ABSTRACT

HUDSON, KATHLEEN. Epigenetic and Transcriptomic Mechanisms Linking Maternal Cadmium Exposure to Disease Susceptibility in the Offspring. (Under the direction of Dr. Michael Cowley).

Adverse exposures during early-life development can program susceptibility to disease later in life, even long after cessation of exposure. Alterations to epigenetic marks such as DNA methylation are strong candidates for a mechanism that provides this memory of early-life exposures. DNA methylation at Imprinting Control Regions (ICRs) controls the expression of imprinted genes and unlike the vast majority of the genome, ICRs persist in their methylation status throughout development and into adult tissues. Regulation of imprinting is tightly controlled and loss of imprinting is associated with a variety of growth and metabolic-related diseases. We hypothesize that ICRs may be more sensitive to early-life exposures and may provide a mechanism of memory to adverse developmental environments.

Maternal exposure to heavy metals can negatively affect the development and health of the offspring. Cadmium (Cd) is a toxic heavy metal found ubiquitously throughout the environment and is one of the top ten chemicals of major public health concern identified by the World Health Organization. Maternal exposure to Cd is associated with a range of adverse outcomes for the offspring, many of which overlap with the growth and metabolic-related diseases seen as a result of loss of imprinting. We have shown in humans that maternal Cd exposure is associated with DNA methylation changes at ICRs in fetal cord blood, and that these changes may be significantly enriched compared to maternal Cd-induced DNA methylation changes at other similarly structured areas of the genome.

To test the hypothesis that DNA methylation at ICRs may be more susceptible to maternal exposures than other similar but non-imprinted loci, we have established a hybrid mouse model of maternal Cd exposure. Our model facilitates the analysis of allele-specific DNA methylation changes at ICRs and allele-specific changes in gene expression of imprinted genes, which was not feasible in the human study. In addition, we can further identify transcriptomic and other molecular changes in tissues not permissible in humans. In **Chapter 2**, molecular and transcriptomic changes are quantified in maternally-exposed offspring that provide insights into how maternal Cd exposure can impact cardiovascular development and disease. **In Chapter 3**, additional molecular and transcriptomic changes are quantified in maternally-exposed offspring that link maternal Cd exposure to altered neurodevelopment and behavior. Finally, in **Chapter 4**, allele-specific DNA methylation and gene expression changes are quantified to test our driving hypothesis. This thesis shows that maternal Cd exposure alone is sufficient to program perturbed cardiovascular health and behavior in adulthood and identifies multiple pathways that may be responsible for programming these outcomes. Ultimately, these studies provide novel insights into molecular mechanisms that may link early-life adverse exposures to disease susceptibility.

Epigenetic and Transcriptomic Mechanisms Linking Maternal Cadmium Exposure to Disease Susceptibility in the Offspring

by
Kathleen Hudson

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Genetics

Raleigh, North Carolina
2019

APPROVED BY:

Dr. Michael Cowley Committee Chair	Dr. Anna Stepanova
Dr. David Aylor	Dr. Carolyn Mattingly

DEDICATION

To our son, Parker Robert Hudson. Although you came into this world only a short time ago, you have already brought so much joy into our lives. It has been amazing watching you turn into your own person and figure out the world around you. May you continue to be curious, inquisitive, and happy.

BIOGRAPHY

Kathleen Marie Hudson (née Behrend) was born in San Diego, California and grew up in Raleigh, North Carolina. She graduated in 2012 from North Carolina State University (NCSU) with a BS in Animal Science and minors in Nutrition and Microbiology. During her undergraduate degree, she volunteered at the NCSU College of Veterinary Medicine's Vector Borne Disease Diagnostic Laboratory and enjoyed working in a molecular biology laboratory on various projects. After she finished her undergraduate degree, she began working as a laboratory technician in the Molecular Genetics department of a health care diagnostics company. There, she realized her passion for genetics and in 2014 began pursuing her Ph.D. in Genetics at NCSU. In 2015, she began her thesis project under the direction of Dr. Michael Cowley.

ACKNOWLEDGMENTS

There are a number of people I would like to thank for making this dissertation possible. First, I would like to thank my husband for his unconditional love, support, and patience, and for enduring through all my years of school, two career changes, and the birth of our child. You always provided an ear to listen to my grievances, doubts, and joys, and kept me motivated to finish. For my family, who were the first people to teach me about unconditional love and support – for keeping that through the rough times, for your unwavering encouragement, and for teaching me the importance of an education. I would especially like to thank my advisor Michael Cowley for welcoming me into his lab and taking me on as his first graduate student. You always made time to listen to me and have helped me blossom as a scientist. I have enjoyed working with and learning from you. To Anna Stepanova, who gave me the confidence to apply to grad school – I would not be where I am if it was not for your support and encouragement. You continue to push me to do better. To my committee – thank you for always listening to my ideas, providing insights, and asking the questions that needed to be asked. And lastly, I would like to thank the Genetics program faculty members and especially the graduate students, both past and present; it has been a pleasure being a part of your family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

Cadmium

Cadmium (Cd) is a toxic heavy metal found ubiquitously throughout the environment and is one of the top ten chemicals of major public health concern according to the World Health Organization (WHO)[1,2]. Cd occurs naturally in the soil at relatively low levels (average 0.3 ppm)[3] and is released into the environment through both natural and human activities. Natural activities include volcanic eruption, erosion, and water transport; human activities include metal mining and refining, fossil fuel combustion, tobacco smoking, municipal waste incineration, and phosphate fertilizer manufacturing. The amount of Cd in the environment has increased significantly over the past century primarily due to human activities[4]. It was recently estimated that human activities contribute 3,000-7,570 tons of Cd/year into the atmosphere, compared to natural fluxes contributing approximately 1,440 tons Cd/year[3]. Cd has many desirable properties for industrial use and is primarily used in nickel-cadmium rechargeable batteries, which remain widely used to this day owing to their superior properties over other rechargeable technologies[3].

Humans are exposed to Cd through inhalation or ingestion. Cd inhalation occurs most frequently through smoking cigarettes, breathing contaminated dust particles, and occupational exposure, such as in the metal mining or refining industries[3–5]. The tobacco plant accumulates Cd from the soil at a high rate compared to average plant species and is the primary source of Cd in cigarettes[6]. Cd ingestion is much more common than inhalation and occurs through the diet or the consumption of Cd-contaminated water[3–5].

The primary dietary sources of Cd are plant-derived, as plants absorb Cd from the soil. Produce that is recommended as part of a healthy diet, such as leafy greens and whole grains, is often the primary source of dietary Cd. In support of this, vegans have the highest Cd burden among non-smokers[7].

Cd is not an essential trace element; it has no known function in biological systems and is toxic to all branches of life[2,3,8–10]. Some microorganisms and plants have developed a resistance to environmental Cd, but do not appear to require Cd[3,8]. In both prokaryotes and eukaryotes, Cd absorption is mediated by proteins that facilitate certain essential trace metal (e.g., zinc [Zn], iron [Fe], copper [Cu], selenium [Se]) absorption from the diet or environment[7,11–13]. The chemical and physical properties of Cd are similar to Zn and many of the toxic effects of Cd are theorized to stem from interfering with Zn-mediated processes[3]. Cd competes with Zn and other essential trace metals in binding to their respective proteins and Cd exposure can induce deficiencies in these metals, even when they are supplied in adequate levels through the diet. Already existing deficiencies in these essential trace metals can exacerbate Cd toxicity, while supplementing these in the diet can alleviate Cd toxicity to varying degrees[7].

When ingested, only a small percentage of Cd is absorbed into the body (average ~6%), while inhalation results in a much higher rate of absorption (up to 50-60% of inhaled Cd)[14,15]. Females tend to absorb and store more Cd in their bodies than males, likely due to their higher dietary requirement for Fe[7,16]. As Cd is not required and there are no bodily processes that have evolved to respond to Cd exposure, Cd is poorly excreted which leads to long-term storage in the body[12]. The primary sites of Cd accumulation are the liver and the renal cortices of the kidneys[5]. Once absorbed, Cd can persist in the human

body for 30 years or more; humans have approximately four times the Cd burden at age 70 than at age 20[5,7]. Chelation therapy, which can be used to remove certain metals from the blood, has not been shown to significantly lower Cd in the body unless the Cd exposure was acute and recent[7,17].

Cd exposure can lead to a variety of adverse effects that vary based on route of exposure, gender, amount and duration of exposure, dietary intake of essential trace elements, and age[3]. Chronic exposure to high levels of Cd can result in death, kidney disease, or disturb calcium (Ca) metabolism, leading to osteoporosis or softening of the bones (e.g., Itai-itai disease)[4,16]. Chronic exposure to low levels of Cd has gained more attention recently as it is associated with multiple noncommunicable diseases reaching epidemic proportions in the world: type 2 diabetes, obesity, hypertension and cardiovascular disease, non-alcoholic fatty liver disease, and cancer[5,16,18,19]. These types of noncommunicable diseases are responsible for roughly two-thirds of deaths globally[18]. Cd has been designated a carcinogen by several health organizations such as the WHO and United States National Toxicology Program. Both animal and human studies have linked Cd exposure with cancer across multiple organs including lung, prostate, renal, liver, hematopoietic system, bladder, and stomach[12].

There are several proposed mechanisms in which Cd exposure may lead to disease. As previously mentioned, Cd can induce essential trace metal deficiencies that are required as cofactors for various biological processes in the body. In addition, Cd can induce oxidative stress and the resulting reactive oxygen species (ROS) can directly damage cells. As Cd is stored in the liver and kidneys, these organs are the primary targets of Cd-induced ROS damage[3]. The molecular mechanisms behind Cd

carcinogenesis are less clear. Unlike other carcinogenic heavy metals, Cd is not considered to be genotoxic. Other heavy metals can bind directly to DNA or form crosslinks between DNA and proteins in the nucleus; Cd has only been shown to weakly bind DNA and has a much higher affinity for other molecules in the cell[12]. It is suspected that the formation of ROS in response to Cd exposure may play a role in Cd carcinogenesis, but there has been evidence that Cd can induce malignant transformation in the absence of ROS[3]. Cd has been shown to induce aberrant expression of oncogenic genes, but the direct mechanism underlying this effect has not been determined. It has been proposed that Cd exposure alters gene expression through perturbation of epigenetic mechanisms, but the molecular mechanisms linking Cd exposure to epigenetic changes are not clear. Cd-induced carcinogenesis might be multifaceted and may have different mechanisms of action in different tissues. Further research is needed to determine the precise mechanisms underlying Cd-triggered carcinogenesis[12].

In mammals, Cd does not appear to efficiently cross the placenta and enter fetal circulation[3,15]. However, maternal exposure to Cd can have negative consequences to the fetus including growth restriction, essential trace element deficiencies, congenital malformations, reproductive system impairments, altered behavior and intelligence, and other long-term consequences[3,15,20–22]. Early-life exposures can lead to disease susceptibility later in life; this is encompassed in the theory known as the Developmental Origins of Health and Disease and will be discussed next.

Developmental Origins of Health and Disease

Exposure to various environmental stressors during early life can impact the health of the individual later in life; this is the principle of the Developmental Origins of Health and Disease (DOHaD) hypothesis, a theory based on epidemiological studies performed by David Barker during the 20th century. Barker found that there was a high correlation within geographical areas for infant mortality and adult deaths due to specific diseases, as well as a correlation between low birthweight and adult death due to heart disease. He proposed that fetal undernutrition leads to reprogramming of glucose, insulin, and growth factor homeostasis that "permanently changes the body's structure, function and metabolism", ultimately programming an increased risk for heart disease during adulthood[23]. The DOHaD hypothesis has become a widely accepted theory that continues to gain supportive evidence through both human epidemiological studies and animal models of exposure. Maternal exposure to heavy metals, endocrine disrupting chemicals, and over- or undernutrition have all been shown to negatively affect the future health of the offspring[24–26].

Heavy Metals

Heavy metals in the environment that are among the WHO's top ten chemicals of major public health concern include Cd, lead (Pb), arsenic (As), and mercury (Hg)[1]. The levels of these toxic heavy metals in the environment, and therefore human exposure, have increased dramatically due to human activities. Humans are often exposed to multiple heavy metals in the environment and the combination may produce effects that are additive, antagonistic, or synergistic[27]. Studies have historically focused on the effects of adult exposure to these metals, but the past several decades have indicated that

maternal heavy metal exposure can negatively impact fetal health as well. Maternal exposure to Pb has been associated with preterm delivery, fetal growth restriction, and lasting neurodevelopmental abnormalities[27]. Prenatal exposure to As is linked to increased risk of stillbirth, child mortality, and congenital heart defects[28,29]. Hg is a potent neurotoxicant and early-life exposure can lead to lasting neurological and behavioral disorders in the offspring; lower levels of exposure have also been associated with impaired cardiovascular and immune function[30]. Pb, As, and Hg have all been shown to cross the placenta effectively and therefore fetal exposure is direct. However, Cd does not efficiently cross the placental barrier. Instead, Cd accumulates in the placenta and can alter essential trace metal transfer to the fetus, interfere with glucocorticoid homeostasis, increase oxidative stress, and alter placental function[31].

Endocrine Disrupting Chemicals

Endocrine disrupting chemicals (EDCs) are naturally-occurring or man-made chemicals that affect the endocrine system and lead to adverse development, reproduction, immune, or neurological responses. EDCs include pharmaceuticals, plasticizers such as bisphenol A (BPA), certain pesticides, and naturally-occurring phytoestrogens such as those found in soy-derived products[32,33]. Within the past decade, Cd has been classified as a metallohormone - a metal that acts as an EDC[34]. EDCs can be found in the environment as well as in a variety of everyday products – bottles, cans, detergents, foods, toys, cosmetics, and flame retardants. Exposure to EDCs during prenatal and early postnatal life may pose the greatest risk due to their effects on organ development[33]. Early-life exposure to EDCs has been associated with reduced fertility, endometriosis, cancer, obesity, metabolic disease, and other chronic noncommunicable

diseases. Cd and other EDCs exert their effects by mimicking naturally occurring sex hormones (e.g., estrogens, androgens), binding to sex hormone receptors to block endogenous hormone production, or inhibiting sex hormones or their receptors from being formed or regulated[33,35].

Maternal Diet

It is well established that proper diet and nutrition are essential to human health. The past two centuries of nutrition science have led to established minimal daily requirements (MDRs) of macronutrients (fat, protein, carbohydrate) and micronutrients (vitamins, minerals) required for humans according to various factors and life stages such as age, sex, physical activity level, and pregnancy. MDRs increase up to 50% for a woman during pregnancy and proper nutrition during fetal development is essential to prevent chronic, noncommunicable diseases through adulthood[36,37]. Single micronutrient deficiencies can lead to severe fetal defects such as cretinism (iodine), neural tube defects (folic acid), night blindness (vitamin A), and stunted growth (Fe)[38]. Malnutrition of multiple macro- and micronutrients has been associated with increased risk of obesity, cardiovascular disease, osteoporosis, type 2 diabetes, and asthma[23]. Maternal undernutrition remains a concern to women living in low-income areas due to their chronically poor diet. Micronutrient deficiencies are less common in developed countries, but may be rising due to higher consumption of the Western diet (high fat, high sugar, high salt, nutrient poor)[38]. In contrast to undernutrition, maternal overconsumption or overnutrition has become a concern in developed countries. Maternal intake of excess macronutrients, sugars, and salts mirrors the increased risks of obesity, cardiovascular disease, and type 2 diabetes seen with maternal undernutrition[23,36]. As previously stated,

maternal diet can also be a source of prenatal Cd exposure through contaminated food; this affects women globally in both low-income and developed countries.

Conclusions

As described in this section, prenatal and early postnatal exposures can lead to susceptibility to disease states later in life; these disease states often happen long after cessation of exposure. Additional environmental exposures after birth or throughout life can activate or exacerbate the disease state – this is known as the "First Hit/Second Hit Framework", where the "first hit" is the gestational exposure and the "second hit" is the environmental insult(s) experienced postnatally by the offspring such as under/overnutrition, social stress, or other environmental toxicants[39]. It has also been shown that some environmental exposures can increase disease risk across multiple generations[40]. These features indicate that there is some type of programming of susceptibility to disease that happens early in life. Epigenetics is a strong candidate for the molecular mechanism behind this programming and will be discussed next.

Epigenetics

The term 'epigenetics' was first used in 1942 by Conrad Waddington to describe changes in a phenotype without changes in the underlying genotype. Epigenetics was initially used to explain features of development for which there was little understanding, such as the differentiation of cells and tissues[41]. The past two decades have been defined as the "modern era of epigenetic research"; we have gained a much better understanding of the molecular mechanisms underlying this phenomenon as it pertains to development and beyond due to the discovery of key epigenetic enzymes and molecules. Epigenetics

is more recently defined as the regulation and/or inheritance of gene expression without alterations to the underlying sequence of DNA. The molecular mechanisms currently encompassed in epigenetics include chromatin state, histone modifications, non-coding RNA (ncRNA), and DNA methylation[42].

Chromatin State and Histone Modifications

Each diploid human cell contains approximately 2 meters of DNA that needs to be packaged into a nucleus that is only a few microns in diameter[43,44]. To accomplish this in human [and other eukaryotic] cells, genomic DNA (gDNA) is wound around proteins called histones and further organized into multiple levels of coils, loops, and folds of chromatin[44]. Each histone is an octamer made of two of each of the H2A, H2B, H3, and H4 proteins. Histones are positively charged and attract negatively charged DNA through nonspecific electrostatic interactions. Around 146 base pairs (bp) of gDNA is wound around one histone octamer to form a nucleosome. gDNA between histones is linked by one H1 protein. The nucleosome is the basic repeating unit of chromatin and due to its appearance under electron microscopy, chromatin has been described as “beads on a string”[44]. Chromatin can be characterized as either heterochromatin or euchromatin. Heterochromatin is tightly packed, transcriptionally inactive areas of the genome, while euchromatin is less condensed and therefore more accessible to the cell’s transcriptional machinery[45]. Due to the non-specific nature of affinity between histones and gDNA, the histones can be repositioned along the gDNA to form euchromatin and allow access to transcriptional machinery[44]. ATP-dependent chromatin remodeling is one mechanism that facilitates this process. The ATP-dependent chromatin remodeling complex

SWI/SNF was first identified in yeast, but roughly two dozen chromatin remodeling complexes have been found in mammals[46].

In addition to the general nonspecific interaction between positively charged histones and negatively charged gDNA, histones can have a vast number of reversible posttranslational modifications (PTMs) that influence gene expression of the gDNA in proximity. The most common histone PTMs are methylation, acetylation, and phosphorylation, but ubiquitylation, sumoylation, glycosylation, carbonylation, 2-hydroxyisobutyrylation, crotonylation, and ADP-ribosylation have also been described[47,48]. To date, over 200 specific PTMs have been defined[47]. Distinct modifications found on specific histone peptide locations are associated with transcriptionally active or silent chromatin. Active gene expression is typically characterized by H3 and H4 lysine acetylation, H3 lysine 4 trimethylation, H3 lysine 79 trimethylation, H2B ubiquitylation, and H3 lysine 36 trimethylation. Histone variants have also been identified (e.g., H3.3 and H2A.Z) and are enriched over transcriptionally active areas. Transcriptionally silent areas of the genome are often characterized by lysine 27 trimethylation, H2A lysine 119 ubiquitylation, and H3 lysine 9 trimethylation[49]. There are numerous combinations of histone PTMs that can occur together, and PTMs that are associated with either gene expression activation or repression can occur simultaneously in “poised” or bivalent chromatin[49,50]. Bivalent chromatin is often found in promoters of key developmental genes in pluripotent cells, allowing the cells to respond quickly to signaling cues during development[50]. Bivalent chromatin has also been implicated as a mechanism underlying the programming of disease susceptibility due to adverse environmental exposures; promoters of genes associated with disease states may be

“poised” as a result of the first hit of exposure in order to respond to the second hit of exposure during adulthood[51].

Non-coding RNA

ncRNA can be defined as RNA that is transcribed from gDNA, but not translated into a protein; this goes against the original belief of the Central Dogma of Biology: one gene gives rise to one messenger RNA molecule (mRNA) which produces one protein. Once thought of as “noise”, ncRNA has been shown to have diverse regulatory functions, and the genes encoding ncRNA outnumber the genes encoding proteins[52]. There are numerous types of ncRNAs encompassing a variety of functions in the cell[53]; however, two types of ncRNA commonly discussed in epigenetics, as they have a role in gene expression regulation, are microRNAs (miRNAs) and long non-coding RNAs (lncRNAs). miRNAs are short (~22 nucleotides in length) ncRNAs that generally act in posttranscriptional gene silencing of mRNAs, although there are a few examples of miRNAs acting to upregulate mRNA translation. miRNAs typically base pair imperfectly to the target mRNA and signal that mRNA to be silenced post-transcriptionally through one of several mechanisms, e.g., degradation[52]. lncRNAs are much longer than miRNAs (>200 nucleotides in length) and have various roles including regulation of chromatin dynamics and gene expression. One of the most well-characterized lncRNAs is *XIST*, which controls the X-inactivation of one of the two X chromosomes in mammalian females. lncRNAs can also act as decoys for transcription factors or repressors by interacting with these proteins and preventing them from binding to their target promoters[54]. Recently, lncRNAs have been implicated as novel biomarkers of responses

to environmental exposures, including Cd exposure. Dysregulation of lncRNAs has been associated with many of the chronic diseases discussed in the previous DOHaD section[55].

DNA Methylation

The most common form of DNA methylation in mammals is the covalent addition of a methyl group to the 5' position of the nucleotide cytosine (5mC for 5'-methyl cytosine). DNA methylation typically occurs on cytosines that are immediately preceding a guanine nucleotide (CpG) in the 5' to 3' direction[47,56]. Roughly 70-80% of all CpGs are methylated in mammalian somatic tissues. Some regions of the genome contain CpG-rich sequences called CpG islands (CGIs). CGIs are commonly found upstream of gene promoters and are typically hypomethylated, in stark contrast to the majority of CpG sites. It is estimated that roughly 60% of human genes contain CGIs in their promoters[56]. In general, DNA methylation is associated with gene silencing[47,56]. DNA methylation can interfere with proteins, e.g., transcription factors, binding to their consensus sequence in DNA. Through this, DNA methylation adds an additional layer of information to DNA that is not encoded in the actual DNA sequence itself[57].

DNA methyltransferases (DNMTs) are the enzymes responsible for adding the methyl group in response to cellular stages or signals. DNMT1 recognizes hemimethylated gDNA that occurs during cell division and methylates the newly formed daughter strand of DNA according to the methylated positions on the parent strand[47]. In this manner, DNA methylation patterns can be stably inherited, although they are not necessarily permanent; they can be added or removed throughout life[56]. While DNMT1 is responsible for the maintenance of DNA methylation, *de novo* DNA methylation occurs through the activity of DNMT3A and DNMT3B; this can occur during routine cellular

differentiation, disease states such as cancer, or in response to environmental exposures such as Cd[57–60]. DNMTs have also been found to directly interact with chromatin through binding to certain histone PTMs, just one example of the complexity of epigenetic marks and their cross-talk[57].

DNA methylation can be removed through active demethylation mediated by TET proteins, or through passive demethylation involving DNMT1 inhibition or absence during replication. TET-mediated demethylation occurs through the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) and other 5mC oxidized derivatives. TET-mediated 5mC oxidation can be followed by thymine-DNA glycosylase (TDG)-mediated excision of the 5mC oxidized derivatives and base excision repair to facilitate active DNA demethylation. Oxidized 5mC derivatives cannot be recognized by DNMT1 and thus leads to passive demethylation during cell division[56].

In mammals, gDNA methylation patterns undergo dynamic changes during early development. During fertilization, gametes containing a haploid maternal or paternal methylated genome fuse to form a zygote. After fertilization, there is a wave of demethylation across the maternal and paternal genomes. Shortly after implantation, *de novo* DNA methylation occurs; this coincides with the occurrence of the inner cell mass cells beginning to differentiate into the three germ layers. *De novo* DNA methylation patterns are then maintained during replication according to their cell lineage[56,57,61].

Select areas of the genome, such as intracisternal A particles, variably erased CGIs, and imprinting control regions (ICRs), are resistant to the post-fertilization wave of demethylation[61] and could provide a mechanism of memory to early-life exposures. ICRs are regions that are differentially methylated based on the chromosomal parent of origin,

e.g., hypomethylated on the paternal chromosome and hypermethylated on the maternal chromosome[62]. ICRs control the allele-specific expression of imprinted genes, which will be discussed next.

Genomic Imprinting

Genomic imprinting is a phenomenon observed in mammals and flowering plants in which there is biased expression from one allele of a gene in a parent-of-origin-dependent manner. This system of gene regulation is dependent upon differential DNA methylation patterns on the maternally- and paternally-inherited alleles, which are established during gametogenesis[63]. The methylation profiles of these ICRs persist through fertilization, development, and the lifespan of the organism, enabling discrimination between the parental alleles. DNMT3L is the third member of the DNMT3 family and lacks the catalytic domain found in DNMT3A and DNMT3B, but has been shown to be essential for establishing methylation in the germline via knockout studies. In addition to other DNMTs, DNMT3L contains a cysteine-rich zinc finger domain (ADD domain), which facilitates binding to various moieties such as DNMT3A and DNMT3B, histone PTMs, and transcription factors. DNMT3A and DNMT3B interact with DNMT3L through their C-terminal domains and this interaction enhances their DNA methylation activity by preventing self-polymerization[57]. Roughly 150 imprinted genes (both ncRNA and protein-coding) have been identified to date, comprising a very small percentage (<1%) of the genes identified thus far in mice and humans[64–67]. Imprinted genes are typically found in clusters, in which one ICR controls the expression of multiple imprinted genes[68]. An example of an ICR is illustrated in Figure 1.1. Most imprinted genes have

roles in essential biological processes during development such as cell division and proliferation, fetal growth, metabolism, placental growth, and neurodevelopment[66]. Misregulation of imprinting can have drastic impacts on human health as seen in Prader-Willi, Angelman, Beckwith-Wiedemann, and Silver-Russell syndromes[68].

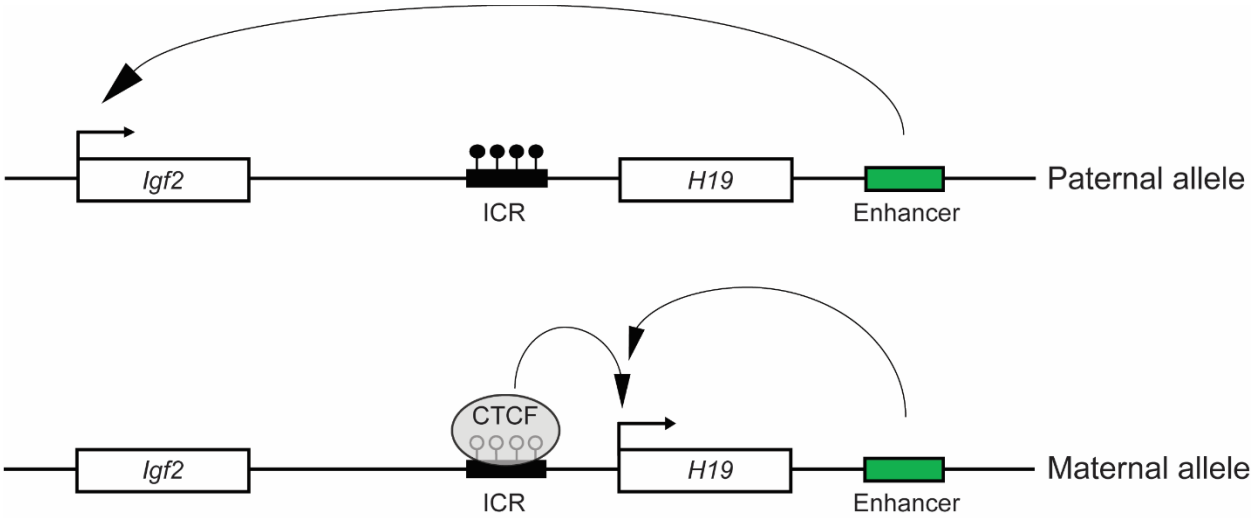


Figure 1.1. *H19/Igf2* ICR. The ICR is methylated on the paternal allele, which leads to expression of *Igf2* and silencing of *H19*. The ICR is hypomethylated on the maternal allele, which permits CTCF to bind and leads to expression of *H19* and blocks the enhancer from activating expression of *Igf2*.

As previously stated, early-life exposures can induce susceptibility to disease states later in life, and epigenetics is suggested to be one molecular mechanism underlying this programming effect. Epigenetic profiles established or altered as a result of an early-life exposure would need to remain stable through development and into adulthood in order to program a disease state in later life. The epigenetic state of many loci is dynamic over time, but ICRs are potential candidates for providing an epigenetic memory of early-life exposures due to their stability throughout the lifetime of the organism. In humans, maternal Cd exposure has been associated with altered DNA

methylation in the offspring and that among the differentially methylated regions (DMRs), ICRs may be more sensitive to maternal Cd exposure than similarly structured non-imprinted loci[58,69].

The aim of this thesis was to test the hypothesis that ICRs are susceptible to early-life exposures and to investigate the effects of these exposures on programming disease states. Maternal Cd exposure was chosen as the exposure model due to its significance to human health and preliminary data in humans showing ICRs are significantly enriched among maternal Cd-associated DMRs. Testing this hypothesis and exploring molecular mechanisms perturbed as result of maternal Cd exposure is not feasible in humans. There are many advantages to using the mouse as a model organism for humans, which will be discussed next.

Mouse as a Model

In an effort to better understand human genetics, the Human Genome Project launched in 1990 and listed the mouse as one of the five central model organisms to use in order to better understand human biology and disease. Humans have been using mice for scientific observations for millennia; however, the past century has yielded a plethora of murine resources in response to the post-Mendelian era of molecular genetics[70]. Of the five central model organisms, the mouse is the closest genetically and physiologically to humans[70–72]. Mice and humans are both placental mammals and share many of the same developmental processes during gestation, including genomic imprinting (See Chapter 4 Table 4.1). Mice have a much shorter gestation than humans (~20 days vs 9 months) and produce multiple animals per pregnancy, making them a much more feasible

model to study the effects of maternal exposures[72]. There are many commercially available inbred strains of mice that have either sequenced reference genomes or known areas of genomic variation[73]. A phenotype can be broadly defined as the sum of the portions resulting from an individual's genotype and the amount due to environmental conditions[74]. When looking at the effect of an environmental exposure, using inbred mice will allow for control of the genetic contribution to phenotype because all of the individuals are genetically homozygous at all loci – any resulting variation in phenotypic response can be attributed to the variation in environment.

In this thesis project, two genetically divergent inbred strains of mice (C57BL/6J and CAST/EiJ) with fully sequenced reference genomes were used as parental strains in order to generate hybrid offspring that were maternally exposed to Cd. This experimental design was employed to facilitate the analysis of allele-specific changes in DNA methylation (e.g., at ICRs) and gene expression (e.g., of imprinted genes) in the offspring due to maternal Cd exposure. In addition to studying the effect of maternal Cd exposure on ICRs and imprinted gene expression, this thesis also identifies molecular pathways in the brain and heart perturbed as a result of maternal Cd exposure, and shows maternal Cd exposure alone can program lasting behavioral and cardiovascular impairments in the offspring during adulthood.

CHAPTER 2

Maternal Cadmium Exposure Perturbs Essential Trace Element Homeostasis and Causes Transcriptome-Wide Changes in the Newborn Mouse Heart Associated with Cardiac Hypertrophy and Susceptibility to Hypertension in Adulthood

The following chapter is the first draft of a publication that will be submitted to a peer-reviewed journal.

Abstract

Cadmium (Cd) is one of ten chemicals of major public health concern identified by the World Health Organization. Maternal exposure to Cd can negatively impact fetal health by inducing fetal growth restriction, trace element deficiencies, and congenital malformations. Cd exposure during adulthood has been associated with cardiovascular disease (CVD) in humans; however, the effects of maternal Cd exposure on cardiovascular development and disease in offspring are not well understood. To address this, we have established a mouse model of maternal Cd exposure. We show that offspring born to Cd-exposed mothers present with cardiac hypertrophy at birth and that maternal Cd exposure programs hypertension during adulthood in a subset of animals. Despite undetectable levels of Cd in the fetus, maternal Cd exposure significantly alters fetal levels of essential trace elements including a deficiency in iron, an element required for cardiovascular system development, oxygen homeostasis, and cellular energy metabolism. Transcriptomic analysis of newborn hearts identifies differentially expressed genes associated with maternal Cd exposure which are enriched for functions in CVD, hypertension, hypertrophy, iron homeostasis, cellular energy, and hypoxic stress. We propose that maternal Cd exposure induces iron deficiency in the offspring, leading to altered cellular metabolic pathways and hypoxic conditions during fetal development. We further propose that this stress may contribute to the cardiac hypertrophy observed at birth and the programing of susceptibility to hypertension during adulthood. Results from these studies will give insights into potential mechanisms through which maternal Cd exposure impacts cardiovascular development and disease.

Introduction

The environment during development can impact health in later life, a concept called developmental programming[23]. Maternal exposure to heavy metals[24], endocrine disruptors[25], and over- or under-nutrition[26] during gestation have all been shown to negatively affect the development and future health of offspring, even after cessation of exposure.

Cadmium (Cd) is a toxic heavy metal found throughout the environment and is designated as one of the ten chemicals of major public health concern by the World Health Organization (WHO)[1,2]. The amount of Cd in the Earth's environment has dramatically increased over the last century due to human activities including fossil fuel combustion, metal mining and refining, tobacco smoking, and municipal waste incineration[4]. Humans are exposed to Cd through inhalation or ingestion. Cd inhalation occurs predominantly through smoking cigarettes, but contaminated dust and occupational exposure are additional sources. Cd ingestion is the main source of exposure among non-smokers and occurs through the consumption of contaminated water or food[8-9], particularly plants that absorb Cd from the soil. Produce that is recommended as part of a healthy diet especially during pregnancy, such as whole grains and leafy greens, is often the primary source of dietary Cd[75].

Chronic exposure to Cd during adulthood has been shown to negatively impact cardiac health[19,76,77]. Several epidemiological studies have demonstrated an association between blood and urinary Cd levels and cardiovascular disease (CVD), which is the leading cause of death and disability in the US[19,78]. Cd exposure increases oxidative

stress[24], a critical event in the pathogenesis of hypertension[79,80]. Uncontrolled hypertension and smoking are two of the three risk factors for CVD[81].

While the association between Cd exposure during adulthood and CVD has been well demonstrated in human cohorts and animal studies, only a few studies have examined the impacts of early-life Cd exposure on cardiac development and the programming of long-term cardiovascular health. Maternal exposure to Cd has been associated with fetal congenital heart defects in humans[29] and altered heart morphology and endothelial function during adulthood in rats[24]. We recently showed in a human cohort that higher maternal blood Cd levels during pregnancy are associated with changes to cord blood DNA methylation at genes involved in cardiometabolic function[58], suggesting that maternal Cd exposure could affect their regulation and potentially influence cardiac development and health. However, it has not been established in humans or rodents whether these molecular and physiological changes are associated with impaired cardiac function or biomarkers of CVD.

To address this outstanding question, we have established a mouse model of maternal Cd exposure. Using this model, we show that offspring born to Cd-exposed mothers present with cardiac hypertrophy at birth. Additionally, we show that maternal Cd exposure alone is sufficient to program hypertension in a subset of offspring during adulthood. To inform on molecular mechanisms that may contribute to the observed phenotypes, we identify differentially expressed genes and altered essential trace element levels associated with maternal Cd exposure.

Materials and Methods

Animal husbandry, Cd exposure, and tissue collection

C57BL/6J (herein referred to as 'B') and CAST/EiJ ('C') mice were obtained from the Jackson Laboratory. Animals were maintained on a 14-hour light, 10-hour dark cycle in Green Line IVC Sealsafe cage housing systems (Tecniplast) and fed LabDiet 5001 rodent diet (Granville Milling Company) ad libitum. 5- to 7-week old female B or C mice (the F_0 generation) were provided unrestricted access to filtered drinking water (Millipore RiOs Essential RO water purification system) containing 0 ppm, 1 ppm, or 50 ppm Cd in the form of $CdCl_2$ (Sigma-Aldrich, catalog number 202908). After 5 weeks, females were mated with previously unexposed males of the opposite strain and of similar age. Cd exposure continued throughout mating and pregnancy, but was discontinued once a litter was observed (Figure 2.1A). Cd exposure was not found to have an effect on water consumption or female body weight during the 5-week exposure window prior to mating (data not shown).

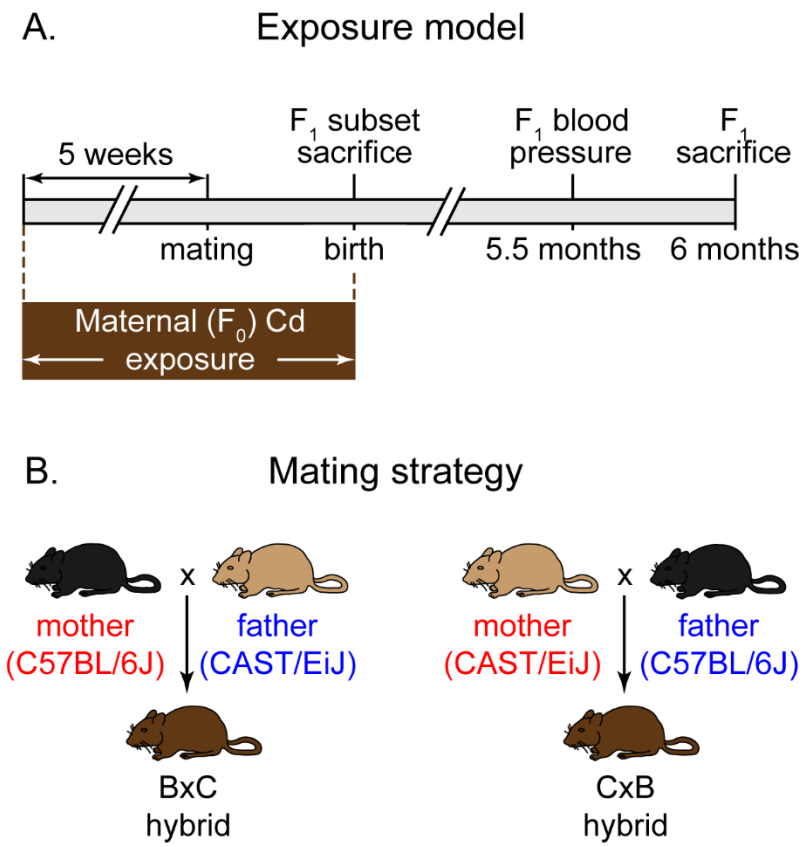


Figure 2.1. Maternal Cd exposure model and mating strategy to produce maternally-exposed hybrid F_1 offspring. **A,** Exposure model. **B,** Mating strategy. B=C57BL/6J, C=CAST/EiJ. F_1 hybrids are designated as maternal x paternal strain.

Hybrid offspring (the F_1 generation) were generated to enable analyses of allele-specific gene expression and DNA methylation for a separate study (KH, MC, manuscript in preparation). The genotypes of F_1 mice are herein referred to as 'BxC' (B mother x C father) and 'CxB' (C mother x B father) (Figure 1B). F_1 animals born to dams exposed to 0 ppm, 1 ppm, or 50 ppm Cd were sacrificed and dissected within 24 hours of birth or at 6 months of age. Tissues were weighed, flash-frozen immediately, and stored at -20°C (for ICP-MS) or -80°C (for all other analyses).

F_1 generation sample sizes for pups dissected within 24 hours of birth were as follows: 0 ppm Cd: 50 BxC females representing 13 litters, 50 BxC males representing 13

litters, 9 CxB females representing 4 litters, 19 CxB males representing 6 litters; 1 ppm Cd: 38 BxC females representing 10 litters, 42 BxC males representing 10 litters, 10 CxB females representing 6 litters, 18 CxB males representing 7 litters; 50 ppm Cd: 53 BxC females representing 12 litters, 40 BxC males representing 12 litters, 5 CxB females representing 3 litters, 10 CxB males representing 4 litters.

F_1 generation sample sizes for pups dissected at 6 months of age were as follows: 0 ppm Cd: 29 BxC females representing 7 litters, 20 BxC males representing 8 litters, 12 CxB females representing 4 litters, 6 CxB males representing 4 litters; 1 ppm Cd: 9 BxC females representing 3 litters, 11 BxC males representing 3 litters, 6 CxB females representing 3 litters, 7 CxB males representing 3 litters; 50 ppm Cd: 19 BxC females representing 5 litters, 16 BxC males representing 5 litters, 4 CxB females representing 2 litters, 4 CxB males representing 3 litters.

All animal work was approved by the North Carolina State University (NCSU) Institutional Animal Care and Use Committee, under protocol 16-045-B. All experiments were conducted in accordance with the Guiding Principles in the Use of Animals in Toxicology.

Blood pressure measurements

Blood pressures were measured on 5.5-month-old F_1 mice by a noninvasive tail-cuff method using a CODA 2-channel system (CODA-HT2, Kent Scientific) as previously described[82].

Numbers of F_1 mice used to obtain blood pressure measurements were as follows: 0 ppm Cd: 13 BxC females representing 7 litters, 9 BxC males representing 5 litters, 11 CxB females representing 4 litters, 6 CxB males representing 4 litters; 1 ppm Cd: 6 BxC

females representing 3 litters, 7 BxC males representing 3 litters, 6 CxB females representing 3 litters, 7 CxB males representing 3 litters; 50 ppm Cd: 10 BxC females representing 5 litters, 10 BxC males representing 5 litters, 4 CxB females representing 2 litters, and 4 CxB males representing 3 litters.

Mice were brought to the testing room and remained in their cages for 30 minutes in order to acclimate to the room. Animals were restrained on a heating pad and covered with a towel to maintain body temperature. Once restrained, animals were left in the restraint for 5 minutes before measurements were taken. Blood pressure data used for analysis were taken on the fourth or fifth consecutive day of testing. Only readings in which a heart rate was detected were used for data analysis.

Serum collection, serum protein quantification, and protein carbonyl ELISA

Trunk blood of 6-month-old F_1 mice was collected after cervical dislocation and decapitation in 1.7 mL microtubes (Genessee Scientific, 22-281), allowed to clot for at least 15 minutes at room temperature, then centrifuged at 4°C for 15 minutes at 2000 rcf to separate the serum. Serum was collected and transferred to a clean 1.7 mL microtube, then stored at -80°C until analysis. The following number of female BxC serum samples were subjected to the bicinchoninic acid (BCA) and protein carbonyl ELISA assays: 0 ppm: 8 females representing 5 litters; 1 ppm: 8 females representing 3 litters; and 50 ppm: 8 females representing 4 litters.

Serum was thawed on ice, then subjected to a BCA assay to determine protein concentration (Pierce BCA Protein Assay Kit, Thermo Scientific, 23227). Briefly, serum was diluted 1/30 in 1X autoclaved PBS (Quality Biological, A611-E404-119) to be within the working range of the BCA kit. Standards for the BCA kit were also diluted in 1X

autoclaved PBS. 3 µL of the standards and serum samples were each tested in triplicate on a Corning Falcon polystyrene 96-well plate (Fisher Scientific, 08-772-3B) according to the BCA kit instructions. Absorbance at 562 nm was measured on a FLUOstar Omega microplate reader (BMG Labtech). Protein concentration was calculated according to the BCA kit instructions.

The 1/30 diluted serum samples were further diluted in 1X autoclaved PBS to reach a final concentration of 10 µg/mL, then subjected to a protein carbonyl ELISA according to the manufacturer's instructions (OxiSelect™ Protein Carbonyl ELISA Kit, Cell Biolabs, STA-310). Absorbance at 450 nm was measured on a FLUOstar Omega microplate reader (BMG Labtech).

Quantification of elements in fetal and adult tissues

Element abbreviations: Cd: cadmium; Co: cobalt; Cu: copper; Mn: manganese; Mo: molybdenum; Se: selenium; Zn: zinc; Ca: calcium; Fe: iron; K: potassium; Mg: magnesium; Na: sodium; P: phosphorus; S: sulfur.

