore	
cg11564239	0.03	4.42E-05	S_Shelf	
cg12652585	0.04	4.45E-05	Island	C6orf195
cg26174583	0.04	4.45E-05	N_Shore	FOXF2
cg05849324	0.05	4.97E-05	Island	NHLRC1
cg21475097	0.05	5.18E-05	Island	PRDM16;PRDM16
cg21386414	0.06	5.51E-05	OpenSea	
cg22162848	0.05	5.52E-05	Island	ITGB4;ITGB4
cg01283343	0.04	5.75E-05	OpenSea	CEACAM5
cg18342462	0.07	5.93E-05	Island	C11orf88;C11orf88
cg24776469	0.05	6.05E-05	N_Shore	CDCA7;CDCA7
cg26987965	0.03	6.46E-05	Island	PSENEN;U2AF1L4;U2AF1L4
cg11967888	0.03	7.12E-05	Island	RNF152
cg08286799	0.05	7.17E-05	Island	TUBB3
cg17250863	0.07	7.27E-05	Island	GGT7
cg23756768	0.04	7.40E-05	N_Shore	
cg24046927	0.05	7.56E-05	Island	LMO4
cg16624076	0.08	7.66E-05	Island	KIFC1;KIFC1
cg16743289	0.04	8.09E-05	N_Shore	CACYBP;CACYBP
cg00850756	0.07	8.23E-05	Island	CCND3;CCND3;CCND3;CCND3
cg26008908	0.03	8.28E-05	N_Shore	MSI1
cg00662002	0.03	8.69E-05	Island	PPP1R13B
cg03744522	0.07	9.08E-05	Island	C1orf174;LOC100133612
cg26160573	0.03	9.33E-05	N_Shore	PTPN18;PTPN18
cg23820661	0.06	9.97E-05	Island	C12orf29
cg18985109	0.07	1.01E-04	N_Shore	ADCK1;ADCK1
cg12819393	0.03	1.04E-04	Island	
cg13443768	0.04	1.05E-04	S_Shore	CDC42EP2
cg14386312	0.04	1.06E-04	S_Shore	ZNF544
cg09122655	0.04	1.19E-04	OpenSea	NPFFR1
cg12435134	0.05	1.28E-04	N_Shore	NAAA;NAAA
cg07621169	0.04	1.30E-04	OpenSea	TKT;TKT;TKT
cg02633398	0.04	1.41E-04	N_Shore	CBFA2T3;CBFA2T3
cg07271561	0.03	1.47E-04	Island	HYAL2;HYAL2
cg04105091	0.04	1.51E-04	N_Shore	PPT2;PPT2;PPT2
cg05895711	0.03	1.52E-04	N_Shelf	

Table A.5. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal co-exposure to Hg and As (Continued)

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg21691116	0.03	1.54E-04	Island	NDUFS2;NDUFS2
cg19037893	0.05	1.58E-04	S_Shore	LRIT1
cg27649764	0.05	1.59E-04	N_Shore	CTDSPL2
cg18076651	0.03	1.68E-04	N_Shore	RARG
cg01188068	0.03	1.68E-04	Island	MLH3;MLH3
cg06025456	0.10	1.75E-04	N_Shore	PPT2;PRRT1;PPT2
cg02323356	0.06	1.75E-04	Island	SPEG
cg23666378	0.04	1.77E-04	Island	KIAA0495;KIAA0495
cg21605781	0.03	1.80E-04	S_Shore	APPL2
cg26584034	0.04	1.80E-04	OpenSea	SYNPO;SYNPO
cg16737517	0.06	1.86E-04	Island	ZBTB46
cg25764534	0.03	1.88E-04	OpenSea	BIN3
cg00679184	0.03	1.91E-04	Island	
cg13388944	0.03	1.94E-04	S_Shore	C1orf123
cg22283195	0.03	2.03E-04	N_Shore	
cg06843231	-0.04	2.19E-04	OpenSea	PDZD2
cg22471742	0.08	2.31E-04	Island	TIMM13;TIMM13
cg04182226	0.03	2.43E-04	OpenSea	TNXB;TNXB
cg01899253	0.04	2.52E-04	S_Shore	FLT1;FLT1;FLT1;FLT1
cg25175654	0.03	2.54E-04	OpenSea	
cg18458353	0.03	2.56E-04	OpenSea	CEACAM5
cg09464263	0.03	2.57E-04	Island	SPOPL
cg13688770	0.05	2.58E-04	S_Shore	ZFYVE28
cg15555527	0.09	2.60E-04	Island	ANTXR2
cg19741660	0.07	2.60E-04	OpenSea	
cg02481307	0.03	2.61E-04	Island	LETM1
cg04015907	0.06	2.62E-04	Island	FAM110A;FAM110A;FAM110A
cg18581445	0.03	2.64E-04	S_Shore	SHD;SHD
cg19014623	0.03	2.70E-04	OpenSea	ACTL6B
cg11905407	0.03	2.74E-04	OpenSea	CASP5;CASP5;CASP5;
cg11521079	0.06	2.87E-04	S_Shore	
cg19827875	0.09	2.90E-04	Island	NAV1
cg07678643	0.03	2.90E-04	N_Shore	C1orf212;C1orf212;C1orf212
cg02962602	0.03	2.92E-04	S_Shore	CCDC57
cg26199857	0.06	2.94E-04	Island	ZNF385A;ZNF385A;ZNF385A
cg02286547	0.06	2.95E-04	Island	
cg11166287	0.03	3.00E-04	OpenSea	
cg13622265	0.07	3.02E-04	OpenSea	TTLL11
cg23754918	0.06	3.02E-04	OpenSea	IQCE;IQCE
cg26825934	0.12	3.05E-04	Island	BRUNOL5
cg26226802	0.09	3.05E-04	Island	GLRA3;GLRA3
cg19243130	0.06	3.11E-04	Island	SIAE;SPA17
cg27096043	0.03	3.15E-04	S_Shore	ATP11A;ATP11A
cg16397071	0.11	3.16E-04	Island	DDX1;DDX1
cg14305313	0.07	3.37E-04	Island	
cg02262025	0.07	3.37E-04	Island	ADAT1
cg21871394	0.05	3.40E-04	Island	TBC1D17;TBC1D17; AKT1S1;AKT1S1;TBC1D17

