Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

________________________________ ________________

Sasha Erin Parets Date

Characterization of DNA Methylation in African Americans with Spontaneous Preterm Birth

By

Sasha Erin Parets
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science

Genetics and Molecular Biology

Alicia K. Smith Ph.D. Advisor	Ramkumar Menon M.S., Ph.D. Committee Member
Karen Conneely Ph.D. Committee Member	Carlos Moreno Ph.D. Committee Member
Elizabeth Corwin Ph.D., RN, FAAN Committee Member	Stephanie Sherman Ph.D. Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.

Dean of the James T. Laney School of Graduate Studies

Date

Characterization of DNA Methylation in African Americans with Spontaneous Preterm Birth

By

Sasha Erin Parets

B.S., Brandeis University, 2010

Advisor: Alicia K. Smith, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2015

Abstract

Characterization of DNA Methylation in African Americans with Spontaneous Preterm Birth

By Sasha Erin Parets

African Americans are at increased risk for spontaneous preterm birth (PTB), but the biological mechanisms underlying PTB are not yet known. Epigenetic factors, such as DNA methylation, may provide insight into the genes that are being actively regulated in those that deliver or are delivered preterm. The objective of this study was to evaluate DNA methylation in paired maternal blood and umbilical cord blood (fetal) samples to identify patterns specific to PTB. Peripheral blood from African American women who delivered preterm (24-34 weeks) or at term (39-41weeks) was assessed for DNA methylation across the genome using the HumanMethylation450 BeadChip. In maternal samples, no sites associated after correction for multiple comparisons though 17,829 CpG sites associated with PTB ($p<.05$). Examination of paired samples, irrespective of PTB status, identified 5,171 CpG sites in which methylation of maternal samples predicted methylation of her respective neonate (false discovery rate (FDR)<.05). The majority of correlated CpG sites could be attributed to one or more nearby genetic variants. However, correlated CpG sites were significantly more likely to be in genes involved in metabolic, cardiovascular and immune pathways, suggesting a role for genetic and environmental contributions to PTB risk. The observation that maternal epigenetic differences predict fetal methylation may provide insight into the heritability of PTB. In umbilical cord blood samples, we identified ~10,000 CpG sites that associate with gestational age (GA), only 29 of which associated with PTB when controlling for GA, suggesting that the majority of CpG sites primarily reflect developmental differences between preterm and term samples. In order to assess the association of DNA methylation with childhood outcomes, we investigated DNA methylation of calcitonin (*CALCA*), which associated with GA in cord blood and PTB in maternal blood. DNA methylation of *CALCA* did not associate with or mediate the relationship between maternal depressive symptoms, a known risk factor for PTB and internalizing behavioral in childhood. These finding show the importance of DNA methylation in understanding the risk and consequences of PTB in African Americans.

Characterization of DNA Methylation in African Americans with Spontaneous Preterm Birth

By

Sasha Erin Parets

B.S., Brandeis University, 2010

Advisor: Alicia K. Smith, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2015

ACKNOWLEDGEMENTS

It is a great pleasure to thank the many people who made this dissertation possible.

I would like to express my sincere gratitude to my advisor Dr. Alicia Smith. She has giving me motivation, guidance, and support throughout graduate school. Without her help, I would never have had the opportunity to attend conferences and flourish as a scientist and a writer. She is the best advisor for me, and I appreciate the countless hours she has put in to training me.

I am extremely grateful to my dissertation committee members: Karen Conneely, Elizabeth Corwin, Ramkumar Menon, Carlos Moreno, and Stephanie Sherman for their mentorship, advice, critical questions and expertise. Each one of them contributed so much to the development of my dissertation and myself as a student. I am also grateful for the support of the GMB and M2M programs for the feedback and encouragement I received throughout this process.

This dissertation would not have been possible without the support of my wonderful colleagues. I wish to thank Alaine Broadway, Kimberly Kerely, Anna Knight, Elizabeth Marder, and all the past and present members of the Smith lab.

I am grateful for all my friends, for their friendship, laughter, dinners, board game parties and support. They have helped me stay grounded throughout this PhD. I am forever grateful for having such wonderful people in my life.

I would like to thank my whole family for their love and never-ending support. Without their encouragement and belief in me, I would never have completed this PhD. My mom always taught be to dream big. My dad always pushed me forward. My Nana

is the reason I am the person who I am today. My Abuela taught me to enjoy life, and my sisters have helped me keep a balance life. I appreciate the addition of the Shinnar clan into my family and their never-ending support, advice, and love throughout my dissertation.

Finally, I thank Eli, for his instrumental support, advice and for always being there for me. He is one of the driving forces behind me and this work. He is my partner, my best friend and my love.

TABLE OF CONTENTS

Chapter 1: Preterm Birth and Its Long-Term Effects: Methylation to Mechanisms 1

Introduction 2

Overview of Pregnancy 3

Types of Preterm Birth 4

Proposed Mechanisms of PTB 6

Consequences of PTB 9

DNA Methylation Studies of PTB 10

DNA Methylation Studies of Long Term Outcomes of PTB 13

Opportunities for Epigenetic Studies of PTB 14

Chapter 2: DNA Methylation Provides Insight into Intergenerational Risk for Preterm Birth in African Americans 34

Introduction 35

Methods 37

Results 41

Discussion 45

References 57

Chapter 3: Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age 67

Introduction 68

Methods 69

Results 73

Discussion 77

References 89

Chapter 4: Fetal DNA Methylation of Calcitonin (*CALCA*) Does Not Associate with Behavioral and Cognitive Outcomes in Childhood 97

Introduction 98

Methods 100

Results 103

Discussion 105

References 111

Chapter 5: A Pilot Study of DNA Methylation in Spontaneous Preterm Birth: Conclusions & Recommendations for Future Studies 116

References 123

Table of Tables

Table 1-1	Overview of DNA methylation studies in PTB	15
Table 2-1	Demographics table for maternal samples	48
Table 2-2	Enrichment for correlation analysis of maternal and fetal blood	49
Table 2-3	Pathway analysis of CpG sites in genes that correlate in maternal-fetal pairs	50
Table 3-1	Clinical and demographic characteristics of the cohort	80
Table 3-2	29 CpG sites that associated with PTB (adjusted for GA)	81
Table 3-3	Enrichment analysis	82
Table 3-4	Gene enrichment of CpG sites that associated with GA	83
Table 3-5	Replication of CpG sites previously associated with GA	84
Table 3-6	Demographic table of the WMHP umbilical cord blood samples	107

Table of Figures

Figure 1-1	A model for the involvement of DNA methylation in the development of chronic disease following PTB	16
Figure 2-1	Manhattan plot of the relationship between maternal and fetal DNA methylation	51
Figure 2-2	Distribution of maternal DNA methylation for CpG sites	52
Figure 2-3	Correlated CpG sites and their enrichment in meQTLs	53
Figure 2-4	Distribution of maternal and cord blood expression	54
Figure 2-5	Association of *MICB* in PTB and its correlation in expression and DNA methylation	55
Figure 3-1	Manhattan plots depicting the association of all CpG sites with PTB and with GA	85
Figure 3-2	Scatter plots of the unmethylated vs. methylated signals for the five PTB-associated CpG sites	86
Figure 3-3	Correlation between the t-statistics depicting association analysis of CpG sites with PTB compared to GA	87
Figure 4-1	Schematic of *CALCA*	108
Figure 4-2	Association between DNA methylation in *CALCA* (cg091889801) with GA	109

Chapter 1

Preterm Birth and Its Long-Term Effects: Methylation to Mechanisms

This chapter has been adapted and was originally published in Biology under a CC-BY 3.0 license (http://creativecommons.org/licenses/by/3.0/). Readers can access the published article at http://www.mdpi.com/2079-7737/3/3/498 or from the citation below:

Parets SE, Bedient C, Menon R, Smith AK. (2014) Preterm Birth and Its Long-Term Effects: Methylation to Mechanisms. Biology 3(3): 498-513.

Introduction

Despite advances in healthcare, preterm birth (birth prior to 37 weeks gestation) remains a major global health problem [1]. Preterm birth increases risk for morbidity and mortality in the first year of life [2], and these consequences extend throughout development. For example, children born preterm have higher levels of neurodevelopmental disability and an increased risk of behavioral problems such as Attention Deficit Hyperactivity Disorder [3-5]. Being born preterm also increases risk for developing chronic diseases such as hypertension, type 2 diabetes, cardiovascular disease, obesity and psychiatric disorders [2,6,7]. Collectively, these observations support the developmental origin of health and disease (DOHaD) hypothesis, which conceptually links the prenatal and early postnatal environments to the development of chronic diseases [8,9].

The molecular mechanisms that underlie the relationship between preterm birth and its developmental consequences are not clear, but advocates of the DOHaD hypothesis believe that epigenetics may play a key role (Figure 1) [10]. Epigenetic modifications, such as histone modification, non-coding RNAs (ncRNA) and DNA methylation, induce changes in gene expression through structural alterations of DNA that are maintained through each round of cell division; they respond to changes in the environment, are potentially reversible and can be targeted for disease therapies [11,12].

The most widely studied epigenetic modification is DNA methylation, which occurs when a methyl group is added to the 5' carbon of a cytosine when it is next to a guanine (CpG site). Many gene promoters contain regions with a high density of CpG sites, called CpG islands, though CpG islands are not restricted to promoters [13]. CpG

islands have lower levels of methylation compared to CpG shores, which exhibit tissue specific methylation patterns and are more commonly identified in disease association studies [14]. The relationship between DNA methylation and gene expression varies based on its genomic context. The general consensus is that increased DNA methylation in the promoter region of a gene has an inverse relationship with gene expression [15]. Intragenic DNA methylation can positively associate with gene expression and may regulate alternative promoters or enhancers that are involved in tissue-specific expression patterns [16]. Subtle differences in the intrauterine environment may influence this tightly controlled process such that environmentally-induced epigenetic changes may result in stable phenotypic differences.

This chapter will provide an overview of the key physiological features of pregnancy and spontaneous preterm birth, a major subset of all preterm births with unknown etiology, with specific emphasis on the emerging role of DNA methylation in the field. It is not meant to serve as a comprehensive overview of preterm birth, as other excellent reviews of its epidemiology, management, prediction, prevention, physiology, and proposed mechanisms have recently been published [2-4,17-22]. Instead, this chapter will introduce key concepts in PTB research and highlight the paucity of DNA methylation studies in this field.