10-30 mg of blood or liver tissue collected during dissections and stored at -20°C was used for this purpose. Samples were digested in high-purity nitric acid (10 µL nitric acid per 1 mg sample, Fisher Scientific, A467-500) overnight at room temperature in 1.7 mL microtubes. Samples were vortexed thoroughly, digested further if needed, and then submitted for analysis through the NCSU Environmental and Agricultural Testing Service Laboratory. Trace elements were determined by inductively coupled plasma mass spectrometry (ICP-MS) using a Perkin Elmer Elan DRCII, while macro elements were determined by ICP-Optical Emission Spectrometry (ICP-OES) using a Perkin Elmer ICP-OES Model 8000. Trace elements (ICP-MS) were run under standard conditions and

collected data were normalized to non-analyte internal standards (115In or 103Rh) to correct for small differences due to instrumental drift, sample/standard solution matrices, and sample transport. A laboratory-check standard of a different stock solution was prepared and used to verify the calibration standard solution (RSD +/- 10%). A continuing laboratory check standard was analyzed every 12 samples, and the calibration curve was repeated if the value was +/- 10%. Samples were run against a 5 or 6 point multi-calibration curve with a linearity $R^2 \geq 0.999$. Macro elements run on the ICP-OES were run using a meinhardt nebulizer/cyclonic spray chamber using the same quality assurance/quality control procedure above except without internal standards. Multi-calibration standard solutions were purchased through Spex Certiprep or Inorganic Ventures, Inc. and used within the expiration date. Detectable limits of all elements tested are listed below.

Method detection limits (reporting limits):

ICP-MS

Cd: 0.050 µg/L (1.0 µg/L)

Co: 0.050 µg/L (0.5 µg/L)

Cu: 0.080 µg/L (0.8 µg/L)

Mn: 0.050 µg/L (0.5 µg/L)

Mo: 0.1 µg/L (1.00 µg/L)

Se: 0.1 µg/L (1.5 µg/L)

Zn: 0.05 µg/L (1.0 µg/L)

ICP-OES

Ca: 0.020 mg/L (0.1 mg/L)

Fe: 0.005 mg/L (0.050 mg/L)

K: 0.020 mg/L (0.2 mg/L)

Mg: 0.020 mg/L (0.1 mg/L)

Na: 0.020 mg/L (0.1 mg/L)

P: 0.050 mg/L (0.2 mg/L)

S: 0.050 mg/L (0.2 mg/L)

The number of newborn F_1 blood samples used to quantify Cd, Zn, and Fe was as follows: 63 0 ppm (47 BxC, 16 CxB), 55 1 ppm (44 BxC, 11 CxB), and 60 50 ppm (50 BxC, 10 CxB). The number of newborn F_1 liver samples used to quantify Cd, Zn, and Fe was as follows: 39 0 ppm (32 BxC, 7 CxB), 38 1 ppm (32 BxC, 6 CxB), and 39 50 ppm (35 BxC, 4 CxB). Male and female pups were equally represented in each treatment group. Sex and direction of cross was determined to not significantly influence the levels of these elements, so all data was pooled within tissue type and treatment group. Due to significant differences seen in Zn and Fe as a result of maternal Cd exposure, additional BxC male and female newborn blood and liver tissues were submitted for analysis for all remaining elements. The number of these additional newborn BxC F_1 blood and liver samples were as follows: 24 0 ppm, 22 1 ppm, and 24 50 ppm, with males and females equally represented in each treatment group.

The number of maternal blood samples tested at delivery for Cd, Zn, and Fe was as follows: 17 0 ppm (10 B, 7 C), 15 1 ppm (9 B, 6 C), and 16 50 ppm (12 B, 4 C). The

number of maternal liver samples tested at delivery for Cd, Zn, and Fe was as follows: 18 0 ppm (10 B, 8 C), 15 1 ppm (9 B, 6 C), and 16 50 ppm (12 B, 4 C). Maternal strain was determined to not significantly influence the levels of these elements, so data was pooled within tissue type and Cd dose. Due to significant differences seen in Zn and Fe as a result of Cd exposure, additional B maternal blood and liver samples were submitted for analysis of all remaining elements. The number of these additional B maternal samples were as follows: 4 0 ppm, 3 1 ppm, and 5 50 ppm.

The number of adult BxC F_1 female blood and liver samples analyzed were as follows: 0 ppm Cd: 16 females representing 7 litters; 1 ppm Cd: 9 females representing 3 litters; 50 ppm Cd: 12 females representing 5 litters.

RNA extraction and quality assessment

RNA was extracted from 24 newborn BxC female whole hearts using the Qiagen AllPrep DNA/RNA/miRNA kit (Qiagen, 80204). Eight 0 ppm females representing 6 litters, eight 1 ppm females representing 4 litters, and eight 50 ppm females representing 5 litters were used. The hearts were removed from -80°C storage and immediately homogenized in the kit's lysis buffer using a microtube homogenizer. RNA was extracted from the hearts following the manufacturer's protocol for 10-30 mg of starting material. RNA was suspended in nuclease-free water and quantified using a Nanodrop 2000. RNA purity and size integrity were determined at the NCSU Genomic Sciences Laboratory (GSL) using an Agilent 2100 Bioanalyzer with an RNA 6000 Nano Chip (Agilent Technologies). All RNA samples had an RNA Integrity Number (RIN) ≥ 9.9.

RNA-seq

Total RNA from the hearts of four 0 ppm and four 50 ppm maternally exposed F_1 BxC female mice (each individual representing a different litter) was submitted to the NCSU GSL for indexed library construction and sequencing. Purification of messenger RNA (mRNA) was performed using oligo-dT beads provided in the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs [NEB]). Complementary DNA (cDNA) libraries for Illumina sequencing were constructed using the NEBNext Ultra Directional RNA Library Prep Kit (NEB) and NEBNext Multiplex Oligos for Illumina (NEB) using the manufacturer-specified protocol. Briefly, the mRNA was chemically fragmented and primed with random oligos for first strand cDNA synthesis. Second strand cDNA synthesis was then carried out with dUTPs to preserve strand orientation information. The double-stranded cDNA was purified, end repaired, and A-tailed for adapter ligation. Following ligation, the samples were selected for a final library size (adapters included) of 400-550 bp using sequential AMPure XP bead isolation (Beckman Coulter). Library enrichment was performed and specific indexes for each sample were added during the protocol-specified PCR amplification. The amplified library fragments were purified and checked for quality and final concentration using an Agilent 2200 Tapestation (D1000 chip, Agilent Technologies) combined with a Qubit fluorometer (ThermoFisher). The final quantified libraries were pooled in equimolar amounts for clustering and sequencing on an Illumina HiSeq 2500 DNA sequencer, utilizing a 125 bp single-end cycle sequencing kit (Illumina). The software package Real Time Analysis (RTA) was used to generate raw bcl, or base call files, which were then de-multiplexed by sample into fastq files using bcl2fastq Conversion Software v2.17 (Illumina).

Raw sequencing reads were transferred to the Cyverse Discovery Environment (DE) using Cyberduck. FastQC was performed using the FastQC 0.11.5 (multi-file) application (app). The Trimmomatic-programmable-0.33 app was used to trim the reads using the parameters HEADCROP:10, SLIDINGWINDOW:4:20, and MINLEN:20. FastQC was performed again to verify the quality of the trimmed reads. The Mouse RNASeq DeSEQ2 pipeline published by Cyverse[83] was used for indexing, aligning, read counts, and differential expression analysis; all analyses were performed using the recommended parameters. The HISAT2-index-align-2.1 app was used to index and align reads to the mouse reference genome (GRCm38.6 assembly). Aligned reads were counted using the HTSeq-count-0.6.1 app. Differential gene expression analysis was performed using the DESeq2 (multifactorial pairwise comparisons) app. A statistical report of this analysis can be found in Supplementary file A2.1 and the comprehensive results of the DESeq2 analysis can be found in Supplementary table A2.3.

The 302 genes found to be significantly differentially expressed between 0 ppm and 50 ppm maternal Cd exposure groups (adjusted p value <0.05) were used as input for enrichment analysis using Ingenuity Pathway Analysis (IPA) software (Qiagen) and Enrichr[84,85]. The KEGG, GO biological process, GO molecular function, GO cellular component, DSigDB, and MGI mammalian phenotype database results were used from the Enrichr analysis. For all enrichment analysis results, only those with significant adjusted p values (<0.05) were considered. 10 genes that frequently occurred across multiple relevant terms and categories were chosen for qRT-PCR validation. The heatmap (Supplementary figure A2.2) was generated using all 302 genes that were differentially expressed with significant adjusted p values <0.05 using

http://www.heatmapper.ca/ . The following settings were used: Scale type: row; Colour brightness: 25; Number of shades: 100; Colour scheme: red/green; Clustering method: average linkage; Distance measurement method: Pearson; Apply clustering to: rows; Plot width (pixels): 500; Plot height (pixels): 3387.

qRT-PCR

100 ng of total RNA extracted as described above from 24 BxC female newborn pup hearts was used as a template to synthesize first strand cDNA. DNase I treatment was performed prior to the RT reaction during RNA extraction as part of the Qiagen AllPrep kit protocol. Random primers (Promega, C118A) were annealed to the DNase I-treated RNA, then the RT reaction was performed according to the manufacturer's protocol (M-MLV RT enzyme, Promega, M170B). qRT-PCR was performed on a Real-Time 7300 machine (Applied Biosystems) using 96 well plates (Genesee Scientific, 24-301) and SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, 1725271). The 20 µL RT-qPCR reaction mix per well was prepared according to the manufacturer's protocol for SYBR Green. 1 µL of the cDNA for the standards or the tested samples was included in the 20 µL reaction. cDNA of the tested samples was diluted 1/10, while the standards were diluted 1/5, 1/10, 1/20, 1/40, and 1/80. Standards, the no template control (NTC), and all other tested samples were each run in triplicate. The cycling conditions were as follows: 95°C for 30 seconds; 40 cycles of 95°C for 15 seconds, 60°C for 30 seconds; dissociation curve 60.0°C-95.0°C. The primer sequences are provided in Supplementary Table A2.1. The dissociation curve confirmed that the primers amplified a single PCR product, or no product in the NTC. Amplification efficiencies were calculated in Microsoft Excel by determining the slope of the regression between the log values of the standard

concentrations and the average Ct value of the standards. *Polr2a* was not significantly differentially expressed between treatment groups and was used as a reference gene. Quantification of expression was calculated using the ΔΔCt method[86].

This study was performed in compliance with MIQE standards[87] (see Supplementary table A2.2).

Statistical analysis

All statistical analyses were performed using a one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparison test using R software, comparing animals from the 1 ppm and 50 ppm groups to 0 ppm controls. Outliers, as defined by a point which falls more than 1.5 times the interquartile range above the third quartile or below the first quartile, were omitted from the analysis if found to be present. BxC litters under n=6 and CxB litters under n=3 were omitted from raw or proportional mass analyses. Data are presented as the mean ± standard error of the mean.

Results

Maternal Cd exposure results in decreased body mass and cardiac hypertrophy at birth

Maternal Cd exposure did not influence litter size or sex ratio of the offspring (data not shown). However, 50 ppm Cd was found to significantly increase the number of problematic births (dystocia, birthing of dead pups) for C mothers ($p=0.02$, chi-square test, expected values based on number of successful or problematic births of 0 ppm C mothers).

The body mass of 50 ppm maternally-exposed F_1 mice was significantly reduced at birth compared to control animals (Figure 2.2A), consistent with previous studies in

mice using the same dose[88]. Reduced body mass was observed in offspring of both sexes and from both crosses. 1 ppm maternally-exposed CxB males also showed a significant reduction in body mass at birth (Figure 2.2A).

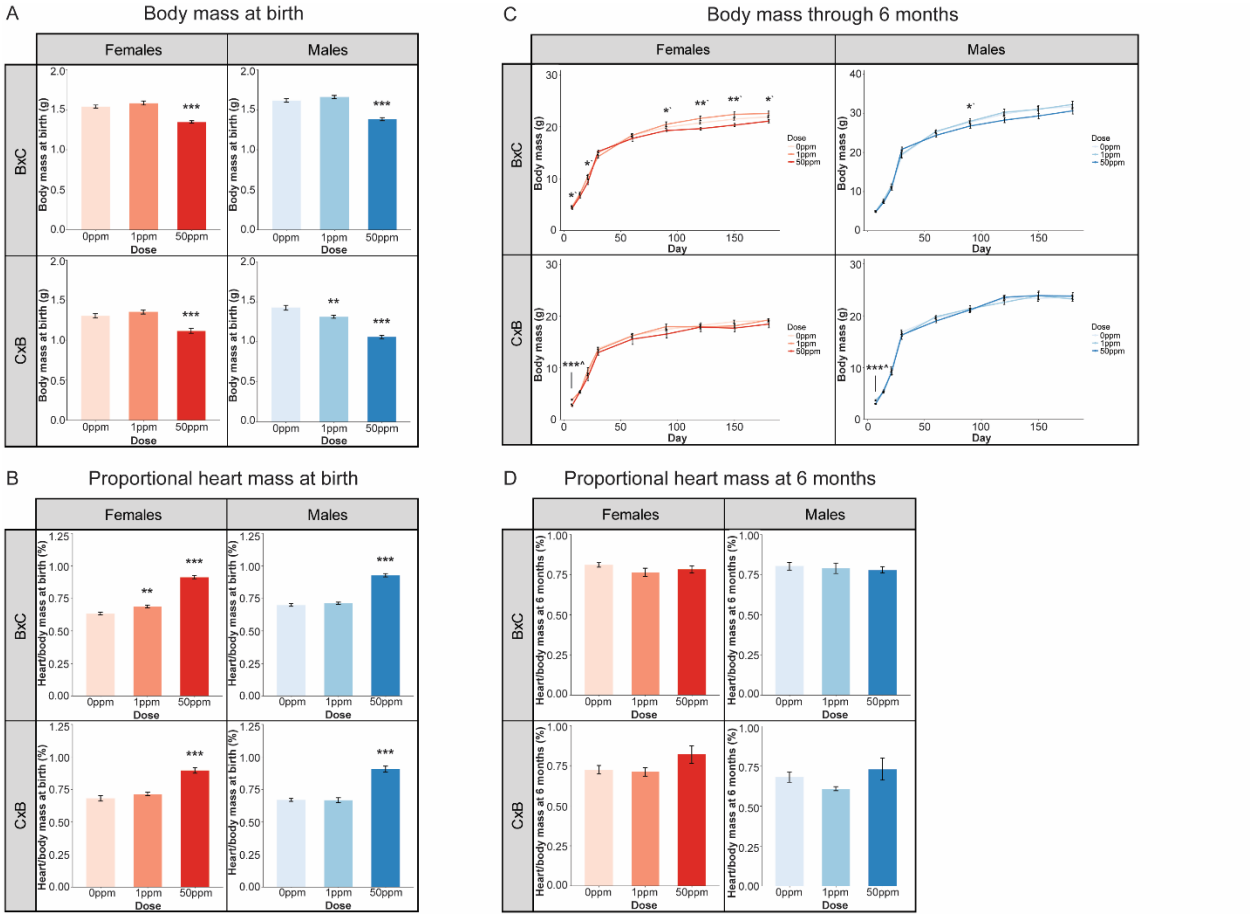


Figure 2.2. The effects of maternal Cd exposure on body mass and normalized heart mass at birth and adulthood. **A,** Body mass at birth. **B,** Proportional heart mass at birth. **C,** Body mass through 6 months. ^=0 ppm vs 1 ppm, `=0 ppm vs 50 ppm. **D,** Proportional heart mass at 6 months. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Mice presented with cardiac hypertrophy at birth due to maternal Cd exposure. A significant increase in heart mass as a proportion of body mass was observed in 50 ppm maternally-exposed F_1 mice compared to controls (Figure 2.2B) in both sexes and crosses. 1 ppm maternally-exposed BxC females also demonstrated a significant

increase in heart mass as a proportion of body mass (Figure 2.2B). Despite the reduction in overall body mass among the 50 ppm maternally-exposed mice, hearts in these animals were still significantly larger than controls in terms of raw mass, with the exception of CxB males (Supplementary figure A2.1A).

Body mass differences at birth due to maternal Cd exposure persisted beyond 7 days of age through 6 months of age for only 50 ppm maternally-exposed BxC females, with the exception of a significant decrease in body mass seen in 50 ppm maternally-exposed BxC males at 90 days of age (Figure 2.2C). At 6 months of age, there were no significant differences in heart mass as a proportion of body mass (Figure 2.2D) or in terms of raw mass, with the exception of a significant reduction in raw heart mass in 50 ppm maternally-exposed BxC females (Supplementary figure A2.1B).

BxC females are susceptible to hypertension and altered blood pressure parameters during adulthood as a result of maternal Cd exposure

Due to the cardiac hypertrophy seen in the maternally-exposed F_1 mice at birth and the previously reported association between adult exposure to Cd and hypertension in humans, blood pressure parameters were measured in 5.5-month-old F_1 mice to test whether hypertension in adulthood can be programmed by maternal Cd exposure alone. 50 ppm maternally-exposed BxC females demonstrated significantly increased systolic, diastolic, and mean arterial blood pressures, consistent with a hypertensive phenotype (Figure 2.3A-C). 50 ppm maternally-exposed BxC males, CxB females, and CxB males displayed increased systolic, diastolic, and mean arterial blood pressures, although these changes were not statistically significant (Figure 2.3A-3C). Both 1 ppm and 50 ppm maternally-exposed BxC females showed a significant increase in tail blood volume and

tail blood flow (Figure 2.3D-E). There were no significant differences in heart rates between any of the groups (Figure 2.3F). Due to the hypertensive phenotype observed in 50 ppm maternally-exposed BxC females, these animals were chosen for further studies.

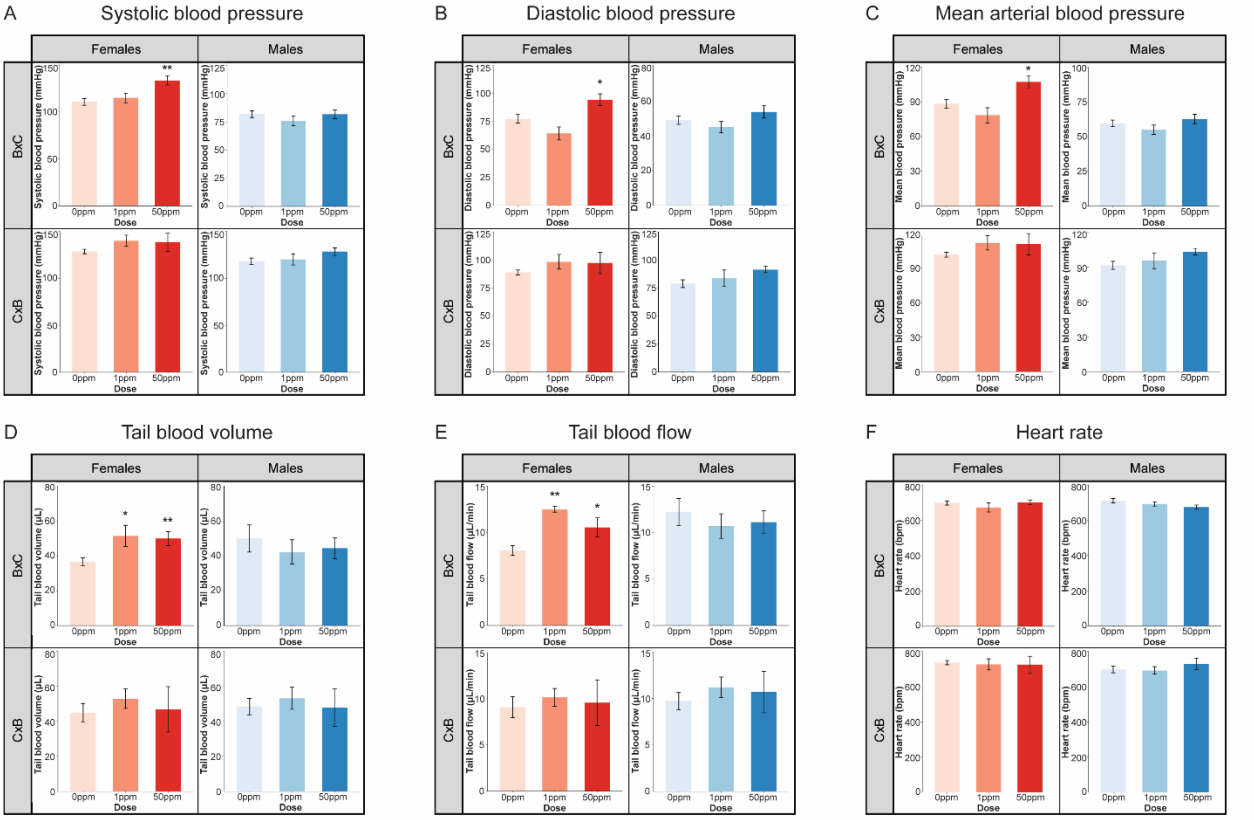


Figure 2.3. Blood pressure parameters in F_1 animals at 5.5 months of age following maternal Cd exposure. **A,** Systolic blood pressure. **B,** Diastolic blood pressure. **C,** Mean arterial blood pressure. **D,** Tail blood volume. **E,** Tail blood flow. **F,** Heart rate. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Circulating ROS levels are not elevated in maternally-exposed BxC females at 6 months of age

Reactive oxygen species (ROS) have been suggested to play a role in the toxicity of Cd exposure during adulthood[89]. In addition, elevated ROS levels in circulation and relevant tissues contribute to hypertension[90]. To determine if the hypertension seen in

maternally-exposed BxC females was associated with elevated ROS levels in circulation, ROS were quantified at 6 months of age by measuring serum levels of protein carbonyl, a stable marker of ROS[91]. No significant differences in serum protein carbonyl content were observed between exposure groups (Figure 2.4).

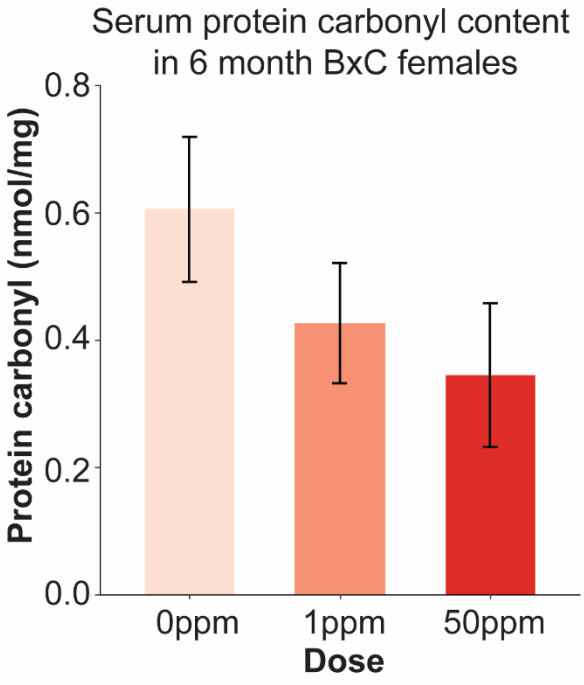


Figure 2.4. Protein carbonyl quantification in serum of 6-month-old BxC females. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Essential trace elements associated with hypertension are not altered in maternally-exposed BxC females at 6 months of age

Maternal exposure to Cd has been shown to affect trace element levels in various tissues in the offspring at birth[92]. To determine if the hypertension observed in adult maternally-exposed BxC females was associated with abnormal homeostasis of essential trace elements known to contribute to blood pressure, levels of Ca, Fe, K, Mb, Na, Cu, Se, and Zn were quantified in 6-month-old BxC female blood and liver tissue (Tables 2.1

and 2.2). No significant differences in any elements were seen at 6 months of age between 0 ppm and 1 ppm or 50 ppm maternal Cd exposure groups, with the exception of a significant reduction of Fe seen in liver tissue of 1 ppm maternally-exposed mice (Table 2.2).

Table 2.1. Blood concentrations of eight blood pressure-associated essential trace elements in 6-month-old BxC females.

	Mean blood concentrations ± SE							
Maternal Cd dose	**Ca (mg/L)**	**Fe (mg/L)**	**K (mg/L)**	**Mg (mg/L)**	**Na (mg/L)**	**Cu (µg/L)**	**Se (µg/L)**	**Zn (µg/L)**
0 ppm	4.84±0.08	49.59±0.38	226.07±1.8	3.41±0.06	174.21±1.5	59.10±1.36	58.45±0.97	1320.3±44.9
1 ppm	4.68±0.09	49.01±0.68	226.53±3.1	3.30±0.04	174.09±1.4	61.14±3.18	58.17±1.11	1310.7±66.5
50 ppm	4.59±0.14	47.73±0.91	222.70±4.0	3.56±0.05	171.21±2.8	58.79±3.34	57.34±1.36	1400.5±49.8

***p<0.05, **p<0.01, ***p<0.001**. One-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Table 2.2. Liver concentrations of eight blood pressure-associated essential trace elements in 6-month-old BxC females.

	Mean liver concentrations ± SE							
Maternal Cd dose	**Ca (mg/L)**	**Fe (mg/L)**	**K (mg/L)**	**Mg (mg/L)**	**Na (mg/L)**	**Cu (µg/L)**	**Se (µg/L)**	**Zn (µg/L)**
0 ppm	3.18±0.06	10.83±0.21	316.53±3.5	22.99±0.23	70.25±1.3	549.2±8.6	186.93±2.7	3731.6±65.9
1 ppm	3.13±0.08	**10.00±0.24***	314.00±5.3	22.87±0.53	68.65±0.8	548.1±13.5	195.22±3.9	3707.2±76.5
50 ppm	3.14±0.05	10.29±0.21	324.42±1.3	23.57±0.19	68.07±1.7	590.2±12.6	194.31±4.3	3917.4±74.9

***p<0.05, **p<0.01, ***p<0.001**. One-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Maternal Cd exposure is associated with significantly altered levels of essential trace elements at birth

Due to the absence of any overt biomarkers of hypertension at adulthood, newborn F_1 tissues were chosen for further analysis to determine if molecular changes in early life could contribute to the programming of hypertension in adulthood. In addition to the 8 elements associated with blood pressure that were quantified in 6-month-old BxC female tissue, 5 additional trace elements were quantified in newborn blood and liver tissue. Maternal Cd exposure resulted in significant differences in levels of all elements tested

(Tables 2.3 and 2.4). Elements that were significantly increased in the pup blood at birth due to maternal Cd exposure were Ca, Na, Zn, Co, Cu, Mn, and Mo, while a decrease was seen in Fe, K, Mg, P, S, and Se. Only Mg, Na, and Mn were significantly increased in the newborn pup livers due to maternal Cd exposure; all others were found to be significantly decreased. Elements that were significantly increased in pup blood, but significantly decreased in pup livers due to maternal Cd exposure were Ca, Zn, Co, Cu, and Mo. Mg was the only element significantly decreased in pup blood, but significantly increased in pup livers due to maternal Cd exposure. Several elements were significantly different from the controls in only the 1 ppm maternally-exposed pups (blood Mn, blood Mo, liver Ca, liver K, liver P, liver Co). No Cd was detected in newborn blood or liver tissue for any of the treatment groups given the reporting limit (<1.0 µg/L), consistent with other studies showing that the placenta acts as a barrier to Cd[92].

Levels of these 13 elements were also quantified in maternal blood and liver tissues at the time of delivery to determine if the levels in F_1 pups correlated with the levels seen in the mothers. 50 ppm Cd significantly affected maternal blood levels of Ca, Fe, K, Mg, Na, P and S; additionally, 50 ppm Cd significantly affected maternal liver levels of Fe and Zn (Tables 2.5 and 2.6). The directional change of the levels (i.e. significantly increased or decreased due to Cd) was the same between pups and mothers, with the exception of liver Zn being decreased in the pups and increased in the mothers. Cd was undetected in maternal blood at delivery for 0 ppm and 1 ppm exposed females, but was detected in the 50 ppm exposed females at 5.23 µg/L ± 0.99 µg/L (average ± SE, <1.0 µg/L reporting limit).

Table 2.3. Blood concentrations of 13 trace elements in newborn F_1 mice maternally-exposed to Cd.

	Mean blood concentrations ± SE												
Maternal Cd dose	**Ca (mg/L)**	**Fe (mg/L)**	**K (mg/L)**	**Mg (mg/L)**	**Na (mg/L)**	**P (mg/L)**	**S (mg/L)**	**Zn (mg/L)**	**Co (µg/L)**	**Cu (µg/L)**	**Mn (µg/L)**	**Mo (µg/L)**	**Se (µg/L)**
0 ppm	5.42 ± 0.05	35.4 ± 0.5	164.6 ± 2.3	4.84 ± 0.04	187.4 ± 1.2	42.0 ± 0.6	85.8 ± 0.9	0.88 ± 0.03	0.40 ± 0.05	143.0 ± 3.6	4.74 ± 0.19	4.92 ± 0.34	30.0 ± 0.8
1 ppm	**5.75 ± 0.07****	**30.9 ± 0.9*****	169.2 ± 2.3	4.99 ± 0.06	**198.2 ± 1.7*****	43.2 ± 0.9	87.7 ± 1.2	0.92 ± 0.02	**0.55 ± 0.01***	152.2 ± 2.1	**6.25 ± 0.54***	**6.03 ± 0.25***	30.2 ± 0.8
50 ppm	**7.22 ± 0.07*****	**15.1 ± 0.2*****	**108.1 ± 1.4*****	**4.30 ± 0.07*****	**255.1 ± 1.9*****	**33.3 ± 0.5*****	**62.0 ± 0.8*****	**1.01 ± 0.04****	**0.75 ± 0.02*****	**186.7 ± 4.0*****	5.99 ± 0.53	5.15 ± 0.25	**26.0 ± 0.7*****

***p<0.05, **p<0.01, ***p<0.001**. One-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Table 2.4. Liver concentrations of 13 trace elements in newborn F_1 mice maternally-exposed to Cd.

	Mean liver concentrations ± SE												
Maternal Cd dose	**Ca (mg/L)**	**Fe (mg/L)**	**K (mg/L)**	**Mg (mg/L)**	**Na (mg/L)**	**P (mg/L)**	**S (mg/L)**	**Zn (mg/L)**	**Co (µg/L)**	**Cu (µg/L)**	**Mn (µg/L)**	**Mo (µg/L)**	**Se (µg/L)**
0 ppm	4.10 ± 0.10	11.1 ± 0.4	354.4 ± 3.4	22.0 ± 0.2	89.1 ± 1.1	308.9 ± 5.2	192.9 ± 1.7	5.08 ± 0.13	2.16 ± 0.06	1489 ± 72	19.0 ± 1.0	23.1 ± 1.3	47.9 ± 1.1
1 ppm	**3.78 ± 0.08***	10.4 ± 0.5	**344.4 ± 2.1***	21.4 ± 0.3	89.4 ± 1.0	**279.3 ± 4.4*****	**186.4 ± 2.1***	**4.63 ± 0.11****	**1.97 ± 0.04***	**1264 ± 37***	17.2 ± 0.4	**18.4 ± 1.1****	46.1 ± 0.8
50 ppm	4.07 ± 0.11	**2.8 ± 0.1*****	349.6 ± 3.3	**22.7 ± 0.2***	**95.0 ± 1.5****	295.1 ± 4.9	**185.9 ± 2.1***	**4.42 ± 0.10*****	2.03 ± 0.04	**1298 ± 52***	**25.5 ± 1.5*****	**14.7 ± 0.3*****	**44.0 ± 1.1***

***p<0.05, **p<0.01, ***p<0.001**. One-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Table 2.5. Blood concentrations of 13 trace elements in mothers at delivery.

	Mean blood concentrations ± SE												
Maternal Cd dose	**Ca (mg/L)**	**Fe (mg/L)**	**K (mg/L)**	**Mg (mg/L)**	**Na (mg/L)**	**P (mg/L)**	**S (mg/L)**	**Zn (mg/L)**	**Co (µg/L)**	**Cu (µg/L)**	**Mn (µg/L)**	**Mo (µg/L)**	**Se (µg/L)**
0 ppm	5.15 ± 0.26	39.16 ± 1.40	215.59 ± 8.37	4.40 ± 0.22	179.15 ± 7.25	58.34 ± 3.50	119.95 ± 2.83	0.58 ± 0.03	1.49 ± 0.61	138.25 ± 7.19	14.13 ± 2.91	5.71 ± 0.74	47.80 ± 3.92
1 ppm	5.64 ± 0.19	37.05 ± 1.33	207.07 ± 14.25	4.11 ± 0.27	185.01 ± 4.09	54.65 ± 3.49	117.84 ± 4.64	0.60 ± 0.04	1.20 ± 0.25	134.67 ± 10.14	11.97 ± 1.53	5.51 ± 0.88	48.00 ± 1.31
50 ppm	**6.66 ± 0.44***	**26.73 ± 0.98*****	**157.33 ± 1.53****	**3.42 ± 0.07****	**227.07 ± 2.27*****	**44.21 ± 0.29****	**104.04 ± 2.39****	0.55 ± 0.06	0.73 ± 0.03	141.50 ± 2.72	8.24 ± 0.50	6.18 ± 0.76	44.42 ± 1.65

***p<0.05, **p<0.01, ***p<0.001**. One-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Table 2.6. Liver concentrations of 13 trace elements in mothers at delivery.

	Mean liver concentrations ± SE												
Maternal Cd dose	**Ca (mg/L)**	**Fe (mg/L)**	**K (mg/L)**	**Mg (mg/L)**	**Na (mg/L)**	**P (mg/L)**	**S (mg/L)**	**Zn (mg/L)**	**Co (µg/L)**	**Cu (µg/L)**	**Mn (µg/L)**	**Mo (µg/L)**	**Se (µg/L)**
0 ppm	4.58 ± 0.54	4.52 ± 0.33	426.77 ± 51.96	31.64 ± 4.21	85.67 ± 11.96	429.91 ± 56.67	365.19 ± 44.85	4.43 ± 0.20	7.84 ± 1.21	918.25 ± 203.21	278.25 ± 36.46	94.23 ± 21.36	145.08 ± 29.30
1 ppm	4.58 ± 0.54	4.86 ± 0.31	358.31 ± 14.47	25.82 ± 1.43	72.65 ± 3.22	355.60 ± 16.80	285.15 ± 29.65	4.57 ± 0.14	6.57 ± 0.26	503.67 ± 24.17	292.00 ± 58.51	72.20 ± 4.55	122.33 ± 11.61
50 ppm	3.86 ± 0.03	**3.27 ± 0.25****	342.58 ± 3.95	25.51 ± 0.83	73.24 ± 1.85	344.36 ± 9.48	285.08 ± 11.83	**5.31 ± 0.26****	6.21 ± 0.14	606.80 ± 54.74	299.20 ± 14.81	63.18 ± 2.02	99.58 ± 3.54

***p<0.05, **p<0.01, ***p<0.001**. One-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm.

Transcriptomic analysis of newborn BxC female hearts shows Cd-associated changes in genes and pathways involved in cardiac hypertrophy and hypertension

To gain a better understanding of molecular pathways that may contribute to cardiac hypertrophy at birth and susceptibility to hypertension during adulthood, RNA-seq was performed on whole hearts obtained from 0 ppm and 50 ppm maternally-exposed BxC female F_1 mice at birth. Out of the 14,698 genes with aligned reads passing our quality control filtering, 302 genes were found to be significantly differentially expressed between the two treatment groups (Figure 2.5A, Supplementary table A2.3, Supplementary figure A2.2). Significantly more genes increased in expression (193) than decreased in expression (109) in the 50 ppm maternal Cd exposure group compared to controls ($p<0.0001$, chi-squared test). Among the differentially expressed genes, there was a significant enrichment for phenotype or disease terms related to hypertension, abnormal heart and cardiovascular system development, cardiac hypertrophy, and nutritional disease (Supplementary table A2.4). Pathways relevant to these phenotypes that had significantly enriched terms included hypoxia, altered cellular energy generation and carbon metabolism, ROS, nitric acid homeostasis, and altered metal homeostasis (Supplementary table A2.4). Additional terms that were significantly enriched included cancer, cell death and survival, connective tissue development, immune response and disorders, lipid metabolism, organismal development and abnormalities, and cell signaling (Supplementary table A2.4). Despite undetected levels of Cd in fetal tissues, the DSigDB (Drug SIGnatures DataBase) analysis showed a significant enrichment for the term ‘Cadmium sulfate’. Cadmium sulfate is one of the highly soluble salt forms of Cd that is more readily absorbed from the intestines compared to other less soluble salt

forms[93], indicating the toxic effects of adult (i.e. direct) and maternal (i.e. indirect) exposure to Cd may act through similar or overlapping pathways.

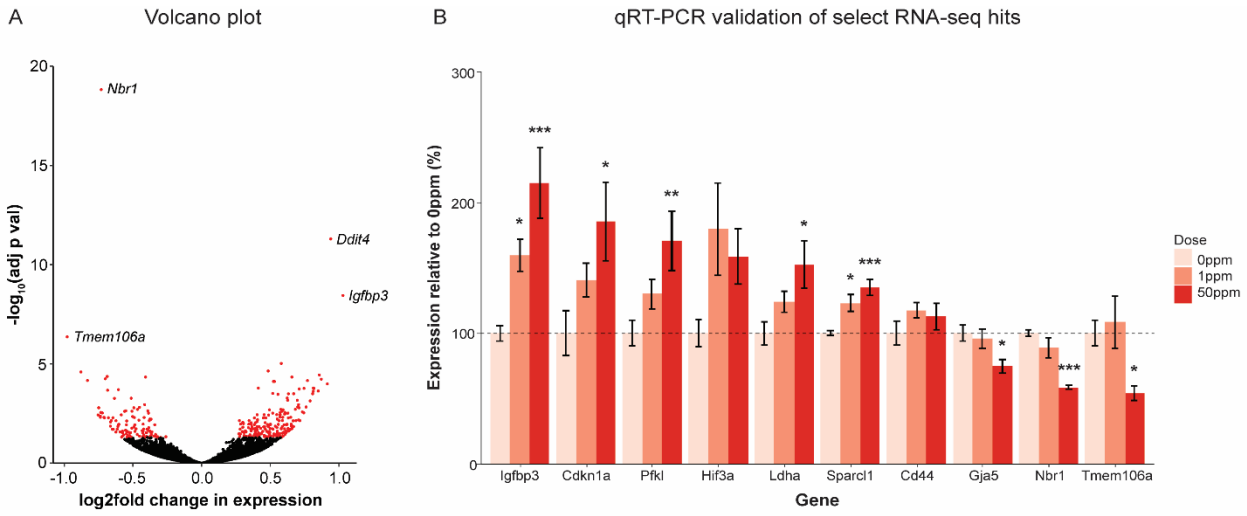


Figure 2.5. Volcano plot and qRT-PCR validation of RNA-seq data in newborn BxC female hearts. **A,** Volcano plot of 14,698 genes identified through RNA-seq. The 302 genes identified as significantly differentially expressed (adj p val<0.05) due to maternal Cd exposure are highlighted in red. The names of the four genes with the highest -log10(adj p val) are indicated to the right of their data points on the graph. **B,** Select RNA-seq hits for qRT-PCR validation. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Ten differentially expressed genes identified by RNA-seq with functions related to cardiac hypertrophy and hypertension were validated by qRT-PCR in 0 ppm, 1 ppm, and 50 ppm maternally-exposed newborn BxC female hearts (Figure 2.5B). All genes except two (*Hif3a*, *Cd44*) were significantly differentially expressed due to maternal Cd exposure, consistent with the RNA-seq data. Most of these showed a dose-dependent response to maternal Cd exposure, although the differences between 0 ppm and 1 ppm maternally-exposed mice were significant only for two genes (*Igfbp3*, *Sparcl1*) (Figure 2.5B).

Discussion

CVD is the leading cause of death and disability worldwide[19,78]. Although genetic risk factors for CVD have been identified, its increasing prevalence has been linked to environmental factors including toxic metal exposure[19]. Exposure to Cd, ranked seventh on the Substance Priority List from the Agency for Toxic Substances and Disease Registry[94], is associated with hypertension and CVD. However, to date, most research in human cohorts and animal models has focused on the significance of chronic Cd exposure during adult life on CVD. Although Cd is inefficiently transferred across the placenta[92], there is growing evidence that maternal exposure to Cd can impact child development and health[22,24,29,95]. However, molecular mechanisms of Cd toxicity in this context have been challenging to study in humans.

To better understand the importance and mechanisms of action of developmental Cd exposure on cardiovascular health, we established a tractable mouse model. Using this model, we have shown that maternal Cd exposure is associated with cardiac hypertrophy at birth and can program hypertension during adulthood. Maternal Cd exposure perturbs a broad range of trace elements during development that are essential for various functions in the body, including cardiovascular development, and causes widespread transcriptional changes in the newborn heart associated with hypertrophy and hypertension. Together, these data provide mechanistic insight into the links between maternal Cd exposure and impaired cardiac function.

Two doses of maternal Cd exposure were chosen to reflect a low and high dose of Cd relevant to human levels of exposure. Outside of human pollution, Cd naturally occurs in the environment at low levels; the average amount of Cd in the Earth's crust is

estimated to be roughly 0.3 ppm[3]. However, geographic clusters can contain higher than average levels of Cd as shown by a study in North Carolina identifying regions with up to 1.6 ppm Cd in the soil[96]. Therefore, exposure to 1 ppm Cd can be considered a realistic low level of human exposure. The mean level of blood Cd in the United States (US) population has been reported to be 0.337 µg/L[96]; this level is below our reporting limits and consistent with our inability to quantify Cd in the blood of 1 ppm exposed F_0 mice. The mean level of Cd detected in the blood of 50 ppm exposed F_0 mice was approximately 15.5 times the mean blood Cd levels of the US population; however, people living in Cd-polluted areas demonstrate blood Cd levels similar to or higher than levels quantified here in 50 ppm exposed F_0 mice[97,98]. Additionally, F_1 mice maternally-exposed to 50 ppm Cd, but not 1 ppm Cd, presented with fetal growth restriction, a phenotype consistent with Cd exposure in humans, providing further support that the high dose is relevant to the study of Cd-associated human health outcomes.

Maternal exposure to Cd has a well-established association with reduced fetal size[69,99]. It has also been shown that Cd exposure can alter levels of certain essential trace elements[7]. The need for adequate nutrition and micronutrient intake during pregnancy is well-established, and malnutrition during development is associated with increased CVD risk[23]. The altered levels of trace elements in newborn pups due to maternal Cd exposure are likely signaling a state of malnutrition and may be early in the mechanistic pathways leading to the phenotypes observed in this study.

Zn and Fe are the most well-studied elements affected by maternal Cd exposure due to their similar chemical properties and shared mechanisms of absorption[16]. One study showed that 50 ppm maternal Cd exposure in the rat resulted in reduced Fe in

maternal liver and whole fetuses, yet increased Zn in maternal liver[15], consistent with the findings in this study. The study also quantified Zn and Cu in whole fetuses, but found no significant differences as a result of 50 ppm maternal Cd exposure. Here, we show a significant increase in blood Zn and Cu, yet a significant decrease in liver Zn and Cu, which could explain that no overall differences were observed in whole rat fetuses as a result of 50 ppm maternal Cd exposure.