Table A.5. Top 100 differentially methylated CpG loci based on lowest p-value for prenatal co-exposure to Hg and As (Continued)

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg01629940	0.04	3.42E-04	S_Shore	TRADD
cg07773769	0.02	3.42E-04	S_Shore	PPM1M;PPM1M;PPM1M;PPM1M
cg17289764	0.03	3.46E-04	S_Shore	
cg07705964	0.06	3.59E-04	OpenSea	MGST2;MGST2
cg12094552	0.05	3.59E-04	S_Shore	

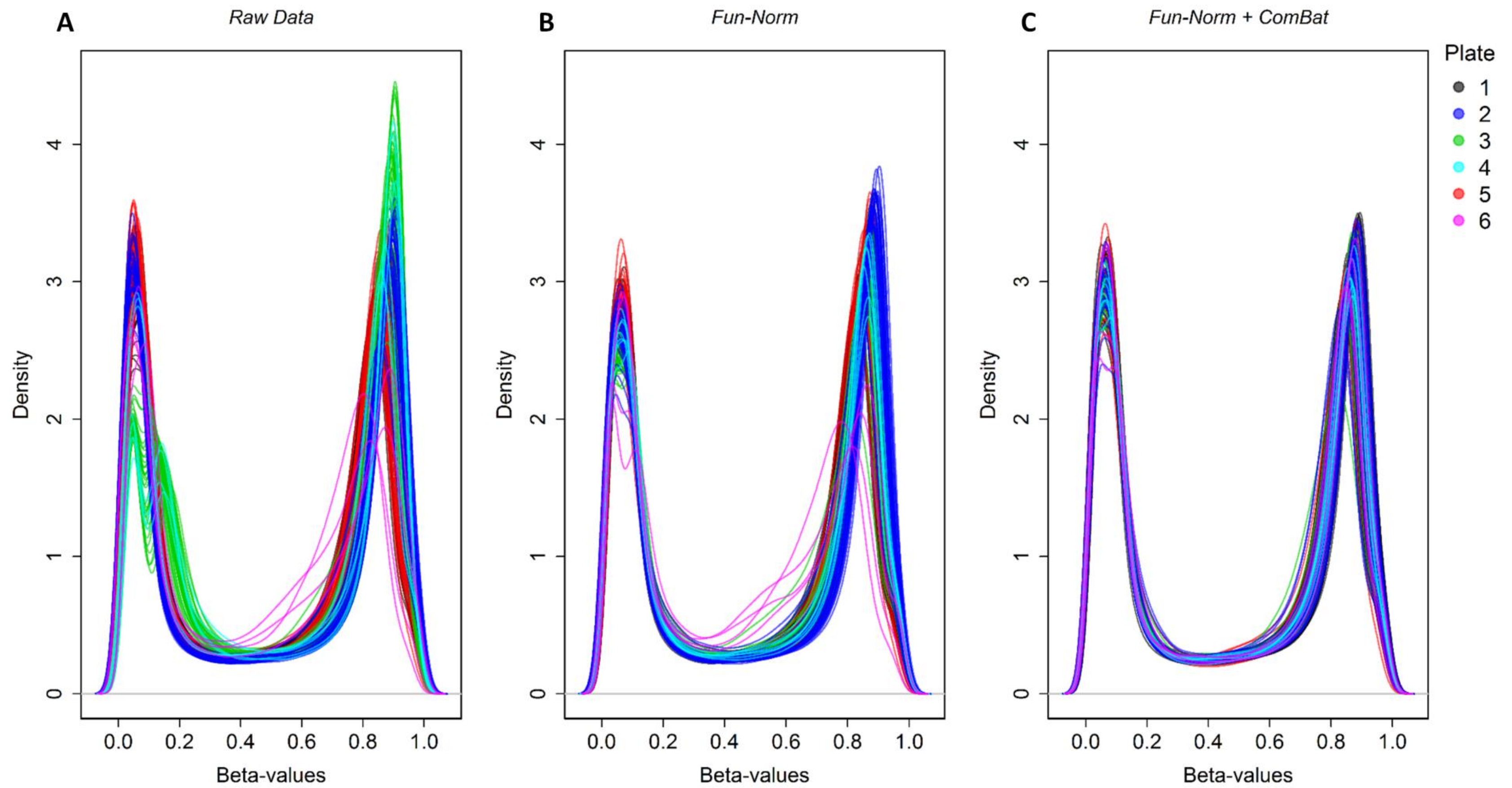

Figure A.1. Density distributions of the DNA methylation by sample