Overview of Pregnancy

Pregnancy initiates after fertilization of the female gamete and implantation of the embryo; it is the gestational period during which fetal development and growth occurs. The typical duration of human pregnancy is 280 days (40 weeks), though the actual length may vary considerably [23,24]. Obstetricians calculate the estimated due date of a

pregnancy starting from the first day of the last menstrual period (LMP) [25]. The LMP marks the beginning of the first trimester, which is characterized by embryonic implantation, organogenesis and maternal adaptations of the cardiovascular, respiratory, renal, endocrine and immune systems [26,27]. These changes are maintained during the second and third trimesters to support rapid growth and development of the fetus.

The relationship between the mother and fetus is considered to be semi-allografic because the fetus is genetically distinct and can be recognized as a foreign body by the maternal immune system. The placental barrier regulates the maternal–fetal interface by controlling substances transmitted to the fetus [27,28]. Both the maternal and fetal circulations access the intervillous space, allowing nutrient and waste exchange, as well as limited immune surveillance [27,29-32]. Thus, the quality of the intrauterine environment is highly dependent on maternal health.

Types of Preterm Birth

While birth occurring before 37 weeks of gestation is considered preterm, the limit of viability is as low as 22 weeks in some medical centers [33]. Thus, preterm birth may be classified as: extreme ($22^{0/7}$–$27^{6/7}$ weeks), very early ($28^{0/7}$–$31^{6/7}$ weeks), moderate ($32^{0/7}$–$33^{6/7}$ weeks) and late ($34^{0/7}$–$36^{6/7}$ weeks) [2]. Multiple mechanisms may result in preterm birth, making it difficult to classify subjects for research [20]. Despite difficulties in classifying preterm birth for research purposes, clinically it may be categorized into three different types: spontaneous preterm birth (PTB), preterm premature rupture of membranes (pPROM), and medically indicated (iatrogenic) preterm birth. Research has identified numerous risk factors for preterm labor that results in preterm deliveries. Women with extreme maternal reproductive ages (<18 or >40),

obesity, risk taking behaviors (tobacco, alcohol or drug use), infections or allergic reactions, or psychosocial stress are at higher risk to have spontaneous preterm labor [2,34].

African American women and those with a lower socioeconomic status are at increased risk to deliver preterm [2,34]. Strikingly, African American women are more likely to deliver preterm and early preterm, independent of socioeconomic status [35-37]. The Nashville Birth Cohort (NBC) was established to evaluate the factors that contribute to racial disparities in PTB rates [38-41]. This unique cohort was leveraged for the pilot studies presented in Chapters 2 and 3, in part, because it was one of the few cohorts available at the time that collected paired biological samples in African American women and fetuses that were appropriate for DNA methylation studies. DNA methylation patterns associate with stress [42,43], diet [44-48], smoking [49-53], inflammatory cytokine levels [54] and medications [55-58], and each of these factors increase risk for PTB. This prompted our hypothesis that DNA methylation differences in African Americans may contribute to the risk of PTB and long-term consequences.

Medically indicated or iatrogenic preterm birth results from obstetrical, fetal or medical complications requiring early delivery for treatment or resolution. Common obstetrical indications for preterm delivery include gestational diabetes and severe preeclampsia (new onset elevated blood pressure associated with inadequate fetal growth or perfusion and maternal organ damage) [59]. Both these illnesses must be treated by delivery and adjunctive therapies to completely resolve the condition [60-63]. Fetal indications for delivery may also occur without maternal pathology. For instance, fetal hydrops, a consequence of severe fetal anemia, requires blood

transfusions in utero, but does not negatively impact the mother's health. If possible iatrogenic preterm delivery occurs after achieving fetal lung maturity [64]. Finally, maternal medical disease, such as a new cancer diagnosis, may dictate an early delivery in order to appropriately treat a life-threatening disease without adversely impacting the fetus [65].

Premature rupture of membranes (PROM) involves a rupture of the fetal membranes prior to the onset of labor. In an uncomplicated term pregnancy, rupture of the membranes immediately precedes or occurs during labor [66], while preterm premature rupture of membranes (pPROM) occurs prior to 37 weeks of gestation. pPROM is characterized by activation of inflammatory mediators, mainly cytokines, chemokines and matrix metalloproteinases, that can weaken fetal membranes via proteolytic damage. The biggest risk factors for pPROM are placental abruption, infection, uterine or cervical abnormalities, and uterine over-distension [20].

Spontaneous PTB can happen either with or without rupture of membranes. Risk factors associated with PTB are similar to that of pPROM and include, but are not limited to, prior preterm birth, intra-amniotic infections, stress, behavior, obesity and inter-pregnancy interval [21]. Although risk factors are well-recognized and many intervention strategies have been developed, (e.g., tocolytics for contractions, antibiotics and steroids for infection/inflammation, progesterone for cervical shortening), none of these interventions have reduced the risk of PTB over three decades. Development of PTB and pPROM treatments will be substantially enhanced by a more in depth understanding of their underlying molecular mechanisms.

Proposed Mechanisms of PTB

There are many hypotheses about the mechanisms that may contribute to PTB. Some note that PTB is heritable and seek to identify the genetic risk factors [17,67,68]. Many note that the factors that increase risk of PTB and pPROM are also fundamental in maintaining a healthy pregnancy, and these hypotheses tend to focus on neuroendocrine and immune systems.

Pregnancy is a period of extensive stress, during which the hypothalamic pituitary adrenal (HPA) axis undergoes extensive changes [69,70]. Hypothalamic corticotropin releasing hormone (CRH) is the primary regulator of pituitary release of adrenocorticotropic hormone (ACTH). In turn, ACTH stimulates the release of glucocorticoids from the adrenal cortex. Glucocorticoids provide negative feedback on the HPA axis, and inhibit at both the hypothalamic and pituitary levels. However, during pregnancy glucocorticoid production stimulates release of placental CRH [71].

The placenta releases CRH into maternal and fetal circulation in significant quantities. Placental CRH stimulates the maternal HPA axis, leading to an increase in total and free cortisol during pregnancy [72], which eventually attenuates, such that pregnancy becomes a period of HPA axis suppression. The HPA axis also regulates placental blood flow and influences the timing of parturition [73,74]. CRH stimulates ACTH release both from the fetal pituitary and the placenta [75,76], which in turn leads to release of cortisol from the fetal adrenal gland [77]. Progressive activation of the fetal HPA axis is important for maturation of organs such as lungs [78,79]. Furthermore, increased levels of placental CRH are associated with onset of labor [73,74].

Due to the above factors, the neuroendocrine system plays an important role in timing of parturition. It is therefore not unexpected that dysregulation of the

neuroendocrine system could be involved in the mechanism of PTB, particularly since stress is a risk factor for PTB [80]. Physical and psychological stress can activate the maternal and fetal HPA axis, which increases production of placental CRH, a vital hormone for fetal maturation as well as the initiation and timing of labor [81]. For this reason, some investigators call it the "placental clock" [18,74].

The neuroendocrine and immune systems are interconnected. Acute stress has an anti-inflammatory response, though chronic mental or physical stress can lead to a pro-inflammatory state and even glucocorticoid resistance [18,82]. This is illustrated by a recent study that showed that pregnant women of low socioeconomic status are more likely to have glucocorticoid resistance and a dysregulated inflammatory response [83]. Perceived stress, cortisol, inflammation and early life socioeconomic status have been associated with DNA methylation differences at a variety of CpG sites [84]. Early life socioeconomic status may continue to impact DNA methylation in women throughout their childbearing years [85].

Over the course of a typical pregnancy, the maternal immune system is characterized by a shift from a pro-inflammatory state that is effective against intracellular pathogens such as viruses and bacteria to an anti-inflammatory state that targets extracellular pathogens using specific antibodies; however, it does so through a dynamic process [86-90]. The first trimester is characterized by a strong pro-inflammatory state, because of the necessity to repair the endometrium after implantation of the blastocyst and to establish placentation [91]. At the beginning of the second trimester, the maternal immune system transitions to an anti-inflammatory state to facilitate rapid fetal growth and development [91]. During the third trimester, transition to

a pro-inflammatory state promotes uterine contractions and delivery of the fetus and placenta [91]. Pro-inflammatory cytokines stimulate prostaglandin and matrix metalloproteinases production, which are involved in cervical ripening, membrane rupture and uterine contractions [81,92]. Consistent with this, there is an increase in pro-inflammatory cytokines in maternal plasma and migration of leukocytes to the myometrium prior to the onset of spontaneous term labor [93,94].

Inflammation is implicated in most PTB. Various infections such as urinary tract infections, bacterial vaginosis, sexually transmitted infections, malaria and even periodontal disease have associated with PTB [95,96]. Even subclinical intrauterine infections stimulate the release of pro-inflammatory proteins that overlap with the mechanism of normal parturition. Inflammation is part of the normal signaling pathway for parturition, and a premature activation of this pathway may lead to premature labor. For example, deliberate infection of mice increases pro-inflammatory cytokines (IL-1 and TNF-alpha) and induces labor [97].

Consequences of PTB

Clinical advancements have significantly reduced the mortality rate of infants delivered preterm, but morbidity remains a substantial concern. It is unclear whether those born preterm are able to meet developmental milestones in a time frame that is comparable to their term-born peers. Even among those that do not have congenital malformations, many infants delivered preterm appear to have distinct developmental trajectories that differentiate them from term infants as they age. For example, some report that preterm infants are able to catch up in both weight and height in the first two

years of life [98-101], though catch up growth could continue into childhood and adolescence [102]. Another report suggests a life-long discrepancy in height [103].

In addition to the question of whether or not preterm infants catch up, it is unclear if there are additional consequences that result from accelerated growth rates. Catch up growth in early and late infancy has been associated with obesity, cardiovascular disease, and insulin resistance in adolescents [6,7]. For example, models of maternal under-nutrition, as well as low protein and high fat diets, support epigenetic modifications and phenotypic changes in the offspring, including alterations in food preferences and cholesterol regulation [104-106]. Others report higher rates of behavioral and emotional problems as well as decreases in cognitive performance during childhood [5,107,108]. Similarly, studies report that children born preterm are more likely to experience slower motor, language and neurological development than children born at term [109]. Finally, children born preterm may be less likely to complete high school or seek higher education [110,111].

DNA Methylation Studies of PTB

Numerous studies report epigenetic differences associated with gestational age and growth patterns [112-115]. For example, one study reported extensive gestational age-associated DNA methylation differences among term births and noted umbilical cord blood DNA methylation differences in genes implicated in labor and delivery [115]. However, despite extensive interest in the biological mechanisms of PTB, there are surprisingly few epigenetic studies of PTB (Table 1) [116-119]. Those that have been conducted are promising and provide insight into the mechanisms underlying PTB risk factors and consequences.