Other element responses shown here are less well-characterized as a result of maternal Cd exposure, but are consistent with their physiological role. Cd exposure during adulthood has been shown to influence Ca homeostasis and bone metabolism through increasing bone resorption and remodeling, leading to increased circulating levels of Ca[98], as was seen here in maternally-exposed newborn F_1 pups. Poor Ca absorption and bone metabolism may contribute to reduced fetal size seen with maternal Cd exposure. P levels in the blood are inversely related to Ca levels[100], consistent with the decrease in P and increase in Ca seen in pup blood as a result of maternal Cd exposure. P has a role in bone development, is an integral part of cell membranes (phospholipids), is part of the energy-carrying molecule ATP, and is part of 2,3-diphosphoglycerate, which is required for oxygen release from the hemoglobin in red blood cells to deliver oxygen to tissues throughout the body[100]. Na was significantly increased in pup blood and livers due to maternal Cd exposure; excess dietary Na is a significant contributing factor to hypertension[101]. Na concentrations also correlate with left ventricular hypertrophy[102], which was previously seen in rats maternally exposed to 30 ppm Cd[24]. Mn is important for regulating blood glucose levels, bone growth and development, and blood clotting. Mn has many overlapping functions with Fe, and Mn levels can increase in response to

decreased Fe levels[103], consistent with the observations in this study. Se levels were significantly decreased in both pup blood and liver in response to maternal Cd exposure. Se is a vital component of antioxidant enzymes and diseases due to Se deficiency present with cardiac hypertrophy, congestive heart failure, and other altered cardiac features[104].

Many of the trace element levels in fetal blood and liver tissue correlated with those seen in maternal tissue; this correlation highlights the importance of including a pre-conception exposure window in animal models of maternal exposure, as maternal exposures in reality often occur for a period of time prior to pregnancy. Some elements appeared to be altered in maternal tissue as a result of 50 ppm Cd exposure, though this was not statistically significant, likely due to small sample sizes. Thus, many of the essential trace elements seen perturbed in the pups as a result of maternal Cd exposure may be due to already existing perturbations in the mother. It is also possible that the trace element levels are further exacerbated at the placental level, as Cd has been shown to interfere with nutrient transport across the placenta[15].

While the effects of maternal Cd exposure are likely due to altered levels of multiple trace elements, Fe deficiency may be one of the most significant contributing factors to reduced birth weight, cardiac hypertrophy at birth, and hypertension. Fe is found in heme, the component of red blood cells that binds and systemically delivers oxygen absorbed from the lungs. Fe is a cofactor for numerous enzymes with roles including cellular energy production, redox reactions, and biological molecule synthesis. A significant deficiency in Fe is associated with an increased risk in CVD. Fe deficiency is associated with increased oxidative stress, potentially damaging cells in the heart or endothelial lining of systemic vasculature[105], which may contribute to the susceptibility to hypertension. Cardiac stress

due to Fe deficiency, as well as the deficiency seen in other essential trace elements that may signal a state of undernutrition, is likely leading to cardiac hypertrophy at birth. The heart responds to chronic stress by increasing its size to increase blood flow, oxygen, and nutrient delivery to tissues[106]. Though circulating ROS levels were not significantly different in 6-month-old BxC females, it is likely that ROS levels are increased during development as a result of maternal Cd exposure, given the Fe deficiency and levels of other trace elements involved with oxidative stress pathways, as well as the RNA-seq data.

In addition to altered essential trace elements, we show global changes in gene expression in the heart at birth due to maternal Cd exposure. To our knowledge, this is the first time any such analysis has been performed in the heart due to Cd exposure. We show changes of expression in genes implicated in hypertension, abnormal heart and cardiovascular system development, and hypertrophy. Our RNA-seq data are further enriched for multiple mechanistic pathways that contribute to these diseases such as hypoxia, altered cellular energy, ROS, nitric acid homeostasis, and altered metal homeostasis, consistent with our trace element analysis. There was a significantly larger proportion of genes that increased in expression in response to maternal Cd exposure; this is likely due to the heart needing to adapt to a new environment (e.g., hypoxia, nutrient-poor).

Mechanisms through which maternal or adult exposure to Cd affects health may differ as the placenta acts as a barrier to Cd, but no such barriers exist when exposed to Cd postnatally. Considering that little, if any, Cd enters fetal circulation, maternal Cd exposure may exert its effects indirectly. However, there was a significant enrichment for

the term ‘Cadmium sulfate’ among differentially expressed genes identified by RNA-seq, indicating that maternal Cd exposure may activate similar pathways in offspring as those activated in adult Cd exposure.

In summary, we propose that maternal Cd exposure leads to a hypoxic and stressful environment in the offspring during development, which leads to cardiac hypertrophy and susceptibility to hypertension during adulthood. Our observations of impaired trace element homeostasis and global gene regulation provide novel insights into mechanisms underlying Cd-associated CVD.

Acknowledgements

This work was funded by the National Institutes of Health under K22ES027510 and P30ES025128. Next-generation sequencing was performed by the North Carolina State University Genomic Sciences Laboratory. Trace metal quantification was performed at the Environmental and Agricultural Testing Service Laboratory (EATS), Department of Crop and Soil Sciences, at North Carolina State University. We would like to thank Scott Belcher for the use of his blood pressure measurement equipment.

CHAPTER 3

Transcriptomic and Proteomic Analyses of Newborn Mouse Brains Identify Multiple Pathways Linking Maternal Cadmium Exposure to Impaired Neurodevelopment and Behavior

The following chapter is the first draft of a publication that will be submitted to a peer-reviewed journal.

Abstract

Cadmium (Cd) is found ubiquitously throughout the environment and is one of ten chemicals of major public concern identified by the World Health Organization. Cd does not efficiently cross the placenta, yet maternal Cd exposure can impair fetal growth and development. There is emerging evidence from human and animal models that maternal Cd exposure can negatively impact fetal neurodevelopment; however, the molecular mechanisms perturbed in the brain as a result of maternal Cd exposure are unclear. To address this, we established a mouse model of maternal Cd exposure. Offspring exposed maternally to Cd have significantly enlarged brains proportional to their body mass at birth and altered behavior at adulthood. RNA-seq performed on whole brains at birth shows an increase in *Hox* gene and myelin marker expression, and suggests disrupted retinoic acid (RA) signaling. Proteomic analyses show an enrichment in proteins involved in cellular energy pathways, hypoxic response, and altered RA signaling. Consistent with indications from RNA-seq and proteomic analyses, we identify increased levels of RA in newborn brains as a result of maternal Cd exposure. Finally, we show that maternal Cd exposure reduces mitochondrial content in the brain at birth, consistent with the proteomic data suggesting alterations to cellular energy pathways. Results from these studies provide insight into multiple pathways perturbed in the fetal brain as a result of maternal Cd exposure that impact neurodevelopment and program behavioral deficits during adulthood.

Introduction

Exposure to heavy metals during pregnancy can impact fetal development and have lasting consequences to the future health of the offspring. Of the ten chemicals of major public health concern identified by the World Health Organization, four are heavy metals: arsenic (As), lead (Pb), mercury (Hg), and cadmium (Cd)[1].

Cd is found naturally-occurring at low levels in the soil, but its levels in the environment have dramatically increased since the early 20th century due to human activities such as fossil fuel combustion and metal mining and refining[4]. Inhalation and ingestion are the primary routes of Cd exposure in humans. Cigarette smoke, contaminated dust, and occupational exposure are the primary sources of inhaled Cd. Among non-smokers, Cd ingestion of contaminated produce or water is the primary route of exposure[75,76]. Exposure to high levels of Cd during adulthood is associated with cancer, hepatotoxicity, renal toxicity, and softening of the bones[2,4,107]; however, it is becoming clear that chronic exposure to low levels of Cd is associated with noncommunicable metabolic disorders prevalent in developed countries such as obesity, type II diabetes, non-alcoholic fatty liver disease, and cardiovascular disease[5,16,18,19].

There is also mounting evidence that maternal exposure to Cd negatively affects fetal health. While some heavy metals such as Pb readily cross the placenta[27], Cd accumulates in the placenta and does not efficiently enter fetal circulation[108]. Despite this, maternal Cd exposure has been associated with reduced fetal growth[22], altered trace metal homeostasis (Chapter 2), congenital malformations[34], cardiac hypertrophy and hypertension (Chapter 2), and impaired neurodevelopment[95].

Human epidemiological studies have linked higher levels of Cd in cord or maternal blood with lowered intelligence and behavioral deficits in children[21,95]. This association has been supported by animal models showing that early-life Cd exposure impacts neurodevelopment. Rats exposed to Cd through their mother's milk exhibit increased motor activity during adulthood[109]. An additional study examined the behavior of rat pups born to mothers who were injected with Cd during the organogenesis period (day 7-15 of gestation) and found delays in sensorimotor development and reduced anxiety during adulthood[110].

Despite strong evidence of early-life exposure to Cd negatively affecting behavior and intelligence, the impacts of maternal Cd exposure alone on the fetal brain at the molecular level are not as clear. Rat pups born to mothers who were exposed to Cd during both pregnancy and lactation had increased proportional brain weights, altered lipid peroxidation, and severely damaged mitochondria in the brain[111]. Studies in mice that were maternally-exposed to Cd from gestational day 1 to postnatal day 10 suggested a disruption to the expression of sex hormone receptor genes in the brain[112], as well as altered thyroid hormone metabolism, which is essential for brain development[113]. Many of the studies of early-life Cd exposure use an exposure window that begins at conception and lasts into early postnatal development, but this may not best reflect the situation in humans in which maternal exposure to Cd may begin before conception. This exposure model also does not differentiate between prenatal and early postnatal effects of Cd exposure. Some models also expose pups prenatally to Cd through subcutaneous injections, which does not mimic the natural route of exposure. Additionally, to our knowledge, global transcriptomic and proteomic changes in the brain as a result of

maternal Cd exposure have not been analyzed, but these data could identify pathways and mechanisms that underpin the previously observed phenotypes.

To address these gaps in knowledge, we established a mouse model of maternal Cd exposure in which females are exposed to Cd through their drinking water for five weeks prior to gestation until the birth of a litter, followed by whole transcriptome and proteome analyses on newborn brains. Through this approach, we identified multiple pathways that are perturbed due to maternal Cd exposure, informing on potential mechanisms through which maternal Cd exposure affects neurodevelopment of the offspring.

Materials and Methods

Animal husbandry, Cd exposure, and tissue collection

Animal work and tissue collection were performed as described in Chapter 2. Briefly, 5- to 7-week-old C57BL/6J ('B') and CAST/EiJ ('C') females (the F_0 generation) were exposed through their drinking water to 0 ppm, 1 ppm, or 50 ppm Cd in the form of $CdCl_2$ (Sigma-Aldrich, catalog number 202908) for a period of 5 weeks, then mated with a male of the opposite strain. Cd exposure continued throughout pregnancy and was discontinued at litter birth. Maternally-exposed hybrid offspring (the F_1 generation) were generated to enable analyses of allele-specific gene expression and DNA methylation for a separate study (Chapter 4). The genotypes of F_1 mice are herein referred to as 'BxC' (B mother x C father) and 'CxB' (C mother x B father). F_1 animals were sacrificed and dissected within 24 hours of birth or at 6 months of age. A detailed account of housing conditions and sample sizes used in this study are the same as described in Chapter 2.

All animal work was approved by the North Carolina State University (NCSU) Institutional Animal Care and Use Committee, under protocol 16-045-B. All experiments were conducted in accordance with the Guiding Principles in the Use of Animals in Toxicology.

Behavioral assessments

F_1 mice were subjected to two behavioral tests at approximately 5.5 months of age: open field (OF) and social interaction (SI). Behavioral testing was performed as previously described[114,115]. All testing was performed during the afternoon between 3pm-6pm under dimmed lighting in a secluded procedure room, video recorded for 30 minutes, and scored using TopScan software (Clever Sys Inc.). A randomly selected subset of 2 videos per treatment group and per sex was scored by hand to validate TopScan scoring.

SI was performed in the same box one day after OF testing using an unexposed B mouse of the same sex and of similar age. Pure B strain mice were used as novel test mice in place of age- and gender-matched hybrids for the SI experiments due to the aggression shown by hybrid male mice.

Sample sizes used in behavioral experiments can be found in Supplementary table B3.5 SI and B3.6 OF.

Nucleic acid isolation

Frozen whole brains obtained from eight 0 ppm, eight 1 ppm, and eight 50 ppm newborn female BxC pups representing four litters per exposure group were pulverized using a hammer and liquid nitrogen. 7-10 mg of tissue was then used to extract DNA and RNA simultaneously using the AllPrep DNA/RNA/miRNA kit (Qiagen, 80204). RNA was extracted using the recommended protocol for fatty tissues. RNA and DNA were

quantified using a Nanodrop 2000. RNA purity and size integrity were determined at the NCSU Genomic Sciences Laboratory (GSL) using an Agilent 2100 Bioanalyzer with an RNA 6000 Nano Chip (Agilent Technologies).

RNA-seq

Total RNA from the brains of four 0 ppm and four 50 ppm maternally exposed F_1 mice (each individual representing a different litter) were submitted to the NCSU GSL for indexed library construction and sequencing as described in Chapter 2.

The quality of raw RNA-seq data was assessed using the FastQC application and the initial 10 bases with poor quality were trimmed. Alignment was performed using STAR short read aligner[116] to the respective mouse strain (C57BL/6J, CAST/EiJ) reference genomes. C57BL/6J and CAST/EiJ RNA-seq data were aligned to mm38 version 87 and CAST/EiJ version 1.86 reference genome downloaded from Ensembl website, respectively. The number of reads mapped to a genome feature was determined using htseq-count script from HTSeq python package[117]. Independent analyses were conducted for the two strains. The two raw count data matrices were imported to R statistical computing environment for further analysis. Genes that had no count in more than 2/3 of the replicate samples were excluded from the analysis. Data normalization based on dispersion and differential analysis was conducted using the DESeq2 R package[118]. A generalized linear model (~Treatment) was fitted between the expression count and a treatment group (0 ppm or 50 ppm maternal Cd exposure). Finally, a list of differentially expressed genes was generated between 50 ppm vs. 0 ppm maternal Cd exposure after applying Benjamini-Hockberg multiple testing correction (padj <0.05). The

entire dataset of all genes with detectable and good quality reads aligned to both genomes can be found in Supplementary table B3.1.

qRT-PCR and qPCR

100 ng of total RNA extracted as described above from 24 BxC female newborn F_1 brains was used to synthesize first strand cDNA as described in Chapter 2. qRT-PCR was performed on a Real-Time 7300 machine (Applied Biosystems) using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, 1725271) using the conditions described in Chapter 2. The primer sequences are provided in Supplementary table B3.2. Analysis was performed as described in Chapter 2.

DNA was extracted from 24 BxC female newborn F_1 brains through the Qiagen kit described above. Mitochondrial DNA (mtDNA) was quantified by qPCR as previously described[119,120].

To provide greater transparency and allow for better reproducibility of results, this study was performed in compliance with MIQE standards[87] (see Supplementary table B3.3).

Proteomics

Frozen homogenized whole brains obtained from newborn BxC maternally exposed pups were submitted to the Molecular Education, Technology, and Research Innovation Center (METRIC) at NCSU for proteomic analysis. The brain samples used in the proteomic analysis originated from the four 0 ppm and four 50 ppm females represented in the RNA-seq analysis; additionally, four 1 ppm females representing four different litters were used for proteomic analysis.

Samples were quantitated for protein content using a Pierce bicinchoninic acid (BCA) kit and normalized to 50 µg of protein by diluting the appropriate amount of sample in 50mM ammonium bicarbonate with 1% (w/w) sodium deoxycholate. Normalized samples were reduced with dithiothreitol (DTT) and alkylated with iodoacetamide (IAM) to break disulfide bonds and prevent reformation. Following this, a filter aided sample preparation (FASP) protocol[121,122] was used to purify the samples which were then treated with trypsin at a 50:1 protein:trypsin concentration. Samples were incubated for 4 hr and then aliquoted for LC-MS analysis.

Chromatographic separation was achieved using a Thermo Scientific EASY nLC 1200 system (Bremen, Germany). Pico-frit columns were purchased from New Objective (Woburn, MA) and bomb packed to a length of 20-30 cm with reverse phase ReproSil-Pur 120 C-18- AQ 3 µm particles (Dr. Maisch, Germany). Two microliters of sample is injected and subsequently separated using a gradient of mobile phase A (98 % water, 2 % acetonitrile, and 0.1 % formic acid) and mobile phase B (80 % acetonitrile, 20% water 0.1 % formic acid). The LC method consists of a 120 minute gradient with a linear ramp from 0 % B to 40 % B, a 1 minute ramp to 100% B which is held for 6 minutes (123-129 minutes), followed by equilibration of the column at 0% B (130-140). A flow rate of 300 nL/min was used.

Orbitrap tandem mass spectrometry was performed using a Thermo Scientific Q-Exactive HF (Bremen, Germany) in a top 20 data dependent acquisition (DDA) mode, where the 20 most abundant precursors are selected for fragmentation per full scan. MS1 and MS2 scans were performed at a resolving power of 120,000 and 15,000 at m/z 200, respectively. A dynamic exclusion window of 20 seconds was used to avoid repeated

interrogation of abundant species. Automatic gain control (AGC) was 1e6 and 1e5 for MS1 and MS2 scans, respectively. Samples were run in random order, and a quality control BSA digest was run every fifth injection to ensure proper LC-MS/MS reproducibility.

Resulting raw data were loaded into Proteome Discoverer (version 2.0). Spectrum files were run through the spectrum selector node to appropriately identify peaks and this data was collated using the Sequest HT node and aligned with a FASTA file which contained all proteins indexed by Uniprot and assigned a taxonomy ID of 10090 (mus musculus). Oxidation (dynamic) and carbamidomethylation (static) modifications were considered in Sequest HT with a maximum of 2 potential missed cleavage sites and peptide lengths between 6 and 144 amino acids. The Percolator node was used for false discovery rate (FDR) calculations.

Data was then filtered to remove the following: abundance ratio adj. p-value (50 ppm vs 0 ppm) >0.05, # unique peptides >1, Abundances (grouped) CV% 50 ppm >20, Abundances (grouped) CV% 0 ppm >25, proteins with a combination of low PEP score and CV% >15, PEP score <7, and proteins that had a 'Found in Sample' value of any value lower than 'High' for at least 2 individuals.

RNA-seq and proteomic enrichment analysis

The lists of 15 significantly differentially expressed genes and 46 proteins significantly different in abundance in the 50 ppm group compared to controls were submitted for separate Enrichr analyses[84,85] at http://amp.pharm.mssm.edu/Enrichr/ . Databases with relevant terms were considered and the comprehensive lists can be found in Supplementary table B3.4.

Retinoid quantification

Whole brains of newborn BxC maternally-exposed pups were removed within 30 seconds of pup sacrifice under yellow light, placed in a black 1.7 ml microcentrifuge tube to prevent exposure to white light, then immediately weighed and flash frozen on dry ice. Sample sizes were as follows: 16 0 ppm females, 5 1 ppm females, 11 50 ppm females, 11 0 ppm males, 7 1 ppm males, and 8 50 ppm males. Samples were kept at -80°C until they were shipped on dry ice for retinoid quantification at the University of Maryland.

Retinoid concentrations were measured as described previously[123–125]. Briefly, an average of 85.2 ± 9.2 mg (mean ± s.d.) (range 67.1 - 117.1 mg) per sample of brain was homogenized in ground glass homogenizers in 1.0 mL normal saline (0.9% NaCl). Extraction of retinoids was performed under yellow lights using a two-step liquid-liquid extraction that has been described in detail previously using 4,4-dimethyl-RA as an internal standard for RA and retinyl acetate as an internal standard for retinol and total retinyl ester[123,125]. RA levels were determined by liquid chromatography-multistage tandem mass spectrometry (LC-MRM3), which is an LC-MS/MS method utilizing two distinct fragmentation events for enhanced selectivity[123]. RA was quantified using a Shimadzu Prominence UFLC XR liquid chromatography system (Shimadzu, Columbia, MD) coupled to an AB Sciex 6500+ QTRAP hybrid triple quadrupole mass spectrometer (AB Sciex, Framingham, MA) using atmospheric pressure chemical ionization (APCI) operated in positive ion mode as previously described[123]. Retinol (ROL) and Retinyl Ester (RE) were quantified via HPLC-UV using an AQUITY H-Class UPLC with a PDA detector (Waters Corporation, Milford, MA) operated in single wavelength monitoring mode at

325nm according to previously published methodology[124,125]. Retinoids in tissues are expressed as mol/g tissue.

Statistical analysis

Unless otherwise noted, all statistical analyses were performed using a one-way analysis of variance (ANOVA) followed by a post-hoc Dunnett's multiple comparison test using R software, comparing animals from the 1 ppm and 50 ppm groups to 0 ppm controls. Outliers, as defined by a point which falls more than 1.5 times the interquartile range above the third quartile or below the first quartile, were omitted from the analysis if found to be present. Data are presented as the mean ± standard error of the mean. *p<0.05, **p<0.01, ***p<0.001.

Results

Maternal Cd exposure increases the proportional mass of the brain at birth

Pups that were exposed maternally to 50 ppm Cd demonstrated a significant reduction in body mass at birth compared to control animals, as reported in our previous publication (Chapter 2). We also previously demonstrated increases in newborn heart mass in these animals associated with a susceptibility to hypertension in adulthood (Chapter 2). In the current study, we focused on the brain because of the known association of Cd with neurobehavioral defects. No significant differences in raw brain masses were seen at birth as a result of maternal Cd exposure, with the exception of a significant increase seen in the 1 ppm BxC males and a significant decrease seen in the 50 ppm CxB males (Figure 3.1A). However, when brain mass was normalized to body mass, a significant increase in the proportional mass of the brain was seen at birth in the

50 ppm maternally-exposed pups, regardless of cross or sex (Figure 3.1B). As previously reported, no body mass differences were observed at 6 months, with the exception of a significant decrease seen in 50 ppm maternally-exposed BxC females (Chapter 2). Differences in proportional or raw brain masses were not observed at 6 months, with the exception of a significant increase in proportional brain mass in CxB females (Supplementary figure B3.1).

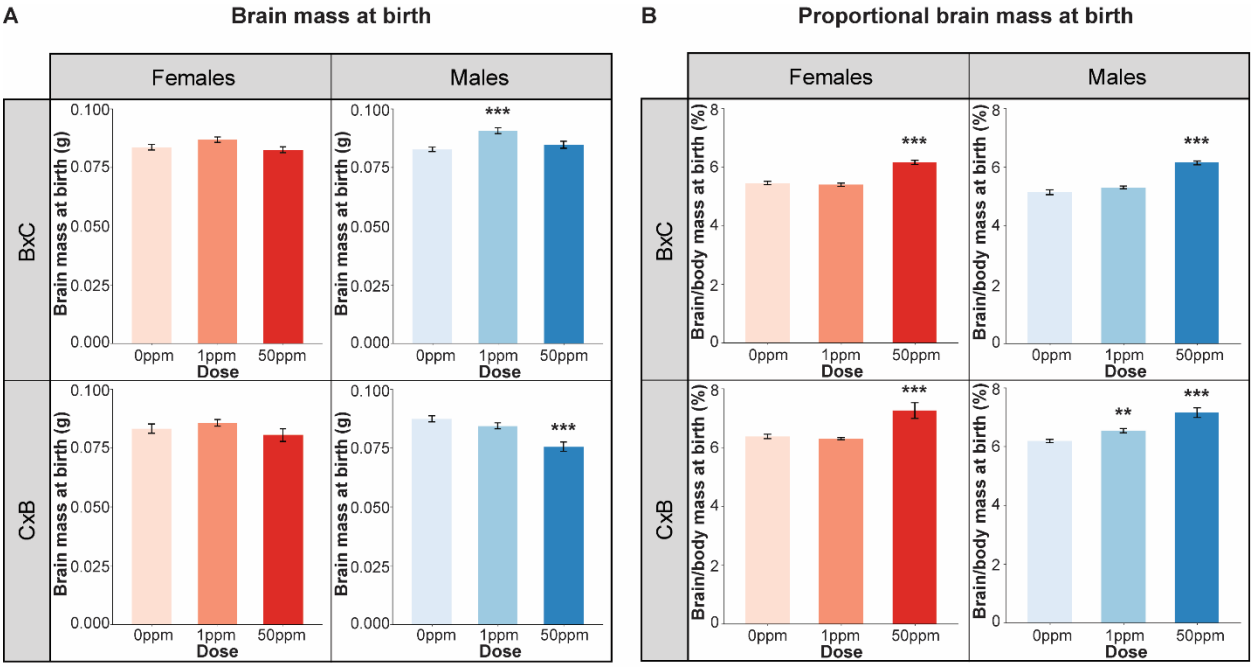


Figure 3.1. The effects of maternal Cd exposure on brain mass and normalized brain mass at birth. **A,** Brain mass at birth. **B,** Proportional brain mass at birth. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Maternal Cd exposure programs perturbed anxiety-related behavior during adulthood

Due to the significant increase in proportional brain mass at birth and the previously reported association with early-life exposure to Cd and altered behavior, Social Interaction (SI) and Open Field (OF) behavioral tests were performed in 5.5-month-old mice to test whether altered behavior in adulthood can be programmed by maternal Cd

exposure alone. The SI test is used to assess sociability in the presence of a novel animal[126], while the OF test is used to measure activity, anxiety, and willingness to explore[127].

Maternal Cd exposure had a more profound and divergent effect on sociability. Maternally-exposed BxC males and females investigated the novel animal more frequently and spent a longer duration of time near the novel animal (Figure 3.2, Supplementary table B3.5 SI). Maternally-exposed BxC animals appeared to initially investigate the novel animal more quickly, though these data were not statistically significant (Supplementary table B3.5 SI). CxB males and females appeared to investigate the novel animal less frequently and took longer to initially investigate, though these data were not statistically significant (Supplementary table B3.5 SI).

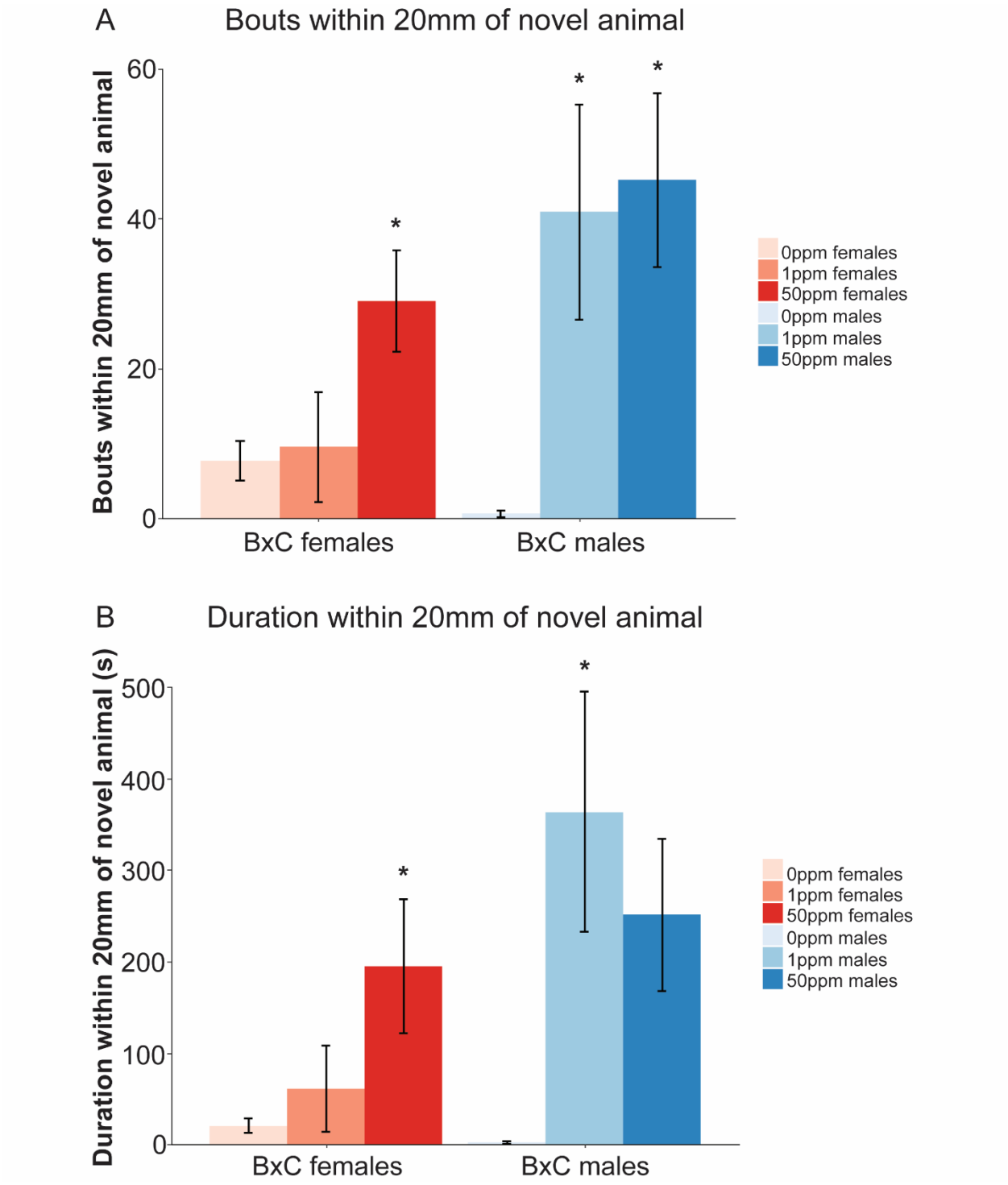


Figure 3.2. Select SI behavioral measurements in BxC mice maternally-exposed to Cd. **A,** Bouts within 20mm of novel animal. **B,** Duration within 20mm of novel animal. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

1 ppm maternally-exposed CxB males showed a significant increase in activity, willingness to explore the center of the box, and took less time to initially explore the center compared to control CxB males (Supplementary table B3.6 OF). The other groups of maternally-exposed mice also displayed an increase in the number of times they investigated the center, although these data were not statistically significant (Supplementary table B3.6 OF). Maternally-exposed BxC females appeared to go to the perimeter of the box more frequently and took less time to initially explore the center, though these data were only significant in the 1 ppm group for the number of times the explored the perimeter (Supplementary table B3.6 OF).

Given their significantly altered behavior at adulthood, BxC females were chosen for further molecular analyses.

Transcriptomic analysis of newborn BxC female brains identifies maternal Cd-associated activation of myelin markers

In order to understand altered pathways that could be contributing to enlarged proportional brain mass at birth and perturbed behavior at adulthood as a result of maternal Cd exposure, RNA-seq was performed on whole brains of newborn control and 50 ppm maternally-exposed BxC females. Over 13,000 genes had detectable reads and reliable alignments considering both parental genomes (14,417 aligned to B genome, 13,277 aligned to C genome, Supplementary table B3.1). Out of these data, 15 genes were found to be significantly differentially expressed when aligned to both B and C genomes (Table 3.1, Supplementary figure B3.2). Twelve out of 15 of these genes are associated with oligodendrocytes or myelination (Table 3.1), a process performed by oligodendrocytes that occurs after birth in the mouse[128]. Eight out of 15 of these genes

are known to be part of the myelin transcriptome[129] (Table 3.1). Through qRT-PCR analyses of 0, 1, and 50 ppm maternally-exposed newborn BxC female brains, expression of two classical myelin markers, *Mag* and *Mbp*, was found to be significantly increased in a dose-dependent manner (Figure 3.3). *Cnp*, which facilitates events early in myelination[130], was also significantly increased in 50 ppm maternally-exposed pup brains (Table 3.1). Term enrichment analysis for the 15 differentially expressed genes (DEGs) showed a significant enrichment for terms related to retinoic acid (RA) signaling (Table 3.2), which is required prior to the onset of myelination[131]. Based on these observations, we hypothesized that the process of myelination is occurring earlier in 50 ppm maternally-exposed F_1 mice. However, we were unable to detect myelin in these animals at birth using histological approaches (data not shown).

Table 3.1. List of 15 significantly differentially expressed genes in newborn BxC female brains as a result of 50 ppm maternal Cd exposure and their relevance to myelin. Numerical values are derived from the alignment to the B genome.

Gene	Log2Fold Change	Adjusted p-value	Associated with oligodendrocytes or myelination	Part of myelin transcriptome[129]
Hoxb8	6.87	6.71E-14		
Bcas1	0.98	2.26E-05	X[129,132]	X
Il33	0.89	3.02E-03	X[129,133]	X
Enpp6	1.66	3.02E-03	X[134]	
Mag	1.16	8.21E-03	X[129,135]	X
Kank1	0.48	8.88E-03	X[136]	
Sparcl1	0.43	8.88E-03		
Aldh1a2	0.44	8.88E-03	X[129,137]	X
Sp9	-0.40	8.88E-03	X[129]	X
Dlx1	-0.42	1.24E-02	X[138]	
Atp1a2	0.29	1.47E-02		
Mbp	0.49	3.13E-02	X[129,139]	X
Kcnj16	0.69	3.75E-02	X[140]	
Cnp	0.52	4.16E-02	X[129,130]	X
Ptgds	0.51	4.88E-02	X[129,141]	X

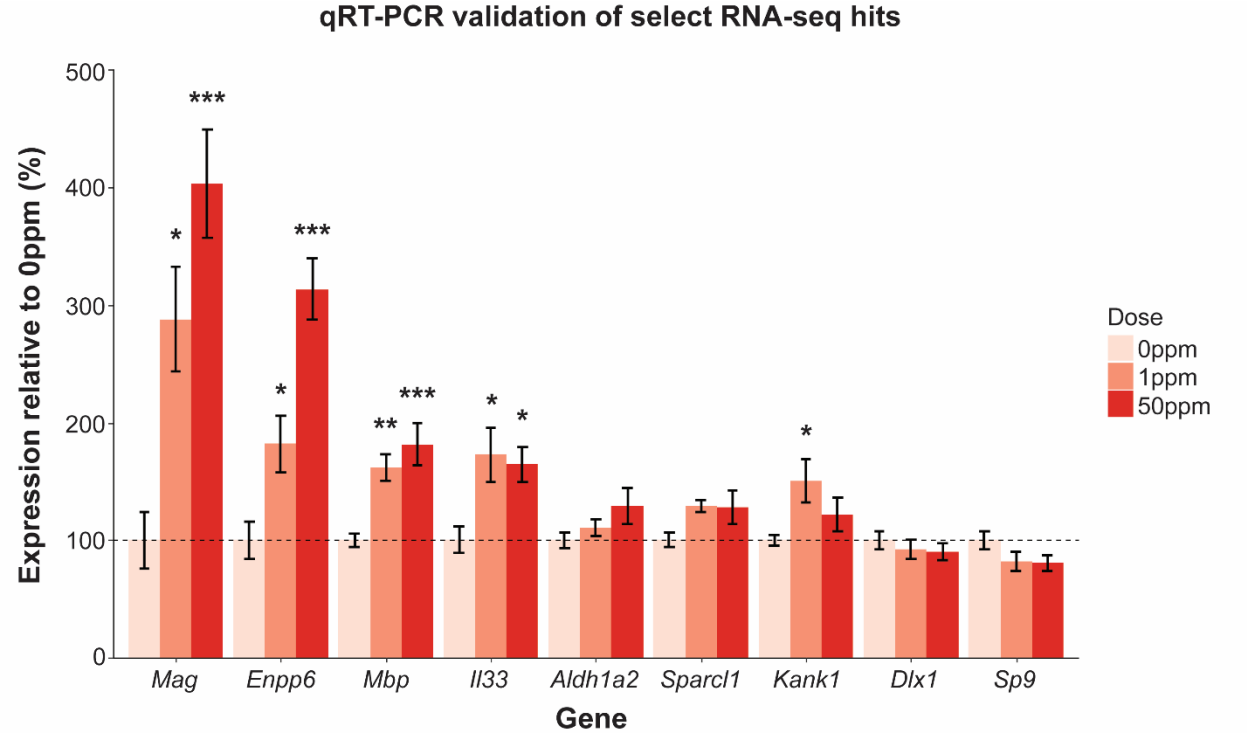


Figure 3.3. qRT-PCR validation of select RNA-seq hits in newborn BxC female brains due to maternal Cd exposure. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Table 3.2. Enrichment for RA terms in RNA-seq data

Term	Database	Adj p val
retinoid binding (GO:0005501)	GO molecular function	4.21E-03
retinoate biosynthesis I_Homo sapiens_PWY-6872	HumanCyc	6.73E-03
retinal dehydrogenase activity (GO:0001758)	GO molecular function	2.91E-02
Retinol metabolism_Mus musculus_WP1259	WikiPathways	4.24E-02
Vitamin A and Carotenoid Metabolism_Homo sapiens_WP716	WikiPathways	4.24E-02

Maternal Cd expression perturbs Hox *gene expression in newborn BxC female brains*

Hoxb8 was identified as the most significantly DEG in the RNA-seq data (Table 3.1). *Hoxb8* appeared to have little to no expression in the controls, consistent with a more posteriorly located expression of this gene during development[142], yet was

significantly increased in expression in the 50 ppm maternally-exposed pup brains (Table 3.1, Supplementary table B3.1). Upon further evaluation of the gene expression levels of transcripts detected through RNA-seq, it appeared that other *Hox* genes were similarly affected, although their differences in expression did not achieve statistical significance (Supplementary table B3.1). To further assess the extent of *Hox* gene misregulation by Cd exposure, *Hox* 1-9 genes contained in the A-D clusters were quantified with qRT-PCR in newborn BxC female brains exposed maternally to 0 ppm, 1 ppm, or 50 ppm (Figure 3.4). Only consistent amplification in at least one of the treatment groups was detected in the *Hoxa* and *Hoxb* clusters. 50 ppm maternal Cd exposure significantly increased the expression of more posteriorly expressed *Hox* genes (i.e. 5-8), and both 1 ppm and 50 ppm maternal Cd exposure significantly increased the expression of *Hoxb2* (Figure 3.4).

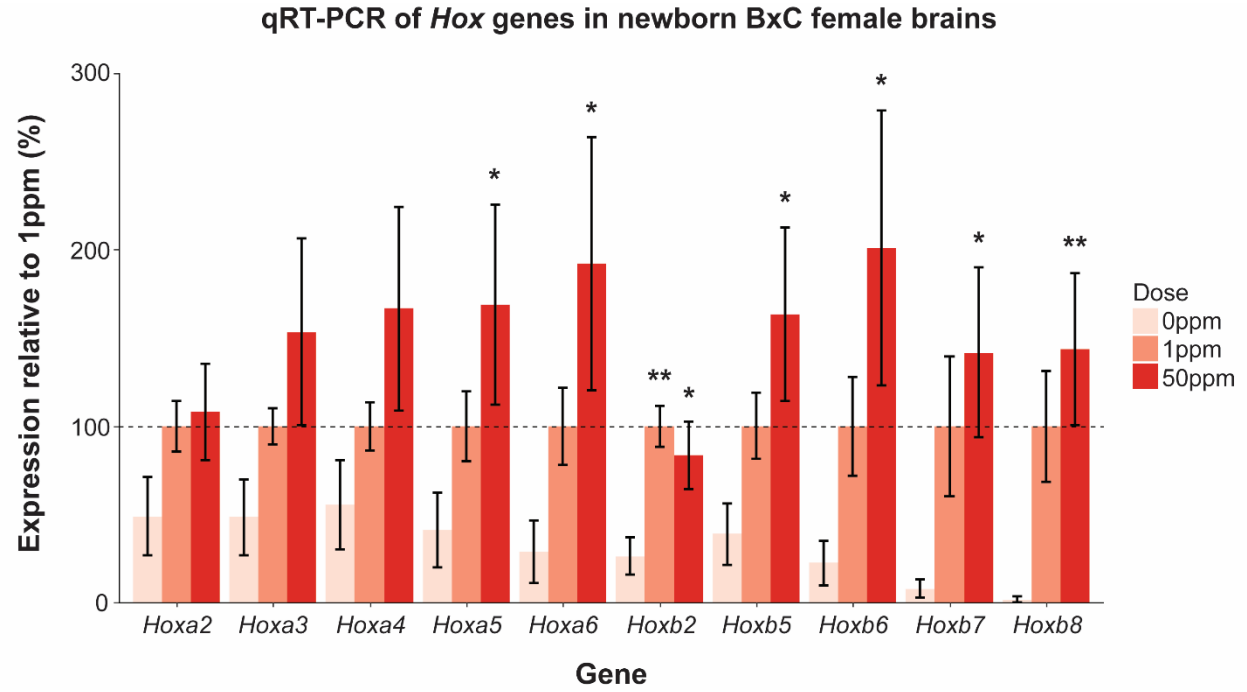


Figure 3.4. qRT-PCR of *Hox* genes in newborn BxC female brains exposed maternally to Cd. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Proteomic analysis of maternally-exposed BxC female newborn brains identifies altered abundances of proteins related to cellular energy, hypoxia, histone 1 subunits, and myelin

To determine whether changes in gene expression correlated with altered abundance at the protein level, a discovery-based mass-spectrometry proteomics approach was performed on newborn BxC female whole brains; this method detects levels of abundant proteins in the tissue of interest. 46 proteins were identified as significantly different in abundance in the 50 ppm group compared to controls after correcting for multiple testing and filtering hits based on criteria described in the Materials and Methods (Table 3.3). In general, a dose-dependent increase or decrease in protein abundance was seen as a result of maternal Cd exposure, although this was only significant for 6/46 proteins in the 1 ppm group compared to controls.

Table 3.3. Proteomics discovery list of significantly differentially abundant proteins in the 50 ppm BxC female brains compared to controls, ranked by adj p val in 50 ppm group. [1]=found in myelin proteome, [2]=RA-related, [3]=histone 1 subunit.