Legend: **(A)** Raw DNA methylation data; **(B)** normalized data using the functional normalization method and **(C)** functional normalized data further adjusted for plate effects using ComBat

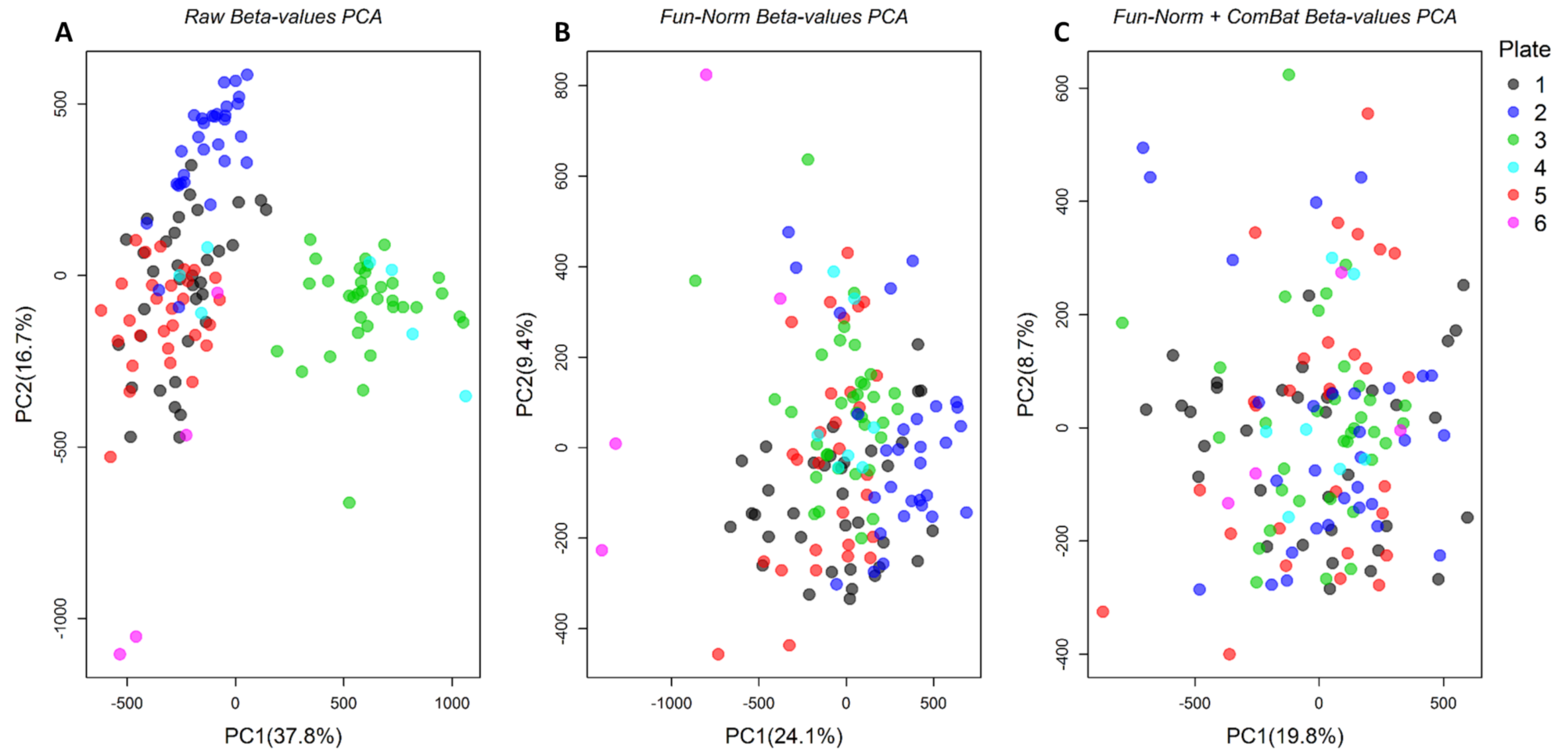

Figure A.2. Principal Component Analysis (PCA) of DNA methylation by plate
Legend: **(A)** Raw DNA methylation data; **(B)** normalized data using the functional normalization method and **(C)** functional normalized data further adjusted for plate effects using ComBat