Because of the complexity of the maternal–fetal interface and the number of tissues involved in pregnancy and delivery, it is not always clear which tissue is the most appropriate for PTB studies. For example, a recent study evaluated myometrium, the middle layer of the uterine wall that induces contractions during labor [120,121]. By comparing DNA methylation of genes that are involved in contraction, the authors sought to identify differences in PTB and term birth samples. The study categorized their 53 samples into six different delivery/labor types and found that DNA methylation of several CpG sites distinguished the groups [121]. A study by Burris and colleagues examined another maternal tissue, the cervix, which separates the uterus and vagina. They evaluated global methylation (long interspersed nuclear elements; LINE-1) using pyrosequencing in cervical swabs collected between 16–19 weeks of gestation and reported increased LINE-1 methylation was associated with shorter gestation. LINE-1 methylation is used as a surrogate for global methylation because these elements are found throughout the genome. In the same report, they also evaluated DNA methylation of *PTGER2* (prostaglandin E receptor 2), which plays a role in response to prostaglandins and labor initiation. They reported associations between DNA methylation of *PTGER2* and both local inflammation and length of gestation [122].

Other studies focus on the role of the placenta. For example, one study examined amnion tissue, the inner layer of the fetal membranes, from 121 term and preterm deliveries [123]. This genome-wide investigation identified CpG sites that associated with both labor and PTB. The authors propose that DNA methylation changes in the amnion may participate in labor and the etiology of preterm birth, which supports the idea that DNA methylation studies can provide insight into the mechanism that contribute

to causes and consequences of PTB. Similarly in a study of 206 placentas, Maccani and colleagues reported the association of DNA methylation of CpG sites in *RUNX3* (runt-related transcription factor 3) with smoking during pregnancy and lower gestational age, which has previously been implicated in long-term exposure to smoking [53].

Most preterm birth studies that examine DNA methylation use blood because of its accessibility. For example, a study using umbilical cord blood examined the association between DNA methylation of imprinted genes and both PTB and infection status [124]. While this study identified no association with PTB, they did note an association between methylation of *PLAGL1* (pleiomorphic adenoma gene-like 1) and chorioamnionitis. DNA methylation of *PLAGL1* has been previously associated with transient neonatal diabetes mellitus and hyperglycemia [124], though it has not yet been linked to immune response. In contrast, a comprehensive evaluation of umbilical cord blood in African Americans by Parets and colleagues identified thousands of CpG sites across the genome that associated with PTB; the associated genes were enriched for numerous development processes [125, Chapter 3]. This study also identified many CpG sites that were associated with PTB [126] and with gestational age in term births [113] in other studies. Finally, a study by Burris and colleagues examined LINE-1 methylation and found that it is more heavily methylated in maternal blood in early pregnancy. They also reported that lower maternal LINE-1 methylation levels in early pregnancy associated with increased risk of PTB while the opposite was true for umbilical cord blood [127]. While this may appear contrary to their previous result [122], DNA methylation is tissue specific, so results from blood and cervical tissue cannot be directly compared. However, each tissue may provide an important window into the biological processes relevant for PTB.

Many hope that DNA methylation studies will yield biomarkers that can be used to screen for preterm birth or its risk factors. One such study explored DNA methylation to diagnose bacterial sepsis, a generalized immune response that is likely to affect children who are born preterm with low birth weight or very low birth weight [128]. A protein encoded by *CALCA* (calcitonin) is proposed as an early detection biomarker for infection status [129]. Therefore, Tendl and colleagues examined CpG sites in the promoter of *CALCA* to assess bacterial sepsis. Though the study was preliminary, they report DNA methylation differences in this region in infants with early onset sepsis and late onset sepsis that were not present in matched controls or neonates with isolated infections. Epigenetic biomarkers have been utilized in a number of diseases that primarily affect adults [130-132]. This promising clinical study supports the role of DNA methylation in obstetrics and neonatal clinical care.

DNA Methylation Studies of Long Term Outcomes of PTB

DNA methylation may provide insight into the long-term effects of PTB. In a comparison of children born preterm and at term, Relton and colleagues measured methylation of numerous genes at birth and reported an association with body size at approximately 9 years [133]; methylation of a CpG site in alkaline phosphatase (*ALPL*) associated with height, and the authors discuss the role of this gene in bone mineralization. Similarly, a longitudinal study comparing DNA methylation across the genome in 12 individuals born preterm to 12 born at term reported numerous methylation differences at birth. Interestingly, some of those CpG sites still distinguished preterm and term birth at 18 years of age [126]. Despite the small sample size, this is an important

preliminary study that shows some evidence that DNA methylation should be further studied in PTB.

Opportunities for Epigenetic Studies of PTB

There are numerous challenges to incorporating DNA methylation into PTB studies, as evidenced by the paucity of literature on the subject. The studies performed to date are conducted with relatively small sample sizes, which may limit a study's power to detect new associations or to replication previously identified associations. While some findings have been reproduced, larger studies will identify differences of more subtle effect sizes. Another challenge is that it is not clear which tissue is most informative in assessing DNA methylation changes in the context of PTB, and therefore this study will focus on whole blood since it is the most readily available. In Chapter 2 I will focus on identifying CpG sites that associate with PTB in maternal blood and DNA methylation patterns shared by maternal-fetal pairs based on PTB status. In Chapter 3, I will identify genes and pathways that are being regulated in umbilical cord blood from fetuses born preterm and at term. Finally in Chapter 4, I will perform a candidate gene study of calcitonin-related polypeptide alpha (*CALCA*) to determine if umbilical cord blood methylation of this gene associates with long-term behavioral consequences. Finally, in Chapter 5, I will summarize and synthesize these studies in the context of the broader PTB field.

Epigenome Wide Association Studies (EWAS)

The Infinium HumanMethylation450 BeadChip, which balances coverage across the genome with per sample cost, is the most widely used method for interrogating DNA methylation in EWAS studies. The array assesses over 485,000 CpG sites covering 99%

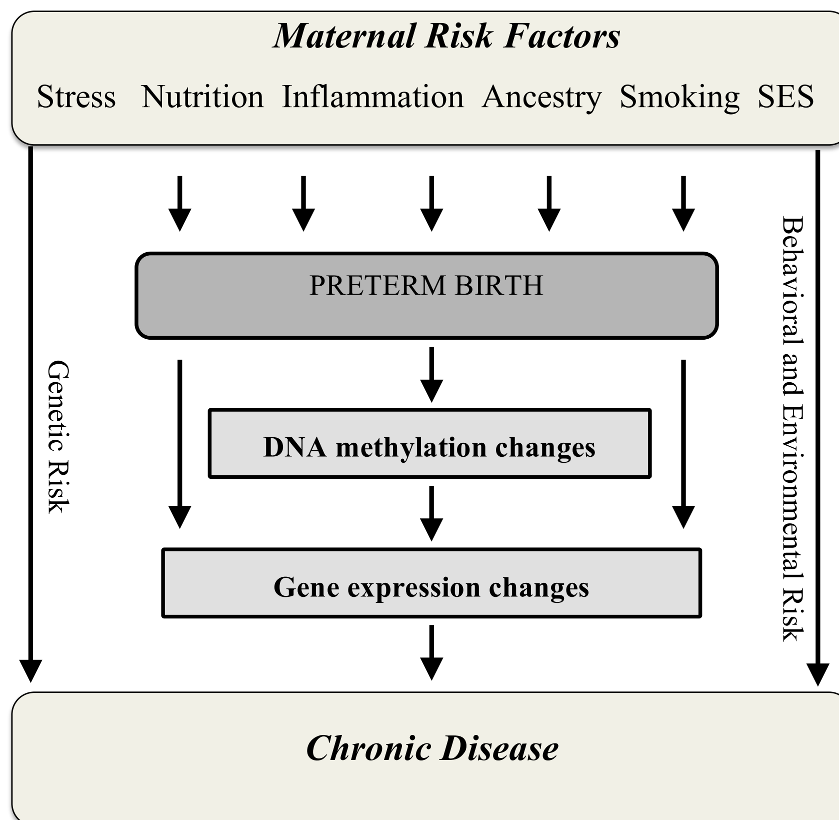
of RefSeq genes. The probes are located throughout the architecture of the gene (transcription start site, 5'UTR, 3'UTR, island, shores, shelves). The Human Methylation450 BeadChip works using two different assays. The first assay type utilizes a two-probe technology in which one oligonucleotide is methylated and the other is not methylated. DNA methylation is then calculated based on the proportion of the amount of bisulfite-converted DNA that binds to each probe type after a single base extension reaction. The other assay uses two-color readout with one bead type. Following a single base extension, it emits a fluorescent signal specific to the incorporated nucleotide. DNA methylation can then be calculated based on the proportion of the methylated or unmethylated fluorescent signal. Utilizing an epigenome-wide approach allows for a broad survey of DNA methylation in the context of the phenotype of interest.

To address the lack of DNA methylation studies for PTB, we preformed a pilot study in the Nashville Birth Cohort to examine genome-wide methylation in paired maternal-fetal samples to lay a foundation and advance our understanding of the causes and consequences of preterm birth.

Table 1-1: Overview of DNA methylation studies of PTB.

Tissue	Design	N	Outcome	Reference
Myometrium	Candidate gene	53	PTB	[121]
Cervical swab	LINE-1 & candidate gene	80	Gestational length	[122]
Amnion	HumanMethylation27	121	PTB & labor	[123]
Placenta	HumanMethylation27	206	Smoking & gestational age	[53]
Cord blood	Candidate gene	181	PTB & infection	[124]
Cord blood	HumanMethylation450	50	PTB & gestational age	[125]
Cord & maternal blood	LINE-1	2393	PTB	[127]
Whole blood at 19 years	Candidate gene	113	SGA	[115]
Blood spots	Candidate gene	49	Bacterial sepsis	[128]
Cord blood	Illumina Cancer Panel 1	178	Child growth	[133]
Blood spots at birth & 18 years	HumanMethylation450	24	PTB	[126]

Figure 1-1: A model for the involvement of DNA methylation in the development of chronic disease following preterm birth (PTB). Multiple maternal risk factors can increase risk of PTB through independent biological mechanisms that may produce changes in DNA methylation or other epigenetic mechanisms. Such risk factors include, but are not limited to, stress, nutrition, immune conditions that produce inflammation, ancestry, smoking and socioeconomic status (SES). These epigenetic changes influence gene expression and thus the developmental trajectory of the neonate. Risk of developing chronic diseases may also be influenced by genetic predisposition independent of PTB or independent of behavioral or environmental factors.