Protein	Abundance Ratio Adj. P-Value: (1ppm) / (0ppm)	Abundance Ratio Adj. P-Value: (50ppm) / (0ppm)	Abundances (Grouped): 0ppm	Abundances (Grouped): 1ppm	Abundances (Grouped): 50ppm	1ppm change to 0ppm (%)	50ppm change to 0ppm (%)
Ldha[1,2]	0.973898894	4.72809E-06	90.9	91.8	117.3	0.990099	29.042904
Pfkl	0.119323304	4.76849E-06	86.1	98.2	115.7	14.053426	34.37863
Slc6a11	0.046109095	1.13451E-05	88.7	97.7	113.6	10.146561	28.072153
Nefm[1,2]	1.45337E-05	1.39145E-05	84.7	104.7	110.6	23.612751	30.578512
Gpi1[1]	0.99117914	0.00012503	90.7	92.5	116.9	1.9845645	28.886439
Aldoc[1,2]	0.014283746	0.0002136	86.9	102.2	110.8	17.606444	27.502877
Hist1h1e[3]	0.680950614	0.000458262	108.9	103.4	87.7	-5.050505	-19.4674
Slc1a3[1,2]	0.026706171	0.000459068	94.3	95	110.7	0.7423118	17.391304
Srsf4[2]	0.254416432	0.000500359	109.8	102	88.2	-7.103825	-19.67213
Tuba4a[1,2]	0.812211612	0.001034191	91	100.3	108.7	10.21978	19.450549
Hpca	0.818093421	0.001676617	92	96.1	111.9	4.4565217	21.630435
Tpi1[1,2]	0.995738881	0.001737399	94.3	93.7	111.9	-0.636267	18.663839
Ahsg[2]	0.259432391	0.001766066	98.4	86.8	114.8	-11.78862	16.666667
Fxyd6	0.084419728	0.0037873	91.9	98	110.1	6.6376496	19.804135
Ckmt1	0.991151472	0.004356406	94.8	94.8	110.3	0	16.350211

Table 3.3 Continued

Hist1h1b[3]	0.995738881	0.004617561	105.6	108.1	86.2	2.3674242	-18.37121
Hprt	0.926735568	0.004685861	95.2	89.9	114.9	-5.567227	20.693277
Aldoa[1,2]	0.937783837	0.004964389	92.4	93.8	113.9	1.5151515	23.268398
Hist1h1d[2,3]	0.606980956	0.004964389	108.6	102.1	89.2	-5.985267	-17.86372
Ptma[2]	0.080791892	0.004964389	112.2	97.6	90.2	-13.01248	-19.60784
Glul[1,2]	0.729109325	0.005102008	91.5	98.8	109.8	7.9781421	20
Prkacb[2]	0.063552234	0.005317616	99.5	90.1	110.4	-9.447236	10.954774
Sncb[1]	0.585421643	0.005422114	93.9	91	115.1	-3.088392	22.57721
Slc12a5	0.970511751	0.006450782	91.6	95.3	113.1	4.0393013	23.471616
Apoe[2]	0.680950614	0.008266455	89.6	96.6	113.7	7.8125	26.897321
Alg2	0.31039933	0.008520425	90.4	98	111.6	8.4070796	23.451327
Apoa1	0.991151472	0.009864999	89.7	94.9	115.4	5.7971014	28.651059
Eno1[1,2]	0.674411211	0.010566996	93.7	92.4	113.9	-1.387407	21.558164
Abat[2]	0.989354576	0.012954966	92.4	94	113.6	1.7316017	22.943723
Crabp1[2]	0.04344094	0.013498454	85.5	103.8	110.7	21.403509	29.473684
Camk2d[1,2]	0.580890155	0.018126666	90.3	96.7	113	7.0874862	25.138427
Satb2	0.973898894	0.01885168	108.5	110.1	81.4	1.4746544	-24.97696
Ndrg2[2]	0.680375748	0.018905298	94.6	95.9	109.5	1.3742072	15.750529
Dnm1[1,2]	0.174291855	0.022293486	93.1	98.7	108.2	6.0150376	16.219119
Trf	0.176950766	0.025109111	96.1	91.7	112.2	-4.578564	16.753382
Vsnl1[1,2]	0.909390199	0.029396495	97.2	93.2	109.6	-4.115226	12.757202
Gstm1[1,2]	0.989354576	0.031417367	93.9	93	113.1	-0.958466	20.447284
Set	0.031217009	0.031978606	106.2	99	94.8	-6.779661	-10.73446
Syn2[1]	0.426588152	0.032509188	88.8	100.9	110.3	13.626126	24.211712
Anxa2[1,2]	0.797814818	0.03380452	95.3	95.6	109.1	0.3147954	14.480588
Trp53i11	0.737170178	0.034474205	107.9	103	89.2	-4.541242	-17.33086
Hpcal1[1,2]	0.818093421	0.038283218	94.2	96.9	108.9	2.866242	15.605096
Stmn2	0.419569789	0.039318421	89.2	102.9	107.9	15.358744	20.964126
Camk2g[1,2]	0.836358682	0.043020031	87.7	102.9	109.3	17.331813	24.629418
Dpp6[1,2]	0.389249664	0.04777657	88.9	97.4	113.7	9.5613048	27.896513
Glud1[2]	0.474755049	0.048280644	92.8	98.4	108.9	6.0344828	17.349138

Among the 46 differentially abundant proteins, there was a significant enrichment of functional terms or pathways related to cellular energy and metabolism pathways and hypoxia (Supplementary table B3.4), consistent with a known Fe deficiency due to maternal Cd exposure (Chapter 2). In addition, three histone 1 subunits were significantly decreased in 50 ppm brains compared to controls (Table 3.3). Finally, 20/46 proteins that are known to be part of the myelin proteome[143] were significantly different in abundance

in 50 ppm brains compared to controls, and all were significantly increased in abundance (Table 3.3). The most significant term in the ESCAPE enrichment analysis predicted an increase in *Sox2* (Supplementary table B3.4), which is essential for the proliferation and differentiation of oligodendrocytes during postnatal myelination[136]. Consistent with the findings of RNA-seq, there was a significant enrichment for terms related to RA (Table 3.4).

Table 3.4. Enrichment for RA terms in proteomics data.

Term	Database	Adj p val
Retinoic acid_CTD_00006918	DSigDB	4.09E-05

Mitochondrial content is reduced in newborn BxC female brains due to maternal Cd exposure

Mitochondrial DNA (mtDNA) content was quantified in newborn BxC female brains through qPCR due to the proteomic analysis indicating alterations in cellular energy pathways due to maternal Cd exposure. mtDNA content was significantly reduced to approximately 75% of controls in newborn BxC female brains as a result of both 1 ppm and 50 ppm maternal Cd exposure (Figure 3.5).

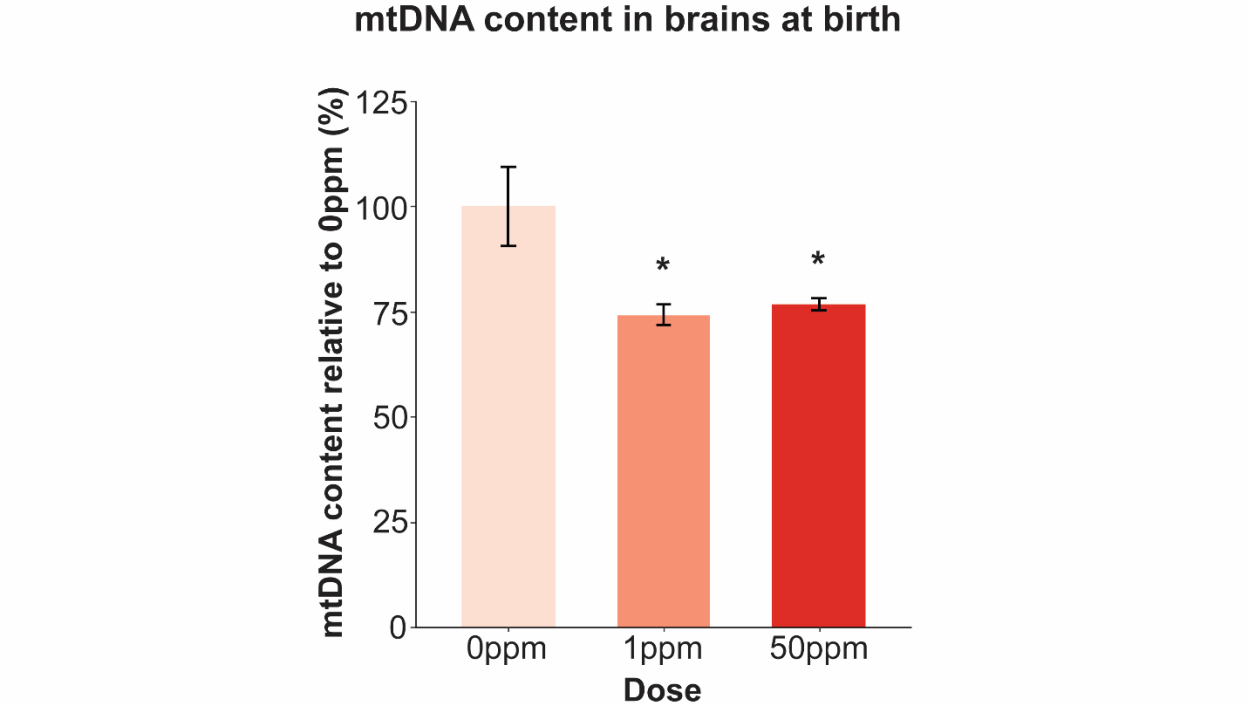


Figure 3.5. mtDNA content in BxC female newborn brains as a result of maternal Cd exposure. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Maternal Cd exposure leads to increased levels of retinoic acid in the brain at birth

Due to significantly altered levels of RA-associated transcripts and proteins (Tables 3.2 and 3.4), and the known role of RA in behavior[144], myelination[131], and *Hox* gene expression regulation[142], we hypothesized that RA levels in the brain were perturbed at birth as a result of maternal Cd exposure. To test this, we quantified three retinoid species in newborn BxC male and female whole brains: RA, retinol (RA substrate), and retinyl ester (retinol storage). Maternal Cd exposure led to significantly increased levels of RA in both 1 ppm and 50 ppm maternally-exposed pups (Figure 3.6B). The 1 ppm maternally-exposed pups tended to have higher levels of RA than the 50 ppm maternally-exposed pups, though this was not statistically significant (p=0.18). Retinol levels were not significantly altered as a result of maternal Cd exposure, except in the 50 ppm

maternally-exposed pups when male and female data were pooled, indicating increased biosynthesis or reduced degradation of RA (Figure 3.6A). Retinyl ester was increased in the brains due to maternal Cd exposure, although this was not significant in the 1 ppm maternally-exposed females (Figure 3.6C).

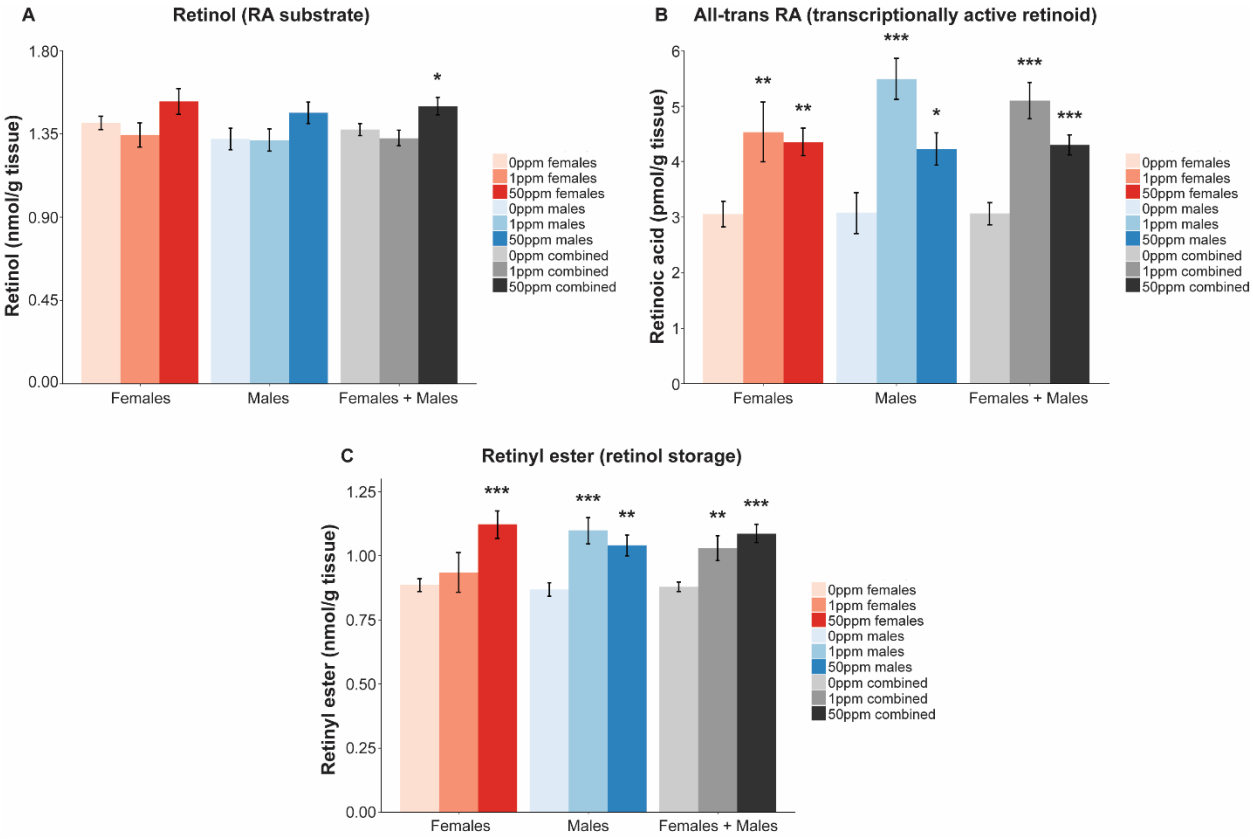


Figure 3.6. Mass-spec quantification of RA and RA metabolites in newborn BxC brains as a result of maternal Cd exposure. **A,** Retinol. **B,** All-trans RA. **C,** Retinyl ester. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Discussion

Exposure to Cd during early life is associated with impaired neurodevelopment and behavior[21,95,109,110], but the underlying molecular mechanisms responsible for these changes are poorly understood. In the present study, we utilize a mouse model of maternal Cd exposure and demonstrate that maternal exposure alone is sufficient to

program altered behavior among F_1 mice during adulthood. To inform on the molecular mechanisms underlying these behavioral defects, we identified exposure-associated transcriptomic and proteomic changes in the brain at birth. Together, these data suggest that maternal Cd exposure may cause premature myelination in the brain and inappropriate regulation of *Hox* gene expression. These data also suggested that Cd perturbs RA signaling in the developing brain, which we subsequently confirmed with targeted analyses. This work provides a foundation for further mechanistic study to understand the mechanisms of Cd action.

Two doses of Cd were used in this study: 1 ppm Cd, to reflect an environmentally relevant level of routine Cd exposure considering the average levels of Cd naturally occurring in the Earth's crust (Chapter 2); and 50 ppm Cd, as this dose results in circulating blood Cd levels consistent with those of people living in heavily Cd polluted areas (Chapter 2). In some of our behavioral assays, the low dose of Cd had more pronounced effects than the high dose; this supports the emerging frequency of non-monotonic dose-responses in toxicological studies[145].

In rodent brains, myelination is an event that primarily occurs during postnatal development[128]. Our RNA-seq and proteomic analyses suggest an earlier onset of myelination due to maternal Cd exposure. 12 genes out of the 15 DEGs identified as a result of maternal Cd exposure are involved in myelination or oligodendrocyte differentiation during myelination, and two classical myelin markers, *Mbp* and *Mag*, were significantly increased in pup brains in a dose-dependent manner. In addition, 20 proteins that are part of the myelin proteome were all found to be significantly increased in abundance due to 50 ppm maternal Cd exposure. Finally, we have shown that maternal

Cd exposure increases circulating fetal copper levels (Chapter 2), and copper is required by proteins involved with myelination[104]. To date, the link between Cd and myelin only considers postnatal or adult exposure to Cd and suggests a negative effect on myelin[146,147]. However, one study in rats that induced a delay in myelination through lysophosphatidylcholine noted opposite behavioral effects as seen in this study (e.g., decreased activity and increased anxiety-related behaviors)[148]. There is emerging evidence that myelin affects behavior[134], thereby supporting a potential mechanistic link through which maternal Cd exposure impacts offspring behavior. Despite that myelin proteins are normally present in the brain as early as 7 days postnatally in the rodent[128], this is difficult to detect histologically, consistent with our inability to detect myelin in newborn maternally-exposed pup brains with black-gold staining (data not shown).

Hox gene expression is also inappropriately regulated following maternal Cd exposure. The *Hox* gene family is conserved across metazoan species and is essential during development to confer anterior to posterior identity to tissues. *Hox* gene expression patterns are uniquely linked to their position along their respective chromosomes: 3' *Hox* genes are expressed in the anterior end of the embryo, while 5' *Hox* genes are expressed more posteriorly. *Hox* expression is regulated in part by a RA gradient in the developing embryo. Mammals have four clusters of *Hox* genes (A, B, C, D), each cluster located on a different chromosome, and the genes are numbered 1-13 based on their 3' to 5' location within their cluster[142]. *Hox* 1-4 paralogs are expressed in the mammalian central nervous system, primarily in the developing hindbrain; the remaining *Hox* 5-13 paralogs are expressed in the spinal cord and other posterior tissues[149]. In this study, we have shown that expression of the *Hoxa* and *Hoxb* clusters is perturbed due to maternal Cd exposure,

with regards to 1-4 paralogs that are normally expressed in the brain as well as inappropriate expression of more posteriorly expressed *Hox* genes. *Hox1-4* paralog expression in the developing central nervous system is essential for linking neurons involved with motor activity to their behavioral output[150]. *Hoxb8*, identified through RNA-seq as the most significantly differentially expressed gene in the brain due to maternal Cd exposure, has been shown to affect behavior; mutant mice display excessive and pathological grooming behaviors, and *Hoxb8* expression has been found in obsessive-compulsive disorder neural circuits in humans[149]. *In vitro*, direct exposure to Cd has been shown to induce expression of *Hoxb8*[151]. Additionally, altered *Hox* expression may explain how maternal Cd exposure can lead to congenital malformations, as mutations in *Hox* genes can lead to drastic developmental malformations[152].

The brain is the most energy expensive organ in the body. Despite only comprising 2% of the body mass, the brain utilizes 20% of the body's energy while at rest. Proper circulating levels of nutrients such as glucose and oxygen are essential to meet this demand[153]. We have previously shown that maternal Cd exposure leads to systemic alterations of essential trace elements in the offspring at birth, including a severe iron deficiency, and that cellular energy pathways may be negatively impacted as a result (Chapter 2). The proteomic analysis and reduced mitochondrial content seen in newborn brains here are consistent with perturbed cellular energy and hypoxic conditions due to maternal Cd exposure. Because of the important role of the brain, this organ is typically spared during intrauterine growth restriction. In response to hypoxic or nutrient-poor conditions, the fetus adapts by rerouting its circulation to conserve oxygen and nutrient

supply for the brain[154]. This brain-sparing phenomenon may at least partially explain why maternal Cd exposure did not significantly alter the raw mass of the brains at birth.

To our knowledge, this study is the first to directly quantify increased levels of RA in newborn brains as a result of maternal Cd exposure. RA is derived from vitamin A (e.g., retinol) and precise RA signaling is essential during development[155]. RA is required for the onset of myelination[131] and *Hox* gene expression is controlled in part by a rostrocaudally decreasing gradient of RA along the body axis[150]. Abnormal levels of RA during development, either excess or reduced, lead to embryological defects, affect hindbrain development, and alter *Hox* gene expression[149]. Excess RA during development can lead to profound behavioral abnormalities and mental retardation in the offspring[144]. *In vitro* and adult exposure to Cd has been shown to induce RA signaling, and Cd teratogenicity is suspected to be at least in part due to increased RA as a result of disruption to RA-metabolizing genes[156], consistent with the data in this study.

The RNA-seq and proteomic analyses of newborn brains maternally exposed to Cd are likely to be affected by the cellular heterogeneity of whole brains. Brain heterogeneity can bias sequencing analyses[157] and may exclude region- or cell-specific changes and limit differential expression analysis to abundant transcripts, consistent with the low number of differentially expressed genes identified in this study. The proteomic analysis can only quantify levels of proteins that are abundant in tissue, and therefore can miss detection of proteins found at lower levels in the cell (e.g., transcription factors). This may explain the lack of overlap between genes identified as differentially expressed through RNA-seq and the proteins found differentially abundant. Despite these caveats, we were able to detect an enrichment for RA-associated terms in both datasets, which

was supported by the retinoid species quantification. In addition, 5 proteins with roles in cellular energy metabolism that were found to be significantly increased in abundance here (Ldha, Pfkl, Aldoc, Anxa2, Aldoa) were also found to be significantly increased at the level of gene expression in newborn hearts as a result of maternal Cd exposure (Chapter 2), suggesting that maternal Cd exposure-associated alterations to cellular energy metabolism may be apparent in multiple organs.

In summary, we have shown global changes to the transcriptome and proteome in newborn mouse brains as a result of maternal Cd exposure. We have identified multiple mechanistic pathways through which maternal Cd exposure alone perturbs neurodevelopment and behavior during adulthood. These results support findings of other studies that suggest a neurotoxic effect of early-life Cd exposure and provide new insights into understanding the link between maternal Cd exposure and impaired neurodevelopment.

Acknowledgements

This work was funded by the National Institutes of Health under K22ES027510 and P30ES025128. We thank Kylie Rock and the Patisaul lab at NCSU for their assistance and use of equipment for the behavioral experiments. Next-generation sequencing was performed by the NCSU Genomics Sciences Laboratory (Raleigh, NC). RNA-Seq data analysis was performed by the Bioinformatics Core associated with the Center for Human Health and Environment (CHHE) at NCSU. We would like to thank the following individuals for their contribution to this study: Emily Shiver, Sanya Mehta, Evan Walsh, Dereje D. Jima, Jeffrey Enders, Jianshi Yu, Maureen Kane, and Heather Patisaul.

CHAPTER 4

Quantifying Allele-Specific Changes in DNA Methylation and Gene Expression in Newborn Pup Brains due to Maternal Cadmium Exposure

Abstract

Adverse exposures during early-life development can lead to susceptibility to disease later in life. Maternal diet and exposure to endocrine disrupting chemicals or heavy metals can impact the health of the offspring. Epigenetic marks, such as DNA methylation, are a strong candidate for a mechanism that provides memory of early-life exposures, due to their stability throughout development and into adulthood. Genomic imprinting is regulated by allele-specific DNA methylation at Imprinting Control Regions (ICRs). ICRs have unique DNA methylation dynamics during development and it has been suggested that they are more sensitive to early-life environmental exposures than other methylated areas of the genome. Allele-specific DNA methylation and imprinted gene expression are intimately related; loss of this relationship is associated with a variety of diseases and may provide a mechanistic link between adverse early-life exposures and susceptibility to disease. We have shown in humans that higher maternal blood Cd is associated with altered DNA methylation in fetal cord blood, with a significant enrichment for ICRs among the differentially methylated regions. As human epidemiological studies are limited in scope with regards to being able to determine allele-specificity of sequencing data, we developed a hybrid mouse model of maternal Cd exposure to further explore the hypothesis that ICRs are sensitive to adverse maternal exposures. In addition to quantifying allele-specific DNA methylation changes and because ICRs control imprinted gene expression, we also quantified allele-specific changes in gene expression due to maternal Cd exposure. Finally, we identified differentially-methylated regions independent of parental genotype due to maternal Cd

exposure. Results from this study provides insights into how maternal exposures can perturb the fetal epigenome and lead to disease susceptibility in the offspring.

Introduction

Early-life exposures to environmental stressors can impact the future health of the individual, even long after cessation of exposure. This is encompassed in the Developmental Origins of Health and Disease (DOHaD) hypothesis, which originates from studies performed by David Barker during the 20th century. He found a correlation between low birthweight in the 1920's and heart disease roughly 50 years later and hypothesized that "an environment which produces poor fetal and infant growth is followed by an adult environment that determines high risk for ischemic heart disease"[23]. The DOHaD hypothesis is widely accepted and continues to be supported through epidemiological studies and animal models of early-life exposure. While Barker's original hypothesis focused on adverse nutrition during early life, we now know that other early-life exposures, such as maternal exposure to endocrine disrupting chemicals (EDCs)[25] and heavy metals[24], can influence susceptibility to disease states later in life.

It is proposed that early-life exposures can alter the epigenome, leading to a lasting memory of exposure that can influence gene expression and susceptibility to disease later in life. There is mounting evidence that maternal diet[26] and maternal exposure to heavy metals[99] or EDCs[35] can alter the epigenome of the offspring. Multiple epigenetic marks have been shown to be altered due to early-life exposures and include chromatin state and histone modifications[35,158–160], non-coding RNA (ncRNA)[161,162], and DNA methylation[35,58,99,161].

The vast majority of mammalian genes are expressed in a biallelic manner. However, a small subset of genes are allelically biased in their expression and depend on the allele's parent-of-origin. This phenomenon called genomic imprinting is regulated by DNA methylation at Imprinting Control Regions (ICRs). Unlike much of the genome, ICRs are differentially methylated on the parental alleles; these methylation marks are established during gametogenesis and persist through fertilization, development, and the lifespan of the organism[63,66]. Because methylation marks are established in gametes, ICRs are often referred to as germline differentially methylated regions (gDMRs)[163].

Imprinted genes make up a very small percentage of the genome - there are approximately 150 imprinted genes (both protein coding and non-coding) identified thus far in mice or humans, whereas there have been over 20,000 protein-coding genes defined for each species and even more non-coding genes identified in humans[64–67,164]. Imprinted genes are often found in close proximity to other imprinted genes and these clusters are often regulated by a single ICR[68]. Approximately 20 ICRs have been identified in humans and mice; most ICRs are found in both species, but several are human- or mouse-specific (Table 4.1)[58].

Table 4.1. Known ICRs in mice and humans and the parental allele that is methylated.

Mouse ICR	**Human ICR**	**Methylated parental allele**
Peg13	*PEG13*	Maternal
Kcnq1ot1	*KvDMR*	Maternal
Igf2r	*IGF2R*	Maternal
Gnasxl	*GNASXL*	Maternal
Nnat	*NNAT*	Maternal
Gnas	*Ex1AGNASEx1A*	Maternal
Mcts2	*MCTS2*	Maternal
Zac1/Plagl1	*PLAGL1*	Maternal
U2af1-rs1	—	Maternal

Table 4.1. Continued

Inpp5f_v2	*INPP5F_V2*	Maternal
Impact	—	Maternal
Grb10	*GRB10*	Maternal
Nap1l5	*NAP1L5*	Maternal
Cdh15	—	Maternal
Mest	*MEST*	Maternal
Zim2/Peg3	*ZIM2/PEG3*	Maternal
Snurf/Snrpn	*SNURF/SNRPN*	Maternal
Peg10	*PEG10*	Maternal
Gpr1/Zdbf2	*GPR1/ZDBF2*	Maternal
—	*RB1*	Maternal
—	*DIRAS3*	Maternal
—	*L3MBTL1*	Maternal
H19/Igf2	*H19/IGF2*	Paternal
Dlk1/Gtl2	*DLK1/MEG3*	Paternal
Rasgrf1	—	Paternal

Adapted from Cowley et al 2018[58].

ICRs are one of the few areas of the genome that resist global demethylation after fertilization in mammals[61]. Due to these unique DNA methylation dynamics, we hypothesize that ICRs may provide lasting memory of early-life exposures and influence susceptibility to disease states later in life. Consistent with this hypothesis, many physiological effects due to dysregulation of imprinted gene expression are seen due to adverse maternal exposures[58].

Cadmium (Cd) is a toxic heavy metal defined by the World Health Organization as one of the top ten chemicals of major public health concern. Cd is ubiquitously found throughout the environment and levels have increased over the last century due to anthropogenic activities[1,2]. Cd does not efficiently cross the placenta, yet maternal Cd exposure is known to negatively impact fetal health[15,20–22]. It is suspected that altered DNA methylation is a mechanism through which Cd exerts its toxic effects[165]. We have

shown that higher blood levels of Cd during the first trimester in humans is associated with altered DNA methylation in fetal cord blood at birth, and that ICRs are enriched in this dataset[58]. However, human epidemiological studies can only provide narrow insights into underlying molecular mechanisms and can be affected by a number of confounding factors.

To test the hypothesis that DNA methylation at ICRs may be more susceptible than other similar but non-imprinted loci to perturbations due to adverse maternal exposures, we have established a hybrid mouse model of maternal Cd exposure. We have utilized two genetically divergent inbred strains of mice (C57BL/6J ['B'] and CAST/EiJ ['C']) as parental strains (F_0) (Chapter 2 Figure 2.1). Both strains have fully sequenced, publicly available reference genomes. This hybrid, maternally-exposed F_1 model will facilitate the analysis of allele-specific DNA methylation changes and gene expression changes due to maternal Cd exposure, which was not feasible in the human study.

Herein, we have performed Reduced Representation Bisulfite Sequencing (RRBS) to quantify DNA methylation changes and developed a novel pipeline for allele-specific gene expression (ASE) analysis in whole brains of newborn BxC (B mother, C father) female mice maternally exposed to 0 ppm or 50 ppm Cd. RRBS is a method to identify areas of differential methylation in CpG-rich areas of the genome, e.g., CpG islands (CGIs), promoters, and ICRs. The RRBS protocol digests gDNA with a restriction enzyme (MspI, cut site CCGG) and size selects for fragments of appropriate sequencing length prior to sequencing; through this method, CpG-rich areas of the genome are enriched prior to sequencing, resulting in a reduction in sequencing costs and allowing increased sequencing depth[166]. Due to the number of ICRs, a sequencing-based approach is more

cost-effective than a targeted approach. Though not all ICRs are covered with RRBS, ICRs and CGIs are both CpG-rich, and therefore through RRBS we can test the hypothesis that DNA methylation changes at ICRs will be enriched among similar but non-imprinted loci (i.e. CGIs). In general, DNA methylation patterns are established at ICRs during gametogenesis and persist throughout cell division and differentiation during development[163]. Therefore, different cell types throughout the body should contain the same methylation patterns at ICRs. Here, we have chosen brain tissue based on the strong neurological phenotype resulting from maternal Cd exposure described in Chapter 3, as well as the raw RNA-seq data generated in Chapter 3.

In addition to exploring allele-specific DNA methylation and gene expression changes in newborn brains due to maternal Cd exposure, we also identified differentially-methylated regions (DMRs) independent of parental genotype.

Materials and Methods

DNA extraction, RRBS library prep, RRBS library QC

Methods for animal exposures, tissue collection, and DNA extraction from newborn BxC female whole brains are described in Chapter 3 Materials and Methods. The protocol for RRBS library prep was adapted from an online protocol provided for the Chromatin Structure and Function course at the Wellcome Genome Campus Oct 29-Nov 8, 2017 that was developed from a previously published method[167]. 500 ng of DNA was digested overnight at 37°C with MspI (NEB R0106S), then the enzyme was inactivated by incubating at 80°C for 20 minutes. End-repairing and A-tailing of MspI-digested DNA was performed with Klenow fragment (3’ to 5’ exo-, NEB, M0212S) and A-tailing dNTPs (10

mM dATP, 1 mM dGTP, 1 mM dCTP; NEB, N0446S) by incubating at 30°C for 20 minutes (to facilitate end repair) and then 37°C for 20 minutes (to facilitate dA-tailing). DNA was cleaned up with 2X AMPure XP beads (Beckman Coulter, A63881) and resuspended in Qiagen EB buffer (10 mM Tris-HCl, pH 8.5). Methylated adaptors (NEBNext Multiplex Oligos for Illumina – Methylated Adaptor, Index primers set 1, NEB E7535S) were ligated to the A-tailed MspI-digested BxC female brain gDNA and 5 pg of HindIII- (NEB R0104S) and SmlI- (NEB R0597S) digested non-methylated pUC19 DNA (Zymo, D5017) in a single reaction. Non-methylated pUC19 DNA was spiked in as a control to test for bisulfite conversion efficiency. Adaptors were ligated to DNA overnight at 16°C. The USER enzyme from the NEB adaptor kit was then incubated with the DNA at 37°C for 15 minutes to cleave the hairpin structure in the adaptors. Reactions were then cleaned up with two AMPure bead cleanup steps to select DNA fragments of desired length: 0.6X beads followed by 1.8X beads. DNA was eluted in buffer EB, then bisulfite-converted according to the manufacturer's protocol (Zymo EZ-96 DNA Methylation MagPrep kit, D5040). Bisulfite-treated DNA libraries were then PCR amplified with indexing primers. Each library was PCR amplified in 8 smaller reactions in order to improve efficiency of the amplification and complexity of the library, then the 8 reactions were pooled into a single tube. PCR-amplified libraries were cleaned up with 1.8X AMPure beads and eluted in EB buffer. Bisulfite conversion was verified through PCR by using primers designed to amplify non-bisulfite-treated and bisulfite-treated pUC19 DNA (Table 4.2).

Table 4.2. Primer sequences used to test bisulfite conversion of pUC19 DNA

	Sequence (5'->3')
pUC19 BSnonconvert_F	GGTCATAGCTGTTTCCTGTGTGAAATTGTTATC
pUC19 BSnonconvert_R	CTGGCCTTTTGCTCACATGTTCTTTCCTGC
pUC19 BSconvert_F	GGTTATAGTTGTTTTTTGTGTGAAATTGTTATT
pUC19 BSconvert_R	CTAACCTTTTACTCACATATTCTTTCCTAC

Prepared libraries were submitted to the Genomic Sciences Laboratory (GSL) at North Carolina State University (NCSU) for fragment size analysis (Tapestation D1000, Agilent) and sequencing. An initial low coverage (1 million reads) sequencing run was performed on a MiSeq (Illumina, v2 250 PE nano) to QC libraries; data analysis for this sequencing run was performed by Kathleen Hudson in the Cyverse Discovery Environment. Fastq files underwent FastQC[168] analysis, had adaptors trimmed using TrimGalore-0.4.5[169], then underwent a final FastQC analysis to ensure adaptor content was trimmed. BSMAP[170] was used to align reads and samtools-1.8[171] was used to sort and index sequences. IGV software (Broad Institute) was used to visualize reads, check bisulfite conversion efficiency, and ensure coverage of desired loci. Prepared libraries then underwent higher-depth sequencing using NextSeq 150PE 300 cycles (Illumina). Sequencing data obtained from the NextSeq run was further analyzed by Heba Saadeh as described below.

Processing RRBS reads

Trim Galore v0.5.0 (parameters: --rrbs --paired) was used to trim the raw RRBS reads and to remove poor quality calls and adapters. Since the RRBS reads are allele-specific, we prepared an N-masked C57BL/6J mouse genome (GRCm38 assembly) in which all of the CAST/EiJ SNP positions compared to the C57BL/6J were replaced with N, so they do not bias the alignment toward either genotype. The reads then were mapped

to this N-masked genome by Bismark[172] v0.20.0 paired-end mode to count overlapping parts of the reads only once. The aligned BAM files produced by bismark were then processed by the allele-specific alignment sorter SNPsplit[173] v0.3.4 to determine the allelic origin of reads that cover known SNP positions into either C57BL/6J or CAST/EiJ. Finally, CG methylation calls were extracted using the Bismark methylation extractor (v0.20.0).

Data analysis and motif finding

The biological replicates for each condition (0 and 50 ppm maternal Cd exposure) were grouped to increase the power, and only the C's that were covered in both conditions were used for further analyses to reduce the effect of methylation differences resulting from un-even coverage. CGIs and gDMRs were defined as published[62,174–176] and lifted over using the UCSC liftover tool into the GRCm38 assembly. The RRBS data were analyzed using SeqMonk v0.34.1. The methylation levels for the gDMRs were quantified based on at least 10 informative Cs, each covered by a minimum of 10 reads. The methylation levels for all the CGIs were quantified based on at least 10 informative Cs but each was covered by a minimum of 1 read; in total 17,781 CGIs were called. Differentially methylated gDMRs/CGIs were called if they pass Chi-Square test with p-value <0.05 and with a methylation difference of at least 10% in one condition relative to the other.

Motif enrichment analysis was performed using DREME[177] using default parameters and MEME[178] within the MEME suite v5.4.0. MEME classic run (with psp and background files) parameters: -revecomp, -mod zoops, -minw 4, -maxw 20, -markov_order 1. MEME discriminative run (with a background file) parameters: similar to the classic run with the additional option of -objfun de. The positive sequences were the

86 differentially methylated CGIs that were identified between the two conditions, in addition to the CGIs associated with the promoters of the 14 DEGs identified in RNA-seq, while the 17,781 CGIs were used as a control sequences. Statistically significant enriched motifs were submitted to Tomtom[179] within the MEME suite using default parameters and searching in JASPAR database.

Allele-specific gene expression analysis

Raw RNA-seq data obtained in Chapter 3 was used for analysis. Adapters were trimmed from sequencing reads using Cutadapt[180]. For the purposes of determining differential expression across treatments, reads were aligned to the mouse reference genome (GRCm38.p4) using TopHat (v.2.1.1) under default settings[181]. In order to evaluate allele-specific expression across all samples, the same sequencing reads were aligned to the CAST/EiJ (C) pseudogenome, and then re-mapped to the reference genome coordinates using the Lapels pipeline[182]. Allele-specific expression of a given gene was assessed using the alignment quality of a given read to each parental strain genome. Reads from each strain-specific alignment file were compared and sorted into three categories based on preferential alignment to 1) the B genome, 2) the C genome, or 3) equal alignment quality to both genomes. This was accomplished using the Suspenders pipeline[183]. In each case, gene expression counts were obtained using HTseq[117]; for differential expression analysis, gene counts were produced from annotated alignments to the reference genome, while in allele-specific expression analysis, gene counts were obtained from each of the three alignment quality bins independently.

Results

Analysis of allele-specific DNA methylation changes at select ICRs

RRBS was performed on DNA extracted from newborn BxC female whole brains maternally exposed to 0 ppm and 50 ppm Cd. 13 CGIs with sufficient coverage in both treatment groups overlapped 12 known gDMRs that are maternally methylated (Figure 4.1, Table 4.3). DNA methylation levels at these gDMRs were consistent with being hypermethylated on the maternal allele (Figure 4.1, B genome), hypomethylated on the paternal allele (Figure 4.1, C genome), and closer to approximately 50% without allele-specific assignment, consistent with one allele being relatively hypermethylated and the other allele being hypomethylated (Figure 4.1, 0 ppm and 50 ppm graphs). Only Zfp777 had >10% change in DNA methylation in one of the parental genomes (C) due to 50 ppm maternal Cd exposure (Table 4.3). The Zfp777 gDMR displays unique DNA methylation dynamics during embryonic development; it is considered a transient maternal gDMR and becomes biallelically methylated early in development, and therefore is not considered an ICR in the brain[174]. Given these data, it was concluded that there were no significant changes in allele-specific DNA methylation at the ICRs with sufficient sequencing coverage in newborn BxC female brains as a result of 50 ppm maternal Cd exposure.

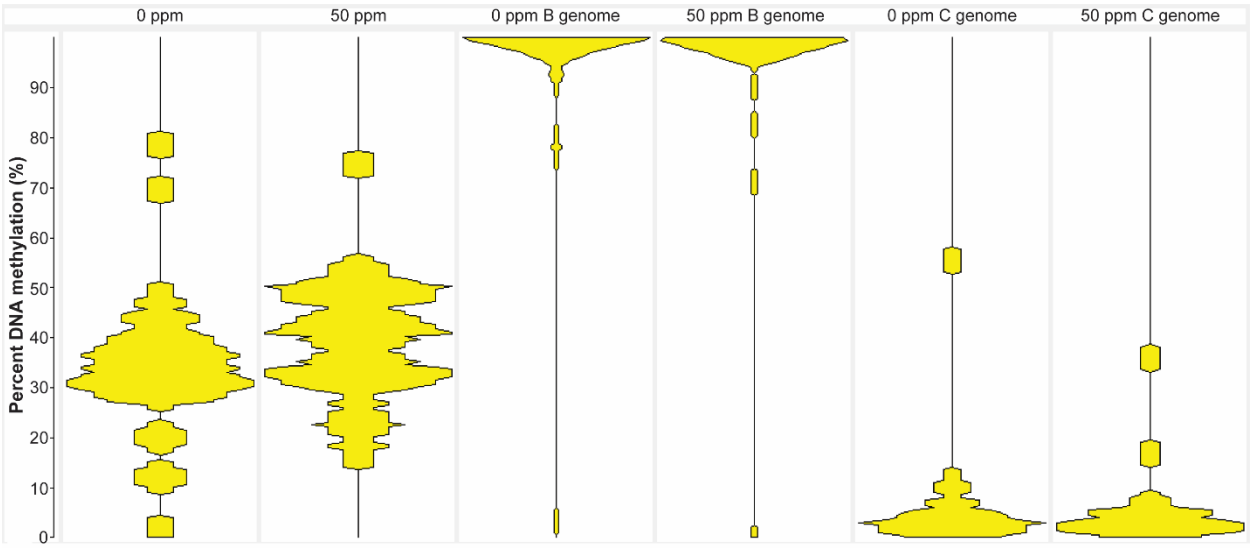


Figure 4.1. Levels of DNA methylation at gDMRs. Two graphs on left (0 ppm and 50 ppm) are not aligned to a specific parental genome. Two graphs in the center (0 ppm B genome and 50 ppm B genome) are aligned to the B (maternal) genome. Two graphs on the right (0 ppm C genome and 50 ppm C genome) are aligned to the C (paternal) genome.

Table 4.3. DNA methylation (%) of 12 gDMRs with sufficient coverage in both treatment groups

Chr	Start	End	gDMR	0 ppm (B)	50 ppm (B)	50-0 (B)	0 ppm (C)	50 ppm (C)	50-0 (C)
2	152686785	152687333	H13 (Mcts2)	97.16014	95.95641	-1.20373	3.22751	3.1653516	-0.062158
2	174298234	174301029	Gnas	98.44078	97.60402	-0.83676	0.8158272	0.5225304	-0.293297
2	174328350	174331182	Gnas	3.125	0	-3.125	8.696846	0.6057016	-8.091144
6	30737764	30738506	Mest	97.1831	99.828766	2.645666	5.14784	1.2701377	-3.877702
6	48024479	48025599	Zfp777	80.362076	71.27601	-9.086066	55.688026	36.018463	-19.66956
7	6729378	6730606	Peg3	100	97.395836	-2.604164	0.8333334	0.54347825	-0.289855
7	60004308	60005228	Snrpn	97.67136	97.474724	-0.196636	5.071725	5.573119	0.501394
7	128687456	128688635	Inpp5f	94.4714	96.75225	2.28085	2.40093	6.9652133	4.5642833
7	143295116	143296665	Kcnq1	90.62639	90.14839	-0.478	3.3244011	3.8876512	0.5632501
8	122864717	122865427	Cdh15	76.28029	82.789444	6.509154	11.361348	16.908487	5.547139
11	12025419	12027039	Grb10	NaN	97.47807	N/A	0.8680555	1.4948704	0.6268149
11	22971887	22973204	Commd1 (U2af1-rs1)	93.56505	97.21009	3.64504	1.3056138	3.4464555	2.1408417
15	72809377	72810812	Trappc9 (Peg13)	98.164116	97.976006	-0.18811	2.7286627	3.4394093	0.7107466

50-0 (B): % methylation change in maternal (B) genome due to maternal Cd exposure
50-0 (C): % methylation change in paternal (C) genome due to maternal Cd exposure

DNA methylation changes due to maternal Cd exposure independent of parental genotypes

Because no significant changes were detected in allele-specific DNA methylation at ICRs, an additional analysis was performed to identify DMRs due to maternal Cd exposure, independent of parental genotype or allele-specificity. 86 CGIs were found to be significantly differentially methylated (chi-square $p<0.05$) in newborn BxC female brains when comparing 50 ppm exposed samples to controls, with 71 CGIs showing at least a 10% change in DNA methylation (Table 4.4). 36 CGIs increased their DNA methylation levels by at least 10%, while 35 CGIs decreased their DNA methylation levels by at least 10% in the 50 ppm samples compared to controls.

Table 4.4. 86 CGIs significantly differentially methylated in newborn BxC female brains due to 50 ppm maternal Cd exposure. CGIs are ranked by decreasing absolute % methylation change (abs(50-0)). Orange highlighted cells: >10% increase in methylation, blue highlighted cells: >10% decrease in methylation. Null values indicate loci is not associated with a gene.