Appendix B: Supplemental Materials for Chapter 3

Table B.1. Top-10 CpGs for unadjusted cell mixture analysis in artery

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg26587014	0.16	4.69E-08	S_Shelf	
cg02197019	0.12	1.43E-07	N_Shore	
cg21535156	-0.12	1.58E-07	S_Shore	MICALL2
cg03569637	0.13	2.08E-07	OpenSea	LOC100233209;FAM113B
cg04548722	0.12	2.67E-07	OpenSea	LOC440335;LOC440335
cg19223106	-0.20	2.76E-07	Island	RPS5
cg10510779	0.17	3.85E-07	OpenSea	MTMR12
cg00171459	0.17	3.91E-07	S_Shore	FBXL18
cg13021192	-0.26	4.82E-07	Island	CTSZ;CTSZ
cg27540799	0.23	4.83E-07	S_Shore	RER1

Table B.2. Top-10 CpGs for adjusted cell mixture analysis in artery

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg11170044	-0.011	3.81E-07	OpenSea	
cg23954655	0.010	9.55E-07	OpenSea	STK24
cg06896987	0.015	1.24E-06	N_Shore	
cg25951430	0.009	4.42E-06	N_Shore	SLC43A3;SLC43A3
cg02197019	0.010	5.14E-06	N_Shore	
cg26587014	0.011	5.67E-06	S_Shelf	
cg07715909	0.008	7.16E-06	OpenSea	HCCA2
cg24503109	-0.014	7.30E-06	Island	KIAA1751
cg09992883	0.000	8.42E-06	Island	
cg04266383	0.008	1.01E-05	S_Shelf	MCM4;MCM4

Table B.3. Top-10 CpGs for unadjusted cell mixture analysis in placenta

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg26390526	0.33	2.11E-07	OpenSea	FLG
cg03857453	0.31	3.03E-07	OpenSea	NR3C1;NR3C1;NR3C1
cg08215036	0.29	7.44E-07	OpenSea	
cg14383062	-0.13	9.50E-07	Island	TBCD
cg05922384	0.30	1.09E-06	OpenSea	CFH
cg11480029	0.25	1.13E-06	OpenSea	OR9A2
cg17942553	-0.14	1.83E-06	S_Shore	ARNT;ARNT;ARNT
cg14560775	0.26	2.11E-06	OpenSea	OR7A10
cg04050696	0.35	2.26E-06	OpenSea	
cg20554753	0.21	2.45E-06	OpenSea	

Table B.4. Top-10 CpGs for adjusted cell mixture analysis in placenta

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg12825509	0.024	5.76E-08	OpenSea	TRA2B
cg20554753	0.023	6.71E-08	OpenSea	
cg23439277	0.025	7.83E-08	OpenSea	PLCE1
cg21055948	0.020	1.18E-07	OpenSea	CD36
cg08860346	0.025	1.59E-07	OpenSea	
cg14352586	-0.014	2.41E-07	OpenSea	GALNT12
cg09044290	0.022	5.16E-07	OpenSea	LYPD6
cg14087209	0.034	5.78E-07	OpenSea	
cg04840339	0.027	8.23E-07	OpenSea	TRDN
cg21021629	0.019	9.19E-07	OpenSea	NRIP1

Table B.5. Top-10 CpGs for unadjusted cell mixture analysis in HUVEC

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg27175016	0.15	3.26E-06	OpenSea	SORCS2
cg05201145	0.14	1.03E-05	OpenSea	IQSEC3;IQSEC3
cg11933267	0.12	1.83E-05	OpenSea	PLA2G2F
cg18087902	-0.13	1.84E-05	N_Shore	MARCH10;MARCH10
cg06080831	-0.13	2.26E-05	Island	
cg01870330	0.17	2.42E-05	Island	
cg09435824	-0.08	2.66E-05	Island	HSDL1;LRRC50;HSDL1
cg19440264	-0.14	2.79E-05	S_Shore	CHRNB1
cg21357088	0.12	3.03E-05	OpenSea	
cg24127106	0.09	3.60E-05	N_Shore	PRMT1;PRMT1;PRMT1