References

1. Hamilton BE, Martin JA, Ventura SJ (2013) Births: preliminary data for 2012. Natl Vital Stat Rep 62: 1-20.

2. Goldenberg RL, Culhane JF, Iams JD, Romero R (2008) Epidemiology and causes of preterm birth. Lancet 371: 75-84.

3. McGowan JE, Alderdice FA, Holmes VA, Johnston L (2011) Early childhood development of late-preterm infants: a systematic review. Pediatrics 127: 1111-1124.

4. Woythaler MA, McCormick MC, Smith VC (2011) Late preterm infants have worse 24-month neurodevelopmental outcomes than term infants. Pediatrics 127: e622-629.

5. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ (2002) Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA 288: 728-737.

6. Kerkhof GF, Willemsen RH, Leunissen RW, Breukhoven PE, Hokken-Koelega AC (2012) Health profile of young adults born preterm: negative effects of rapid weight gain in early life. J Clin Endocrinol Metab 97: 4498-4506.

7. Singhal A, Fewtrell M, Cole TJ, Lucas A (2003) Low nutrient intake and early growth for later insulin resistance in adolescents born preterm. Lancet 361: 1089-1097.

8. Gluckman PD, Hanson MA, Cooper C, Thornburg KL (2008) Effect of in utero and early-life conditions on adult health and disease. N Engl J Med 359: 61-73.

9. Barker DJ (2007) The origins of the developmental origins theory. J Intern Med 261: 412-417.

10. de Boo HA, Harding JE (2006) The developmental origins of adult disease (Barker) hypothesis. Aust N Z J Obstet Gynaecol 46: 4-14.

11. Conaway JW (2012) Introduction to theme "Chromatin, epigenetics, and transcription". Annu Rev Biochem 81: 61-64.

12. Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. Nature 447: 433-440.

13. Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. Genes Dev 25: 1010-1022.

14. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, et al. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41: 178-186.

15. Bonasio R, Tu S, Reinberg D (2010) Molecular signals of epigenetic states. Science 330: 612-616.

16. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, et al. (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 466: 253-257.

17. Menon R (2008) Spontaneous preterm birth, a clinical dilemma: etiologic, pathophysiologic and genetic heterogeneities and racial disparity. Acta Obstet Gynecol Scand 87: 590-600.

18. (2007) Preterm birth causes, consequences, and prevention. . In: Behrman RE BA, editor. Institute of Medicine (US) Understanding premature birth and assuring healthy outcomes. Washington (DC): National Academies Press (US).

19. Ananth CV, Vintzileos AM (2006) Epidemiology of preterm birth and its clinical subtypes. J Matern Fetal Neonatal Med 19: 773-782.

20. Moutquin JM (2003) Classification and heterogeneity of preterm birth. BJOG 110 Suppl 20: 30-33.

21. Committee on Practice Bulletins-Obstetrics TACoO, Gynecologists (2012) Practice bulletin no. 130: prediction and prevention of preterm birth. Obstet Gynecol 120: 964-973.

22. (2013) Practice bulletins No. 139: premature rupture of membranes. Obstet Gynecol 122: 918-930.

23. Mittendorf R, Williams MA, Berkey CS, Cotter PF (1990) The length of uncomplicated human gestation. Obstet Gynecol 75: 929-932.

24. Jukic AM, Baird DD, Weinberg CR, McConnaughey DR, Wilcox AJ (2013) Length of human pregnancy and contributors to its natural variation. Hum Reprod 28: 2848-2855.

25. (2013) ACOG Committee Opinion No 579: Definition of term pregnancy. Obstet Gynecol 122: 1139-1140.

26. Carlin A, Alfirevic Z (2008) Physiological changes of pregnancy and monitoring. Best Pract Res Clin Obstet Gynaecol 22: 801-823.

27. Murphy VE, Smith R, Giles WB, Clifton VL (2006) Endocrine regulation of human fetal growth: the role of the mother, placenta, and fetus. Endocr Rev 27: 141-169.

28. Iqbal M, Audette MC, Petropoulos S, Gibb W, Matthews SG (2012) Placental drug transporters and their role in fetal protection. Placenta 33: 137-142.

29. Bryant-Greenwood GD (1998) The extracellular matrix of the human fetal membranes: structure and function. Placenta 19: 1-11.

30. Jaffe R, Jauniaux E, Hustin J (1997) Maternal circulation in the first-trimester human placenta--myth or reality? Am J Obstet Gynecol 176: 695-705.

31. Palmeira P, Quinello C, Silveira-Lessa AL, Zago CA, Carneiro-Sampaio M (2012) IgG placental transfer in healthy and pathological pregnancies. Clin Dev Immunol 2012: 985646.

32. Arck PC, Hecher K (2013) Fetomaternal immune cross-talk and its consequences for maternal and offspring's health. Nat Med 19: 548-556.

33. Arzuaga BH, Lee BH (2011) Limits of human viability in the United States: a medicolegal review. Pediatrics 128: 1047-1052.

34. Smith LK, Draper ES, Manktelow BN, Dorling JS, Field DJ (2007) Socioeconomic inequalities in very preterm birth rates. Arch Dis Child Fetal Neonatal Ed 92: F11-14.

35. Goldenberg RL, Cliver SP, Mulvihill FX, Hickey CA, Hoffman HJ, et al. (1996) Medical, psychosocial, and behavioral risk factors do not explain the increased risk for low birth weight among black women. Am J Obstet Gynecol 175: 1317-1324.

36. McGrady GA, Sung JF, Rowley DL, Hogue CJ (1992) Preterm delivery and low birth weight among first-born infants of black and white college graduates. Am J Epidemiol 136: 266-276.

37. Hamilton BE, Martin JA, Ventura SJ (2012) Births: preliminary data for 2011. Natl Vital Stat Rep 61: 1-18.

38. Torloni MR, Fortunato SJ, Betran AP, Williams S, Brou L, et al. (2012) Ethnic disparity in spontaneous preterm birth and maternal pre-pregnancy body mass index. Arch Gynecol Obstet 285: 959-966.

39. Menon R, Fortunato SJ, Edwards DR, Williams SM (2010) Association of genetic variants, ethnicity and preterm birth with amniotic fluid cytokine concentrations. Ann Hum Genet 74: 165-183.

40. Menon R, Yu J, Basanta-Henry P, Brou L, Berga SL, et al. (2012) Short fetal leukocyte telomere length and preterm prelabor rupture of the membranes. PLoS One 7: e31136.

41. Fortunato SJ, Menon R, Velez DR, Thorsen P, Williams SM (2008) Racial disparity in maternal-fetal genetic epistasis in spontaneous preterm birth. Am J Obstet Gynecol 198: 666 e661-669; discussion 666 e669-610.

42. Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, et al. (2008) Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. Epigenetics 3: 97-106.

43. McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonte B, et al. (2009) Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. Nat Neurosci 12: 342-348.

44. McKay JA, Waltham KJ, Williams EA, Mathers JC (2011) Folate depletion during pregnancy and lactation reduces genomic DNA methylation in murine adult offspring. Genes Nutr 6: 189-196.

45. McKay JA, Wong YK, Relton CL, Ford D, Mathers JC (2011) Maternal folate supply and sex influence gene-specific DNA methylation in the fetal gut. Mol Nutr Food Res 55: 1717-1723.

46. McKay JA, Xie L, Harris S, Wong YK, Ford D, et al. (2011) Blood as a surrogate marker for tissue-specific DNA methylation and changes due to folate depletion in post-partum female mice. Mol Nutr Food Res 55: 1026-1035.

47. Maloney CA, Hay SM, Rees WD (2007) Folate deficiency during pregnancy impacts on methyl metabolism without affecting global DNA methylation in the rat fetus. Br J Nutr 97: 1090-1098.

48. Hoyo C, Murtha AP, Schildkraut JM, Jirtle RL, Demark-Wahnefried W, et al. (2011) Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. Epigenetics 6: 928-936.

49. Flom JD, Ferris JS, Liao Y, Tehranifar P, Richards CB, et al. (2011) Prenatal smoke exposure and genomic DNA methylation in a multiethnic birth cohort. Cancer Epidemiol Biomarkers Prev 20: 2518-2523.

50. Sun YV, Smith AK, Conneely KN, Chang Q, Li W, et al. (2013) Epigenomic association analysis identifies smoking-related DNA methylation sites in African Americans. Hum Genet 132: 1027-1037.

51. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, et al. (2012) 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ Health Perspect 120: 1425-1431.

52. Breton CV, Byun HM, Wenten M, Pan F, Yang A, et al. (2009) Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med 180: 462-467.

53. Maccani JZ, Koestler DC, Houseman EA, Marsit CJ, Kelsey KT (2013) Placental DNA methylation alterations associated with maternal tobacco smoking at the RUNX3 gene are also associated with gestational age. Epigenomics 5: 619-630.

54. Sun YV, Lazarus A, Smith JA, Chuang YH, Zhao W, et al. (2013) Gene-specific DNA methylation association with serum levels of C-reactive protein in African Americans. PLoS One 8: e73480.

55. Gottlicher M, Minucci S, Zhu P, Kramer OH, Schimpf A, et al. (2001) Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. EMBO J 20: 6969-6978.

56. Phiel CJ, Zhang F, Huang EY, Guenther MG, Lazar MA, et al. (2001) Histone deacetylase is a direct target of valproic acid, a potent anticonvulsant, mood stabilizer, and teratogen. J Biol Chem 276: 36734-36741.

57. Eyal S, Yagen B, Sobol E, Altschuler Y, Shmuel M, et al. (2004) The activity of antiepileptic drugs as histone deacetylase inhibitors. Epilepsia 45: 737-744.

58. Smith AK, Conneely KN, Newport DJ, Kilaru V, Schroeder JW, et al. (2012) Prenatal antiepileptic exposure associates with neonatal DNA methylation differences. Epigenetics 7: 458-463.

59. American College of O, Gynecologists, Task Force on Hypertension in P (2013) Hypertension in pregnancy. Report of the American College of Obstetricians and

Gynecologists' Task Force on Hypertension in Pregnancy. Obstet Gynecol 122: 1122-1131.