Chr	Start	End	p-value	Feature	0ppm (% methylation)	50 ppm (% methylation)	50-0	abs(50-0)
19	5087728	5089033	7.33E-08	Yif1a	0	100	**100**	**100**
10	88336752	88337034	1.73E-05	Dram1	100	0	**-100**	**100**
15	102510449	102511039	1.94E-06	Map3k12	67.64706	2.0100443	**-65.63702**	**65.637016**
7	99858508	99859329	5.03E-07	Spcs2	95.652176	30.434786	**-65.21739**	**65.21739**
1	93401244	93401974	0.0010588	Ano7	0	64.28571	**64.28571**	**64.28571**
8	61928109	61928265	1.30E-04	Ddx60	0	62.2807	**62.2807**	**62.2807**
9	21337869	21338095	0.0041016	Slc44a2	7.142857	64.28571	**57.142853**	**57.142853**
19	44247577	44248498	2.02E-09	null	0.7407407	57.037037	**56.296296**	**56.296296**
18	77198107	77198549	0.0155477	St8sia5	90.90909	36.363636	**-54.54545**	**54.545454**
11	85834594	85834829	0.0023672	Tbx2	66.99197	16.294949	**-50.69702**	**50.697021**
2	77589674	77589878	1.05E-07	Zfp385b	68.92857	18.333334	**-50.59524**	**50.595236**
14	41013470	41014000	2.50E-11	5730469M10Rik	48.88889	0.4938272	**-48.39506**	**48.395063**
10	42635785	42636263	0.010488	Ostm1	45.833332	0	**-45.83333**	**45.833332**
12	72082459	72082711	5.02E-07	2810055F11Rik	51.8	6	**-45.8**	**45.8**
11	50887313	50887692	0.0120905	Zfp454	42.857143	0	**-42.85714**	**42.857143**
7	16296874	16297202	0.0214291	null	41.666668	0	**-41.66667**	**41.666668**
15	25779874	25780073	0.0017067	Myo10	89.28571	51.530613	**-37.7551**	**37.755097**

Table 4.4 Continued

1	188714184	188714350	3.85E-04	Ush2a	53.57143	91.08844	**37.51701**	**37.51701**
4	115972474	115972717	4.72E-07	Faah	85.11905	48.61111	**-36.50794**	**36.50794**
16	17734019	17734441	1.44E-05	null	54.924244	18.939394	**-35.98485**	**35.98485**
4	141382157	141382884	0.0291697	Fam131c	0	34.583336	**34.583336**	**34.583336**
5	137293126	137293527	2.20E-05	Ache	34.126984	1.904762	**-32.22222**	**32.222222**
16	90520353	90520679	0.0301996	null	40.909092	72.72727	**31.818178**	**31.818178**
13	57282223	57282403	0.003723	null	31.944445	63.594276	**31.649831**	**31.649831**
X	52744247	52744660	0.0156665	Kis2	0	30.303032	**30.303032**	**30.303032**
7	24175067	24175301	5.25E-04	Zfp114	28.88889	0.89031345	**-27.99858**	**27.998577**
13	3457408	3457618	1.63E-08	null	19.722223	47.675648	**27.953425**	**27.953425**
5	122644047	122644288	0.00257	P2rx7	92.10526	64.47369	**-27.63157**	**27.63157**
11	78501581	78501989	0.0490034	Vtn	100	73.333336	**-26.66666**	**26.666664**
7	16993991	16995116	0.0016704	Ccdc8	24.242424	49.69697	**25.454546**	**25.454546**
5	147294093	147294578	0.0221098	null	33.333332	8.333334	**-25**	**24.999998**
14	55062756	55063057	0.0301996	Gm20687	37.5	13	**-24.5**	**24.5**
7	142654285	142654450	9.89E-05	Igf2	43.969307	19.536823	**-24.43248**	**24.432484**
1	176997930	176998315	0.020481	Sdccag8	5.769231	28.834707	**23.065476**	**23.065476**
11	63888932	63889247	1.94E-06	Hs3st3b1	43.67236	66.216896	**22.544536**	**22.544536**
X	169986621	169987870	0	Mid1	13.383652	35.68239	**22.298738**	**22.298738**
1	178917298	178918101	0.0214291	Kif26b	16.731602	38.360214	**21.628612**	**21.628612**
12	104606085	104606370	0.0049432	null	78.57143	100	**21.42857**	**21.42857**
X	8096583	8096858	0.022818	Gm6787	22.987421	43.4696	**20.482179**	**20.482179**
9	96536881	96537228	0.002111	null	53.982685	34.013943	**-19.96874**	**19.968742**
10	128821454	128822167	0.004987	Ormdl2	0	19.52381	**19.52381**	**19.52381**
X	57503364	57503944	0.0019776	Gpr101	35.949474	16.585129	**-19.36435**	**19.364345**
1	87620512	87620687	0.0226034	Inpp5d	78.78788	97.9798	**19.19192**	**19.19192**
12	111360861	111361184	0.003247	Cdc42bpb	95.454544	76.36364	**-19.0909**	**19.090904**
10	34127324	34127994	0.0360146	Fam26f	28.571428	10.317461	**-18.25397**	**18.253967**
8	105293196	105293780	0.0110606	Exoc3l	55.701756	73.81579	**18.114034**	**18.114034**
17	83231654	83231884	0.0140349	null	25.079365	42.51701	**17.437645**	**17.437645**
7	80587364	80587668	0.0016704	Crtc3	50.79772	33.938156	**-16.85956**	**16.859564**
6	48024479	48025599	0	Zfp777	69.47702	52.657326	**-16.81969**	**16.819694**
8	87742129	87742367	0.0282194	Zfp423	31.111113	46.289474	**15.178361**	**15.178361**
X	56455229	56456174	5.25E-11	Ddx26b	33.632107	18.624384	**-15.00772**	**15.007723**
4	130173911	130174162	0.0069574	Tinagl1	34.505497	49.20292	**14.697423**	**14.697423**
4	139338191	139338454	0.0029549	AL807811.1	31.39881	45.942783	**14.543973**	**14.543973**
7	18962796	18963674	0.0214291	Nova2	0	14.199135	**14.199135**	**14.199135**
3	37063148	37063905	1.51E-06	Adad1	54.379185	68.2356	**13.856415**	**13.856415**
1	14918589	14918922	0.0182343	Trpa1	1.0204083	14.39394	**13.373532**	**13.373532**
2	152703803	152704198	0.0030471	H13	87.5	74.30736	**-13.19264**	**13.19264**
6	135526636	135527005	0.0041016	null	27.407408	40.583775	**13.176367**	**13.176367**
5	149749200	149749440	0.0024514	B3galtl	86.90476	73.809525	**-13.09524**	**13.095235**
8	48416852	48417088	0.0110194	Odz3	94.871796	82.1978	**-12.674**	**12.673996**
X	147553872	147554029	0.0440946	Lrch2	12.937063	25.60332	**12.666257**	**12.666257**

Table 4.4 Continued

12	24493221	24493796	3.38E-08	Gm16372	51.221348	63.53657	**12.315222**	**12.315222**
13	34916604	34917111	8.89E-05	1700026J04Rik	18.849207	30.952383	**12.103176**	**12.103176**
10	112271620	112272356	1.48E-14	Kcnc2	12.638889	0.6111111	**-12.02778**	**12.027778**
3	130754746	130755047	0.0320832	null	98.611115	86.9943	**-11.61682**	**11.616815**
4	8865760	8866532	7.15E-04	Chd7	70.56103	59.330315	**-11.23072**	**11.230715**
X	7422651	7422880	0.0312335	2010204K13Rik	50	60.876625	**10.876625**	**10.876625**
2	22588552	22589766	0.02614	Myo3a	78.64759	89.18367	**10.53608**	**10.53608**
4	127195004	127195209	0.0010018	Dlgap3	85.20584	74.78742	**-10.41842**	**10.41842**
2	77411928	77412106	0.0017067	Zfp385b	70.8658	81.236336	**10.370536**	**10.370536**
2	85036806	85037785	0.0024198	P2rx3	1.0526316	11.35965	**10.307018**	**10.307018**
1	37687115	37687538	4.57E-05	2010300C02Rik	59.392437	69.0099	**9.617463**	**9.617463**
9	122310803	122310979	0.0058215	null	10.565476	1.1904763	**-9.375**	**9.3749997**
2	35720302	35720529	0.0010856	Dab2ip	46.073914	54.67086	**8.596946**	**8.596946**
16	14292335	14292608	0.0051747	Myh11	8.333334	0	**-8.333334**	**8.333334**
4	32864134	32864583	2.47E-06	Ankrd6	5.8068786	13.646086	**7.8392074**	**7.8392074**
3	68572274	68572664	0.0013821	Schip1	8.530021	0.69676703	**-7.833254**	**7.833254**
3	5860625	5860830	0.0068912	null	27.916668	20.987486	**-6.929182**	**6.929182**
14	63947910	63948160	0.0092725	Sox7	70.34105	64.27888	**-6.06217**	**6.06217**
15	98003149	98003334	1.69E-04	Col2a1	6.4200683	11.529582	**5.1095137**	**5.1095137**
15	78799104	78799377	0.0029369	Card10	51.40801	56.45299	**5.04498**	**5.04498**
13	58668741	58668933	1.48E-14	null	50.3804	54.269447	**3.889047**	**3.889047**
1	38358854	38359432	2.64E-07	Aff3	56.335686	52.718605	**-3.617081**	**3.617081**
19	6420613	6420807	2.21E-04	Nrxn2	35.651268	38.84371	**3.192442**	**3.192442**
19	16871269	16871504	0.0047349	Foxb2	8.201058	9.291187	**1.090129**	**1.090129**
10	117023784	117024447	5.29E-05	Best3	28.688042	29.024107	**0.336065**	**0.336065**

To determine if there were enriched motifs contained within these 86 CGIs that may explain a shared pathway or binding site of DNA methylation-associated proteins, an enrichment analysis using the MEME suite was performed. There were no significantly enriched motifs (E-value<0.05) identified. Enrichment analysis was performed on the gene names associated with these CGIs using Enrichr[84,85]; however, no relevant phenotype or disease-related terms were significantly enriched (data not shown).

Allele-specific expression (ASE) analysis

Although the allele-specific expression (ASE) of imprinted genes is regulated by DNA methylation at ICRs, other mechanisms can control their expression. To identify any

potential shifts in allele-specific or overall expression of imprinted genes that were independent of DNA methylation changes, we utilized the RNA-seq data that was generated from newborn BxC female brains maternally exposed to Cd (Chapter 3). A separate RNA-seq analysis was performed on the raw sequencing reads to facilitate ASE analysis. 14 genes were found to be significantly differentially expressed after accounting for multiple testing (adj p val <0.05) and setting an FDR threshold of 0.05 (Table 4.5). Out of these 14 differentially expressed genes (DEGs), 11 were identified to be significantly differentially expressed in the Chapter 3 RNA-seq analysis, and an additional gene (*Hoxb7*) was found to be significantly differentially expressed through qRT-PCR (Chapter 3).

Table 4.5. 14 DEGs identified through RNA-seq analysis for ASE

Gene	log2FoldChange	adj p val	FDR	Identified as DEG in Ch 3
Hoxb8	7.229149431	1.22E-11	0.05	x
Hoxb7	5.171587427	2.10E-05	0.05	x
Mag	1.255775442	0.0009292	0.05	x
Il33	0.939363331	0.0014341	0.05	x
Bcas1	0.926384253	0.0014341	0.05	x
Aldh1a2	0.465329229	0.0150621	0.05	x
Capn11	3.50586075	0.0161906	0.05	
Enpp6	1.608276447	0.016564	0.05	x
Kcnj16	0.703587392	0.016564	0.05	x
Sp9	-0.409669337	0.016564	0.05	x
Sparcl1	0.421085897	0.0176273	0.05	x
Mobp	4.044722743	0.0353295	0.05	
Kank1	0.440716079	0.0451981	0.05	x
Mbp	0.498282668	0.0477232	0.05	x

Normalized read counts of these 14 DEGs were then assigned to either the B or C genome based on polymorphisms known to be present in either the B or C genome. If

there were no identifying polymorphisms in the sequencing reads to correctly identify the parental genome of origin, the reads were considered ‘unassigned’. None of the 14 DEGs identified in this RNA-seq analysis were found to have a significant change in allelic expression due to maternal Cd exposure (Figure 4.2).

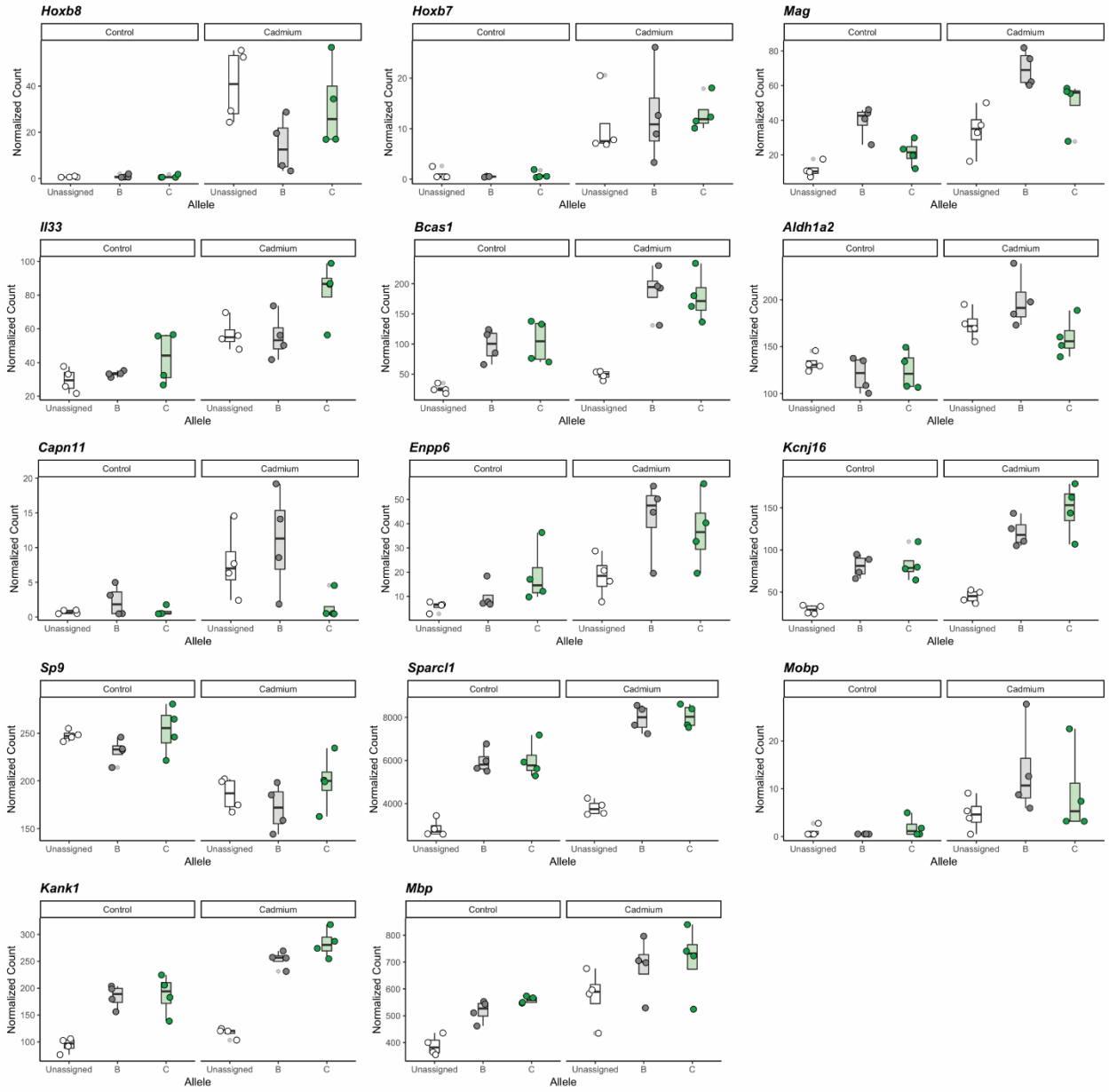


Figure 4.2. ASE of 14 DEGs identified in RNA-seq.

No imprinted genes were identified as significantly differentially expressed in either the Chapter 3 or ASE RNA-seq analysis. The 5 imprinted genes with the lowest p-values (Table 4.6) were analyzed for ASE. None of these genes were found to have a significant change in ASE due to maternal Cd exposure (Figure 4.3). *Zim3* and *Klrb1f* had low overall normalized read counts (<10), and therefore were not considered reliable data for ASE analysis. *Grb10* had significantly more unassigned reads than reads that contained identifying SNPs, thereby hindering ASE analysis. *Peg3* (Paternally Expressed Gene 3) is a paternally-expressed gene[184], consistent with the higher number of reads aligned to the C genome. In humans, *Casd1* is biallelically expressed in the brain and other tissues and therefore is not considered an imprinted gene in humans[185], consistent with what was observed here.

Table 4.6. 5 imprinted genes with lowest p-values in RNA-seq data

Gene	log2FoldChange	p-value	adj p val	FDR
Zim3	-1.467784	0.1875643	0.9999019	NS
Grb10	-0.065016	0.3630265	0.9999019	NS
Klrb1f	-0.618552	0.4340119	0.9999019	NS
Peg3	-0.066754	0.590436	0.9999019	NS
Casd1	-0.026683	0.7652265	0.9999019	NS

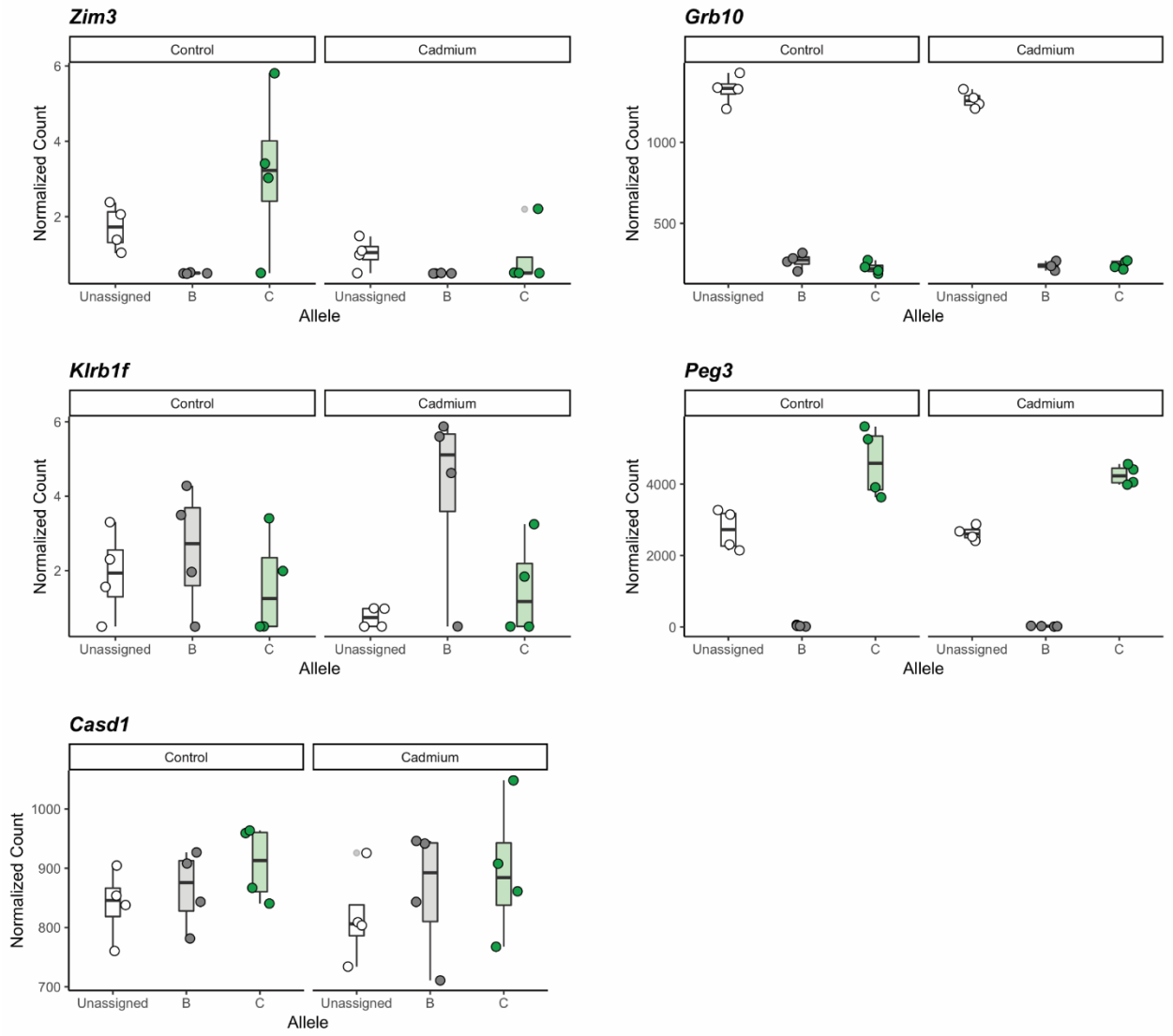


Figure 4.3. ASE of 5 imprinted genes with lowest p-values in RNA-seq data

Discussion

In this model, we aimed to test the hypothesis that ICRs are more susceptible than other similar but non-imprinted loci to DNA methylation changes due to environmental exposures during early-life development. We utilized a hybrid mouse model of maternal Cd exposure to facilitate the analysis of allele-specific DNA methylation and gene expression changes in newborn brains due to maternal Cd exposure. We did not identify any ICRs that were significantly differentially methylated due to maternal Cd exposure.

Additionally, we did not identify any imprinted genes or non-imprinted genes that had a significant shift in allele-specific expression. However, we did identify DMRs resulting from maternal Cd exposure that may contribute to the neurodevelopmental impacts of maternal Cd exposure.

Of the gDMRs we evaluated, only *Zfp777* was found to have a significant change in DNA methylation in one of the parental genomes (increased methylation in C genome) as a result of 50 ppm maternal Cd exposure. However, as previously discussed, *Zfp777* becomes biallelically methylated after implantation and allele-specific differences do not persist past birth. It is therefore not considered to be a persisting gDMR (i.e. ICR) in the brain at birth[174]. *Zfp777* is a zinc finger protein that is not well-studied, but it would be interesting to further investigate downstream effects of this DMR in the brain observed in this study. It was not identified here as a significantly differentially expressed gene (adj p val = 0.9999, log2fold change = -0.05).

Of the four ICRs identified as Cd-associated DMRs in newborn human cord blood (*KvDMR, IGF2R, GNASXL, SNURF/SNRPN*)[58], two had sufficient coverage in the RRBS data in this study (*Kcnq1* [human *KvDMR* equivalent in mouse]*, Snurf/Snrpn*), but were not significantly differentially methylated in newborn mouse brains due to maternal Cd exposure. Although the human cohort separated fetal cord blood samples into two groups based on high or low maternal blood Cd levels, it is possible that other environmental exposures could have contributed to these ICRs being identified as DMRs. The mothers lived in a region known to have clustered geographical hotspots of other heavy metals such as mercury, lead, and arsenic, though attempts were made to exclude individuals who lived in these areas[96]. Human epidemiological studies cannot completely control for

phenotypic variance to the environment and this may represent a confounding factor. It is also possible that maternal Cd exposure can have different effects in mice and humans. Cd accumulates in the placenta and poorly enters fetal circulation in mammals, and is suspected to exert many of its toxic effects at the placental level[3,15]. While mice and humans are both placental mammals, murine placentas lack many features that developed in primate placentas during evolution[186].

Despite no significant difference in DNA methylation at ICRs, we were still able to identify 86 CGIs as significantly differentially methylated in newborn mouse brains as a result of maternal Cd exposure. Only 2 of these (*Chd7*, *Inpp5d*) were found in the 444 genes overlapping DMRs in newborn cord blood[58]. *Chd7* was found to be hypomethylated in newborn cord blood, consistent with the hypomethylation observed here due to maternal Cd exposure; however, *Inpp5d* was hypomethylated in newborn cord blood yet hypermethylated in newborn mouse brains as a result of maternal Cd exposure. This discrepancy and the lack of conserved DMRs is likely due to the difference in tissue type, as individual cell types have different DNA methylation and epigenetic profiles[157]. However, there was no overlap of the 86 DMRs with regulatory areas of the 14 DEGs identified through RNA-seq, indicating their changes in expression may be regulated by a mechanism other than DNA methylation. There was also no significant enrichment for DNA sequence motifs in the 86 DMRs, indicating that there is not a shared recognition site that attracts DNA methyltransferases (Dnmts) or Dnmt-associated DNA binding proteins to induce DNA methylation changes in these areas. Finally, there were no relevant phenotype or disease-related categorical terms enriched among the genes

associated with these DMRs, suggesting that DNA methylation changes may not be a major driver of Cd-associated changes in gene expression.

The RNA-seq analysis performed in this chapter is concordant with the analysis performed in Chapter 3 that used slightly different methods. However, due to the low number of DEGs (14), our ASE analysis was limited. As discussed in Chapter 3, the cellular heterogeneity of whole brains is likely contributing to the low number of DEGs identified through RNA-seq. Our method of ASE analysis can only identify allele-specific changes in genes that are significantly differentially expressed. Despite this, we still examined the ratio of allele-specific expression in 5 imprinted genes that had the lowest p values, but were not determined to be significant DEGs. *Zim3* and *Klrb1f* appear to have a shift in allelic expression due to maternal Cd exposure; however, they both had extremely low normalized read counts (<10) and therefore cannot be considered reliable measurements.

Although imprinted gene expression has been shown to be influenced by early-life exposures, it is possible that their expression is influenced by mechanisms other than DNA methylation at ICRs. A study in mice that restricted maternal protein during gestation found changes in imprinted gene expression in the offspring's livers, but this was not accompanied by changes in ICR methylation, indicating a different mechanism governed changes in imprinted gene expression[163]. It is also possible that we were unable to detect ICR DNA methylation changes due to our experimental design. While RRBS performed here encompassed 12 gDMRs, these gDMRs only comprise approximately half of the ICRs known in mice and it is possible that ICRs that were not included in the analysis performed here may have had differential changes to their DNA methylation. Future work

will need to be done to examine the DNA methylation at excluded ICRs in newborn brains as a result of maternal Cd exposure, as well as perform similar studies in other tissues to see if this is a conserved phenomenon.

Acknowledgements

We would like to thank the following individuals for their contribution to this chapter: Samuel Widmayer, David Aylor, and Heba Saadeh.

CHAPTER 5

Conclusions and Future Directions

In this work, we have tested the hypothesis that DNA methylation at Imprinting Control Regions (ICRs) may be more susceptible to maternal exposures than other similar but non-imprinted loci. We used a hybrid mouse model of maternal cadmium (Cd) exposure due to the increasing evidence in human and animal studies that maternal Cd exposure negatively impacts offspring health. We have tested two doses of maternal Cd exposure: a low, environmentally-relevant dose, and a high dose similar to what is experienced by people living in Cd-polluted areas. By using the mouse as a model, we were able to identify perturbed molecular mechanisms in tissues of interest that would not be feasible in human studies.

In **Chapter 2**, we studied the effect of maternal Cd exposure on fetal cardiovascular health and development. We showed that maternal Cd exposure leads to cardiac hypertrophy at birth and is sufficient to program hypertension at adulthood in a subset of animals who appeared to be more sensitive to maternal Cd exposure. We quantified significant perturbations to levels of all essential trace elements in newborn blood and liver tissues. This work highlights the need for a pre-conception window of exposure, as many essential trace elements that were altered in the offspring were also altered in the mothers due to maternal Cd exposure. This suggests that some of the mechanisms through which maternal Cd exposure negatively affects the offspring may be a consequence of Cd-induced alterations to maternal metabolism or trace element homeostasis. Finally, we identified transcriptome-wide changes in newborn hearts as a result of maternal Cd exposure to provide insights into molecular mechanisms behind the

observed cardiovascular phenotypes. We hypothesize that maternal Cd exposure creates a hypoxic and stressful environment during fetal development, and this adverse environment leads to cardiac hypertrophy at birth and susceptibility to hypertension during adulthood.

In **Chapter 3**, we studied the effect of maternal Cd exposure on fetal neurodevelopment and behavior. We showed that maternal Cd exposure leads to a significant increase in proportional brain mass at birth and altered behaviors at adulthood that are suggestive of reduced anxiety or inhibition. We identified multiple pathways perturbed in newborn brains at the transcript, protein, metabolite, and organelle levels as a result of maternal Cd exposure that likely have profound effects on neurodevelopment and behavior. We hypothesize that maternal Cd exposure disrupts retinoic acid (RA)-metabolizing pathways, leading to an increase in RA in the brain at birth. This increase in RA leads to perturbed expression of *Hox* genes and myelin markers at birth, and leads to altered behavior at adulthood. In addition, we show reduced mitochondrial content in the brains at birth due to maternal Cd exposure. Consistent with our findings in **Chapter 2**, we hypothesize that altered trace element levels in the blood are leading to a hypoxic environment in the brain and that cellular energy pathways may be negatively impacted due to nutrient- and oxygen-poor conditions during neurodevelopment.

Finally, in **Chapter 4**, we quantify allele-specific DNA methylation and gene expression changes to test our hypothesis that DNA methylation at ICRs may be more susceptible to adverse early-life exposures than similar, non-imprinted loci. Although we did not identify any significant changes in allele-specific DNA methylation or gene expression of ICRs or imprinted genes, respectively, there are several limitations to our

experimental design that may have hindered our ability to detect such changes. Due to RRBS coverage of the genome, we only quantified DNA methylation changes at approximately half of the known ICRs in mice; it is possible that other ICRs had significant changes in allele-specific DNA methylation that we were unable to identify. Future work would be needed to quantify allele-specific DNA methylation changes at ICRs excluded from the analysis in this study. In addition, future work may be done in other tissues or in the context of other adverse maternal exposures to see if this is a conserved observation. Our analysis of allele-specific gene expression (ASE) changes was limited by the low number of differentially-expressed genes (DEGs) in the brain due to maternal Cd exposure. Performing RNA-seq in whole brains may not be ideal due to the high level of cellular heterogeneity, and this may limit the number of DEGs identified as a result of maternal Cd exposure. To address this, future ASE analysis in maternally-exposed offspring may be done with a specific region or cell type of the brain, or in a more homogeneous tissue. Imprinted genes are expressed in a very cell type-specific manner in the brain, which may have hindered our analysis[187,188].

Future work will need to be done to further elucidate how maternal Cd exposure leads to the perturbed molecular pathways identified in this work. As we have identified altered trace element levels in newborn blood and liver due to maternal Cd exposure, the logical next step may be to test which elements are contributing to the observed cardiovascular and neurodevelopmental phenotypes. We suspect that the iron (Fe) deficiency induced by maternal Cd exposure is a significant contributing factor to the observed phenotypes due to its role in oxygen homeostasis and cellular energy metabolism. A future experiment may be to induce an Fe deficiency during development

in the absence of maternal Cd exposure, and identify shared consequences to fetal neuro- and cardiovascular development. In addition, we could supplement maternal Cd exposure with additional Fe to determine if Fe can rescue or at least partially alleviate some of the impacts of maternal Cd exposure.

Additional work is also needed to identify the mechanisms linking maternal Cd exposure with increased RA in the brain. We suspect that RA-metabolizing pathways are disrupted due to maternal Cd exposure; future work may identify specific areas in the brain where this is happening, as RA synthesis tends to be localized to the limbic system[144]. Once identified, these areas should have *Hox* and myelin markers quantified to link increased RA with increased *Hox* and myelin marker gene expression. Future experiments are also needed to identify other areas of the body where RA levels may be altered due to maternal Cd exposure. The DEGs identified in newborn hearts as a result of maternal Cd exposure were significantly enriched for RA-related terms, indicating that RA levels may be perturbed in other areas of the body or systemically due to maternal Cd exposure.

In summary, this work has identified multiple molecular pathways perturbed as a result of maternal Cd exposure that may link early-life Cd exposure to impaired neurodevelopment and cardiovascular health. We have shown that even a low, environmentally-relevant dose of maternal Cd exposure can impact fetal health, supporting the emerging concern that chronic low levels of Cd exposure are impacting human health. There was not always a dose-dependent response to maternal Cd exposure, suggesting that low and high levels of maternal Cd exposure may impact offspring health in different ways. Future studies are needed to dissect the precise

molecular mechanisms in which maternal Cd exposure at various doses impacts fetal health and development. Although the work performed here does not support our central hypothesis, future work is still needed to further test that DNA methylation at ICRs may be more susceptible than other similar, non-imprinted loci to early-life exposures.

REFERENCES

(1) W.H.O. Ten chemicals of major public health concern https://www.who.int/ipcs/assessment/public_health/chemicals_phc/en/ (accessed Nov 26, 2018).

(2) Filipic, M. Mechanisms of Cadmium Induced Genomic Instability. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **2012**, *733* (1–2), 69–77.

(3) Andresen, E. & Kupper, H. Cadmium: From Toxicity to Essentiality. In *Cadmium: From Toxicity to Essentiality*; 2013; Vol. 11, pp 395–413.

(4) WHO. EXPOSURE TO CADMIUM : A MAJOR PUBLIC HEALTH CONCERN. **2010**.

(5) Hyder, O.; Chung, M.; Cosgrove, D.; Herman, J. M.; Li, Z.; Firoozmand, A.; Gurakar, A.; Koteish, A.; Pawlik, T. M. Cadmium Exposure and Liver Disease among US Adults. *J. Gastrointest. Surg.* **2013**, *17* (7), 1265–1273.

(6) Satarug, S.; Baker, J. R.; Urbenjapol, S.; Haswell-elkins, M.; Reilly, P. E. B.; Williams, J.; Moore, R. A Global Perspecti v e on Cadmium Pollution and Toxicity in Non-Occupationally Exposed Population. **2003**, *137*.

(7) McCarty, M. F. Zinc and Multi-Mineral Supplementation Should Mitigate the Pathogenic Impact of Cadmium Exposure. *Med. Hypotheses* **2012**, *79* (5), 642–648.

(8) Trevors, J.T., Stratton, G.W. & Gadd, G. M. Cadmium Transport, Resistance, and Toxicity in Bacteria, Algae, and Fungi1. *Can. J. Microbiol.* **1986**, *32* (6), 447–464.

(9) Orell, A., Remonsellez, F., Arancibia, R. & Jerez, C. Molecular Characterization of Copper and Cadmium Resistance Determinants in the Biomining Thermoacidophilic Archaeon Sulfolobus Metallicus. *Archaea* **2013**, *2013*, 1–16.

(10) Jacobson, K. B.; Opresko, L.; Owenby, R. K.; Christie, N. T. Effects of Cadmium on Drosophila: Toxicity, Proteins, and Transfer RNAs. *Toxicol. Appl. Pharmacol.* **1981**, *60*, 368–378.

(11) Coyle, P.; Philcox, J. C.; Carey, L. C.; Rofe, A. M. Metallothionein: The Multipurpose Protein. *Cell. Mol. Life Sci.* **2002**, *59* (4), 627–647.

(12) Waalkes, M. P. Cadmium Carcinogenesis. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.* **2003**, *533* (1–2), 107–120.

(13) Waalkes, M. P.; Goering, P. L. Metallothionein and Other Cadmium-Binding Proteins: Recent Developments. *Res. Toxicol.* **1990**, *3* (4), 281–288.

(14) United States Department of Labor - Occupational Safety and Health

Administration. Occupational Exposure to Cadmium https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=preambles&p_id=819.

(15) Mikolic, A.; Piasek, M.; Sulimanec Grgec, A.; Varnai, V. M.; Stasenko, S.; Kralik Oguic, S. Oral Cadmium Exposure during Rat Pregnancy: Assessment of Transplacental Micronutrient Transport and Steroidogenesis at Term. *J. Appl. Toxicol.* **2015**, *35* (5), 508–519.

(16) Satarug, S.; Garrett, S. H.; Sens, M. A.; Sens, D. A. Cadmium , Environmental Exposure , and Health Outcomes. **2010**, *182* (2), 182–190.

(17) Cao, Y.; Aralt, M.; Andersen, O.; Aaseth, J. Journal of Trace Elements in Medicine and Biology Chelation Therapy in Intoxications with Mercury , Lead and Copper. *J. Trace Elem. Med. Biol.* **2015**, *31*, 188–192.

(18) WHO. Global Status Report on Noncommunicable Diseases 2010. **2011**.

(19) Tellez-Plaza, M.; Navas-Acien, A.; Menke, A.; Crainiceanu, C. M.; Pastor-Barriuso, R.; Guallar, E. Cadmium Exposure and All Cause and Cardiovascular Mortality in the US General Population. *Environ. Health Perspect.* **2012**, *120* (April), 1017–1022.

(20) Chmielnicka, J.; Sowa, B. Cadmium Interaction with Essential Metals (Zn, Cu, Fe), Metabolism Metallothionein, and Ceruloplasmin in Pregnant Rats and Fetuses. *Ecotoxicol. Environ. Saf.* **1996**, *35* (3), 277–281.

(21) Tian, L. L.; Zhao, Y. C.; Wang, X. C.; Gu, J. L.; Sun, Z. J.; Zhang, Y. L.; Wang, J. X. Effects of Gestational Cadmium Exposure on Pregnancy Outcome and Development in the Offspring at Age 4.5 Years. *Biol. Trace Elem. Res.* **2009**, *132* (1–3), 51–59.

(22) Ronco, A. M.; Urrutia, M.; Montenegro, M.; Llanos, M. N. Cadmium Exposure during Pregnancy Reduces Birth Weight and Increases Maternal and Foetal Glucocorticoids. *Toxicol. Lett.* **2009**, *188* (3), 186–191.

(23) Wadhwa, P. D. Buss, C. Entringer, S. Swanson, J. M.; Ph, D. Developmental Origins of Health and Disease:Brief History of the Approach and Current Focus on Epigenetic Mechanisms. *Semin. Reprod. Med.* **2010**, *27* (5), 358–368.

(24) Ronco, A. M.; Montenegro, M.; Castillo, P.; Urrutia, M.; Saez, D.; Hirsch, S.; Zepeda, R.; Llanos, M. N. Maternal Exposure to Cadmium during Gestation Perturbs the Vascular System of the Adult Rat Offspring. *Toxicol. Appl. Pharmacol.* **2011**, *251* (2), 137–145.

(25) García-Arevalo, M.; Alonso-Magdalena, P.; Santos, J. R. Dos; Quesada, I.; Carneiro, E. M.; Nadal, A. Exposure to Bisphenol-A during Pregnancy Partially Mimics the Effects of a High-Fat Diet Altering Glucose Homeostasis and Gene

Expression in Adult Male Mice. *PLoS One* **2014**, *9* (6).

(26) Lee, H. S. Impact of Maternal Diet on the Epigenome during in Utero Life and the Developmental Programming of Diseases in Childhood and Adulthood. *Nutrients* **2015**, *7* (11), 9492–9507.

(27) Tchounwou, P. B.; Yedjou, C. G.; Patlolla, A. K.; Sutton, D. J. Heavy Metals Toxicity and the Environment. **2014**, 1–30.

(28) Shih, Y.; Islam, T.; Hore, S. K.; Sarwar, G.; Hasan, M.; Parvez, F.; Ahsan, H.; Argos, M. Associations between Prenatal Arsenic Exposure with Adverse Pregnancy Outcome and Child Mortality. **2018**, 456–461.

(29) Jin, X.; Tian, X.; Liu, Z.; Hu, H.; Li, X.; Deng, Y.; Li, N.; Zhu, J. Maternal Exposure to Arsenic and Cadmium and the Risk of Congenital Heart Defects in Offspring. *Reprod. Toxicol.* **2016**, *59*, 109–116.

(30) Taylor, C. M.; Golding, J.; Emond, A. M. Blood Mercury Levels and Fish Consumption in Pregnancy: Risks and Benefits for Birth Outcomes in a Prospective Observational Birth Cohort. *Int. J. Hyg. Environ. Health* **2016**, *219* (6), 513–520.

(31) Vilahur, N.; Vahter, M.; Broberg, K. The Epigenetic Effects of Prenatal Cadmium Exposure. *Curr. Environ. Heal. Reports* **2015**, *2* (2), 195–203.

(32) Wei, J.; Lin, Y.; Li, Y.; Ying, C.; Chen, J.; Song, L.; Zhou, Z.; Lv, Z.; Xia, W.; Chen, X.; Xu, S. Perinatal Exposure to Bisphenol A at Reference Dose Predisposes Offspring to Metabolic Syndrome in Adult Rats on a High-Fat Diet. *Endocrinology* **2011**, *152* (8), 3049–3061.

(33) National Institute of Environmental Health Sciences. Endocrine Disruptors https://www.niehs.nih.gov/health/topics/agents/endocrine/index.cfm (accessed Nov 27, 2018).

(34) del Díaz, C. M.; González, N.; Gómez, S.; Quiroga, M.; Najle, R.; CG., B. Effect of a Single Dose of Cadmium on Pregnant Wistar Rats and Their Offspring. *Reprod. Domest Anim.* **2014**, *49* (6), 1049–1056.

(35) Strakovsky, R. S.; Wang, H.; Engeseth, N. J.; Flaws, J. A.; Helferich, W. G.; Pan, Y. X.; Lezmi, S. Developmental Bisphenol A (BPA) Exposure Leads to Sex-Specific Modification of Hepatic Gene Expression and Epigenome at Birth That May Exacerbate High-Fat Diet-Induced Hepatic Steatosis. *Toxicol. Appl. Pharmacol.* **2015**, *284* (2), 101–112.

(36) Raubenheimer, D.; Simpson, S. J. Nutritional Ecology and Human Health. **2016**.

(37) Department of Health and Human Services. USDA Dietary Guidelines.

(38) Gernand, A. D.; Schulze, K. J.; Stewart, C. P.; Jr, K. P. W.; Bloomberg, J. H.

Micronutrient Deficiencies in Pregnancy Worldwide: Health Effects and Prevention. **2016**, *12* (5), 274–289.

(39) Winett, Liana; Wallack, Lawrence; Richardson, Dawn; Boone-Heinonen, Janne; Messer, L. A Framework to Address Challenges in Communicating the Developmental Origins of Health and Disease. *Curr Env. Heal. Rep* **2016**, *3* (3), 169–177.

(40) Heindel, J. J.; Vandenberg, L. N. Developmental Origins of Health and Disease: A Paradigm for Understanding Disease Etiology and Prevention. *Curr Opin Pediatr* **2015**, *27* (2), 248–253.

(41) Waddington, C. H. The Epigenotype. *Endeavour* **1942**, *1*, 18–20.

(42) Allis, C. D.; Jenuwein, T. The Molecular Hallmarks of Epigenetic Control. *Nat. Publ. Gr.* **2016**, *17* (8), 487–500.

(43) Li, G.; Zhu, P. Structure and Organization of Chromatin Fiber in the Nucleus. **2015**, *589*, 2893–2904.

(44) Annunziato, A. T. DNA Packaging: Nucleosomes and Chromatin. *Nat. Educ.* **2008**, *1* (1), 26.

(45) Tamaru, H. Confining Euchromatin / Heterochromatin Territory : Jumonji Crosses the Line. *Genes Dev.* **2010**, *24*, 1465–1478.

(46) Kadoch, C.; Crabtree, G. R. Mammalian SWI/SNF Chromatin Remodeling Complexes and Cancer: Mechanistic Insights Gained from Human Genomics. *Sci Adv 1* (5), 1–17.

(47) Feinberg, A. P. The Key Role of Epigenetics in Human Disease Prevention and Mitigation. *N. Engl. J. Med.* **2018**, *378* (14), 1323–1334.

(48) Fan, J.; Krautkramer, K. A.; Feldman, J. L.; Denu, J. M. Metabolic Regulation of Histone Post-Translational Modifications. *ACS Chem. Biol.* **2015**, *10* (1), 95–108.

(49) Zhang, T.; Cooper, S.; Brockdorff, N.; Ash, L.; Dot, L. The Interplay of Histone Modifications – Writers That Read. *EMBO Rep.* **2015**, *16* (11), 1467–1481.

(50) Lesch, B. J.; Page, D. C. Poised Chromatin in the Mammalian Germ Line. *Development* **2014**, *141*, 3619–3626.

(51) Bowers, E. C.; Mccullough, S. D. Linking the Epigenome with Exposure Effects and Susceptibility : The Epigenetic Seed and Soil Model. *Toxicol. Sci.* **2017**, *155* (2), 302–314.

(52) Guil, S.; Esteller, M. RNA – RNA Interactions in Gene Regulation : The Coding and Noncoding Players. *Trends Biochem. Sci.* **2015**, *40* (5), 248–256.

(53) Cech, T. R.; Steitz, J. A. The Noncoding RNA Revolution — Trashing Old Rules to Forge New Ones. *Cell* **2014**, *157* (1), 77–94.

(54) Bhan, A.; Soleimani, M.; Mandal, S. S. Long Noncoding RNA and Cancer : A New Paradigm. *Cancer Res.* **2017**, *77* (15), 3965–3982.

(55) Dempsey, J. L.; Cui, J. Y. Long Non-Coding RNAs : A Novel Paradigm for Toxicology. *Toxicol. Sci.* **2017**, *155* (1), 3–21.

(56) Li, E.; Zhang, Y. DNA Methylation in Mammals. *Cold Spring Harb. Perspect. Biol. 6* (5), a019133.

(57) Jeltsch, A.; Jurkowska, R. Z. *DNA Methyltransferases – Role and Function*, Volume 945.; Springer International Publishing Switzerland, 2016.

(58) Cowley, M.; Skaar, D. A.; Jima, D. D.; Maguire, R. L.; Hudson, K. M.; Park, S. S.; Sorrow, P.; Hoyo, C. Effects of Cadmium Exposure on DNA Methylation at Imprinting Control Regions and Genome-Wide in Mothers and Newborn Children. *Environ. Health Perspect. 126* (3), 0370031-12.

(59) Turdi, S.; Sun, W.; Tan, Y.; Yang, X.; Cai, L.; Ren, J. Inhibition of DNA Methylation Attenuates Low-Dose Cadmium-Induced Cardiac Contractile and Intracellular Ca2+ Anomalies. *Clin. Exp. Pharmacol. Physiol.* **2013**, *40* (10), 706–712.