Table B.6. Top-10 CpGs for adjusted cell mixture analysis in HUVEC

CpG Name	Coefficient	P-value	Relation to CpG Island	Gene Name
cg21149548	-0.001	3.86E-06	Island	UXS1
cg09435824	-0.001	1.12E-05	Island	HSDL1;LRRC50;HSDL1
cg23250157	-0.030	1.26E-05	OpenSea	SYNE2;SYNE2;SYNE2
cg27175016	0.017	1.86E-05	OpenSea	SORCS2
cg00490543	0.012	2.15E-05	OpenSea	HVCN1;HVCN1
cg12747469	0.013	3.13E-05	S_Shore	CCDC108;CCDC108
cg02092906	-0.001	3.20E-05	Island	PGM3;RWDD2A
cg18087902	-0.002	4.01E-05	N_Shore	MARCH10;MARCH10
cg03635649	-0.002	4.06E-05	S_Shore	PLCXD3
cg19528338	0.013	4.31E-05	N_Shore	BMP2

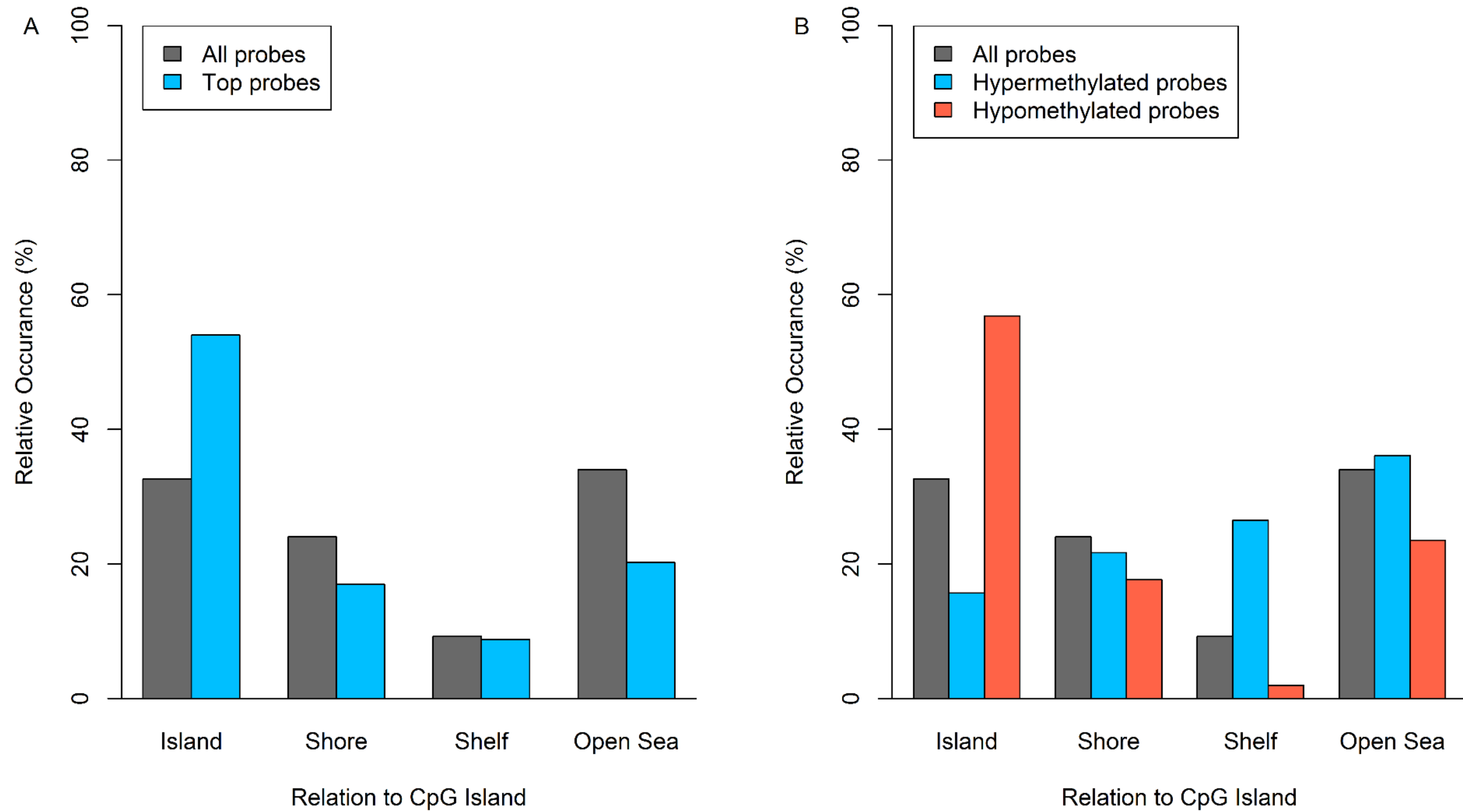

Figure B.1. Relative occurrence of the top differentially methylated loci relative to CpG island location in umbilical artery Legend: (**A**) top unadjusted loci (q-value<0.05) and (**B**) top hyper and hypo methylated CpGs with a nominal $P<1x10^{-4}$ after adjusting for cellular heterogeneity.