60. Mercer BM (2003) Preterm premature rupture of the membranes. Obstet Gynecol 101: 178-193.
61. Lyell DJ, Pullen K, Fuh K, Zamah AM, Caughey AB, et al. (2010) Daily compared with 8-hour gentamicin for the treatment of intrapartum chorioamnionitis: a randomized controlled trial. Obstet Gynecol 115: 344-349.
62. Sibai B, Dekker G, Kupferminc M (2005) Pre-eclampsia. Lancet 365: 785-799.
63. Pennington KA, Schlitt JM, Jackson DL, Schulz LC, Schust DJ (2012) Preeclampsia: multiple approaches for a multifactorial disease. Dis Model Mech 5: 9-18.
64. Cacciatore A, Rapiti S, Carrara S, Cavaliere A, Ermito S, et al. (2009) Obstetric management in Rh alloimmunizated pregnancy. J Prenat Med 3: 25-27.
65. Rodriguez AO, Chew H, Cress R, Xing G, McElvy S, et al. (2008) Evidence of poorer survival in pregnancy-associated breast cancer. Obstet Gynecol 112: 71-78.
66. Simhan HN, Canavan TP (2005) Preterm premature rupture of membranes: diagnosis, evaluation and management strategies. BJOG 112 Suppl 1: 32-37.
67. Latendresse G (2011) Perinatal genomics: Current research on genetic contributions to preterm birth and placental phenotype. Annu Rev Nurs Res 29: 331-351.
68. York TP, Eaves LJ, Neale MC, Strauss JF, 3rd (2014) The contribution of genetic and environmental factors to the duration of pregnancy. Am J Obstet Gynecol 210: 398-405.

69. Mulder EJ, Robles de Medina PG, Huizink AC, Van den Bergh BR, Buitelaar JK, et al. (2002) Prenatal maternal stress: effects on pregnancy and the (unborn) child. Early Hum Dev 70: 3-14.

70. Tsigos C, Chrousos GP (2002) Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress. J Psychosom Res 53: 865-871.

71. Sasaki A, Shinkawa O, Yoshinaga K (1989) Placental corticotropin-releasing hormone may be a stimulator of maternal pituitary adrenocorticotropic hormone secretion in humans. J Clin Invest 84: 1997-2001.

72. Mastorakos G, Ilias I (2003) Maternal and fetal hypothalamic-pituitary-adrenal axes during pregnancy and postpartum. Ann N Y Acad Sci 997: 136-149.

73. Wadhwa PD, Garite TJ, Porto M, Glynn L, Chicz-DeMet A, et al. (2004) Placental corticotropin-releasing hormone (CRH), spontaneous preterm birth, and fetal growth restriction: a prospective investigation. Am J Obstet Gynecol 191: 1063-1069.

74. McLean M, Bisits A, Davies J, Woods R, Lowry P, et al. (1995) A placental clock controlling the length of human pregnancy. Nat Med 1: 460-463.

75. Petraglia F, Sawchenko PE, Rivier J, Vale W (1987) Evidence for local stimulation of ACTH secretion by corticotropin-releasing factor in human placenta. Nature 328: 717-719.

76. Margioris AN, Grino M, Rabin D, Chrousos GP (1988) Human placenta and the hypothalamic-pituitary-adrenal axis. Adv Exp Med Biol 245: 389-398.

77. Smith R, Mesiano S, Chan EC, Brown S, Jaffe RB (1998) Corticotropin-releasing hormone directly and preferentially stimulates dehydroepiandrosterone sulfate

secretion by human fetal adrenal cortical cells. J Clin Endocrinol Metab 83: 2916-2920.

78. Fencl MD, Stillman RJ, Cohen J, Tulchinsky D (1980) Direct evidence of sudden rise in fetal corticoids late in human gestation. Nature 287: 225-226.

79. Muglia L, Jacobson L, Dikkes P, Majzoub JA (1995) Corticotropin-releasing hormone deficiency reveals major fetal but not adult glucocorticoid need. Nature 373: 427-432.

80. Dole N, Savitz DA, Hertz-Picciotto I, Siega-Riz AM, McMahon MJ, et al. (2003) Maternal stress and preterm birth. Am J Epidemiol 157: 14-24.

81. Christiaens I, Zaragoza DB, Guilbert L, Robertson SA, Mitchell BF, et al. (2008) Inflammatory processes in preterm and term parturition. J Reprod Immunol 79: 50-57.

82. Cohen S, Janicki-Deverts D, Doyle WJ, Miller GE, Frank E, et al. (2012) Chronic stress, glucocorticoid receptor resistance, inflammation, and disease risk. Proc Natl Acad Sci U S A 109: 5995-5999.

83. Corwin EJ, Guo Y, Pajer K, Lowe N, McCarthy D, et al. (2013) Immune dysregulation and glucocorticoid resistance in minority and low income pregnant women. Psychoneuroendocrinology 38: 1786-1796.

84. Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, et al. (2012) Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A 109 Suppl 2: 17253-17260.

85. Tehranifar P, Wu HC, Fan X, Flom JD, Ferris JS, et al. (2013) Early life socioeconomic factors and genomic DNA methylation in mid-life. Epigenetics 8: 23-27.

86. Reinhard G, Noll A, Schlebusch H, Mallmann P, Ruecker AV (1998) Shifts in the TH1/TH2 balance during human pregnancy correlate with apoptotic changes. Biochem Biophys Res Commun 245: 933-938.

87. Mor G, Cardenas I (2010) The immune system in pregnancy: a unique complexity. Am J Reprod Immunol 63: 425-433.

88. Sykes L, MacIntyre DA, Yap XJ, Teoh TG, Bennett PR (2012) The Th1:th2 dichotomy of pregnancy and preterm labour. Mediators Inflamm 2012: 967629.

89. Berger A (2000) Th1 and Th2 responses: what are they? BMJ 321: 424.

90. Veenstra van Nieuwenhoven AL, Heineman MJ, Faas MM (2003) The immunology of successful pregnancy. Hum Reprod Update 9: 347-357.

91. Mor G, Cardenas I, Abrahams V, Guller S (2011) Inflammation and pregnancy: the role of the immune system at the implantation site. Ann N Y Acad Sci 1221: 80-87.

92. Olson DM (2003) The role of prostaglandins in the initiation of parturition. Best Pract Res Clin Obstet Gynaecol 17: 717-730.

93. Unal ER, Cierny JT, Roedner C, Newman R, Goetzl L (2011) Maternal inflammation in spontaneous term labor. Am J Obstet Gynecol 204: 223 e221-225.

94. Shynlova O, Lee YH, Srikhajon K, Lye SJ (2013) Physiologic uterine inflammation and labor onset: integration of endocrine and mechanical signals. Reprod Sci 20: 154-167.

95. Gibbs RS, Romero R, Hillier SL, Eschenbach DA, Sweet RL (1992) A review of premature birth and subclinical infection. Am J Obstet Gynecol 166: 1515-1528.

96. Jeffcoat MK, Geurs NC, Reddy MS, Cliver SP, Goldenberg RL, et al. (2001) Periodontal infection and preterm birth: results of a prospective study. J Am Dent Assoc 132: 875-880.

97. Mussalli GM, Blanchard R, Brunnert SR, Hirsch E (1999) Inflammatory cytokines in a murine model of infection-induced preterm labor: cause or effect? J Soc Gynecol Investig 6: 188-195.

98. Forslund M, Bjerre I (1985) Growth and development in preterm infants during the first 18 months. Early Hum Dev 10: 201-216.

99. Euser AM, de Wit CC, Finken MJ, Rijken M, Wit JM (2008) Growth of preterm born children. Horm Res 70: 319-328.

100. Itabashi K, Mishina J, Tada H, Sakurai M, Nanri Y, et al. (2007) Longitudinal follow-up of height up to five years of age in infants born preterm small for gestational age; comparison to full-term small for gestational age infants. Early Hum Dev 83: 327-333.

101. Gong YH, Ji CY, Shan JP (2013) A Longitudinal Study on the Catch-Up Growth of Preterm and Term Infants of Low, Appropriate, and High Birth Weight. Asia Pac J Public Health.

102. Knops NB, Sneeuw KC, Brand R, Hille ET, den Ouden AL, et al. (2005) Catch-up growth up to ten years of age in children born very preterm or with very low birth weight. BMC Pediatr 5: 26.

103. Stein AD, Barros FC, Bhargava SK, Hao W, Horta BL, et al. (2013) Birth status, child growth, and adult outcomes in low- and middle-income countries. J Pediatr 163: 1740-1746 e1744.

104. Vucetic Z, Kimmel J, Totoki K, Hollenbeck E, Reyes TM (2010) Maternal high-fat diet alters methylation and gene expression of dopamine and opioid-related genes. Endocrinology 151: 4756-4764.

105. Gallou-Kabani C, Gabory A, Tost J, Karimi M, Mayeur S, et al. (2010) Sex- and diet-specific changes of imprinted gene expression and DNA methylation in mouse placenta under a high-fat diet. PLoS One 5: e14398.

106. Sohi G, Marchand K, Revesz A, Arany E, Hardy DB (2011) Maternal protein restriction elevates cholesterol in adult rat offspring due to repressive changes in histone modifications at the cholesterol 7alpha-hydroxylase promoter. Mol Endocrinol 25: 785-798.

107. El Marroun H, Zeegers M, Steegers EA, van der Ende J, Schenk JJ, et al. (2012) Post-term birth and the risk of behavioural and emotional problems in early childhood. Int J Epidemiol 41: 773-781.

108. Moore GP, Lemyre B, Barrowman N, Daboval T (2013) Neurodevelopmental outcomes at 4 to 8 years of children born at 22 to 25 weeks' gestational age: a meta-analysis. JAMA Pediatr 167: 967-974.

109. Wocadlo C, Rieger I (2008) Motor impairment and low achievement in very preterm children at eight years of age. Early Hum Dev 84: 769-776.

110. Moster D, Lie RT, Markestad T (2008) Long-term medical and social consequences of preterm birth. N Engl J Med 359: 262-273.

111. Saigal S, Doyle LW (2008) An overview of mortality and sequelae of preterm birth from infancy to adulthood. Lancet 371: 261-269.

112. Novakovic B, Yuen RK, Gordon L, Penaherrera MS, Sharkey A, et al. (2011) Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors. BMC Genomics 12: 529.

113. Schroeder JW, Conneely KN, Cubells JC, Kilaru V, Newport DJ, et al. Neonatal DNA methylation patterns associate with gestational age. Epigenetics 6: 1498-1504.

114. Banister CE, Koestler DC, Maccani MA, Padbury JF, Houseman EA, et al. (2011) Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. Epigenetics 6: 920-927.

115. Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, et al. (2011) DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. Epigenetics 6: 171-176.

116. Menon R, Conneely KN, Smith AK (2012) DNA methylation: an epigenetic risk factor in preterm birth. Reprod Sci 19: 6-13.