(60) Takiguchi, M.; Achanzar, W. E.; Qu, W.; Li, G.; Waalkes, M. P. Effects of Cadmium on DNA-(Cytosine-5) Methyltransferase Activity and DNA Methylation Status during Cadmium-Induced Cellular Transformation. *Exp. Cell Res.* **2003**, *286* (2), 355–365.

(61) Cowley, M.; Oakey, R. J. Resetting for the next Generation. *Mol. Cell* **2012**, *48* (December 28), 819–821.

(62) Wang, L.; Zhang, J.; Duan, J.; Gao, X.; Zhu, W.; Lu, X.; Yang, L.; Zhang, J.; Li, G.; Ci, W.; Li, W.; Zhou, Q.; Aluru, N.; Tang, F.; He, C.; Huang, X.; Liu, J. Programming and Inheritance of Parental DNA Methylomes in Mammals. *Cell* **2014**, *157* (4), 979–991.

(63) Rodrigues, J. A.; Zilberman, D. Evolution and Function of Genomic Imprinting in Plants. *Genes Dev. 29* (24), 2517–2531.

(64) Ezkurdia, I.; Juan, D.; Rodriguez, J. M.; Frankish, A.; Diekhans, M.; Harrow, J.; Vazquez, J.; Valencia, A.; Tress, M. L. Multiple Evidence Strands Suggest That There May Be as Few as 19 000 Human Protein-Coding Genes. *Hum. Mol. Genet.* **2014**, *23* (22), 5866–5878.

(65) Bult, C. J.; Eppig, J. T.; Blake, J. A.; Kadin, J. A.; Richardson, J. E.; Mouse Genome Database. Mouse Genome Database 2016. *Nucleic Acids Res.* **2016**, *44* (November 2015), 840–847.

(66) Montibus, B.; Nury, D.; Tayama, C.; Wassef, M.; Kota, S. K.; Fogli, A.; Campos, F. C.; Hata, K.; Feil, R.; Margueron, R.; Nakabayashi, K.; Court, F.; Arnaud, P. Imprinting Control Regions (ICRs) Are Marked by Mono-Allelic Bivalent Chromatin When Transcriptionally Inactive. *Nucleic Acids Res.* **2016**, *44* (2), 621–635.

(67) Koerner, M. V; Pauler, F. M.; Huang, R.; Barlow, D. P. The Function of Non-Coding RNAs in Genomic Imprinting. *Development 136* (11), 1771–1783.

(68) Ideraabdullah, F. Y.; Thorvaldsen, J. L.; Myers, J. A.; Bartolomei, M. S. Tissue-Specific Insulator Function at H19/Igf2 Revealed by Deletions at the Imprinting Control Region. *Hum. Mol. Genet.* **2014**, *23* (23), 6246–6259.

(69) Vidal, A. C.; Semenova, V.; Darrah, T.; Vengosh, A.; Huang, Z.; King, K.; Nye, M. D.; Fry, R.; Skaar, D.; Maguire, R.; Murtha, A.; Schildkraut, J.; Murphy, S.; Hoyo, C. Maternal Cadmium, Iron and Zinc Levels, DNA Methylation and Birth Weight. *BMC Pharmacol. Toxicol.* **2015**, *16*, 20.

(70) National Human Genome Research Institute. Background on the History of the Mouse https://www.genome.gov/10005832/background-on-the-history-of-the-mouse/.

(71) National Human Genome Research Institute. Understanding Our Genetic Inheritance https://www.genome.gov/10001477/human-genome-projects-fiveyear-plan-19911995/.

(72) Perlman, R. L. Mouse Models of Human Disease - An Evolutionary Perspective. *Evol. Med. Public Heal.* **2016**, 170–176.

(73) Wellcome Sanger Institute. Mouse Genomes Project https://www.sanger.ac.uk/science/data/mouse-genomes-project (accessed Jan 14, 2019).

(74) Falconer, D. S.; Mackay, T. F. C. *Introduction to Quantitative Genetics*, 4th ed.; Harlow, England:Prentice Hall, 1996.

(75) Järup, L.; Åkesson, A. Current Status of Cadmium as an Environmental Health Problem. *Toxicol. Appl. Pharmacol.* **2009**, *238* (3), 201–208.

(76) Ferramola, M. L.; Antón, R. I.; Anzulovich, A. C.; Giménez, M. S. Myocardial Oxidative Stress Following Sub-Chronic and Chronic Oral Cadmium Exposure in Rats. *Environ. Toxicol. Pharmacol.* **2011**, *32* (1), 17–26.

(77) Tellez-Plaza, M.; Guallar, E.; Fabsitz, R. R.; Howard, B. V.; Umans, J. G.; Francesconi, K. A.; Goessler, W.; Devereux, R. B.; Navas-Acien, A. Cadmium Exposure and Incident Peripheral Arterial Disease. *Circ. Cardiovasc. Qual. Outcomes* **2013**, *6* (6), 626–633.

(78) Thomas Farley; DeMaria, A. N.; Wright, J.; Conway, P. H.; Valderrama, A. L.; Blair,

N. A.; Kelly, M.; Popovic, T. CDC Grand Rounds: The Million Hearts Initiative. *MMWR, Morb. Mortal. Wkly. Rep.* **2012**, *61* (50), 1022–1024.

(79) Eum, K. D.; Lee, M. S.; Paek, D. Cadmium in Blood and Hypertension. *Sci. Total Environ.* **2008**, *407* (1), 147–153.

(80) Tellez-Plaza, M.; Navas-Acien, A.; Crainiceanu, C. M.; Guallar, E. Cadmium Exposure and Hypertension in the 1999-2004 National Health and Nutrition Examination Survey (NHANES). *Environ. Health Perspect.* **2008**, *116* (1), 51–56.

(81) Fryar, C. D.; Chen, T.-C.; Li, X. Prevalence of Uncontrolled Risk Factors for Cardiovascular Disease: United States, 1999-2010. *NCHS Data Brief* **2012**, *August* (103), 1–8.

(82) Daugherty, A.; Rateri, D.; Hong, L.; Balakrishnan, A. Measuring Blood Pressure in Mice Using Volume Pressure Recording, a Tail-Cuff Method. *J. Vis. Exp.* **2009**, *27* (e1291).

(83) Amcooksey. Mouse RNASeq DeSEQ2 pipe https://wiki.cyverse.org/wiki/display/TUT/Mouse+RNASeq+DeSEQ2+pipe (accessed Dec 12, 2018).

(84) Chen, E. Y.; Tan, C. M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G. V.; Clark, N. R.; Ma'ayan, A. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool. *BMC Bioinformatics* **2013**, *14*, 128.

(85) Kuleshov, M. V; Jones, M. R.; Rouillard, A. D.; Fernandez, N. F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S. L.; Jagodnik, K. M.; Lachmann, A.; McDermott, M. G.; Monteiro, C. D.; Gundersen, G. W.; Ma'ayan, A. Enrichr: A Comprehensive Gene Set Enrichment Analysis Web Server 2016 Update. *Nucleic Acids Res.* **2016**, *44* (W1), W90-7.

(86) Livak, K. J.; Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods 25* (4), 402–408.

(87) Bustin, S. A.; Benes, V.; Garson, J.; Hellemans, J.; Huggett, J.; Kubista, M.; Mueller, R.; Nolan, T.; Pfaffl, M. W.; Shipley, G.; Vandesompele, J.; Wittwer, C. T. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **2009**, *55* (4), 1–12.

(88) Whelton, B.; Bhattacharyya, M.; Carnes, B.; Moretti, E.; Peterson, D. Female Reproduction and Pup Survival and Growth for Mice Fed a Cadmium-Containing Purified Diet through Six Consecutive Rounds of Gestation and Lactation. *J Toxicol Env. Heal.* **1988**, *24* (3), 321–343.

(89) Liu, J.; Qu, W.; Kadiiska, M. B. Role of Oxidative Stress in Cadmium Toxicity and Carcinogenesis. *Toxicol. Appl. Pharmacol.* **2009**, *238* (3), 209–214.

(90) Lassègue, B.; Griendling, K. Reactive Oxygen Species in Hypertension; An Update. *Am. J. Hypertens.* **2004**, *17* (9), 852–860.

(91) Dalle-Donne, I.; Giustarini, D.; Colombo, R.; Rossi, R.; Milzani, A. Protein Carbonylation in Human Diseases. *Trends Mol. Med.* **2003**, *9* (4), 169–176.

(92) Piasek, M.; Blanuša, M.; Kostial, K.; Laskey, J. W. Low Iron Diet and Parenteral Cadmium Exposure in Pregnant Rats: The Effects on Trace Elements and Fetal Viability. *BioMetals* **2004**, *17* (1), 1–14.

(93) Gwaltney-Brant, S. M. Chapter 41 - Heavy Metals. In *Haschek and Rousseaux's Handbook of Toxicologic Pathology (Third Edition)*; Haschek, W. M., Rousseaux, C. G., Wallig, M. A., Eds.; Academic Press: Boston, 2013; pp 1315–1347.

(94) ATSDR. Substance Priority List. U.S. Department of Health and Human Services 2017.

(95) Wang, Y.; Chen, L.; Gao, Y.; Zhang, Y.; Wang, C.; Zhou, Y.; Hu, Y.; Shi, R.; Tian, Y. Effects of Prenatal Exposure to Cadmium on Neurodevelopment of Infants in Shandong, China. *Environ. Pollut.* **2016**, *211*, 67–73.

(96) King, K. E.; Darrah, T. H.; Money, E.; Meentemeyer, R.; Maguire, R. L.; Nye, M. D.; Michener, L.; Murtha, A. P.; Jirtle, R.; Murphy, S. K.; Mendez, M. A.; Robarge, W.; Vengosh, A.; Hoyo, C. Geographic Clustering of Elevated Blood Heavy Metal Levels in Pregnant Women Environmental Health. *BMC Public Health* **2015**, *15* (1035), 1–12.

(97) Sasaki, T.; Horiguchi, H.; Arakawa, A.; Oguma, E.; Komatsuda, A.; Sawada, K.; Murata, K.; Yokoyama, K.; Matsukawa, T.; Chiba, M.; Omori, Y.; Kamikomaki, N. Hospital-Based Screening to Detect Patients with Cadmium Nephropathy in Cadmium-Polluted Areas in Japan. *Environ. Health Prev. Med.* **2019**, *24* (8), 1–8.

(98) Nishijo, M.; Nambunmee, K.; Suvagandha, D.; Swaddiwudhipong, W.; Ruangyuttikarn, W.; Nishino, Y. Gender-Specific Impact of Cadmium Exposure on Bone Metabolism in Older People Living in a Cadmium-Polluted Area in Thailand. *Int. J. Environ. Res. Public Health* **2017**, *14* (401), 1–11.

(99) Castillo, P.; Ibáñez, F.; Guajardo, A.; Llanos, M. N.; Ronco, A. M. Impact of Cadmium Exposure during Pregnancy on Hepatic Glucocorticoid Receptor Methylation and Expression in Rat Fetus. *PLoS One* **2012**, *7* (9), 1–9.

(100) Kraft, M. D. Phosphorus and Calcium: A Review for the Adult Nutrition Support Clinician. *Nutr. Clin. Pract.* **2015**, *30* (1), 21–33.

(101) Ando, K.; Fujita, T. Pathophysiology of Salt Sensitivity Hypertension. *Ann Med* **2012**, *44 Suppl 1*, S119-126.

(102) Schneider, M. P.; Raff, U.; Kopp, C.; Scheppach, J. B.; Toncar, S.; Wanner, C.;

Schlieper, G.; Saritas, T.; Floege, J.; Schmid, M.; Birukov, A.; Dahlmann, A.; Linz, P.; Janka, R.; Uder, M.; Schmieder, R. E.; Titze, J. M.; Eckardt, K.-U. Skin Sodium Concentration Correlates with Left Ventricular Hypertrophy in CKD. *J Am Soc Nephrol* **2017**, *28* (6), 1867–1876.

(103) Ye, Q.; Park, J. E.; Gugnani, K.; Betharia, S.; Pino-figueroa, A.; Kim, J. Influence of Iron Metabolism on Manganese Transport and Toxicity. *Metallomics* **2017**, *9* (8), 1028–1046.

(104) Bhattacharya, P. T.; Misra, S. R.; Hussain, M. Nutritional Aspects of Essential Trace Elements in Oral Health and Disease: An Extensive Review. *Scientifica (Cairo).* **2016**, *2016*, 1–12.

(105) Lapice, E.; Masulli, M.; Vaccaro, O. Iron Deficiency and Cardiovascular Disease : An Updated Review of the Evidence. *Curr Atheroscler Rep* **2013**, *15* (358), 1–14.

(106) Tardiff, J. C. Cardiac Hypertrophy: Stressing out the Heart. *J. Clin. Invest.* **2006**, *116* (6), 1467–1470.

(107) Hu, H.; Lu, X.; Cen, X.; Chen, X.; Li, F.; Zhong, S. RNA-Seq Identifies Key Reproductive Gene Expression Alterations in Response to Cadmium Exposure. *Biomed Res. Int.* **2014**, *2014*.

(108) Caserta, D.; Graziano, A.; Monte, G. LO; Bordi, G.; Moscarini, M. Heavy Metals and Placental Fetal-Maternal Barrier: A Mini-Review on the Major Concerns. *Eur. Rev. Med. Pharmacol. Sci.* **2013**, *17*, 2198–2206.

(109) Grawé, K. P.; Teiling-Gårdlund, A.; Jalkesten, E.; Oskarsson, A. Increased Spontaneous Motor Activity in Offspring after Maternal Cadmium Exposure during Lactation. *Environ. Toxicol. Pharmacol.* **2004**, *17*, 35–43.

(110) Minetti, A.; Reale, C. A. Sensorimotor Developmental Delays and Lower Anxiety in Rats Prenatally Exposed to Cadmium. *J. Appl. Toxicol.* **2006**, *26*, 35–41.

(111) Zhang, Y. M.; Liu, X. Z.; Lu, H.; Mei, L.; Liu, Z. P. Lipid Peroxidation and Ultrastructural Modifications in Brain after Perinatal Exposure to Lead and/or Cadmium in Rat Pups. *Biomed. Environ. Sci.* **2009**, *22*, 423–429.

(112) Ishitobi, H.; Mori, K.; Yoshida, K.; Watanabe, C. Effects of Perinatal Exposure to Low-Dose Cadmium on Thyroid Hormone-Related and Sex Hormone Receptor Gene Expressions in Brain of Offspring. *Neurotoxicology* **2007**, *28*, 790–797.

(113) Mori, K.; Yoshida, K.; Hoshikawa, S.; Ito, S.; Yoshida, M.; Satoh, M.; Watanabe, C. Effects of Perinatal Exposure to Low Doses of Cadmium or Methylmercury on Thyroid Hormone Metabolism in Metallothionein-Deficient Mouse Neonates. *Toxicology* **2006**, *228*, 77–84.

(114) Hicks, K. D.; Sullivan, A. W.; Cao, J.; Sluzas, E.; Rebuli, M.; Patisaul, H. B.

Interaction of Bisphenol A (BPA) and Soy Phytoestrogens on Sexually Dimorphic Sociosexual Behaviors in Male and Female Rats. *Horm. Behav.* **2016**, *84*, 121–126.

(115) Sullivan, A. W.; Beach, E. C.; Stetzik, L. A.; Perry, A.; D'Addezio, A. S.; Cushing, B. S.; Patisaul, H. B. A Novel Model for Neuroendocrine Toxicology: Neurobehavioral Effects of BPA Exposure in a Prosocial Species, the Prairie Vole (Microtus Ochrogaster). *Endocrinology 155* (10), 3867–3881.

(116) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29* (1), 15–21.

(117) Anders, S.; Pyl, P. T.; Huber, W. HTSeq - A Python Framework to Work with High-Throughput Sequencing Data. *Bioinformatics* **2015**, *31* (2), 166–169.

(118) Love, M. I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15* (550), 1–21.

(119) Santos, J. H.; Mandavilli, B. S.; Van Houten, B. Measuring Oxidative MtDNA Damage and Repair Using Quantitative PCR. In *Mitochondrial DNA: Methods and Protocols*; Copeland, W. C., Ed.; Humana Press: Totowa, NJ, 2002; pp 159–176.

(120) Quiros, P. M.; Goyal, A.; Jha, P.; Auwerx, J. Analysis of MtDNA/NDNA Ratio in Mice. *Curr. Protoc. Mouse Biol.* **2017**, *7* (1), 47–54.

(121) Manza, L. L.; Stamer, S. L.; Ham, A.-J. L.; Codreanu, S. G.; Liebler, D. C. Sample Preparation and Digestion for Proteomic Analyses Using Spin Filters. *Proteomics* **2005**, *5* (7), 1742–1745.

(122) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal Sample Preparation Method for Proteome Analysis. *Nat. Methods* **2009**, *6* (5), 359–362.

(123) Jones, J. W.; Pierzchalski, K.; Yu, J.; Kane, M. A. Use of Fast HPLC Multiple Reaction Monitoring Cubed for Endogenous Retinoic Acid Quantification in Complex Matrices. *Anal. Chem.* **2015**, *87* (6), 3222–3230.

(124) Kane, M. A.; Folias, A. E.; Napoli, J. L. HPLC/UV Quantitation of Retinal, Retinol, and Retinyl Esters in Serum and Tissues. *Anal. Biochem.* **2008**, *378* (1), 71–79.

(125) Kane, M. A.; Napoli, J. L. Quantification of Endogenous Retinoids. *Methods Mol. Biol.* **2010**, *652*, 1–54.

(126) Kaidanovich-Beilin, O.; Lipina, T.; Vukobradovic, I.; Roder, J.; Woodgett, J. R. Assessment of Social Interaction Behaviors. *J. Vis. Exp.* **2011**, *0* (48), 1–6.

(127) Seibenhener, M. L.; Wooten, M. C. Use of the Open Field Maze to Measure Locomotor and Anxiety-like Behavior in Mice. *J. Vis. Exp.* **2015**, *February* (e52434),

1–6.

(128) Salvati, S.; Attorri, L.; Avellino, C.; Di Biase, A.; Sanchez, M. Diet, Lipids and Brain Development. *Dev. Neurosci.* **2000**, *22* (5–6), 481–487.

(129) Thakurela, S.; Garding, A.; Jung, R. B.; Müller, C.; Goebbels, S.; White, R.; Werner, H. B.; Tiwari, V. K. The Transcriptome of Mouse Central Nervous System Myelin. *Sci. Rep.* **2016**, *6* (April), 1–12.

(130) Gravel, M.; Trapp, B.; Peterson, J.; Braun, P. E. CNP in Myelination - Overexpression Alters Oligodendrocyte Morphogenesis. *Cell Biol. Pathol. Myelin* **1997**, 75–82.

(131) Kazakova, N.; Li, H.; Mora, A.; Jessen, K. R.; Mirsky, R.; Richardson, W. D.; Smith, H. K. A Screen for Mutations in Zebrafish That Affect Myelin Gene Expression in Schwann Cells and Oligodendrocytes. *Dev. Biol.* **2006**, *297* (1), 1–13.

(132) Fard, M. K.; van der Meer, F.; Sanchez, P.; Cantuti-Castelvetri, L.; Mandad, S.; Jakel, S.; Fornasiero, E. F.; Schmitt, S.; Ehrlich, M.; Starost, L.; Kuhlmann, T.; Sergiou, C.; Schultz, V.; Wrzos, C.; Bruck, W.; Urlaub, H.; Dimou, L.; Stadelmann, C.; Simons, M. BCAS1 Expression Defines a Population of Early Myelinating Oligodendrocytes in Multiple Sclerosis Lesions. *Sci. Transl. Med.* **2017**, *9* (419).

(133) Wicher, G.; Husic, E.; Nilsson, G.; Forsberg-Nilsson, K. Developmental Expression of IL-33 in the Mouse Brain. *Neurosci. Lett.* **2013**, *555*, 171–176.

(134) Kaller, M. S.; Lazari, A.; Blanco-Duque, C.; Sampaio-Baptista, C.; Johansen-Berg, H. Myelin Plasticity and Behaviour — Connecting the Dots. *Curr. Opin. Neurobiol.* **2017**, *47*, 86–92.

(135) Noronha, A. B.; Hammer, J. A.; Lai, C.; Kiel, M.; Milner, R. J.; Sutcliffe, J. G.; Quarles, R. H. Myelin-Associated Glycoprotein (MAG) and Rat Brain-Specific 1B236 Protein: Mapping of Epitopes and Demonstration of Immunological Identity. *J. Mol. Neurosci.* **1989**, *1* (3), 159–170.

(136) Zhang, S.; Zhu, X.; Gui, X.; Croteau, C.; Song, L.; Xu, J.; Wang, A.; Bannerman, P.; Guo, F. Sox2 Is Essential for Oligodendroglial Proliferation and Differentiation during Postnatal Brain Myelination and CNS Remyelination. *J. Neurosci. 38* (7), 1802–1820.

(137) Monk, K. R.; Talbot, W. S. Genetic Dissection of Myelinated Axons in Zebrafish. *Curr. Opin. Neurobiol.* **2009**, *19* (5), 486–490.

(138) Petryniak, M. A.; Potter, G. B.; Rowitch, D. H.; Rubenstein, J. L. R. Dlx1 and Dlx2 Control Neuronal versus Oligodendroglial Cell Fate Acquisition in the Developing Forebrain. *Neuron* **2007**, *55* (3), 417–433.

(139) Boggs, J. M. Myelin Basic Protein: A Multifunctional Protein. *Cell. Mol. Life Sci.*

2006, *63* (17), 1945–1961.

(140) Brasko, C.; Hawkins, V.; De La Rocha, I. C.; Butt, A. M. Expression of Kir4.1 and Kir5.1 Inwardly Rectifying Potassium Channels in Oligodendrocytes, the Myelinating Cells of the CNS. *Brain Struct. Funct.* **2017**, *222* (1), 41–59.

(141) Trimarco, A.; Forese, M. G.; Alfieri, V.; Lucente, A.; Brambilla, P.; Dina, G.; Pieragostino, D.; Sacchetta, P.; Urade, Y.; Boizet-Bonhoure, B.; Martinelli Boneschi, F.; Quattrini, A.; Taveggia, C. Prostaglandin D2 Synthase/GPR44: A Signaling Axis in PNS Myelination. *Nat. Neurosci.* **2014**, *17* (12), 1682–1692.

(142) Deschamps, J. Developmental Regulation of the Hox Genes during Axial Morphogenesis in the Mouse. *Development* **2005**, *132* (13), 2931–2942.

(143) Jahn, O.; Tenzer, S.; Werner, H. B. Myelin Proteomics: Molecular Anatomy of an Insulating Sheath. *Mol. Neurobiol.* **2009**, *40*, 55–72.

(144) Luo, T.; Wagner, E.; Crandall, J. E.; Dr̈ger, U. C. A Retinoic-Acid Critical Period in the Early Postnatal Mouse Brain. *Biol. Psychiatry* **2004**, *56* (12), 971–980.

(145) Lagarde, F.; Beausoleil, C.; Belcher, S. M.; Belzunces, L. P.; Emond, C.; Guerbet, M.; Rousselle, C. Non-Monotonic Dose-Response Relationships and Endocrine Disruptors: A Qualitative Method of Assessment. *Environ. Heal.* **2015**, *14* (13), 1–15.

(146) Monaco, A.; Grimaldi, M. C.; Ferrandino, I. Neuroglial Alterations in the Zebrafish Brain Exposed to Cadmium Chloride. *J. Appl. Toxicol.* **2016**, *36* (12), 1629–1638.

(147) Rai, N. K.; Ashok, A.; Rai, A.; Tripathi, S.; Nagar, G. K.; Mitra, K.; Bandyopadhyay, S. Exposure to As, Cd and Pb-Mixture Impairs Myelin and Axon Development in Rat Brain, Optic Nerve and Retina. *Toxicol. Appl. Pharmacol.* **2013**, *273*, 242–258.

(148) Makinodan, M.; Tatsumi, K.; Okuda, H.; Manabe, T.; Yamauchi, T.; Noriyama, Y.; Kishimoto, T.; Wanaka, A. Lysophosphatidylcholine Induces Delayed Myelination in the Juvenile Ventral Hippocampus and Behavioral Alterations in Adulthood. *Neurochem. Int.* **2008**, *53*, 374–381.

(149) Nolte, C.; Krumlauf, R. Expression of Hox Genes in the Nervous System of Vertebrates. In *Madame Curie Bioscience Database [Internet]*; Landes Bioscience: Austin, TX, 2013; pp 1–30.

(150) Dasen, J. S.; Jessell, T. M. *Chapter Six: Hox Networks and the Origins of Motor Neuron Diversity*, 1st ed.; Elsevier Inc., 2009; Vol. 88.

(151) Nakagawa, K.; Lee, M. J.; Sasaki, N.; Hayashi, C.; Nishio, H. Cadmium Exposure Induces Expression of the HOXB8 Gene in COS-7 Cells. *Toxicol. Vitr.* **2008**, *22* (6), 1447–1451.

(152) Quinonez, S. C.; Innis, J. W. Human HOX Gene Disorders. *Mol. Genet. Metab.* **2014**, *111*, 4–15.

(153) Andreone, B. J.; Lacoste, B.; Gu, C. Neuronal and Vascular Interactions. *Annu. Rev. Neurosci.* **2015**, *38*, 25–46.

(154) Cohen, E.; Baerts, W.; Van Bel, F. Brain-Sparing in Intrauterine Growth Restriction: Considerations for the Neonatologist. *Neonatology* **2015**, *108*, 269–276.

(155) Rhinn, M.; Dolle, P. Retinoic Acid Signalling during Development. *Development* **2012**, *139*, 843–858.

(156) Cui, Y.; Freedman, J. H. Cadmium Induces Retinoic Acid Signaling by Regulating Retinoic Acid Metabolic Gene Expression. *J. Biol. Chem.* **2009**, *284* (37), 24925–24932.

(157) Guintivano, J.; Aryee, M. J.; Kaminsky, Z. A. A Cell Epigenotype Specific Model for the Correction of Brain Cellular Heterogeneity Bias and Its Application to Age, Brain Region and Major Depression. *Epigenetics 8* (3), 290–302.

(158) Baroncelli, L.; Scali, M.; Sansevero, G.; Olimpico, F.; Manno, I.; Costa, M.; Sale, A. Experience Affects Critical Period Plasticity in the Visual Cortex through an Epigenetic Regulation of Histone Post-Translational Modifications. *J. Neurosci.* **2016**, *36* (12), 3430–3440.

(159) Zhang, H.; Li, X.; Nie, J.; Niu, Q. Lactation Exposure to BDE-153 Damages Learning and Memory, Disrupts Spontaneous Behavior and Induces Hippocampus Neuron Death in Adult Rats. *Brain Res.* **2013**, *1517*, 44–56.

(160) Gavin, D. P.; Grayson, D. R.; Varghese, S. P.; Guizzetti, M. Chromatin Switches during Neural Cell Differentiation and Their Dysregulation by Prenatal Alcohol Exposure. *Genes (Basel).* **2017**, *8* (5).

(161) Bommarito, P. A.; Martin, E.; Fry, R. C. Effects of Prenatal Exposure to Endocrine Disruptors and Toxic Metals on the Fetal Epigenome. *Epigenomics* **2017**, *9* (3), 333–350.

(162) Hara, Y.; Ago, Y.; Takano, E.; Hasebe, S.; Nakazawa, T.; Hashimoto, H.; Matsuda, T.; Takuma, K. Prenatal Exposure to Valproic Acid Increases MiR-132 Levels in the Mouse Embryonic Brain. *Mol. Autism* **2017**, *8*, 33.

(163) Ivanova, E.; Chen, J.-H.; Segonds-Pichon, A.; Ozanne, S. E.; Kelsey, G. DNA Methylation at Differentially Methylated Regions of Imprinted Genes Is Resistant to Developmental Programming by Maternal Nutrition. *Epigenetics* **2012**, *7* (10), 1200–1210.

(164) Salzberg, S. L. Open Questions: How Many Genes Do We Have? *BMC Biol.* **2018**, *16* (1), 94.

(165) Poirier, L. A.; Vlasova, T. I. The Prospective Role of Abnormal Methyl Metabolism in Cadmium Toxicity. *Environ. Health Perspect.* **2002**, *110* (SUPPL. 5), 793–795.

(166) Fouse, S. D.; Nagarajan, R. P.; Costello, J. F. Genome-Scale DNA Methylation Analysis. *Epigenomics 2* (1), 105–117.

(167) Gu, H.; Smith, Z. D.; Bock, C.; Boyle, P.; Gnirke, A.; Meissner, A. Preparation of Reduced Representation Bisulfite Sequencing Libraries for Genome-Scale DNA Methylation Profiling. *Nat. Protoc.* **2011**, *6* (4), 468–481.

(168) FastQC, Babraham Bioinformatics https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

(169) Trim Galore, Babraham Bioinformatics http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

(170) Xi, Y.; Li, W. BSMAP: Whole Genome Bisulfite Sequencing MAPping Program. *BMC Bioinformatics* **2009**, *10* (232).

(171) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25* (16), 2078–2079.

(172) Krueger, F.; Andrews, S. R. Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications. *Bioinformatics* **2011**, *27* (11), 1571–1572.

(173) Krueger, F.; Andrews, S. R. SNPsplit: Allele-Specific Splitting of Alignments between Genomes with Known SNP Genotypes [Version 1; Referees: 3 Approved]. *F1000Research* **2016**, *5* (1479).

(174) Proudhon, C.; Duffié, R.; Ajjan, S.; Cowley, M.; Iranzo, J.; Carbajosa, G.; Saadeh, H.; Holland, M. L.; Oakey, R. J.; Rakyan, V. K.; Schulz, R.; Bourc'his, D. Protection against De Novo Methylation Is Instrumental in Maintaining Parent-of-Origin Methylation Inherited from the Gametes. *Mol. Cell* **2012**, *47*, 909–920.

(175) Illingworth, R. S.; Gruenewald-Schneider, U.; Webb, S.; Kerr, A. R. W.; James, K. D.; Turner, D. J.; Smith, C.; Harrison, D. J.; Andrews, R.; Bird, A. P. Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. *PLoS Genet.* **2010**, *6* (9), e1001134.

(176) Tomizawa, S.; Kobayashi, H.; Watanabe, T.; Andrews, S.; Hata, K.; Kelsey, G.; Sasaki, H. Dynamic Stage-Specific Changes in Imprinted Differentially Methylated Regions during Early Mammalian Development and Prevalence of Non-CpG Methylation in Oocytes. *Development* **2011**, *138* (5), 811–820.

(177) Bailey, T. L. DREME: Motif Discovery in Transcription Factor ChIP-Seq Data. *Bioinformatics* **2011**, *27* (12), 1653–1659.

(178) Bailey, T. L.; Williams, N.; Misleh, C.; Li, W. W. MEME: Discovering and Analyzing DNA and Protein Sequence Motifs. *Nucleic Acids Res.* **2006**, *34* (Web Server issue), W369-73.

(179) Gupta, S.; Stamatoyannopoulos, J. A.; Bailey, T. L.; Noble, W. S. Quantifying Similarity between Motifs. *Genome Biol.* **2007**, *8* (2), R24.

(180) Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal* **2011**, *17* (1), 10–12.

(181) Trapnell, C.; Pachter, L.; Salzberg, S. L. TopHat: Discovering Splice Junctions with RNA-Seq. *Bioinformatics* **2009**, *25* (9), 1105–1111.

(182) Huang, S.; Holt, J.; Kao, C.-Y.; McMillan, L.; Wang, W. A Novel Multi-Alignment Pipeline for High-Throughput Sequencing Data. *Database (Oxford).* **2014**, *2014* (bau057).

(183) Holt, J.; Huang, S.; McMillan, L.; Wang, W. Read Annotation Pipeline for High-Throughput Sequencing Data. *Proc. Int. Conf. Bioinforma.* **2013**, *Computatio* (ACM, New York), 605–612.

(184) He, H.; Perera, B. P. U.; Ye, A.; Kim, J. Parental and Sexual Conflicts over the Peg3 Imprinted Domain. *Sci. Rep.* **2016**, *6*, 38136.

(185) Monk, D.; Wagschal, A.; Arnaud, P.; Müller, P.-S.; Parker-Katiraee, L.; Bourc'his, D.; Scherer, S. W.; Feil, R.; Stanier, P.; Moore, G. E. Comparative Analysis of Human Chromosome 7q21 and Mouse Proximal Chromosome 6 Reveals a Placental-Specific Imprinted Gene, TFPI2/Tfpi2, Which Requires EHMT2 and EED for Allelic-Silencing. *Genome Res.* **2008**, *18* (8), 1270–1281.

(186) Schmidt, A.; Morales-Prieto, D. M.; Pastuschek, J.; Fröhlich, K.; Markert, U. R. Only Humans Have Human Placentas: Molecular Differences between Mice and Humans. *J. Reprod. Immunol.* **2015**, *108*, 65–71.

(187) Ferron, S. R.; Charalambous, M.; Radford, E.; McEwen, K.; Wildner, H.; Hind, E.; Morante-Redolat, J. M.; Laborda, J.; Guillemot, F.; Bauer, S. R.; Farinas, I.; Ferguson-Smith, A. C. Postnatal Loss of Dlk1 Imprinting in Stem Cells and Niche Astrocytes Regulates Neurogenesis. *Nature* **2011**, *475* (7356), 381–385.

(188) Garfield, A. S.; Cowley, M.; Smith, F. M.; Moorwood, K.; Stewart-Cox, J. E.; Gilroy, K.; Baker, S.; Xia, J.; Dalley, J. W.; Hurst, L. D.; Wilkinson, L. S.; Isles, A. R.; Ward, A. Distinct Physiological and Behavioural Functions for Parental Alleles of Imprinted Grb10. *Nature* **2011**, *469* (7331), 534–538.

APPENDICES

Appendix A

Chapter 2 Supplemental Data

Supplementary file A2.1. DESeq2.pdf

Statistical report of project RNA-seq PND1 BxC female hearts: pairwise comparison(s) of conditions with DESeq2

Author: Kathleen M Hudson Date: 2018-12-14

Table of contents

1 Introduction

The analyses reported in this document are part of the RNA-seq PND1 BxC female hearts project. The aim is to find features that are differentially expressed between 0ppm and 50ppm. The statistical analysis process includes data normalization, graphical exploration of raw and normalized data, test for differential expression for each feature between the conditions, raw p- value adjustment and export of lists of features having a significant differential expression between the conditions.

The analysis is performed using the R software [R Core Team, 2014], Bioconductor [Gentleman, 2004] packages including DESeq2 [Anders, 2010 and Love, 2014] and the SARTools package developed at PF2 - Institut Pasteur. Normalization and differential analysis are carried out according to the DESeq2 model and package. This report comes with additional tab-delimited text files that contain lists of differentially expressed features.

For more details about the DESeq2 methodology, please refer to its related publications [Anders, 2010 and Love, 2014].

Supplementary file A2.1. DESeq2.pdf continued

2 Description of raw data

The count data files and associated biological conditions are listed in the following table.

label	files	group
300-7	trimS300_7_S39_L007_R1_001.sorted.txt	0ppm
302-2	trimS302_2_S40_L007_R1_001.sorted.txt	0ppm
304-2a	trimS304_2a_S41_L007_R1_001.sorted.txt	0ppm
305-4	trimS305_4_S42_L007_R1_001.sorted.txt	0ppm
312-2	trimS312_2_S43_L007_R1_001.sorted.txt	50ppm
317-4a	trimS317_4a_S46_L007_R1_001.sorted.txt	50ppm
314-4	trimS314_4_S44_L007_R1_001.sorted.txt	50ppm
316-5	trimS316_5_S45_L007_R1_001.sorted.txt	50ppm

Table 1: Data files and associated biological conditions.

After loading the data we first have a look at the raw data table itself. The data table contains one row per annotated feature and one column per sequenced sample. Row names of this table are feature IDs (unique identifiers). The table contains raw count values representing the number of reads that map onto the features. For this project, there are 54535 features in the count data table.

	300-7	302-2	304-2a	305-4	312-2	317-4a	314-4	316-5
__ambiguous	1400240	1668300	1470845	1092383	1448140	1821658	1706559	2002669
__no_feature	969809	1005092	994718	869465	929517	1263211	1226020	1398831
__not_aligned	739409	1071316	758294	587358	802247	1026542	1120002	1237595
__too_low_aQual	0	0	0	0	0	0	0	0
ENSMUSG00000000003	0	0	0	0	0	0	0	0
ENSMUSG00000000028	298	354	207	161	478	437	354	427

Table 2: Partial view of the count data table.

Looking at the summary of the count table provides a basic description of these raw counts (min and max values, median, etc).

	300-7	302-2	304-2a	305-4	312-2	317-4a	314-4	316-5
Min.	0	0	0	0	0	0	0	0
1st Qu.	0	0	0	0	0	0	0	0
Median	0	0	0	0	0	0	0	0
Mean	465	555	503	382	475	604	592	675
3rd Qu.	22	25	26	21	22	33	29	31
Max.	1400240	1668300	1470845	1092383	1448140	1821658	1706559	2002669

Table 3: Summary of the raw counts.

Figure 1 shows the total number of mapped reads for each sample. Reads that map on multiple locations on the transcriptome are counted more than once, as far as they are mapped on less than 50 different loci. We expect total read counts to be similar within conditions, they may be different across conditions. Total counts sometimes vary widely between replicates. This may happen for several reasons, including:

- different rRNA contamination levels between samples (even between biological replicates);
- slight differences between library concentrations, since they may be difficult to measure with high precision.

Supplementary file A2.1. DESeq2.pdf continued

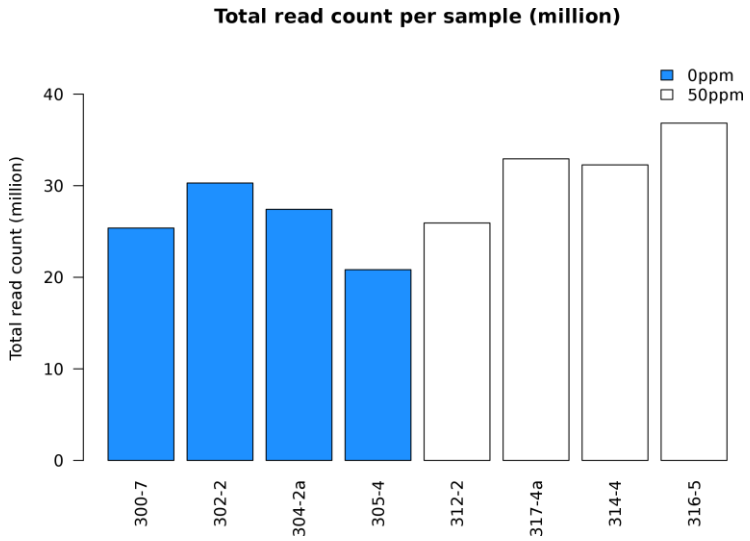


Figure 1: Number of mapped reads per sample. Colors refer to the biological condition of the sample.

Figure 2 shows the proportion of features with no read count in each sample. We expect this proportion to be similar within conditions. Features with null read counts in the 8 samples are left in the data but are not taken into account for the analysis with DESeq2. Here, 23839 features (43.71%) are in this situation (dashed line). Results for those features (fold-change and p-values) are set to NA in the results files.

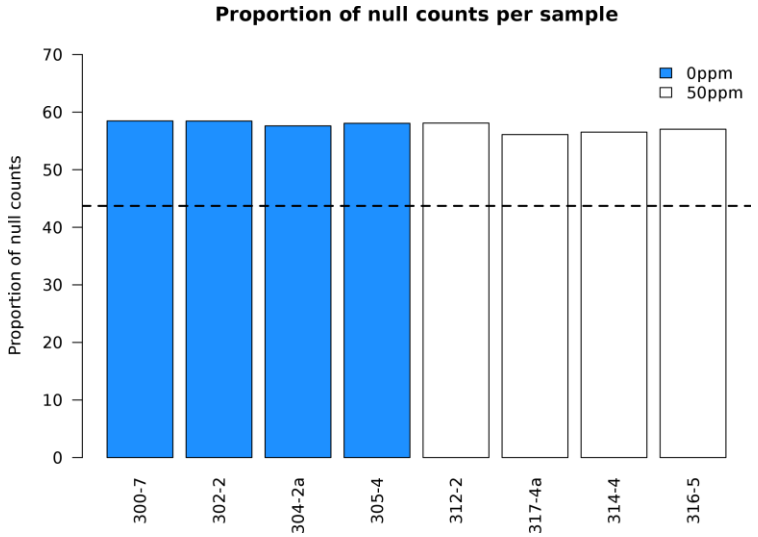


Figure 2: Proportion of features with null read counts in each sample.

Figure 3 shows the distribution of read counts for each sample. For sake of readability, $\log_2(counts + 1)$ are used instead of raw counts. Again we expect replicates to have similar distributions. In addition, this figure shows if read counts are preferably low, medium or high. This depends on the organisms as well as the biological conditions under consideration.

Supplementary file A2.1. DESeq2.pdf continued

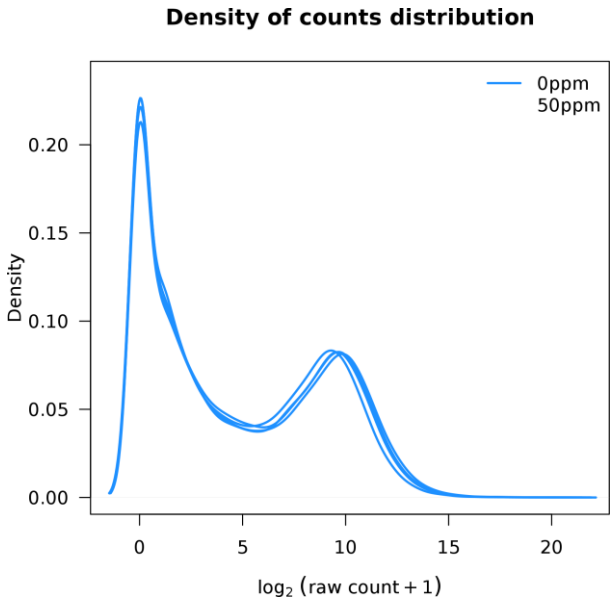


Figure 3: Density distribution of read counts.

It may happen that one or a few features capture a high proportion of reads (up to 20% or more). This phenomenon should not influence the normalization process. The DESeq2 normalization has proved to be robust to this situation [Dillies, 2012]. Anyway, we expect these high count features to be the same across replicates. They are not necessarily the same across conditions. Figure 4 illustrate the possible presence of such high count features in the data set.

Supplementary file A2.1. DESeq2.pdf continued

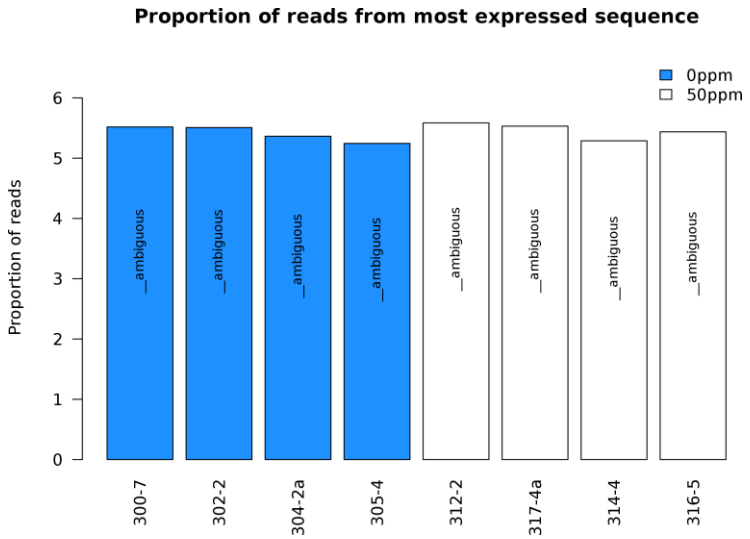


Figure 4: Percentage of reads associated with the sequence having the highest count (provided in each box on the graph) for each sample.