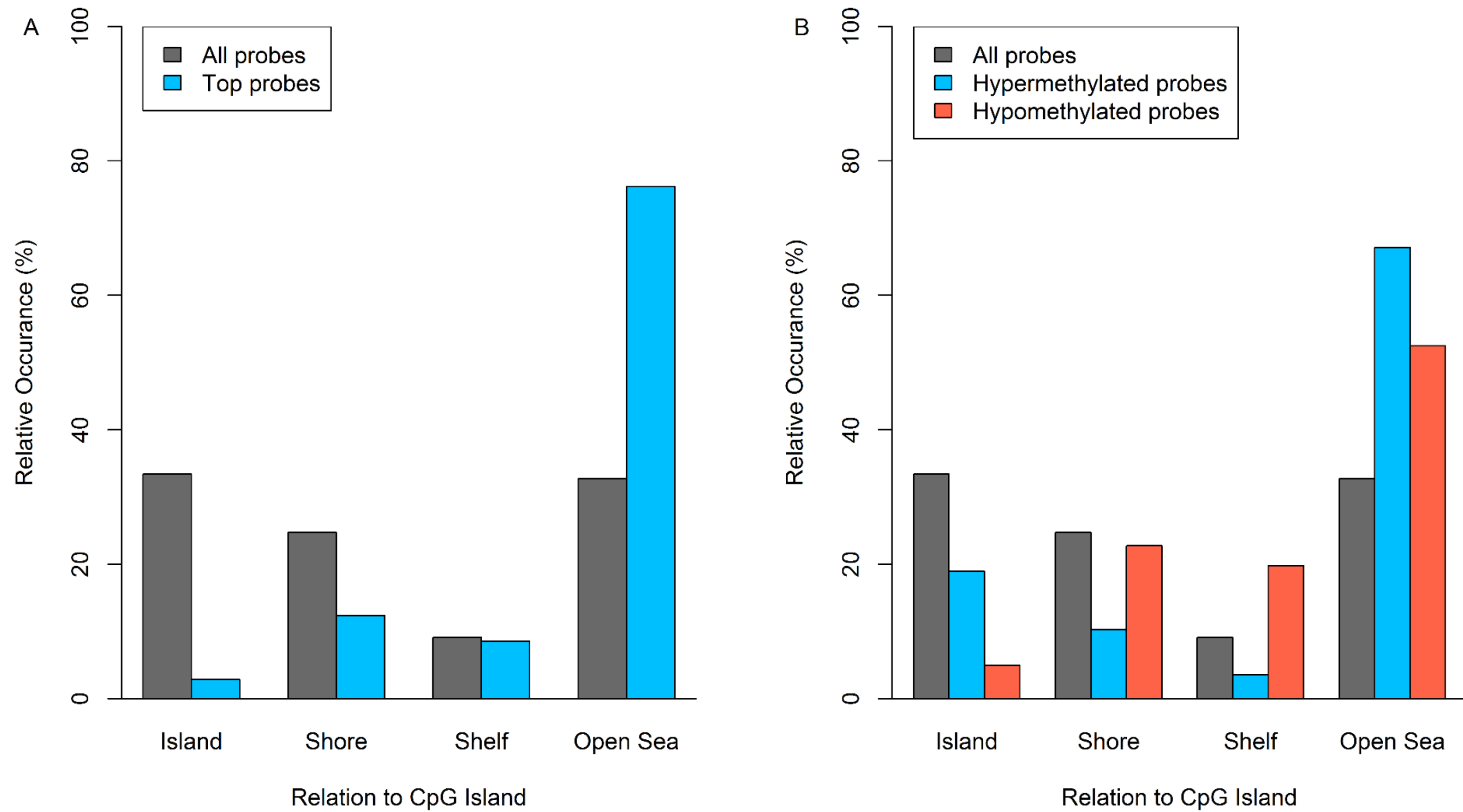

Figure B.2. Relative occurrence of the top differentially methylated loci relative to CpG island location in placenta
Legend: (**A**) top unadjusted loci (P<1x10^{-4}) and (**B**) top hyper and hypo methylated CpGs with a q-value<0.05 after adjusting for cellular heterogeneity.