117. Dhobale M, Joshi S (2012) Altered maternal micronutrients (folic acid, vitamin B(12)) and omega 3 fatty acids through oxidative stress may reduce neurotrophic factors in preterm pregnancy. J Matern Fetal Neonatal Med 25: 317-323.

118. Robins JC, Marsit CJ, Padbury JF, Sharma SS (2011) Endocrine disruptors, environmental oxygen, epigenetics and pregnancy. Front Biosci (Elite Ed) 3: 690-700.

119. Cutfield WS, Hofman PL, Mitchell M, Morison IM (2007) Could epigenetics play a role in the developmental origins of health and disease? Pediatr Res 61: 68R-75R.

120. Aguilar HN, Mitchell BF (2010) Physiological pathways and molecular mechanisms regulating uterine contractility. Hum Reprod Update 16: 725-744.

121. Mitsuya K, Singh N, Sooranna SR, Johnson MR, Myatt L (2014) Epigenetics of Human Myometrium: DNA Methylation of Genes Encoding Contraction-Associated Proteins in Term and Preterm Labor. Biol Reprod.

122. Burris HH, Baccarelli AA, Motta V, Byun HM, Just AC, et al. (2014) Association between length of gestation and cervical DNA methylation of PTGER2 and LINE 1-HS. Epigenetics 9.

123. Kim J, Pitlick MM, Christine PJ, Schaefer AR, Saleme C, et al. (2013) Genome-Wide Analysis of DNA Methylation in Human Amnion. The Scientific World Journal 2013: 11.

124. Liu Y HC, Murphy S, Huang Z, Overcash F, Thompson J, Brown H, Murtha AP (2013) DNA methylation at imprint regulatory regions in preterm birth and infection. Am J Obst Gynecol pii S0002-9378: 00142–00147.

125. Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, et al. (2013) Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age. PLoS One 8: e67489.

126. Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, et al. (2013) Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. Genome Med 5: 96.

127. Burris HH, Rifas-Shiman SL, Baccarelli A, Tarantini L, Boeke CE, et al. (2012) Associations of LINE-1 DNA Methylation with Preterm Birth in a Prospective Cohort Study. J Dev Orig Health Dis 3: 173-181.

128. Tendl KA, Schulz SM, Mechtler TP, Bohn A, Metz T, et al. (2013) DNA methylation pattern of CALCA in preterm neonates with bacterial sepsis as a putative epigenetic biomarker. Epigenetics 8: 1261-1267.

129. Matwiyoff GN, Prahl JD, Miller RJ, Carmichael JJ, Amundson DE, et al. (2012) Immune regulation of procalcitonin: a biomarker and mediator of infection. Inflamm Res 61: 401-409.

130. Fuchikami M, Morinobu S, Segawa M, Okamoto Y, Yamawaki S, et al. (2011) DNA methylation profiles of the brain-derived neurotrophic factor (BDNF) gene as a potent diagnostic biomarker in major depression. PLoS One 6: e23881.

131. Mikeska T, Bock C, Do H, Dobrovic A (2012) DNA methylation biomarkers in cancer: progress towards clinical implementation. Expert Rev Mol Diagn 12: 473-487.

132. Philibert RA, Beach SR, Lei MK, Brody GH (2013) Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. Clin Epigenetics 5: 19.

133. Relton CL, Groom A, St Pourcain B, Sayers AE, Swan DC, et al. (2012) DNA methylation patterns in cord blood DNA and body size in childhood. PLoS One 7: e31821.

Chapter 2

DNA Methylation Provides Insight into Intergenerational Risk for Preterm Birth in African Americans

Information presented in this chapter is currently under review as cited below:

Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, Saade G, Menon R, Smith AK. (2015) DNA Methylation in African American Women: Intergenerational Risk for Preterm Delivery. Submitted to Epigenetics

Introduction

Though the overall rate of preterm birth has slightly decreased in recent years, African Americans have more than 1.5 times the risk of spontaneous preterm birth (PTB; <37 weeks gestation) and more than twice the risk of early PTB (< 32 weeks) when compared to Caucasians [1-3]. Studies to date have identified numerous maternal risk factors for PTB [1,4-7], such as low socioeconomic status (SES), but less than half of the increased risk in African Americans is explained by SES and other known risk factors (Figure 1-1) [8-10].

Personal and family history of PTB are the greatest risk factors for PTB, and studies estimate its heritability from 17- 30% [11-17]. However, genetic studies have not identified variants that account for this intergenerational risk [1,11,18-21], prompting the hypothesis that epigenetic factors may also contribute to PTB [17,22-24]. Few studies have evaluated the epigenetics of PTB, and those that have focus mostly on those born preterm [25-28]. Other studies focus on the short-term and long-term consequences of PTB for the neonate [1,29-32], in part, because of interest in the developmental origin of health and disease (DOHaD) hypothesis [32,33]. Because of the complexity of the maternal-fetal relationship it is not clear if DNA methylation patterns predictive of PTB may be identified in maternal or fetal samples. It is equally possible that mother who delivers PTB may share common epigenetic patterns with their fetuses.

We recently evaluated DNA methylation in African American cord blood samples and identified thousands of DNA methylation differences between preterm (N=22) and term fetuses (N=28) [25]. These DNA methylation differences may underlie some of the consequences associated with being born preterm, though longitudinal studies will be

more informative for determining whether methylation differences observed at birth have long-term consequences or whether they simply reflect developmental differences. Cruickshank and colleagues performed one such study in 12 PTB cases and 12 matched controls [28]. They evaluated DNA methylation from blood at birth and at 18 years and observed substantial overlap with the PTB-associated CpGs reported birth for both studies. However, the majority DNA methylation differences observed at birth were no longer associated with PTB status at 18 years for the majority of CpG sites examined. They did find 10 CpGs that continue to differ in methylation at both time points, suggesting the potential for a long-term epigenetic signature of PTB. These CpG sites should be further evaluated to determine if they have any predictive value for consequences of being born preterm. If they are replicated, these CpGs may serve as a biological indicator of early risk.

No study has examined genome-wide DNA methylation in the blood of women who deliver preterm. However, studies have recently begun documenting the long-term implications of delivering preterm for maternal health. Women who deliver preterm are at increased risk to develop cardiovascular and other chronic disorders as they age [34-42]. For example, a series of studies demonstrate that mothers who deliver very preterm are at subsequent risk for Type 2 diabetes [38,43]. The first, conducted in primarily Caucasian women from the Nurses' Health Study II, reports that women who deliver very early preterm are more likely to be diagnosed with Type 2 diabetes in the decade following pregnancy [43]. A second investigation of ~30,000 women from the Black Women's Health Study also reports that early preterm birth associates with a higher risk of developing Type 2 diabetes even after correcting for age at first birth, family history of

diabetes, education, personal history of preterm birth, and body mass index [38]. They then demonstrated that the increased diabetes risk was independent of gestational diabetes history, consistent with other studies [39,40].

The mechanism underlying the relationship between PTB and the development of chronic disorders later in life is not yet clear, but some suggest that inflammation or immune dysregulation may increase risk for PTB and other chronic disorders [38,44]. DNA methylation patterns regulate the functional properties of immune cells [45,46] and associate with inflammatory markers [47], chronic disorders [48-50] and PTB [27,51]. It is also possible that mothers and fetuses delivered preterm may share correlated DNA methylation patterns that provide insight into common genetic or environmental risk. We hypothesize that DNA methylation patterns may reveal genes whose regulation is unique to women who deliver preterm or provide insight into its intergenerational risk.

Methods

Nashville Birth Cohort

All subjects were recruited at Centennial Women's Hospital and the Perinatal Research Center in Nashville, TN beginning in 2003 as part of the Nashville Birth Cohort (NBC) that was established to examine biological risk factors that distinguish spontaneous preterm from term labor. Pregnant women were enrolled at the time of admission for labor at preterm or term after obtaining informed consent. Maternal demographic and clinical data (race, socioeconomic education, household income, marital status, cigarette smoking) were recorded from medical records or by interviews during the consenting process. Demographic and clinical data specific to the fetus were collected from clinical records. Gestational age (GA) was determined by maternal

reporting of the last menstrual period and corroboration by ultrasound dating. Birth weight percentile was based on GA in accordance with the United States national reference [52]. Race was identified by self-reporting that traced back to three generations from maternal and paternal side of the fetus. Only African Americans of non-Hispanic ethnicity were included in this study.

Subjects were included in this study if they had contractions (rate of 2 contractions/10 minutes) leading to delivery either at preterm or at term. Cases delivered preterm with intact membranes between $24^{1/7}$ weeks and $34^{0/7}$ weeks. Controls delivered ($>39^{0/7}$ weeks) with spontaneous term labor and delivery and no current or history of pregnancy-related complications including preterm birth and preterm or prelabor rupture of the membranes (pPROM). Subjects who had multiple gestations, preeclampsia, placenta-previa, fetal anomalies, and/or medical or surgical complications during pregnancy were excluded from the study. Subjects with any surgical procedures during pregnancy were treated for preterm labor or for suspected intra-amniotic infection and delivered at term were excluded from the control group. This study was conducted in accordance with the Helsinki Declaration of 1975.

Biological Sample Collection and DNA Extraction

Maternal peripheral blood samples were collected in EDTA tubes at time of admission for labor. Blood samples were centrifuged at 3,000 RPM to separate plasma, and buffy coats were aliquoted and stored at −80°C. DNA was extracted using the Autopure automated system (Gentra Systems, Minneapolis, MN).

DNA Methylation Analysis

For each subject, >485,000 CpG sites across the genome were interrogated using the HumanMethylation450 BeadChip (Illumina, San Diego, CA). Briefly, 1 ug of DNA was converted with sodium bisulfite, amplified, fragmented, and hybridized on the BeadChip according to the manufacturer's instructions. CpGassoc was used to perform quality control and calculate ß values [53]. Data points with probe detection p-values >.001 were set to missing, and CpG sites with missing data for >10% of samples were excluded from analysis; 479,808 CpG sites passed the above criteria. Samples with probe detection call rates <90% and those with an average intensity value of either <50% of the experiment-wide sample mean or <2,000 arbitrary units (AU) were excluded from further analysis. One sample of female DNA was included on each BeadChip as a technical control throughout the experiment and assessed for reproducibility using the Pearson correlation coefficient, to ensure that Pearson correlation coefficient >0.99 for all pairwise comparisons of technical replicates. For each individual sample and CpG site, the signals from methylated (M) and unmethylated (U) bead types were used to calculate a beta value as ß = M/(U+M).