We may wish to assess the similarity between samples across conditions. A pairwise scatter plot is produced (figure 5) to show how replicates and samples from different biological conditions are similar or different ($\log_2(\text{counts} + 1)$ are used instead of raw count values). Moreover, as the Pearson correlation has been shown not to be relevant to measure the similarity between replicates, the SERE statistic has been proposed as a similarity index between RNA-Seq samples [Schulze, 2012]. It measures whether the variability between samples is random Poisson variability or higher. Pairwise SERE values are printed in the lower triangle of the pairwise scatter plot. The value of the SERE statistic is:

- 0 when samples are identical (no variability at all: this may happen in the case of a sample duplication);
- 1 for technical replicates (technical variability follows a Poisson distribution);
- greater than 1 for biological replicates and samples from different biological conditions (biological variability is higher than technical one, data are over-dispersed with respect to Poisson). The higher the SERE value, the lower the similarity. It is expected to be lower between biological replicates than between samples of different biological conditions. Hence, the SERE statistic can be used to detect inversions between samples.

Supplementary file A2.1. DESeq2.pdf continued

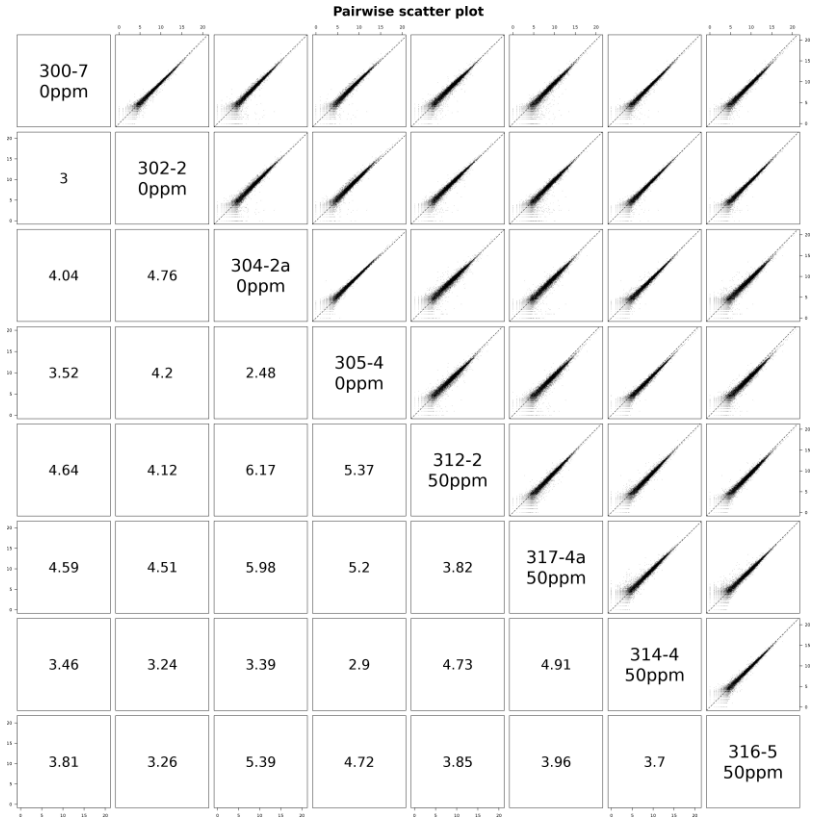


Figure 5: Pairwise comparison of samples.

3 Variability within the experiment: data exploration

The main variability within the experiment is expected to come from biological differences between the samples. This can be checked in two ways. The first one is to perform a hierarchical clustering of the whole sample set. This is performed after a transformation of the count data which can be either a Variance Stabilizing Transformation (VST) or a regularized log transformation (rlog) [Anders, 2010 and Love, 2014].

A VST is a transformation of the data that makes them homoscedastic, meaning that the variance is then independent of the mean. It is performed in two steps: (i) a mean-variance relationship is estimated from the data with the same function that is used to normalize count data and (ii) from this relationship, a transformation of the data is performed in order to get a dataset in which the variance is independent of the mean. The homoscedasticity is a prerequisite for the use of some data analysis methods, such as hierarchical clustering or Principal Component Analysis (PCA). The regularized log transformation is based on a GLM (Generalized Linear Model) on the counts and has the same goal as a VST but is more robust in the case when the size factors vary widely.

Figure 6 shows the dendrogram obtained from VST-transformed data. An euclidean distance is computed between samples, and the dendrogram is built upon the Ward criterion. We expect this dendrogram to group replicates and separate biological conditions.

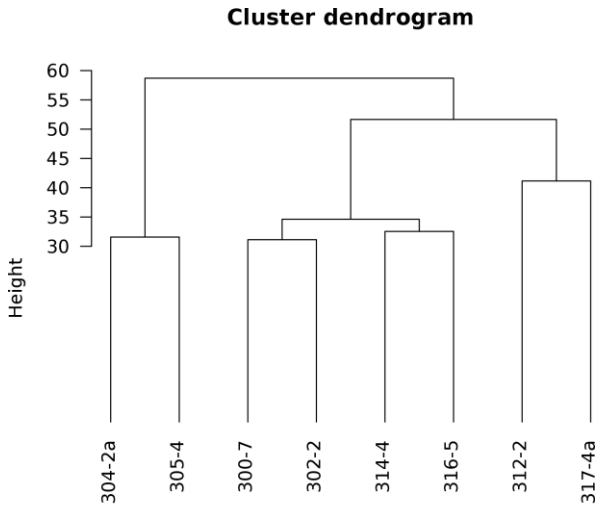


Figure 6: Sample clustering based on normalized data.

Another way of visualizing the experiment variability is to look at the first principal components of the PCA, as shown on the figure 7. On this figure, the first principal component (PC1) is expected to separate samples from the different biological conditions, meaning that the biological variability is the main source of variance in the data.

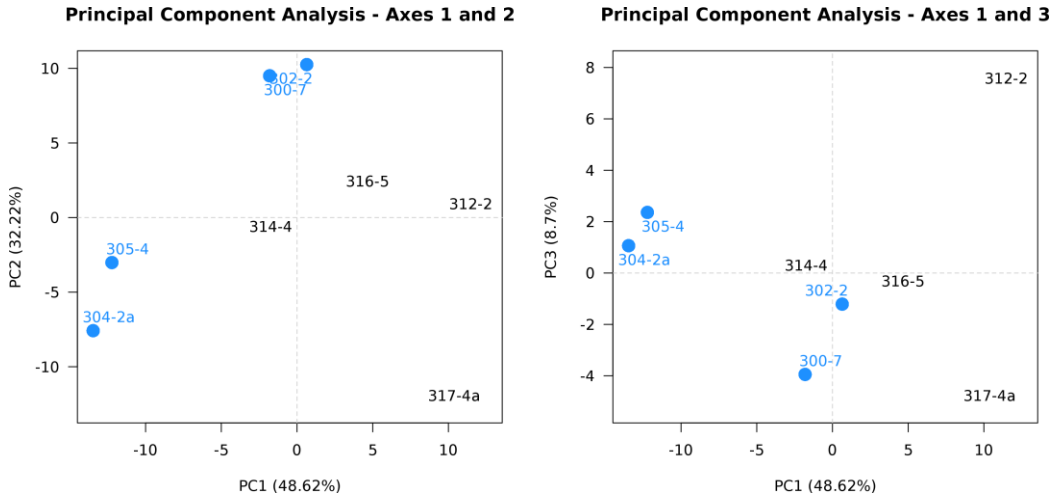


Figure 7: First two components of a Principal Component Analysis, with percentages of variance associated with each axis.

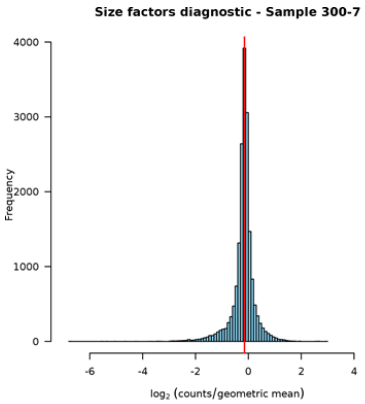
4 Normalization

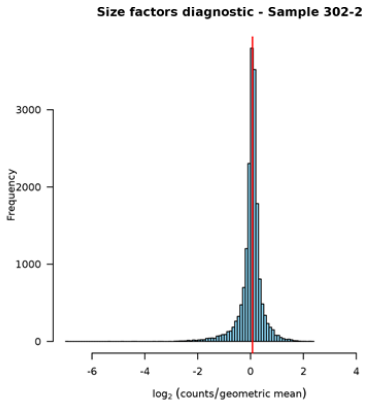
Normalization aims at correcting systematic technical biases in the data, in order to make read counts comparable across samples. The normalization proposed by DESeq2 relies on the hypothesis that most features are not differentially expressed. It computes a scaling factor for each sample. Normalized read counts are obtained by dividing raw read counts by the scaling factor associated with the sample they belong to. Scaling factors around 1 mean (almost) no normalization is performed. Scaling factors lower than 1 will produce normalized counts higher than raw ones, and the other way around. Two options are available to compute scaling factors: locfunc="median" (default) or locfunc="shorth". Here, the normalization was performed with locfunc="median".

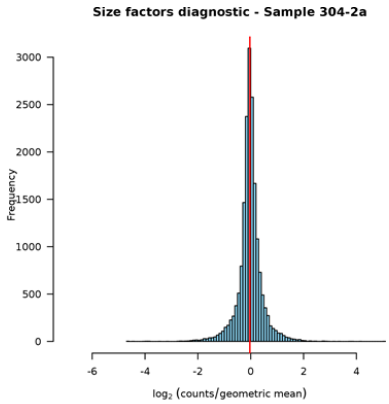
	300-7	302-2	304-2a	305-4	312-2	317-4a	314-4	316-5
Size factor	0.90	1.05	0.98	0.76	0.89	1.16	1.14	1.28

Table 5: Normalization factors.

The histograms (figure 8) can help to validate the choice of the normalization parameter ("median" or "shorth"). Under the hypothesis that most features are not differentially expressed, each size factor represented by a red line is expected to be close to the mode of the distribution of the counts divided by their geometric means across samples.







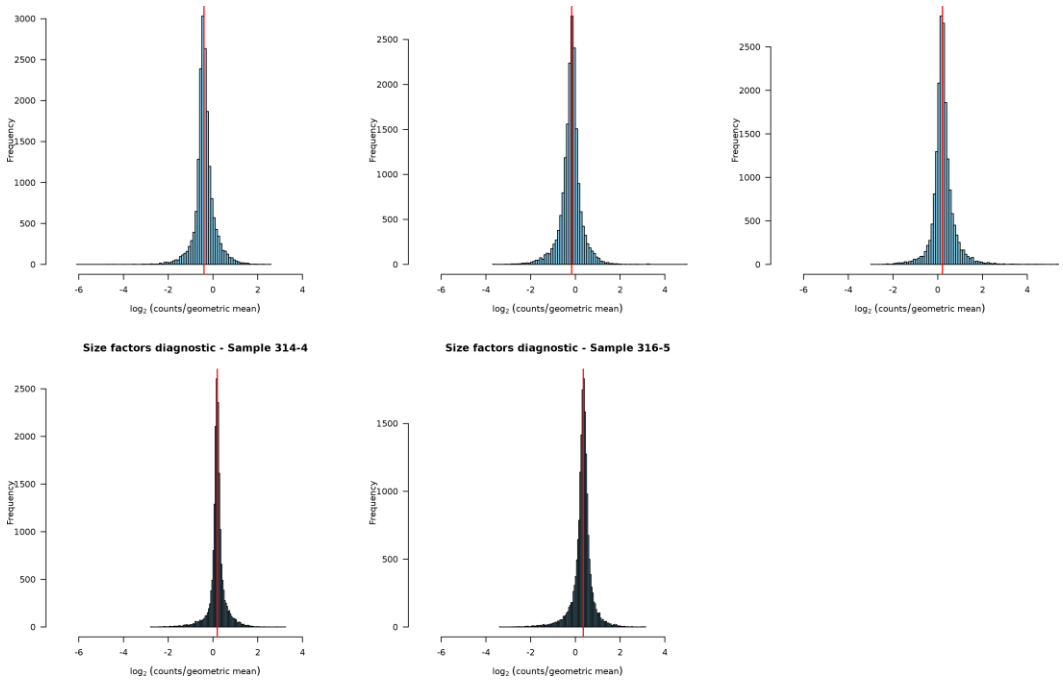


Figure 8: Diagnostic of the estimation of the size factors.

The figure 9 shows that the scaling factors of DESeq2 and the total count normalization factors may not perform similarly.

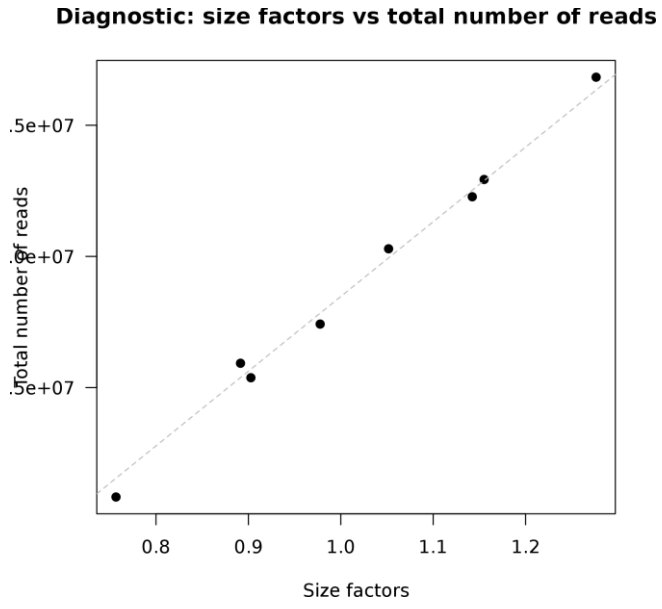


Figure 9: Plot of the estimated size factors and the total number of reads per sample.

Boxplots are often used as a qualitative measure of the quality of the normalization process, as they show how distributions are globally affected during this process. We expect normalization to stabilize distributions across samples. Figure 10 shows boxplots of raw (left) and normalized (right) data respectively.

Supplementary file A2.1. DESeq2.pdf continued

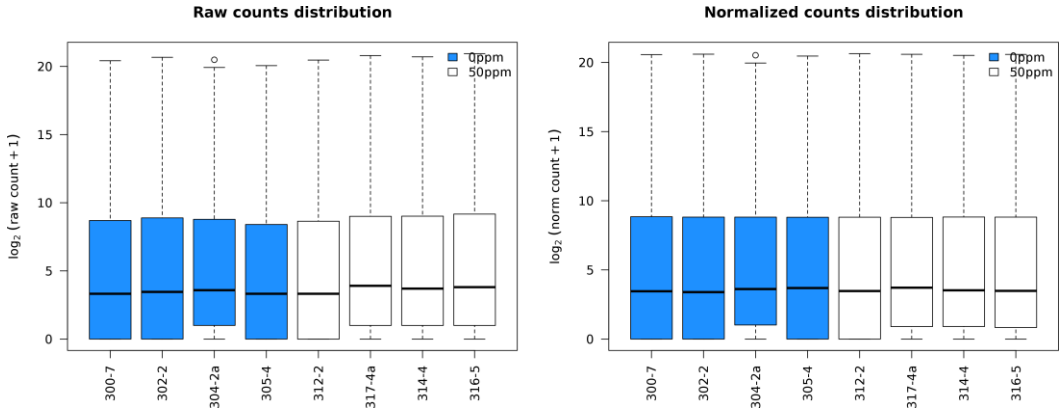


Figure 10: Boxplots of raw (left) and normalized (right) read counts.

5 Differential analysis

5.1 Modelisation

DESeq2 aims at fitting one linear model per feature. For this project, the design used is counts ~ group and the goal is to estimate the models' coefficients which can be interpreted as $\log_2(FC)$. These coefficients will then be tested to get p-values and adjusted p-values.

5.2 Outlier detection

Model outliers are features for which at least one sample seems unrelated to the experimental or study design. For every feature and for every sample, the Cook's distance [Cook, 1977] reflects how the sample matches the model. A large value of the Cook's distance indicates an outlier count and p-values are not computed for the corresponding feature.

5.3 Dispersions estimation

The DESeq2 model assumes that the count data follow a negative binomial distribution which is a robust alternative to the Poisson law when data are over-dispersed (the variance is higher than the mean). The first step of the statistical procedure is to estimate the dispersion of the data. Its purpose is to determine the shape of the mean-variance relationship. The default is to apply a GLM (Generalized Linear Model) based method (fitType="parametric"), which can handle complex designs but may not converge in some cases. The alternative is to use fitType="local" as described in the original paper [Anders, 2010]. The parameter used for this project is fitType="". Then, DESeq2 imposes a Cox Reid-adjusted profile likelihood maximization [Cox, 1987 and McCarthy, 2012] and uses the maximum *a posteriori* (MAP) of the dispersion [Wu, 2013].

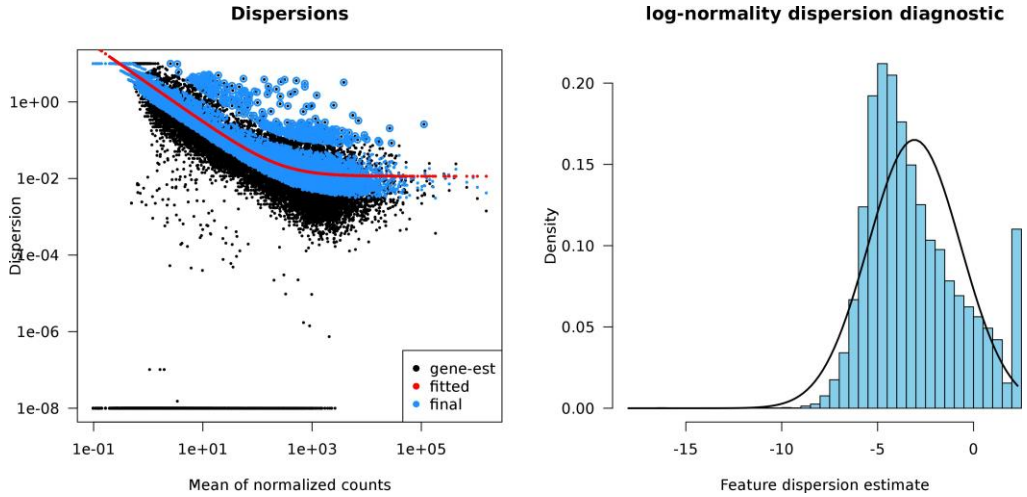


Figure 11: Dispersion estimates (left) and diagnostic of log-normality (right).

The left panel on figure 11 shows the result of the dispersion estimation step. The x- and y-axes represent the mean count value and the estimated dispersion respectively. Black dots represent empirical dispersion estimates for each feature (from the observed counts). The red dots show the mean-variance relationship function (fitted dispersion value) as estimated by the model. The blue dots are the final estimates from the maximum *a posteriori* and are used to perform the statistical test. Blue circles (if any) point out dispersion outliers. These are features with a very high empirical variance (computed from observed counts). These high dispersion values fall far from the model estimation. For these features, the statistical test is based on the empirical variance in order to be more conservative than with the MAP dispersion. These features will have low chance to be declared significant. The figure on the right panel allows to check the hypothesis of log-normality of the dispersions.

5.4 Statistical test for differential expression

Once the dispersion estimation and the model fitting have been done, DESeq2 can perform the statistical testing. Figure 12 shows the distributions of raw p-values computed by the statistical test for the comparison(s) done. This distribution is expected to be a mixture of a uniform distribution on $[0, 1]$ and a peak around 0 corresponding to the differentially expressed features.

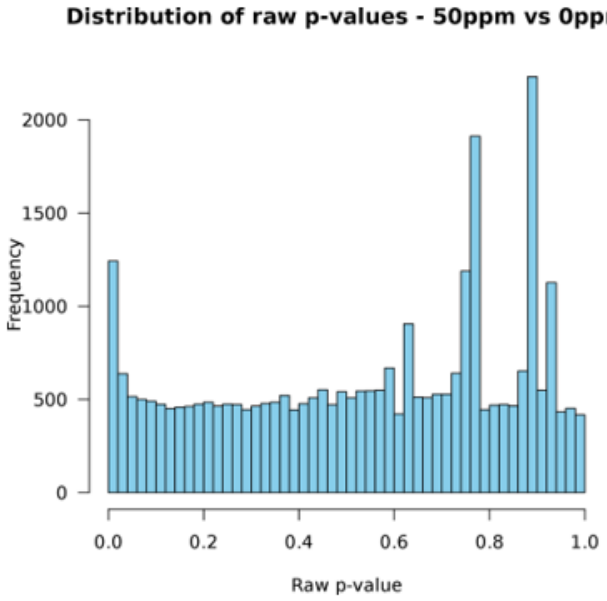


Figure 12: Distribution(s) of raw p-values.

Supplementary file A2.1. DESeq2.pdf continued

5.5 Independent filtering

DESeq2 can perform an independent filtering to increase the detection power of differentially expressed features at the same experiment-wide type I error. Since features with very low counts are not likely to see significant differences typically due to high dispersion, it defines a threshold on the mean of the normalized counts irrespective of the biological condition. This procedure is independent because the information about the variables in the design formula is not used [Love, 2014].

Table 6 reports the thresholds used for each comparison and the number of features discarded by the independent filtering. Adjusted p-values of discarded features are then set to NA.

Test vs Ref	Threshold	# discarded
50ppm vs 0ppm	14.5339821	

Table 6: Number of features discarded by the independent filtering for each comparison.

5.6 Final results

A p-value adjustment is performed to take into account multiple testing and control the false positive rate to a chosen level α. For this analysis, a BH p-value adjustment was performed [Benjamini, 1995 and 2001] and the level of controlled false positive rate was set to 0.5.

Test vs Ref	# down	# up	# total
50ppm vs 0ppm	1494	1673	3167

Table 7: Number of up-, down- and total number of differentially expressed features for each comparison.

Figure 13 represents the MA-plot of the data for the comparisons done, where differentially expressed features are highlighted in red. A MA-plot represents the log ratio of differential expression as a function of the mean intensity for each feature. Triangles correspond to features having a too low/high $\log_2$ (FC) to be displayed on the plot.

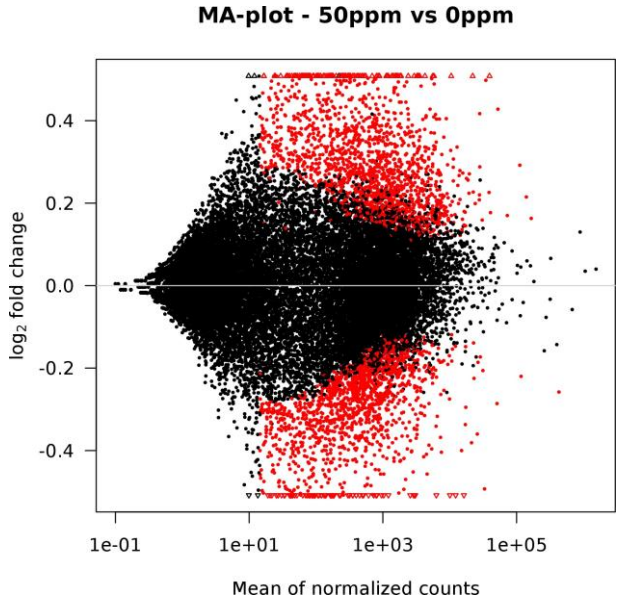


Figure 13: MA-plot(s) of each comparison. Red dots represent significantly differentially expressed features.

Supplementary file A2.1. DESeq2.pdf continued

Figure 14 shows the volcano plots for the comparisons performed and differentially expressed features are still highlighted in red. A volcano plot represents the log of the adjusted P value as a function of the log ratio of differential expression.

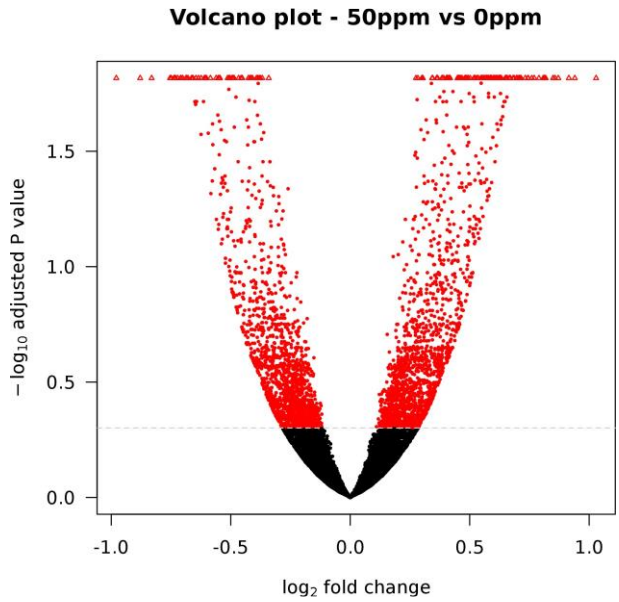


Figure 14: Volcano plot(s) of each comparison. Red dots represent significantly differentially expressed features.

Full results as well as lists of differentially expressed features are provided in the following text files which can be easily read in a spreadsheet. For each comparison:

- TestVsRef.complete.txt contains results for all the features;
- TestVsRef.up.txt contains results for significantly up-regulated features. Features are ordered from the most significant adjusted p-value to the less significant one;
- TestVsRef.down.txt contains results for significantly down-regulated features. Features are ordered from the most significant adjusted p-value to the less significant one.

These files contain the following columns:

- Id: unique feature identifier;
- sampleName: raw counts per sample;
- norm.sampleName: rounded normalized counts per sample;
- baseMean: base mean over all samples;
- 0ppm and 50ppm: means (rounded) of normalized counts of the biological conditions;
- FoldChange: fold change of expression, calculated as $2^{\log_2(FC)}$;
- log2FoldChange: $\log_2(FC)$ as estimated by the GLM model. It reflects the differential expression between Test and Ref and can be interpreted as $\log_2\left(\frac{Test}{Ref}\right)$ If this value is:

Supplementary file A2.1. DESeq2.pdf continued

6 R session information and parameters

The versions of the R software and Bioconductor packages used for this analysis are listed below. It is important to save them if one wants to re-perform the analysis in the same conditions.

- R version 3.3.1 RC (2016-06-14 r70789), x86_64-pc-linux-gnu
- Locale: LC_CTYPE=en_US.UTF-8, LC_NUMERIC=C, LC_TIME=en_US.UTF-8, LC_COLLATE=en_US.UTF-8, LC_MONETARY=en_US.UTF-8, LC_MESSAGES=en_US.UTF-8, LC_PAPER=en_US.UTF-8, LC_NAME=C, LC_ADDRESS=C, LC_TELEPHONE=C, LC_MEASUREMENT=en_US.UTF-8, LC_IDENTIFICATION=C
- Base packages: base, datasets, graphics, grDevices, methods, parallel, stats, stats4, utils
- Other packages: Biobase 2.32.0, BiocGenerics 0.18.0, DESeq2 1.12.3, devtools 1.12.0, edgeR 3.14.0, genefilter 1.54.2, GenomeInfoDb 1.8.1, GenomicRanges 1.24.2, getopt 1.20.0, IRanges 2.6.1, knitr 1.13, limma 3.28.11, RColorBrewer 1.1-2, S4Vectors 0.10.1, SARTools 1.3.2, SummarizedExperiment 1.2.3, xtable 1.8-2
- Loaded via a namespace (and not attached): acepack 1.3-3.3, annotate 1.50.0, AnnotationDbi 1.34.3, BiocParallel 1.6.2, chron 2.3-47, cluster 2.0.2, codetools 0.2-14, colorspace 1.2-6, data.table 1.9.6, DBI 0.4-1, digest 0.6.9, evaluate 0.9, foreign 0.8-65, formatR 1.4, Formula 1.2-1, geneplotter 1.50.0, ggplot2 2.1.0, grid 3.3.1, gridExtra 2.2.1, gtable 0.2.0, Hmisc 3.17-4, lattice 0.20-33, latticeExtra 0.6-28, locfit 1.5-9.1, magrittr 1.5, memoise 1.0.0, munsell 0.4.3, nnet 7.3-10, plyr 1.8.4, Rcpp 0.12.5, rpart 4.1-10, RSQLite 1.0.0, scales 0.4.0, splines 3.3.1, stringi 0.5-5, stringr 1.0.0, survival 2.38-3, tools 3.3.1, withr 1.0.2, XML 3.98-1.4, XVector 0.12.0, zlibbioc 1.18.0

Parameter values used for this analysis are:

- projectName: RNA-seq PND1 BxC female hearts
- author: Kathleen M Hudson
- targetFile: DEseq_target.txt
- rawDir: Feature_counts
- featuresToRemove: alignment_not_unique
- varInt: group
- condRef: 0ppm
- batch: NULL
- fitType:
- cooksCutoff: TRUE
- independentFiltering: TRUE

- alpha: 0.5
- pAdjustMethod: BH
- typeTrans: VST
- locfunc: median
- colors: dodgerblue

7 Bibliography

- R Core Team, **R: A Language and Environment for Statistical Computing**, *R Foundation for Statistical Computing*, 2014
- Gentleman, Carey, Bates et al, **Bioconductor: Open software development for computational biology and bioinformatics**, *Genome Biology*, 2004
- Anders and Huber, **Differential expression analysis for sequence count data**, *Genome Biology*, 2010
- Love, Huber and Anders, **Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2**, *Genome Biology*, 2014
- Dillies, Rau, Aubert et al, **A comprehensive evaluation of normalization methods for Illumina RNA-seq data analysis**, *Briefings in Bioinformatics*, 2012
- Schulze, Kanwar, Golzenleuchter et al, **SERE: Single-parameter quality control and sample comparison for RNA-Seq**, *BMC Genomics*, 2012
- Cook, **Detection of Influential Observation in Linear Regression**, *Technometrics*, 1977
- Cox and Reid, **Parameter orthogonality and approximate conditional inference**, *Journal of the Royal Statistical Society*, 1987
- McCarthy, Chen and Smyth, **Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation**, *Nucleic Acids Research*, 2012
- Wu, Wang and Wu, **A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data**, *Biostatistics*, 2013
- Benjamini and Hochberg, **Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing**, *Journal of the Royal Statistical Society*, 1995
- Benjamini and Yekutieli, **The control of the false discovery rate in multiple testing under dependency**, *The Annals of Statistics*, 2001

Supplementary Table A2.1. qRT-PCR primer list.

Gene	Ensembl ID	Primer	Sequence (5'->3')	Amplicon length
Polr2a	ENSMUSG00000005198	Forward primer	CGGTTGAATCTTAGTGTGAC	113
		Reverse primer	ATAGCCAACTCTTGGATCTC	
Igfbp3	ENSMUSG00000020427	Forward primer	CTGCTGGTGTGTGGACAAGT	158
		Reverse primer	TTGTTGGCAGTCTTTTGTGC	
Cdkn1a	ENSMUSG00000023067	Forward primer	GCAGATCCACAGCGATATCC	205
		Reverse primer	CGTGACGAAGTCAAAGTTCC	
Pfkl	ENSMUSG00000020277	Forward primer	TGCAGCCTACAATCTGCTCC	171
		Reverse primer	GTCAAGTGTGCGTAGTTCTGA	
Hif3a	ENSMUSG00000004328	Forward primer	GAAGTTCACATACTGCGACGA	75
		Reverse primer	GCAGAACAGCCAATCAGGTC	
Ldha	ENSMUSG00000063229	Forward primer	AGGCTCCCCAGAACAAGATT	238
		Reverse primer	GCGGTGATAATGACCAGCTT	
Sparcl1	ENSMUSG00000029309	Forward primer	GGCAATCCCGACAAGTACAAG	82
		Reverse primer	TGTAGCGTCTTCCGGTGTCA	
Cd44	ENSMUSG00000005087	Forward primer	ACTTTGCCTCTTGCAGTTGAG	94
		Reverse primer	TTTCTCCACATGGAATACACCTG	
Gja5	ENSMUSG00000057123	Forward primer	GGTCCACAAGCACTCCACAG	136
		Reverse primer	CTGAATGGTATCGCACCGGAA	
Nbr1	ENSMUSG00000017119	Forward primer	CAGAGTCCCAAGTCCCCTTT	236
		Reverse primer	CCCCACATGAACTTGAGCTT	
Tmem106a	ENSMUSG00000034947	Forward primer	GCTCCCCACTTTCAACACAT	245
		Reverse primer	TGGCTGGGTTATCAGGTAGG	

Supplementary Table A2.2. MIQE checklist

ITEM TO CHECK	IMPORTANCE	CHECKLIST
EXPERIMENTAL DESIGN		
Definition of experimental and control groups	**E**	Experimental groups: mice maternally exposed to 1 ppm or 50 ppm Cd. Control group: mice maternally exposed to 0 ppm Cd
Number within each group	**E**	n = 8 for each group
Assay carried out by core lab or investigator's lab?	D	Assays set up in investigator's lab and run on a shared instrument
Acknowledgement of authors' contributions	D	
SAMPLE		
Description	**E**	Experimental samples: whole hearts from mice within 24 hours of birth that were maternally exposed to 0 ppm, 1 ppm, or 50 ppm Cd
Volume/mass of sample processed	D	Entire hearts, ~9-13 mg
Microdissection or macrodissection	**E**	Macrodissection
Processing procedure	**E**	Pups were decapitated, trunk blood collected, then chest cavities opened and hearts removed.
If frozen - how and how quickly?	**E**	Hearts were removed, placed in a 1.7 ml microcentrifuge tube to obtain the weight, then immediately flash frozen by placing in a cold tube rack embedded in dry ice.
If fixed - with what, how quickly?	**E**	Not fixed
Sample storage conditions and duration (especially for FFPE samples)	**E**	Samples were maintained at -80°C for <2 months before RNA isolation
NUCLEIC ACID EXTRACTION		
Procedure and/or instrumentation	**E**	RNA was extracted using the Qiagen AllPrep DNA/RNA/miRNA kit (Qiagen, 80204) following manufacturer's protocol for 10-30mg of starting material. Whole hearts were homogenized in the kit's lysis buffer. RNA was resuspended in nuclease-free water.
Name of kit and details of any modifications	**E**	Qiagen AllPrep DNA/RNA/miRNA kit (Qiagen, 80204). No modifications to the manufacturer's protocol.
Source of additional reagents used	D	Ethanol (molecular grade, 200 proof, Fisher Scientific BP2818500); nuclease-free water (Fisher Scientific BP5611)
Details of DNase or RNAse treatment	**E**	DNase treatment included as step in Qiagen RNA extraction
Contamination assessment (DNA or RNA)	**E**	Nanodrop 2000 260/280 ratio, Agilent 2100 Bioanalyzer. Primers were designed to amplify only RNA.
Nucleic acid quantification	**E**	RNA concentration was determined using a Nanodrop 2000
Instrument and method	**E**	Nanodrop 2000
Purity (A260/A280)	D	Assessed by the Nanodrop 2000 (range: 2.04-2.09, avg: 2.05625, median: 2.05)
Yield	D	Average yield of RNA per whole heart was 25 µg
RNA integrity method/instrument	**E**	Agilent 2100 Bioanalyzer with an RNA 6000 Nano Chip
RIN/RQI or Cq of 3' and 5' transcripts	**E**	≥9.9
Electrophoresis traces	D	
Inhibition testing (Cq dilutions, spike or other)	**E**	Standard curves were used and were efficient, thus ruling out the possibility of inhibition
REVERSE TRANSCRIPTION		
Complete reaction conditions	**E**	RT reaction was performed according to manufacturer's protocol (Promega M170B)
Amount of RNA and reaction volume	**E**	100 ng of RNA in a 25 µl RT reaction
Priming oligonucleotide (if using GSP) and concentration	**E**	500 ng of random primers (Promega C118A) used per 25 µl RT reaction
Reverse transcriptase and concentration	**E**	200 units of M-MLV RT enzyme (Promega M170B)
Temperature and time	**E**	37°C for 60 minutes
Manufacturer of reagents and catalogue numbers	D	
Cqs with and without RT	D*	
Storage conditions of cDNA	D	cDNA was kept at -20°C while not in use, or at 4°C while preparing qRT-PCR reactions
qPCR TARGET INFORMATION		
If multiplex, efficiency and LOD of each assay.	**E**	N/A
Sequence accession number	**E**	See supplementary table
Location of amplicon	D	
Amplicon length	**E**	See supplementary table
In silico specificity screen (BLAST, etc)	**E**	NCBI Primer BLAST
Pseudogenes, retropseudogenes or other homologs?	D	
Sequence alignment	D	100% to both C57BL/6J and CAST/EiJ genomes
Secondary structure analysis of amplicon	D	
Location of each primer by exon or intron (if applicable)	**E**	Primers were designed to flank exon-exon junctions, or extend over an exon-exon junction when possible
What splice variants are targeted?	**E**	All (if applicable)
qPCR OLIGONUCLEOTIDES		
Primer sequences	**E**	See supplementary table
RTPrimerDB Identification Number	D	
Probe sequences	D**	
Location and identity of any modifications	**E**	No modifications
Manufacturer of oligonucleotides	D	IDT
Purification method	D	
qPCR PROTOCOL		
Complete reaction conditions	**E**	Described in methods section
Reaction volume and amount of cDNA/DNA	**E**	Described in methods section
Primer, (probe), Mg++ and dNTP concentrations	**E**	Described in methods section
Polymerase identity and concentration	**E**	Described in methods section
Buffer/kit identity and manufacturer	**E**	Described in methods section
Exact chemical constitution of the buffer	D	
Additives (SYBR Green I, DMSO, etc.)	**E**	Described in methods section
Manufacturer of plates/tubes and catalog number	D	
Complete thermocycling parameters	**E**	Described in methods section
Reaction setup (manual/robotic)	D	Manual
Manufacturer of qPCR instrument	**E**	Applied Biosystems
qPCR VALIDATION		
Evidence of optimisation (from gradients)	D	
Specificity (gel, sequence, melt, or digest)	**E**	Specificity was determined through melting curve analysis
For SYBR Green I, Cq of the NTC	**E**	All NTC values were undetermined
Standard curves with slope and y-intercept	**E**	Polr2a: -3.4617x + 33.005; Igfbp3: y = -3.4654x + 32.091; Cdkn1a: y = -3.871x + 35.954; Pfkl: y = -3.746x + 31.677; Hif3a: y = -3.5404x + 32.74; Ldha: y = -3.7358x + 28.32; Sparcl1: y = -3.1658x + 31.587; Cd44: y = -3.3034x + 33.713; Gja5: y = -3.5949x + 31.956; Nbr1: y = -3.9006x + 32.074; Tmem106a: y = -3.7572x + 38.22
PCR efficiency calculated from slope	**E**	Polr2a: 94.48%, Igfbp3: 94.34%, Cdkn1a: 81.27%, Pfkl: 84.91%, Hif3a: 91.63%, Ldha: 85.22%, Sparcl1: 106.96%, Cd44: 100.78%, Gja5: 89.75%, Nbr1: 80.46%, Tmem106a: 84.57%
Confidence interval for PCR efficiency or standard error	D	
r2 of standard curve	**E**	Polr2a: 0.984, Igfbp3: 0.9845, Cdkn1a: 0.9608, Pfkl: 0.9709, Hif3a: 0.9559, Ldha: 0.9866, Sparcl1: 0.9675, Cd44: 0.9684, Gja5: 0.9904, Nbr1: 0.9821, Tmem106a: 0.8025
Linear dynamic range	**E**	Linear dynamic range of primers was consistent with an RNA input of 0-500 ng into the RT reaction
Cq variation at lower limit	**E**	Standard curves were performed using 1/5 to 1/80 dilutions of cDNA to encompass the 1/10 dilution of cDNA tested
Confidence intervals throughout range	D	
Evidence for limit of detection	**E**	Was not performed
If multiplex, efficiency and LOD of each assay.	**E**	N/A
DATA ANALYSIS		
qPCR analysis program (source, version)	**E**	Microsoft Excel, Applied Biosystems 7300 System SDS Software
Cq method determination	**E**	Determined through the SDS software
Outlier identification and disposition	**E**	Described in methods section
Results of NTCs	**E**	Undetermined Ct values
Justification of number and choice of reference genes	**E**	Described in methods section
Description of normalisation method	**E**	Described in methods section
Number and concordance of biological replicates	D	
Number and stage (RT or qPCR) of technical replicates	**E**	3 technical replicates were used for each sample
Repeatability (Intra-assay variation)	E	Standard deviation of triplicates were <0.5
Reproducibility (inter-assay variation, %CV)	D	
Power analysis	D	
Statistical methods for result significance	**E**	Described in methods section
Software (source, version)	E	Described in methods section
Cq or raw data submission using RDML	**D**	

Table 1. MIQE checklist for authors, reviewers and editors. All essential information (E) must be submitted with the manuscript. Desirable information (D) should be submitted if available. If using primers obtained from RTPrimerDB, information on qPCR target, oligonucleotides, protocols and validation is available from that source.

*: Assessing the absence of DNA using a no RT assay is essential when first extracting RNA. Once the sample has been validated as RDNA-free, inclusion of a no-RT control is desirable, but no longer essential.

**: Disclosure of the probe sequence is highly desirable and strongly encouraged. However, since not all commercial pre-designed assay vendors provide this information, it cannot be an essential requirement. Use of such assays is advised against.

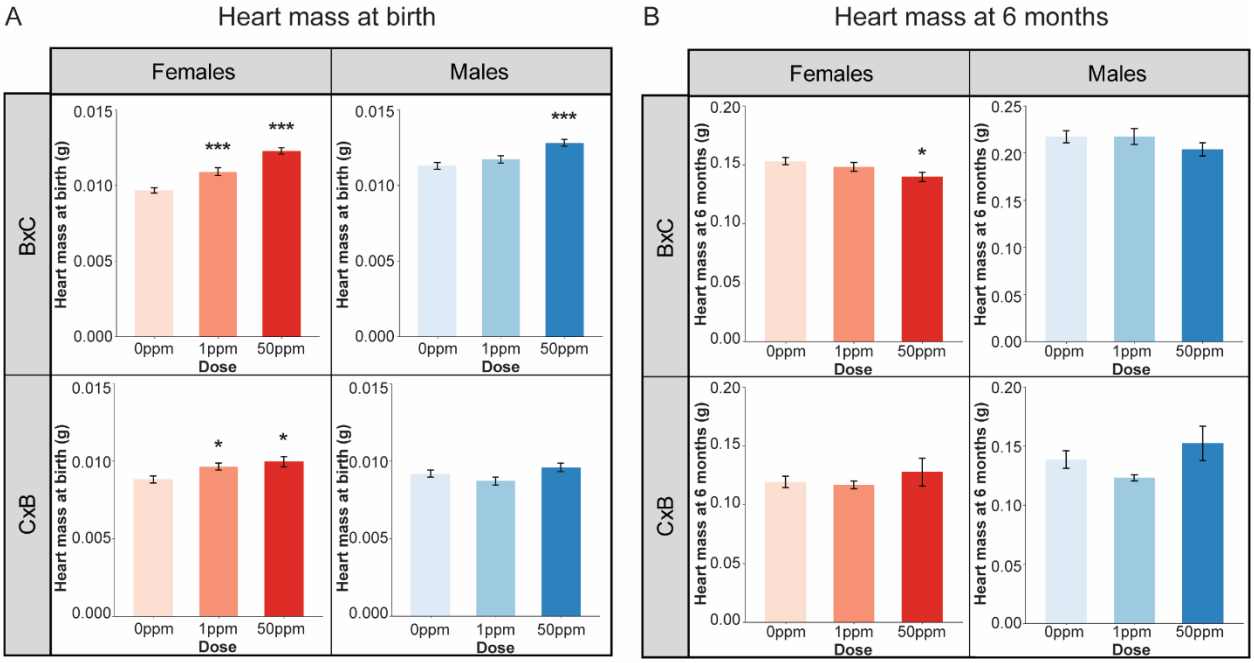
Supplementary table A2.3. Comprehensive RNA-seq data

This table is provided as an additional document that is downloadable on the repository page for this ETD.