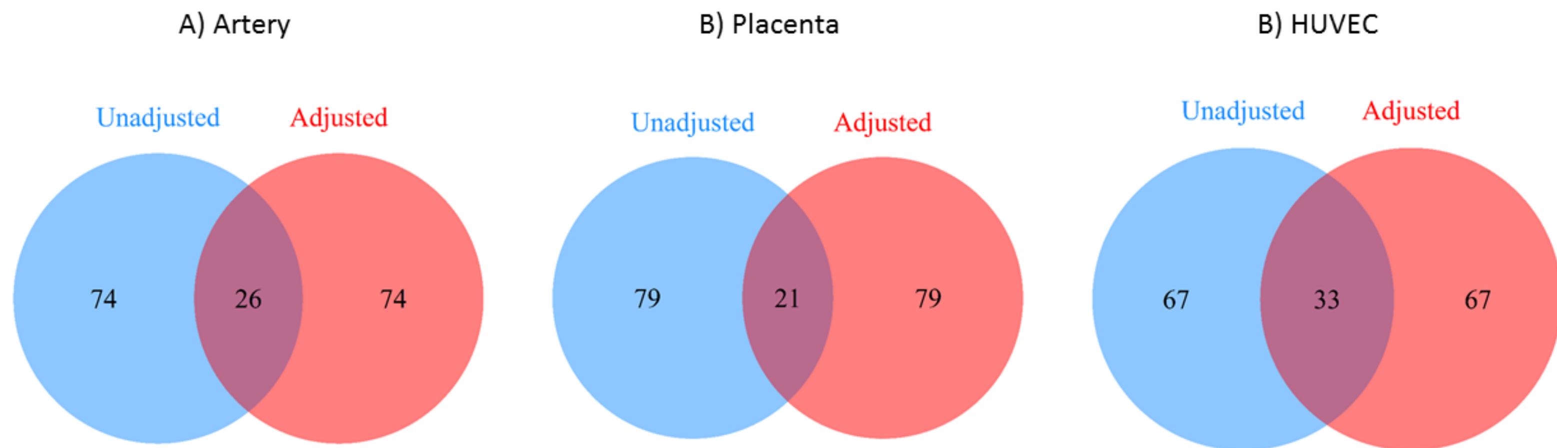

Figure B.3. Overlap for the unadjusted and cell mixture adjusted results among the top 100
Legend: Top-100 CpGs (ranked by lowest *p*-value) that were differentially methylated in: **A**) Artery, **B**) Placenta and **C**) HUVEC