Statistical Analysis

MethLAB was used to test for association with PTB via linear regressions that modeled β-values as the outcome and PTB as the independent variable, adjusting for maternal age, cell composition and positional effects in the array as covariates [54]. Cell type proportions were estimated using publically available data (GSE36069) as a reference panel for applying the method described by Houseman and colleagues [55,56] to our data. We examined the association between methylation of each CpG site and potential confounding factors including: birth weight percentile, gravidity, parity,

infection and smoking. In a univariate analysis, these factors did not associate with methylation of any CpG site after correction for multiple testing, thus, they were not included as covariates in the final models. Logit transformation of the β values (i.e. M values) did not substantially alter the results so analyses of untransformed β are presented to ease biological interpretation and to make comparisons to our previous study [25]. For all genome-wide analyses, the False Discovery Rate (FDR) was controlled at 5% using Storey's q-value [57]. For all replication analyses, we set the significance threshold at a nominal $p<.05$.

To evaluate the relationship between DNA methylation in maternal and fetal samples, linear regressions compared β values for each maternal sample (predictor) to those of her fetus (outcome) for each CpG site while accounting for positional effects on the array and cellular proportions. To compare the observed correlations to what would be observed if maternal and fetal methylation were completely independent, we repeated the analysis comparing each mother to an unrelated fetus that was matched for case status, positional effects, and fetal sex. We also performed an exploratory analysis that evaluated the relationship between methylation in preterm and term pairs separately.

The location of each CpG site was determined using the Illumina array annotation for the HumanMethylation450 BeadChip based on build 37 of the human genome. Chi-square tests were used to compare the number of correlated CpG sites that did or did not occur in a particular gene region (e.g. promoter, 5′UTR, Body, 1st exon, 3′UTR, or intragenic regions) to the sites not associated with PTB in that gene region. We performed similar tests of enrichment for regions characterized by CpG density (islands, shores, shelves and open seas). DAVID was used to evaluate whether groups of CpG

sites were in genes enriched for any specific biological pathways and focused specifically on KEGG pathways [58,59].

To determine if genetic variation influenced DNA methylation, methylation quantitative trait loci (meQTL) were identified by applying the approach described previously [60] to the methylation data from African American subject in the Grady Trauma Project [47,61-63]. Briefly, the relationship between the proportion of methylation at each CpG site and each SNP within 50 kb of that site was examined via linear regression, where methylation was modeled as a linear function of the number of reference alleles (0, 1, or 2). CpG sites were excluded from the meQTL analysis if the probe sequence contained a SNP with a minor allele frequency greater than 1% in any population, as identified from the 1000 genomes project (TGP). In total, 98,741 CpG sites had a TGP SNP within its probe sequence, and an additional 74,712 were meQTLs in an African American cohort (FDR<.05). We then plotted the odds ratio of whether correlated CpG sites were enriched in meQTLs at varying significance levels.

Gene Expression

To evaluate whether correlated CpG sites were located in genes whose expression is also correlated, we used publicly available gene expression data from maternal-fetal pairs (GSE27272) [64]. Expression of total RNA in umbilical cord blood and maternal peripheral blood was evaluated using the HumanRef-8v3.0 BeadChip (Illumina). The data were extracted using Illumina's BeadStudio Software v3 and then quantile normalized using Lumi [65]. For each gene containing a correlated CpG site, linear regressions were used to compare expression of each maternal sample to that of her fetus.

Results

The cohort is comprised of African American women who deliver early preterm (GA range 24.1–34.0 weeks) and at term (39.0–40.9 weeks). As expected, the groups differed by GA and birthweight, but did not differ significantly by any other demographic or clinical factor (Table 2-1).

Association between maternal DNA methylation and PTB

First, we examined the association between DNA methylation at each CpG site and PTB. Overall, 17,829 CpG sites associated with PTB ($1.83 \times 10^{-6} < p < .05$), but none remained significantly associated after correction for multiple testing (FDR<.05). Among the CpG sites with the strongest association were two (cg22486214, cg16980736) in regulatory associated protein of MTOR (*RPTOR*; $2.20 \times 10^{-5} < p < 1.03 \times 10^{-4}$).

We have previously reported 9,637 CpG sites that associate with gestational age (GA) in the umbilical cord blood of fetuses born to this cohort of women [25]. CpG sites that associate with PTB in maternal blood are more likely to associate with GA in fetuses when compared to those that do not associate with PTB (5.2% vs. 3.5%; $p < 2.2 \times 10^{-16}$), suggesting that there may be epigenetic factors shared between mothers who deliver preterm and their fetuses.

Correlation between maternal and fetal methylation

We next evaluated the relationship between maternal and fetal DNA methylation across the genome. We identified 5,171 CpG sites in which maternal methylation predicted umbilical cord blood methylation (FDR<.05; Figure 2-1), 98.8% of which occurred in the same direction. To empirically assess whether these differences are greater than expected by chance, we repeated the analysis comparing methylation from each mother to methylation from an unrelated fetus that was matched for PTB status and

sex. Only 35 CpG sites associated between unrelated pairs (FDR<.05), suggesting that the high degree of correlation observed between a mother and her fetus is substantially greater than expected by chance.

A recent study by McRae and colleagues [66] evaluated heritability of DNA methylation patterns in a multigenerational cohort and observed that sequence variation accounts for the majority of intergenerational inheritance. The heritability of non-genetic CpG sites may provide insight into environmental factors shared between a mother and her fetus. Of the 7,013 CpG sites with heritability of 0.8 or greater that they identified, 52.3% also demonstrated heritability in this cohort ($p<0.05$). Among the heritable CpG sites reported not to be influenced by sequence variants (N=3,078), 45.8% replicated in this cohort. However, 2,819 (91.6%) could be attributed to sequence variation in African Americans.

There is a wide range of variation in methylation levels across the 5,171 correlated CpG sites (Figure 2-2A), that is consistent with distribution of variation in all CpG sites assessed on the array (Figure 2-2B). Correlated CpG sites were more likely to occur in regions of low CpG density (i.e. shelves and open seas) and less likely occur in regulatory regions near the transcription start site (i.e. promoters, 1st exon, and 5'UTRs) when compared to uncorrelated CpG sites ($1.5\times10^{-6}<p<2.2\times10^{-16;}$ Table 2-2). They were also more likely to be located in genes involved in metabolic (i.e. type 1 and type 2 diabetes mellitus), cardiovascular (i.e. viral myocarditis, arrhythmogenic right ventricular cardiomyopathy) and immune (i.e. graft-versus host disease, allograft rejection) pathways (Table 3). The majority of correlated CpG sites (3,857; 74.6%) could be attributed to one or more genetic variants, defined as a SNP that either overlapped with the CpG's probe

sequence (1393; 26.9%) or associated with the CpG as a methylation quantitative trait locus (meQTL; 2464; 47.7%). We also found that CpG sites that correlated between a mother and her fetus were enriched for meQTLs at increasing levels of significance (Figure 2-3), consistent with the results of McRae et al.. The remaining 1,314 CpG sites (25.4%) could not be attributed to genetic factors and may reflect the shared intrauterine environment. Both classes of CpG sites support the pathways identified in the combined analysis when evaluated individually (Table 2-3). Evaluation of the 200 genes containing CpG sites whose methylation levels are both predictive of cord blood methylation and associate with PTB in maternal samples also reveal enrichment for genes in the type 2 diabetes mellitus pathway (p=.015).

Correlated CpG sites were more likely than uncorrelated sites to associate with PTB in maternal samples (Table 2-2; OR=1.7; $p<2.2\times10^{-16}$), but there was no difference in the rates of CpG sites influenced by genetic versus non-genetic factors among those associated with PTB (p=.41). Thus, we re-examined the relationship between maternal and fetal DNA methylation separately in PTB and term birth pairs. There were 79 CpG sites that correlated (FDR<.05) in the PTB pairs, 57 (72.2%) of which were unique to the PTB samples. PTB-specific sites were enriched in genes involved in vascular smooth muscle contraction (KEGG:04270; p=.037). Although, this may simply reflect common changes to the myometrium based on gestational weeks. Consistent with the results of the combined analysis, the majority (70.2%) of the correlated CpG sites specific to PTB were attributable to genetic variation.

Correlation Between Maternal and Fetal Gene Expression

We next examined whether genes containing correlated CpG sites have correlated gene expression levels. Correlated CpG sites (N=5,171) were located in 3,297 expression probes, representing 2,282 unique genes and 1015 transcript variants. Maternal expression of 738 transcripts (22.3%; $2.3\times10^{-14}<p<.05$) predicted fetal expression, 357 (10.8%) of which remained associated after correction for multiple tests (FDR<.05). The distribution of the expression of both maternal blood and cord blood is similar to what was seen in our methylation data (Figure 2-4). For example, methylation of CpG sites in *MICB* associate with PTB in maternal blood and correlate with methylation in fetal blood (Figure 2-5A, 2-5B). Maternal *MICB* expression also predicts fetal *MICB* expression ($p=1.09\times10^{-3}$; Figure 2-5C). Pathway analysis of genes with correlated expression levels (FDR<.05) were consistent with the results of those performed with correlated CpG sites (data not shown).

Discussion

In this study, no CpG site of large effect size was associated with PTB in African American women who delivered preterm, though thousand of CpG sites were nominally significant. Only 5.2% of the CpG sites that were associated with GA in the cord blood of fetuses born preterm also associated with PTB in maternal samples. This is only slightly more than what would be expected by chance, suggesting that the majority of CpG sites associated with GA in fetuses may reflect developmental differences. Nevertheless, our study does provide insight into correlated methylation and expression patterns in maternal-fetal pairs that were enriched in genes involved in PTB and chronic disease risk.

One of the biggest risk factors for PTB is a prior history or a family history. We identified 5,171 CpG sites in which maternal methylation predicts fetal methylation. The

vast majority of correlated CpG sites (98.8%) occurred in the same direction, consistent with a high degree of genetic and environmental similarity in these pairs. These correlated CpG sites were enriched in areas of low CpG density, regions of high inter-individual variation that are more likely to associate with environmental factors and complex diseases [67]. Though this study did not specifically evaluate sequence variation, methylation of almost 75% of correlated CpG sites could be attributed to genetic variation such as a SNP or meQTL. The results of this study were consistent with those reported in a large multigenerational cohort of Caucasians [66], which determined that heritable CpG sites were primarily under genetic influence. African Americans and Caucasians have distinct patterns of genetic-epigenetic correlation [60] that may contribute to the increased risk for PTB and other disorders more common in African Americans.