Supplementary table A2.4. RNA-seq enrichment tables

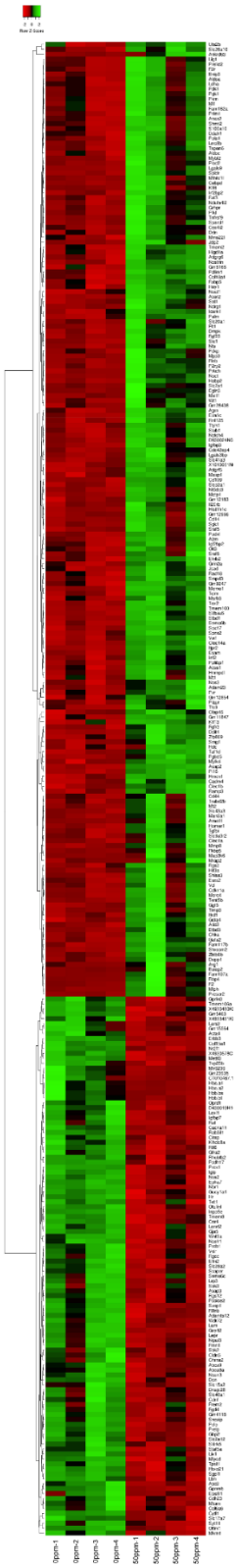
These tables are provided as an additional document that is downloadable on the repository page for this ETD.

Supplementary figure A2.1.



Supplementary figure A2.1. Raw heart mass at birth and 6 months of age following maternal Cd exposure. **A,** Heart mass at birth. **B,** Heart mass at 6 months of age. *p<0.05, **p<0.01, ***p<0.0001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Supplementary figure A2.2



Supplementary figure A2.2. Heatmap of 302 significantly differentially expressed genes in newborn BxC female whole hearts as a result of 50 ppm maternal Cd exposure.

Appendix B

Chapter 3 Supplemental Data

Supplementary table B3.1. Comprehensive RNA-seq data

This table is provided as an additional document that is downloadable on the repository page for this ETD.

Supplementary table B3.2. qRT-PCR and qPCR primer sequences.

Gene	Ensembl or NCBI ID	Primer	Sequence (5'->3')	Amplicon length
Polr2a	ENSMUSG00000005198	Forward primer	CGGTTGAATCTTAGTGTGAC	113
		Reverse primer	ATAGCCAACTCTTGGATCTC	
Mag	ENSMUSG00000036634	Forward primer	GGTACATGGCGTCTGGTATTTC	78
		Reverse primer	ACTTGTGTGCGGGACTTGAAG	
Enpp6	ENSMUSG00000038173	Forward primer	TTGTGAACAGAGGCGTCAAAG	118
		Reverse primer	TTGCCGATCATCTGGTGGAC	
Mbp	ENSMUSG00000041607	Forward primer	TCACAGCGATCCAAGTACCTG	125
		Reverse primer	CCCCTGTCACCGCTAAAGAA	
Il33	ENSMUSG00000024810	Forward primer	TCCAACTCCAAGATTTCCCCG	120
		Reverse primer	CATGCAGTAGACATGGCAGAA	
Aldh1a2	ENSMUSG00000013584	Forward primer	CAGAGAGTGGGAGAGTGTTCC	136
		Reverse primer	CACACAGAACCAAGAGAGAAGG	
Sparcl1	ENSMUSG00000029309	Forward primer	GGCAATCCCGACAAGTACAAG	82
		Reverse primer	TGTAGCGTCTTCCGGTGTCA	
Kank1	ENSMUSG00000032702	Forward primer	ACCACAAAAGTTAATGGCTGTGC	96
		Reverse primer	GGTTTCCACAAAATACGGGTCTT	
Dlx1	ENSMUSG00000041911	Forward primer	ATGCCAGAAAGTCTCAACAGC	118
		Reverse primer	AACAGTGCATGGAGTAGTGCC	
Sp9	ENSMUSG00000068859	Forward primer	CGTCTATACTTGGGGAAGAG	82
		Reverse primer	GTGTTGCCGATCTTGTTG	
Hoxa2	ENSMUSG00000014704	Forward primer	TACGAATTTGAGCGAGAGATTGG	116
		Reverse primer	GTCGAGGTCTTGATTGATGAACT	
Hoxa3	ENSMUSG00000079560	Forward primer	ATGCAAAAAGCGACCTACTACG	80
		Reverse primer	TTGTAAGCGAACCCATTGGCT	
Hoxa4	ENSMUSG00000000942	Forward primer	CGGTGGTGTACCCCTGGAT	75
		Reverse primer	GCTTAGGTTCGCCTCCGTTAT	
Hoxa5	ENSMUSG00000038253	Forward primer	CTCATTTTGCGGTCGCTATCC	139
		Reverse primer	ATCCATGCCATTGTAGCCGTA	

Supplementary table B3.2 Continued

Hoxa6	ENSMUSG00000043219	Forward primer	TTACCAACAGTCCAACTCGGT	92
		Reverse primer	CGCCACTGAGGTCTTTATCAGA	
Hoxb2	ENSMUSG00000075588	Forward primer	GCCTACACCAACACGCAACT	65
		Reverse primer	CGGCACAGGTACTTATTGAAGTG	
Hoxb5	ENSMUSG00000038700	Forward primer	CCTGCACTAACGGCGACAG	108
		Reverse primer	TGGCCTCGTCTATTTCGGTGA	
Hoxb6	ENSMUSG00000000690	Forward primer	TCCTATTTCGTGAACTCCACCT	92
		Reverse primer	GCATAGCCAGACGAGTAGAGC	
Hoxb7	ENSMUSG00000038721	Forward primer	GCCGCAAGTTCGGTTTTCG	59
		Reverse primer	GCAAAGGCGCAAGAAGTTTGT	
Hoxb8	ENSMUSG00000056648	Forward primer	CCTGCGCCCCAATTATTATGA	117
		Reverse primer	AACTCCTGGATTTGCGAAGGG	
mtDNA	NCBI ID: 17724	Forward primer	CCCAGCTACTACCATCATTCAAGT	117
		Reverse primer	GATGGTTTGGGAGATTGGTTGATGT	
Hk2 (gDNA)	NCBI ID: 15277	Forward primer	GCCAGCCTCTCCTGATTTTAGTGT	116
		Reverse primer	GGGAACACAAAAGACCTCTTCTGG	

Supplementary table B3.3. MIQE checklist.

ITEM TO CHECK	IMPORTANCE	CHECKLIST
EXPERIMENTAL DESIGN		
Definition of experimental and control groups	**E**	Experimental groups: mice maternally exposed to 1 ppm or 50 ppm Cd. Control group: mice maternally exposed to 0 ppm Cd
Number within each group	**E**	n = 8 for each group
Assay carried out by core lab or investigator's lab?	D	Assays set up in investigator's lab and run on a shared instrument
Acknowledgement of authors' contributions	D	
SAMPLE		
Description	**E**	Experimental samples: whole brains from mice within 24 hours of birth that were maternally exposed to 0 ppm, 1 ppm, or 50 ppm Cd
Volume/mass of sample processed	D	Fraction of frozen homogenized brains, ~7-10 mg
Microdissection or macrodissection	**E**	Macrodissection
Processing procedure	**E**	Pups were decapitated, trunk blood collected, then dissected.
If frozen - how and how quickly?	**E**	Brains were removed from decapitated head, placed in a 1.7 ml microcentrifuge tube to obtain the weight, then immediately flash frozen by placing in a cold tube rack embedded in dry ice.
If fixed - with what, how quickly?	**E**	Not fixed
Sample storage conditions and duration (especially for FFPE samples)	**E**	Samples were maintained at -80°C for several months before RNA isolation
NUCLEIC ACID EXTRACTION		
Procedure and/or instrumentation	**E**	RNA was extracted using the Qiagen AllPrep DNA/RNA/miRNA kit (Qiagen, 80204) following manufacturer's protocol for fatty tissue. Whole brains were pulverized under liquid N temps, then 7-10mg was added to the kit's lysis buffer. RNA was resuspended in nuclease-free water.
Name of kit and details of any modifications	**E**	Qiagen AllPrep DNA/RNA/miRNA kit (Qiagen, 80204). No modifications to the manufacturer's protocol.
Source of additional reagents used	D	Ethanol (molecular grade, 200 proof, Fisher Scientific BP2818500); nuclease-free water (Fisher Scientific BP5611)
Details of DNase or RNAse treatment	**E**	DNase treatment included as step in Qiagen RNA extraction
Contamination assessment (DNA or RNA)	**E**	Nanodrop 2000 260/280 ratio, Agilent 2100 Bioanalyzer.
Nucleic acid quantification	**E**	RNA concentration was determined using a Nanodrop 2000
Instrument and method	**E**	Nanodrop 2000
Purity (A260/A280)	D	Assessed by the Nanodrop 2000 (range: 2.05-2.13, avg: 2.092, median: 2.09)
Yield	D	Average yield of RNA per 7-10mg of whole brain was 18.9 µg
RNA integrity method/instrument	**E**	Agilent 2100 Bioanalyzer with an RNA 6000 Nano Chip
RIN/RQI or Cq of 3' and 5' transcripts	**E**	≥9.9
Electrophoresis traces	D	
Inhibition testing (Cq dilutions, spike or other)	**E**	Standard curves were used and were efficient, thus ruling out the possibility of inhibition
REVERSE TRANSCRIPTION		
Complete reaction conditions	**E**	RT reaction was performed according to manufacturer's protocol (Promega M170B)
Amount of RNA and reaction volume	**E**	100 ng of RNA in a 25 µl RT reaction
Priming oligonucleotide (if using GSP) and concentration	**E**	500 ng of random primers (Promega C118A) used per 25 µl RT reaction
Reverse transcriptase and concentration	**E**	200 units of M-MLV RT enzyme (Promega M170B)
Temperature and time	**E**	37°C for 60 minutes
Manufacturer of reagents and catalogue numbers	D	
Cqs with and without RT	D*	
Storage conditions of cDNA	D	cDNA was kept at -20°C while not in use, or at 4°C while preparing qRT-PCR reactions
qPCR TARGET INFORMATION		
If multiplex, efficiency and LOD of each assay.	**E**	N/A
Sequence accession number	**E**	See supplementary table
Location of amplicon	D	
Amplicon length	**E**	See supplementary table
In silico specificity screen (BLAST, etc)	**E**	NCBI Primer BLAST
Pseudogenes, retropseudogenes or other homologs?	D	
Sequence alignment	D	100% to both C57BL/6J and CAST/EiJ genomes
Secondary structure analysis of amplicon	D	
Location of each primer by exon or intron (if applicable)	**E**	Primers were designed to flank exon-exon junctions, or extend over an exon-exon junction when possible
What splice variants are targeted?	**E**	All (if applicable)
qPCR OLIGONUCLEOTIDES		
Primer sequences	**E**	See supplementary table
RTPrimerDB Identification Number	D	
Probe sequences	D**	
Location and identity of any modifications	**E**	No modifications
Manufacturer of oligonucleotides	D	IDT
Purification method	D	
qPCR PROTOCOL		
Complete reaction conditions	**E**	Described in methods section
Reaction volume and amount of cDNA/DNA	**E**	Described in methods section
Primer, (probe), Mg++ and dNTP concentrations	**E**	Described in methods section
Polymerase identity and concentration	**E**	Described in methods section
Buffer/kit identity and manufacturer	**E**	Described in methods section
Exact chemical constitution of the buffer	D	
Additives (SYBR Green I, DMSO, etc.)	**E**	Described in methods section
Manufacturer of plates/tubes and catalog number	D	
Complete thermocycling parameters	**E**	Described in methods section
Reaction setup (manual/robotic)	D	Manual
Manufacturer of qPCR instrument	**E**	Applied Biosystems
qPCR VALIDATION		
Evidence of optimisation (from gradients)	D	
Specificity (gel, sequence, melt, or digest)	**E**	Specificity was determined through melting curve analysis
For SYBR Green I, Cq of the NTC	**E**	All NTC values were undetermined
Standard curves with slope and y-intercept	**E**	See next tab
PCR efficiency calculated from slope	**E**	See next tab
Confidence interval for PCR efficiency or standard error	D	
r2 of standard curve	**E**	See next tab
Linear dynamic range	**E**	Linear dynamic range of primers was consistent with an RNA input of 0-500 ng into the RT reaction
Cq variation at lower limit	**E**	Standard curves were performed using 1/5 to 1/80 dilutions of cDNA to encompass the 1/10 dilution of cDNA tested
Confidence intervals throughout range	D	
Evidence for limit of detection	**E**	Was not performed
If multiplex, efficiency and LOD of each assay.	**E**	N/A
DATA ANALYSIS		
qPCR analysis program (source, version)	**E**	Microsoft Excel, Applied Biosystems 7300 System SDS Software
Cq method determination	**E**	Determined through the SDS software
Outlier identification and disposition	**E**	Described in methods section
Results of NTCs	**E**	Undetermined Ct values
Justification of number and choice of reference genes	**E**	Described in methods section
Description of normalisation method	**E**	Described in methods section
Number and concordance of biological replicates	D	
Number and stage (RT or qPCR) of technical replicates	**E**	3 technical replicates were used for each sample
Repeatability (intra-assay variation)	E	Standard deviation of triplicates were <0.5
Reproducibility (inter-assay variation, %CV)	D	
Power analysis	D	
Statistical methods for result significance	**E**	Described in methods section
Software (source, version)	E	Described in methods section
Cq or raw data submission using RDML	**D**	

Table 1. MIQE checklist for authors, reviewers and editors. All essential information (E) must be submitted with the manuscript. Desirable information (D) should be submitted if available. If using primers obtained from RTPrimerDB, information on qPCR target, oligonucleotides, protocols and validation is available from that source.

*: Assessing the absence of DNA using a no RT assay is essential when first extracting RNA. Once the sample has been validated as RDNA-free, inclusion of a no-RT control is desirable, but no longer essential.

**: Disclosure of the probe sequence is highly desirable and strongly encouraged. However, since not all commercial pre-designed assay vendors provide this information, it cannot be an essential requirement. Use of such assays is advised against.

Supplementary table B3.4. Proteomic enrichment analysis.

This table is provided as an additional document that is downloadable on the repository page for this ETD.

Supplementary table B3.5 SI.

Social interaction behavioral testing testing. *p<0.05, *p<0.01, *p<0.001 (one way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm). Significant p values are in bold.

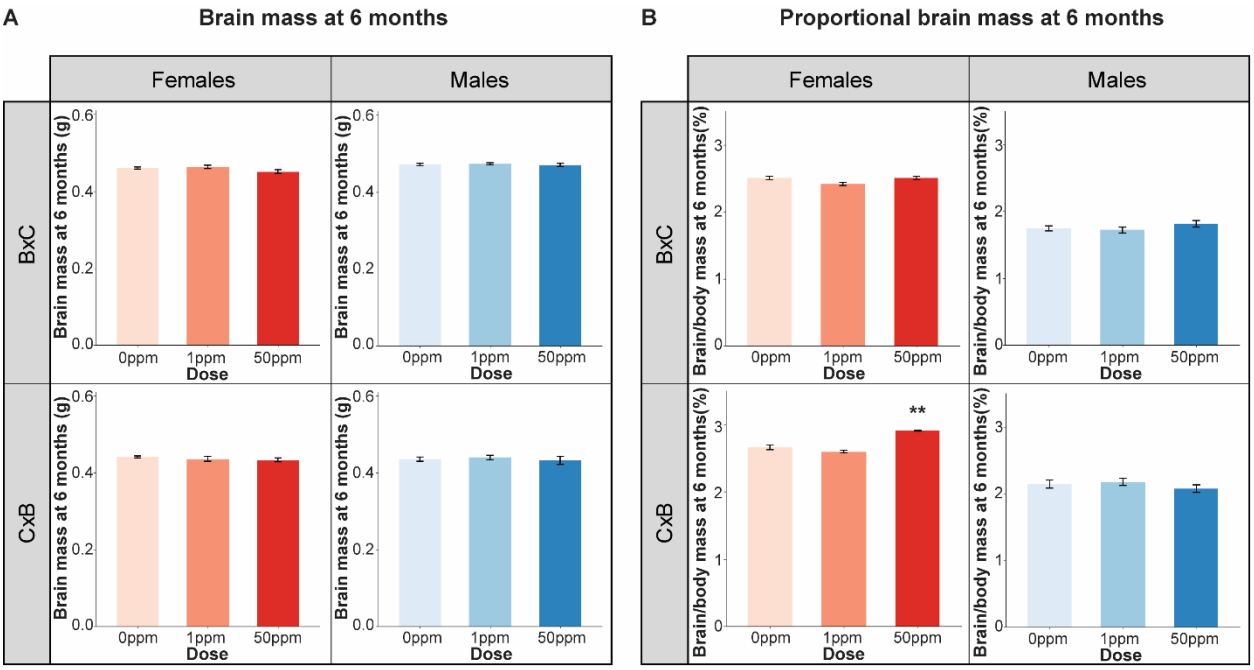
Cross	Sex	Dose	Sample size (n)	Area:Mouse 1 Center In Novel Animal					Mouse 1 Body In Novel Animal Over 50.0%					nuse 1 distance to point 1 Less Than 45.00 n					Mouse 1 distance to point 1 Less Than 55.00 mm					Mouse body distance to area Novel Animal <10.00 mm					Mouse body distance to area Novel Animal <20.00 mm				
				Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)
AVG																																
BxC	F	0ppm	13	0.61538	2.77692	710.44	739.84	5.44769	0.30769	2.09154	713.28	743.34	2.70538	3.30769	14.1992	463.277	497.153	45.1462	7	28.0138	788.3766667	826.56	4.904	6.53846	14.5867	320.474	360.816	154.73	7.69231	20.61	200.224	240.177	248.022
BxC	F	1ppm	9	3.11111	14.8589	514.9	544.825	57.1356	2.22222	13.9367	514.815	544.745	53.9967	11.7778	56.9978	442.768	472.634	492.013	25.2222	111.368	419.8071429	453.8014286	598.739	10	53.6311	68.17	101.513	754.344	9.55556	61.3567	68.0467	101.59	978.798
BxC	F	50ppm	7	4.57143	31.5371	528.66	573.697	264.621	2.57143	30.1414	502.62	548.07	169.72	12.2857	31.9471	398.9	440.14	264.169	25.1429	74.0314	327.452	366.98	603.073	27.8571	163.523	75.3217	119.452	1773.65	29	194.597	59.37	103.93	2468.52
SE																																
BxC	F	0ppm		0.61538	2.77692	N/A	N/A	5.44769	0.30769	2.09154	N/A	N/A	2.70538	1.78808	11.8965	151.559	147.469	38.2442	3.31082	17.5697	229.4197375	230.0265584	2.27942	2.1739	5.65864	117.924	120.762	57.1392	2.5879	7.60591	83.0965	86.1288	88.0637
BxC	F	1ppm		2.75603	13.1539	484.6	486.335	48.4157	1.87659	12.2785	484.715	486.315	45.6457	5.33536	29.8466	156.08	154.773	307.468	10.073	50.5612	254.993319	257.0546328	328.495	7.56637	42.1373	54.2917	52.9584	623.061	7.33354	47.0526	54.2858	52.9567	807.183
BxC	F	50ppm		2.84402	23.1348	286.341	285.391	153.643	1.71627	23.0859	501.82	495.51	141.058	8.07617	25.8599	398.9	396.7	173.778	11.1493	45.0946	292.9179434	294.1880621	343.782	6.91248	68.06	48.4247	48.2089	476.314	6.733	73.1603	49.389	49.953	632.709
ANOVA w/dunnetts		1vs0		0.554	0.691	0.934	0.934	0.808	0.44	0.685	0.957	0.956	0.781	0.315	0.261	0.996	0.994	0.145	0.153	0.169	0.477	0.473	0.161	0.8495	0.6714	0.231	0.227	0.4587	0.953	0.696	0.462	0.45	0.5046
		50vs0		0.301	0.2	0.936	0.946	**0.0282***	0.38	0.196	0.952	0.958	0.144	0.328	0.803	0.973	0.979	0.638	0.197	0.598	0.396	0.401	0.178	**0.0183***	**0.0213***	0.135	0.151	**0.0202***	**0.018***	**0.014***	0.281	0.32	**0.0147***
AVG																																
BxC	M	0ppm	10	5.6	13.256	400.8	443.06	74.94	2.4	9.366	601.905	644.68	63.099	13.375	18.3138	329.483	370.838	149.628	57.1	172.979	118.7016667	167.4883333	1849.21	12.6	64.995	103.338	144.513	16.2475	0.625	2.085	103.213	144.58	26.53
BxC	M	1ppm	7	55.1429	139.967	153.638	183.658	2949.74	25.8571	109.163	158.983	189.573	2264.92	68.8571	87.95	249.418	279.055	1716.19	75.4286	238.494	135.9	168.8116667	2330.87	41	341.917	82.602	118.324	7644.67	40.8571	363.177	81.956	118.38	8528.11
BxC	M	50ppm	7	9.5	43.2714	211.064	271.886	6342.72	5.28571	18.6329	568.5	630.19	1665.3	37.8571	73.2486	45.3875	104.205	523.851	32.1429	133.58	14.505	76.0075	1394.66	44.4286	217.816	33.34	93.6533	9628.97	45.1429	250.633	33.2217	94.105	11146
SE																																
BxC	M	0ppm		4.68615	12.4587	219.15	218.03	67.5446	1.99555	8.94554	420.255	419.72	57.4571	5.86435	8.86962	146.37	146.299	66.601	19.2766	64.3177	61.77170981	60.46635651	731.083	8.20867	44.2804	26.3828	26.0434	10.981	0.41993	1.38069	26.3771	26.0551	17.8002
BxC	M	1ppm		20.913	55.7616	103.028	99.9316	2145.04	9.18443	45.0834	102.345	99.0961	1704.18	27.968	37.8893	101.476	98.4487	1276.95	26.5535	86.3096	45.06438905	45.16221241	1077.93	14.3162	124.366	50.3788	51.2763	3250.17	14.3401	131.386	50.3694	51.2924	3465.55
BxC	M	50ppm		5.56028	24.6094	142.244	143.053	4092.48	2.6069	10.265	246.118	250.011	1062.83	23.5346	46.0174	17.867	10.9356	337.194	15.9261	76.8513	8.847534967	14.07968358	853.529	12.3208	75.3349	15.8818	11.6818	4725.91	11.5932	83.474	15.8262	11.5814	5083.48
ANOVA w/dunnetts		1vs0		**0.0113***	**0.0179***	0.488	0.458	0.593	**0.00557****	**0.0136***	0.384	0.372	0.229	0.12	0.257	0.862	0.822	0.24	0.766	0.764	0.955	1	0.901	0.149	**0.0393***	0.878	0.811	0.1751	**0.0199***	**0.0135***	0.872	0.811	0.1557
		50vs0		0.9637	0.7348	0.616	0.666	0.118	0.88838	0.9482	0.993	0.999	0.436	0.612	0.411	0.217	0.249	0.909	0.618	0.906	0.317	0.395	0.903	0.0989	0.3122	0.277	0.467	0.0766	**0.0103***	0.0933	0.277	0.472	0.0546
AVG																																
CxB	F	0ppm	10	8.44444	30.9689	14.182	56.38	238.682	5.88889	27.6433	104.944	148.5	161.879	13.4444	33.053	423.346	457.217	164.936	41.4	138.467	160.766	202.03	980.429	28.2	211.587	2.14167	43.5317	2906.24	28.2	231.986	2.1	43.7	3451.46
CxB	F	1ppm	5	0	0	N/A	N/A	0	0	0	N/A	N/A	0	1.6	0.376	982.12	1013.24	0.242	4.8	7.582	971.3	1013.295	15.0225	10.6	34.478	151.96	201.753	366.652	10.2	48.006	109.31	160.887	537.232
CxB	F	50ppm	4	11.75	38.94	663.12	746.003	239.318	6.75	36.3375	882.35	932.18	245.178	12.5	26.41	1266.27	931.523	83.5875	21	61.71	197.9633333	279.4666667	419.468	21	116.015	111.91	166.5	1488.48	21.5	134.2	111.547	166.923	1844.42
SE																																
CxB	F	0ppm		3.23226	12.0388	12.2565	17.4844	115.039	2.32406	11.1681	55.9964	55.6396	80.8867	5.89282	12.0656	233.596	232.894	68.7937	11.0707	48.2692	72.91399771	71.8424942	348.593	7.86243	74.64	0.87793	4.65477	1113.59	7.7943	79.7232	0.85927	4.76395	1298.06
CxB	F	1ppm		0	0	N/A	N/A	0	0	0	N/A	N/A	0	1.6	0.376	N/A	N/A	0.242	2.00998	3.42958	36.26258885	41.73753157	6.67127	4.55631	15.3035	91.6535	92.2122	150.112	4.38634	20.7949	87.7452	84.196	219.678
CxB	F	50ppm		7.576	31.1395	374.998	351.437	220.704	5.12144	29.9753	383.899	386.179	235.425	8.95824	24.4063	50.75	385.883	82.3952	11.3211	42.4959	91.27198006	59.78196421	323.705	9.33631	59.7349	66.8062	61.2184	770.152	8.56835	69.4292	66.5802	60.8988	956.171
ANOVA w/dunnetts		1vs0		0.252	0.293	N/A	N/A	0.345	0.254	0.331	N/A	N/A	0.501	0.32	0.203	0.611	0.661	0.199	0.0572	0.119	**0.0000134*****	**0.0000895*****	0.165	0.256	0.181	0.0876	0.066	0.196	0.23	0.202	0.221	0.15	0.208
		50vs0		0.813	0.923	0.0559	**0.0367***	1	0.971	0.9	**0.0366***	**0.0362***	0.846	0.993	0.936	0.195	0.501	0.683	0.4023	0.488	0.948	0.783	0.506	0.802	0.618	0.2243	0.157	0.614	0.819	0.647	0.21	0.129	0.632
AVG																																
CxB	M	0ppm	6	4.4	10.92	195.493	232.36	570.88	9.5	41.8733	231.907	270.4	53.018	21.8	126.027	164.99	199.413	830.434	49.8333	258.798	128.0183333	166.4666667	3907.62	23.8333	93.112	98.8	144.78	892.078	24.3333	109.216	98.58	144.85	1239.69
CxB	M	1ppm	5	2.8	3.912	383.765	417.185	62.582	1	3.37	384.965	422.56	56.276	26.2	50.764	342.978	374.023	161.83	33.75	48.79	12.158	41.5425	388.015	11.8	68.368	12.175	49.435	736.886	12	77.75	12.175	49.485	954.32
CxB	M	50ppm	4	4.25	5.83	100.44	147.96	35.9725	1.75	3.3	269.04	320.2	14.88	11.25	23.68	97.23	142.63	65.8275	18.5	84.7875	70.91	117.4166667	413.505	16.5	129.818	106.33	136.273	467.543	16	143.678	105.807	136.363	660.767
SE																																
CxB	M	0ppm		2.97658	9.18251	10.0756	8.36593	513.418	7.63653	33.8807	24.2898	28.242	41.8817	12.6507	70.2181	39.3765	40.2222	744.543	21.6817	129.405	23.54321309	27.84137541	2290.65	10.6063	43.7961	42.8053	46.8299	399.669	10.7971	52.5532	42.7661	46.8354	561.818
CxB	M	1ppm		1.71464	2.42547	362.545	369.435	59.3685	0.63246	2.1763	361.345	363.93	53.7356	19.187	42.0863	182.934	186.085	93.5877	9.52519	20.8558	7.087398253	2.771074927	139.748	7.70973	44.1946	11.445	9.725	460.496	8	50.4762	11.445	9.745	609.875
CxB	M	50ppm		4.25	5.83	N/A	N/A	35.9725	1.75	3.3	N/A	N/A	14.88	6.6505	18.1068	3.17	5.13	39.625	11.0793	47.53	30.93210684	30.79385779	248.113	4.80451	56.6748	8.13831	18.3021	116.356	4.60072	59.8289	7.8495	18.3077	169.77
ANOVA w/dunnetts		1vs0		0.896	0.665	0.749	0.763	0.512	0.442	0.423	0.821	0.826	0.998	0.966	0.533	0.528	0.549	0.567	0.762	0.287	**0.00309****	**0.00933****	0.303	0.536	0.902	0.261	0.237	0.946	0.533	0.879	0.262	0.237	0.904
		50vs0		0.999	0.82	0.951	0.962	0.52	0.541	0.463	0.992	0.986	0.774	0.843	0.377	0.932	0.953	0.486	0.393	0.406	0.18924	0.37863	0.308	0.805	0.821	0.983	0.977	0.741	0.765	0.872	0.984	0.977	0.74

Supplementary table B3.6 OF.

Open field behavioral testing testing. *p<0.05, *p<0.01, *p<0.001 (one way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm). Significant p values are in bold.

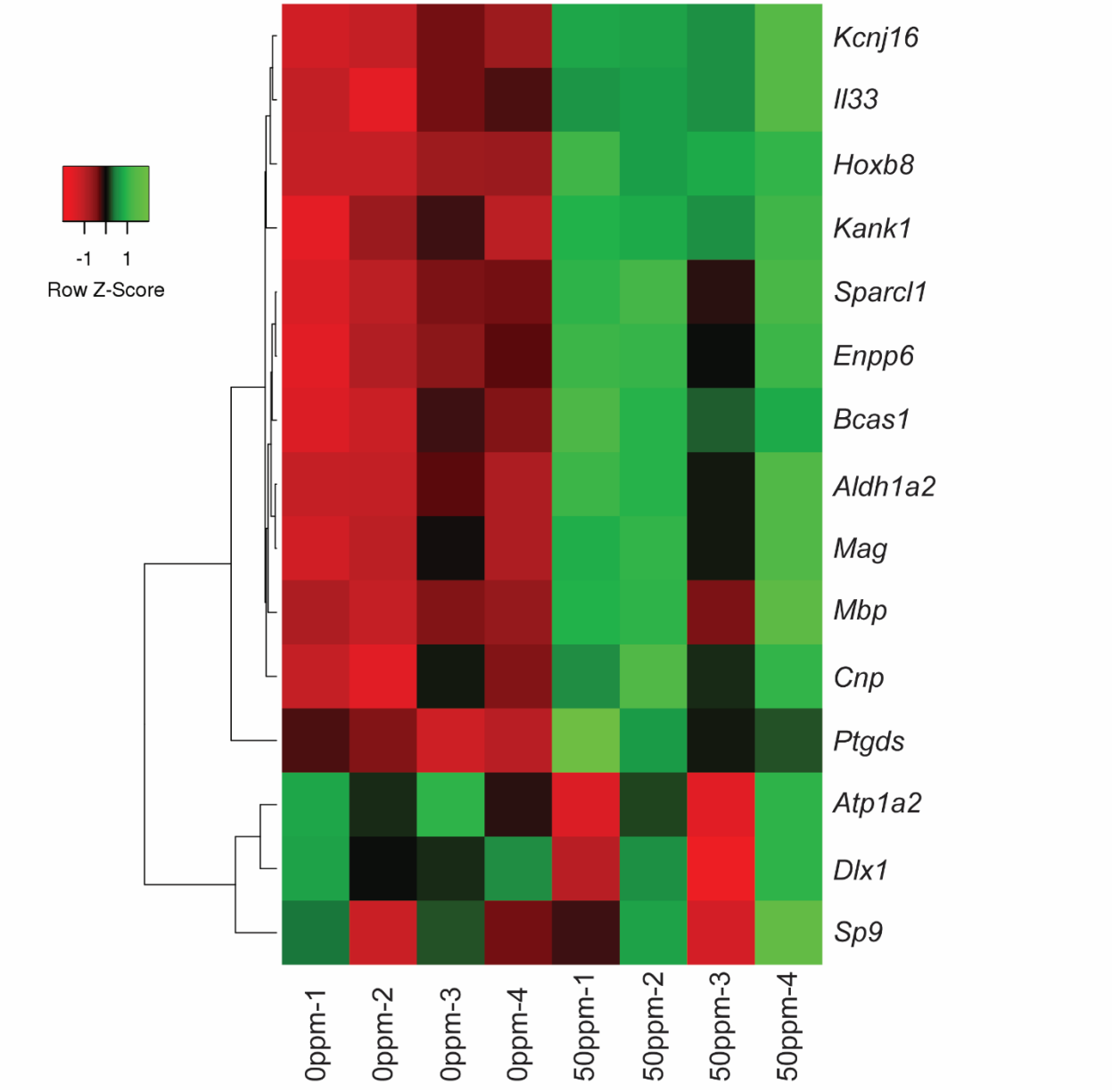
Cross	Sex	Dose	Sample size (n)	Area:Mouse 1 Center In Perimeter					Area:Mouse 1 Center In Center					Area:Mouse 1 Nose In Center					Mouse 1 Body In Perimeter Over 80.0%					Mouse 1 Body In Center Over 80.0%					Total distance traveled in box during 30 minutes (mm)
				Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	Bouts	Duration (Second)	Latency (Second)	End Time (Second) of First Bout	Distance (mm)	
AVG																													
BxC	F	0ppm	13	251.9231	514.221538	1.929091	39.80333	21141.95	45.38462	77.72692	110.7515	150.6769	3067.258	61.53846	78.74769	78.57154	117.0769	2989.198	122.3077	327.7738	7.715455	46.09182	11658.25	20.69231	47.39615	190.2362	229.8185	1659.365	38914.396
BxC	F	1ppm	9	295	563.753333	2.873333	35.53111	25844.45	58.11111	90.46556	109.6389	142.5722	2746.654	80	94.10889	31.22143	64.36429	3321.838	156.5	401.0975	2.88125	36.79556	15357.66	19.75	55.92556	165.6889	198.6067	1642.656	36014.862
BxC	F	50ppm	5	339.6	522.065	1.334	42.114	26464.57	62.8	100.814	58.818	99.044	4136.918	94.4	91.792	30.89	71.038	3966.08	149.2	364.244	4.426	45.846	15082.66	23.5	61.688	142.6775	178.1925	2252.658	39451.4
SE																													
BxC	F	0ppm		15.45442	32.6049573	0.554398	1.872826	1710.209	5.989883	11.78464	36.67772	36.0472	513.776	7.541862	11.78783	26.92409	27.00221	477.8415	9.127575	28.55488	2.160327	2.849079	1190.44	3.393781	7.584366	57.02335	55.79844	308.9726	728.23502
BxC	F	1ppm		28.25528	38.1220209	1.137169	2.213852	2841.118	12.19112	16.71494	50.65403	50.55109	437.6055	14.19018	15.59115	14.92948	16.17103	579.4207	7.711309	14.8189	0.965241	2.89328	829.8964	3.080758	10.55142	48.74079	49.26236	238.7341	1847.0623
BxC	F	50ppm		37.69164	28.7928835	0.855947	6.853431	2772.593	11.2712	16.48781	26.94974	26.35668	695.6212	14.66492	10.75799	19.4971	21.17689	590.996	15.23286	40.52004	1.829275	7.585074	1973.592	5.484828	11.87573	89.44237	89.8438	441.8691	4034.9178
ANOVA w/dunnetts		1vs0		0.3106	0.51	0.639	0.461	0.243	0.499	0.75	1	0.986	0.891	0.378	0.626	0.35	0.288	0.873	**0.0366***	0.133	0.131	0.132	0.0993	0.976	0.735	0.939	0.901	0.999	0.666
		50vs0		0.0545	0.99	0.88	0.848	0.288	0.419	0.536	0.689	0.687	0.388	0.144	0.791	0.424	0.457	0.447	0.1886	0.661	0.46	0.999	0.1923	0.878	0.564	0.873	0.849	0.437	0.146
AVG																													
BxC	M	0ppm	10	233.6	528.021	2.156667	49.119	20242.96	43.5	72.528	76.648	122.203	3049.92	62	72.119	69.11556	113.2567	2964.608	111.5	294.2656	26.022	71.262	10551.2	17.6	45.928	160.204	206.502	1404.714	31545.906
BxC	M	1ppm	7	232	547.188571	6.664286	50.26571	20755.84	44.28571	82.74857	79.54714	124.9057	3084.934	80.57143	81.50429	78.05429	120.6771	2944.8	115.8571	348.0757	7.174286	50.24143	10406.81	19.28571	50.99143	146.2257	190.92	1582.637	33398.23
BxC	M	50ppm	11	227.8182	450.24	0.798889	53.11364	17450.25	42.36364	78.92333	118.5445	165.3682	2958.775	66.27273	82.55182	105.54	151.3236	2860.701	105	283.7291	12.97273	58.79818	9495.729	17	48.24444	170.4309	217.0845	1410.399	28621.599
SE																													
BxC	M	0ppm		22.13504	33.0315125	0.659598	3.234172	2014.52	6.947821	10.55199	34.62843	34.72117	535.8717	8.514041	9.935742	26.84094	28.79719	502.9724	8.546929	23.10446	10.1463	11.60407	1124.562	2.753584	7.319538	58.45694	59.84068	225.0728	1592.7238
BxC	M	1ppm		22.54308	34.6713436	3.42059	5.424842	1636.294	7.821564	14.36235	37.52534	36.85031	481.2572	17.42311	12.64005	37.86531	37.98453	425.2624	11.11076	33.07949	3.404403	5.423056	852.4491	2.925376	8.92119	52.60731	52.48946	220.2211	3214.7326
BxC	M	50ppm		21.66808	43.5302444	0.37624	5.058657	1595.816	5.848501	7.565656	46.77463	45.0101	423.1868	10.72981	13.61574	39.44049	39.46955	387.3383	9.83038	29.80459	4.434609	5.771431	967.4097	2.371123	5.8421	57.18828	55.39432	201.2642	2193.4853
ANOVA w/dunnetts		1vs0		0.998	0.927	0.147	0.981	0.975	0.996	0.744	0.998	0.999	0.998	0.482	0.838	0.981	0.987	0.999	0.94	0.38	0.162	0.2	0.994	0.88	0.853	0.982	0.977	0.806	0.813
		50vs0		0.974	0.255	0.788	0.749	0.418	0.989	0.873	0.68	0.652	0.986	0.948	0.76	0.68	0.666	0.979	0.841	0.949	0.309	0.458	0.673	0.98	0.962	0.988	0.987	1	0.536
AVG																													
CxB	F	0ppm	11	192.8182	407.697778	1.682	38.45818	16724.69	23.2	48.13455	18.36222	54.69111	1639.893	39.36364	48.27636	16.43111	52.30889	1656.307	96.55556	230.2167	7.57	45.236	7979.134	11	32.59182	118.648	155.788	928.987	28668.261
CxB	F	1ppm	6	253.6	479.23	4.236667	41.07833	19856.97	30.66667	65.33833	100.5	139.0433	1998.952	46	66.11667	98.64167	136.4217	2076.242	123.3333	315.745	4.076	43.144	11374.89	15.66667	46.41833	171.21	209.8417	1257.138	27877
CxB	F	50ppm	3	224.3333	404.25	2.113333	51.24	19636.77	38.33333	53.49667	1.09	50.80667	2611.287	52.66667	50.22667	24.53667	74.71667	2539.997	111.3333	277.8767	13.68	63.97333	11870.17	18	31.59667	24.54667	74.79667	1541.32	34222.91
SE																													
CxB	F	0ppm		25.42862	23.0739255	0.694006	3.218224	2169.879	4.170798	8.54664	9.477998	11.05821	297.4179	7.686201	7.906086	9.058528	10.81412	278.8161	8.241344	16.40644	2.390278	5.382305	1147.425	2.027588	6.265864	50.27996	48.46931	187.5925	2983.6773
CxB	F	1ppm		11.70726	44.0260978	1.443664	4.419379	1049.067	6.790516	13.6678	62.13961	61.99093	397.3761	9.549869	12.52777	62.46843	62.66368	369.4249	11.46783	42.05955	1.568466	5.743711	1263.994	2.577682	8.75426	71.74525	68.7679	234.4411	2145.5502
CxB	F	50ppm		35.31918	73.8828372	0.566108	24.94503	2998.336	10.58825	18.42001	0.553263	24.79668	801.5346	16.29247	16.19034	22.76733	24.48428	762.2832	18.88856	57.37631	7.263948	31.95102	2100.432	5.773503	10.89809	22.91186	24.48948	475.7391	4789.7786
ANOVA w/dunnetts		1vs0		0.24	0.287	0.138	0.947	0.569	0.561	0.466	0.189	0.187	0.74	0.84	0.383	0.197	0.192	0.632	0.146	0.104	0.646	0.985	0.141	0.362	0.358	0.757	0.729	0.523	0.978
		50vs0		0.745	0.998	0.959	0.479	0.71	0.259	0.953	0.947	0.997	0.292	0.66	0.992	0.988	0.917	0.312	0.652	0.592	0.414	0.456	0.195	0.257	0.996	0.587	0.65	0.272	0.542
AVG																													
CxB	M	0ppm	5	198	465.948	1.345	39.16333	14707.86	23.2	51.4	298.524	336.52	1558.284	38.8	57.51	221.154	258.084	1669.228	94	296.508	10.432	49.332	8785.878	12.2	37.13	203.8	243.4125	921.258	24107.158
CxB	M	1ppm	7	224	526.117143	2.5	38.044	20137.63	44.42857	95.22429	0.247143	35.24667	3484.531	74	96.15714	0.238571	35.06	3479.884	132.1429	359.2	10.52	46.64714	12995.33	23.57143	71.96167	0.3	37.24167	2136.994	28315.707
CxB	M	50ppm	3	176.3333	463.71	3.836667	59.21333	14708.39	29.33333	88.82	76.25333	132.0533	2412.483	47.33333	92.22667	76.04	131.82	2530.867	101	287.92	32.33333	88.2	8626.26	17.66667	65.45667	171.8367	229.47	1670.33	27042.17
SE																													
CxB	M	0ppm		12.32883	20.3445859	0.569525	0.688097	1092.39	6.216108	20.64073	118.7743	115.9158	491.4643	8.748714	20.31434	90.90277	88.81261	450.1603	9.027735	31.71556	3.835924	3.168428	1243.794	4.140048	16.93168	92.25173	89.55368	320.835	923.89306
CxB	M	1ppm		18.68154	38.5803418	0.706022	0.565744	1991.446	5.027814	5.575004	0.120232	0.882457	386.7234	8.403514	6.156645	0.114672	0.656331	363.5387	10.22469	35.51618	2.245823	4.304565	1350.343	1.950057	2.083426	0.146133	2.481484	199.3417	2894.3022
CxB	M	50ppm		20.41514	52.8123	3.555476	5.525849	1465.242	6.009252	26.86765	42.98949	45.60738	399.1093	7.688375	22.33992	42.85489	45.59313	368.7059	14.22439	63.37663	28.3437	29.32611	1439.115	2.962731	25.50634	38.4227	41.28815	431.6011	1836.1981
ANOVA w/dunnetts		1vs0		0.476	0.408	0.77	0.92921	0.0785	**0.031***	0.0892	**0.0128***	**0.0184***	**0.0123***	**0.0209***	0.118	**0.0167***	**0.0406***	**0.0114***	**0.0385***	0.415	1	0.971	0.0756	**0.0267***	0.14	**0.0245***	**0.0207***	**0.0126***	0.0678
		50vs0		0.702	0.999	0.473	**0.00189****	1	0.754	0.272	0.1324	0.1865	0.4279	0.7921	0.288	0.2174	0.3547	0.3749	0.898	0.988	0.301	0.0658	0.9966	0.4651	0.361	0.8926	0.9773	0.226	0.6668

Supplementary figure B3.1.



Supplementary figure B3.1. The effects of maternal Cd exposure on brain mass and normalized brain mass at 6 months of age. A, Brain mass at 6 months. B, Proportional brain mass at 6 months. *p<0.05, **p<0.01, ***p<0.001 (one-way ANOVA with post-hoc Dunnett's test comparing 1 ppm and 50 ppm to 0 ppm).

Supplementary figure B3.2.



Supplementary figure B3.2. Heatmap of 15 significantly differentially expressed genes in newborn BxC female whole brains as a result of 50 ppm maternal Cd exposure.