Appendix C: Supplemental Materials for Chapter 4

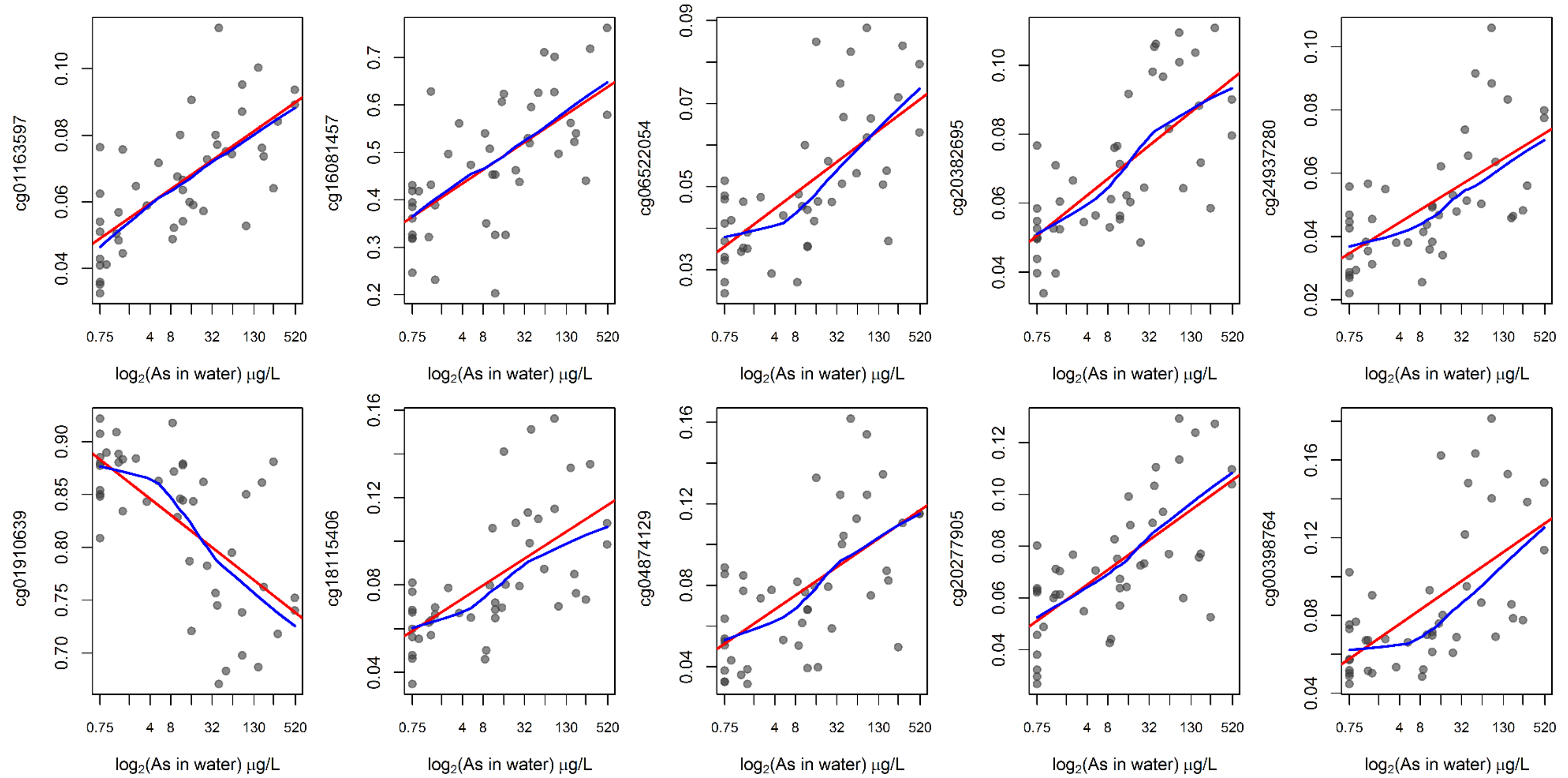

Figure C.1. Association between prenatal As exposure and CpG methylation of selected loci
Legend: red: linear regression line; blue: locally weighted scatter plot smoothing

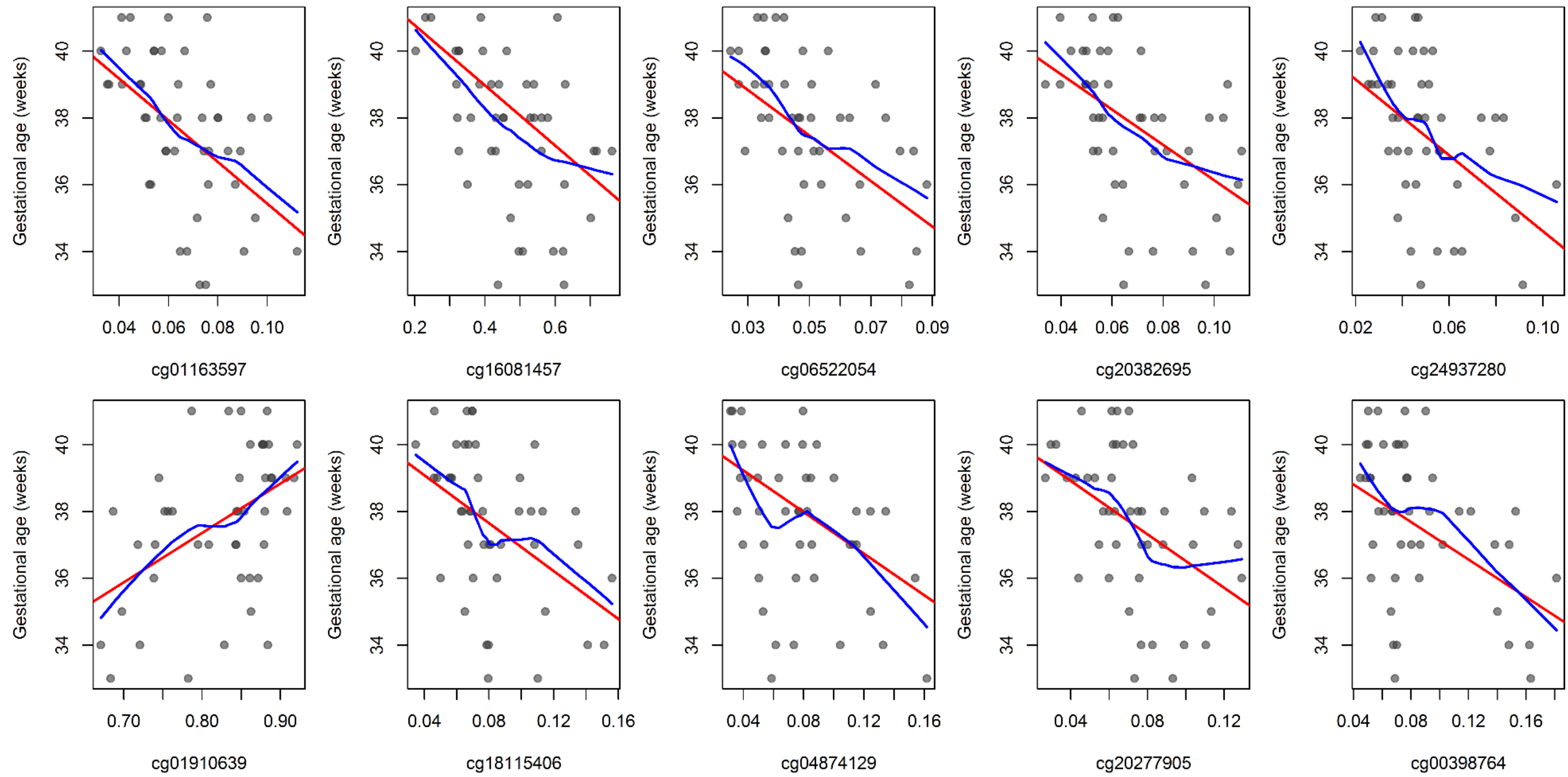

Figure C.2. Associations between birth gestational age and CpG methylation of selected loci
Legend: red: linear regression line; blue: locally weighted scatter plot smoothing

INTENTIONALLY LEFT BLANK