Correlated CpG sites were also enriched among genes whose expression levels were correlated in maternal-fetal pairs, providing a potential mechanism linking correlated methylation in women who deliver preterm to biological differences. For example, *MICB* is part of the MHC class I chain, and is induced by cellular stress to initiate an immune response [68]. Activation of inflammatory pathways has been implicated in the timing of parturition specifically in PTB [11], and we observed lower methylation of CpGs in *MICB* in PTB as well as correlated methylation and expression patterns in maternal-fetal pairs. These results suggest a complex relationship between sequence variation, DNA methylation and gene expression that should be considered in future studies of PTB.

Correlated CpG sites were enriched in genes involved in chronic disorders that are common in African Americans [1,34-36,69-72], suggesting an epigenetic link between PTB and metabolic, cardiovascular and immune dysregulation across generations. Though we made every effort to limit inclusion of clinical factors that could influence these results, such as gestational diabetes or preeclampsia, the implications of these findings are difficult to interpret. However, the results suggest that chronic disorders diagnosed subsequent to PTB may not be limited to those indicated by gestational diabetes or hypertension. On the contrary, spontaneous PTB with unknown etiology may also increase lifetime risk for chronic disorders. In general, African American women have higher levels of inflammation when compared to Caucasian women, and chronic inflammation has been presented as a potential mechanism through which PTB and other chronic conditions occur [11,17,24,38,44,70,73]. For example, Liu and colleagues examined DNA methylation of 8 imprinted genes in umbilical cord blood samples for association with PTB and infection status [26]. Though they did not find any association with PTB, they reported that *PLAGL1* DNA methylation associates with chorioamnionitis. Consistent with these results, correlated CpG sites exclusive to PTB pairs were also identified in genes involved in immune regulation. Other environmental factors such as high BMI, stress, smoking and infection also increase inflammation and PTB risk, and DNA methylation may mediate these relationships or serve as a potential marker of underlying inflammation. These factors did not associate with DNA methylation in this cohort although future studies should include them when evaluating DNA methylation and inflammation.

This study has a number of strengths and limitations. The primary limitation is the sample size, which is in part due to the fact that we restricted the design to only African Americans with spontaneous PTB prior to 34 weeks gestation and uncomplicated controls. In this study, we did not identify any individual CpG sites that associated with PTB in maternal blood. Thus, it is reasonable to conclude that there are no CpG sites on the array that associate strongly with PTB, though evaluations of larger cohorts may reveal associations of more subtle effect. However, our study was well powered to detect CpG sites whose methylation and expression levels correlated in maternal-fetal pairs.

This is the first epigenetic study of maternal-fetal pairs for PTB and the first study of heritable CpG sites in African Americans, an understudied population with an increased risk of PTB. The results of this study support a complex genetic and environmental relationship underlying the intergenerational risk for PTB and are consistent with the hypothesis that pregnancy complications, including spontaneous PTB, may be an early indicator of future risk for mothers as well as their fetuses. Future studies should prospectively examine women who are at high risk for PTB throughout pregnancy and beyond.

Table 2-1: Demographics table for maternal samples

Phenotype	PTB (N=16) Mean ± SD	TB (N=24) Mean ± SD	p-value
Maternal Age	24.69 ± 4.7	24.13 ± 6.2	NS
Weeks Gestation	30.2 ±3 .6	39.9 ± 0.4	<.0001
Parity	1.2 ± 1.4	1.4 ± 1.7	NS
BMI	29.93 ± 8.4	26.46 ± 5.9	NS
Household Income			
<15K	(10) 66.7%	(11) 45.8%	NS
15-30K	(2) 13.3%	(7) 29.2%	NS
>30K	(3) 20%	(6) 25%	NS
Smoking	(4) 25%	(5) 21%	NS
Married	(5) 33.3%	(5) 20.8%	NS
Employed	(1) 7.1%	(6) 42.9%	NS
Chorioamnionitis	(9) 56%	NA	NS
Granulocytes	74.7 ± 9.2	69.6 ± 7.8	NS
Monocytes	9.8 ± 3.7	10.7 ± 3.2	NS

Table 2-2: Enrichment for correlation analysis of maternal and fetal blood

	Correlated	Not Correlated	p-value
CpG Islands	17.1%	31.3%	$< 2.2\times10^{-16}$
Shores	21.9%	23.2%	1.5×10^{-6}
Shelves	12.1%	9.7%	9.0×10^{-9}
Open Sea	49.0%	35.7%	$< 2.2\times10^{-16}$
Promoter	17.3%	25.2%	$< 2.2\times10^{-16}$
5'UTR	6.0%	8.9%	9.9×10^{-13}
1st Exon	2.4%	4.7%	8.9×10^{-15}
Gene Body	37.0%	33.2%	7.2×10^{-9}
3'UTR	4.9%	3.6%	1.2×10^{-6}
Intragenic	32.4%	24.3%	$< 2.2\times10^{-16}$
PTB	6.5%	3.8%	$<2.2\times10^{-16}$

Table 2-3: Pathway analysis of CpG sites in genes that correlate in maternal-fetal pairs. The enrichment p-value indicates whether genes with correlated CpG sites are more likely to occur in the indicated biological pathway. The group of correlated CpGs is also stratified by whether or not they can be attributed to genetic variation. NS indicates that the p-value is not significant.

	KEGG ID	Enrichment p-value	Enrichment p-value	Enrichment p-value
		Correlated N=2639	Genetic N=1985	Non-genetic N=831
Metabolic				
Type 1 diabetes mellitus	04940	.006	.018	NS
Fructose and mannose metabolism	00051	.028	.043	NS
Ether lipid metabolism	00565	.034	.017	NS
Type 2 diabetes mellitus	04930	.038		NS
Glycolysis/Gluconeogenesis	00010	.045	NS	NS
Cardiovascular				
Arrhythmogenic right ventricular cardiomyopathy	05412	.004	.031	.012
Hypertrophic cardiomyopathy	05410	.013	NS	.022
Dilated cardiomyopathy	05414	.015		.033
Viral myocarditis	05416	.020	.041	NS
Immune				
Natural killer cell mediated cytotoxicity	04650	.013	.018	NS
Fc gamma R-mediated phagocytosis	04666	.021	.021	NS
Graft-versus-host disease	05332	.026	.011	NS
Allograft rejection	05330	.040	NS	NS
Other				
Tight Junction	04530	$9.8x10^{-4}$	.001	NS
Endocytosis	04144	.001	.006	.036
Focal adhesion	04510	.012	.011	.034
Cell adhesion molecules	04514	.021	.030	NS
ECM-receptor interaction	04512	NS	NS	.020
Regulation of actin cytoskeleton	04810	NS	NS	.028

Figure 2-1: Manhattan plot of the relationship between maternal and fetal DNA methylation. The x-axis is the position of each CpG site by chromosome. The y-axis is the negative $\log_{10}$ of the p-value for association between maternal and fetal methylation. The red line indicates experiment-wide significance (FDR < .05).

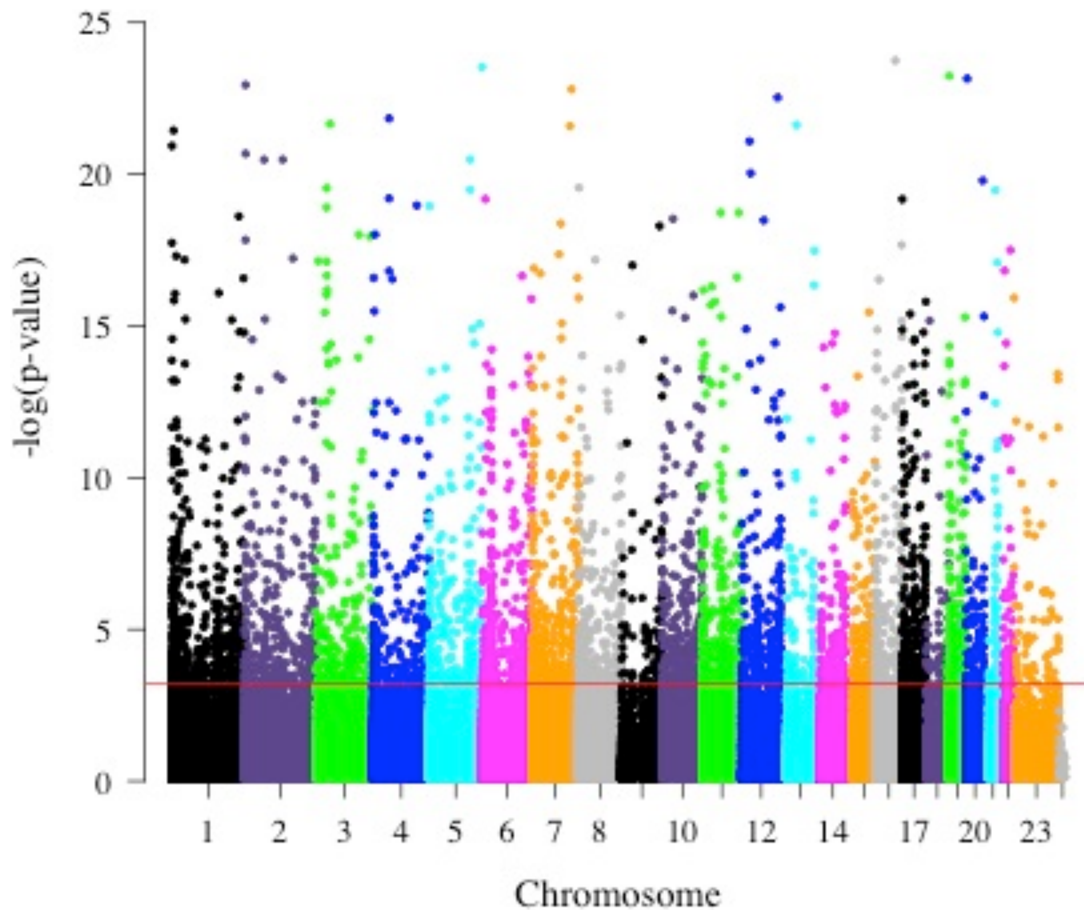


Figure 2-2: Distribution of maternal DNA methylation for CpG sites. The x-axis is the standard deviation (SD) of each CpG site's beta values in the maternal samples. The y-axis indicates the proportion of CpG sites in each SD category. Black represents CpG sites that may be attributed to genetic variation while grey represents CpG sites that cannot be attributed to genetic variation. Graph (A) depicts correlated CpG sites (n=5,171) and their distribution. Graph (B) depicts all CpG sites (n=479,808) and their distribution.

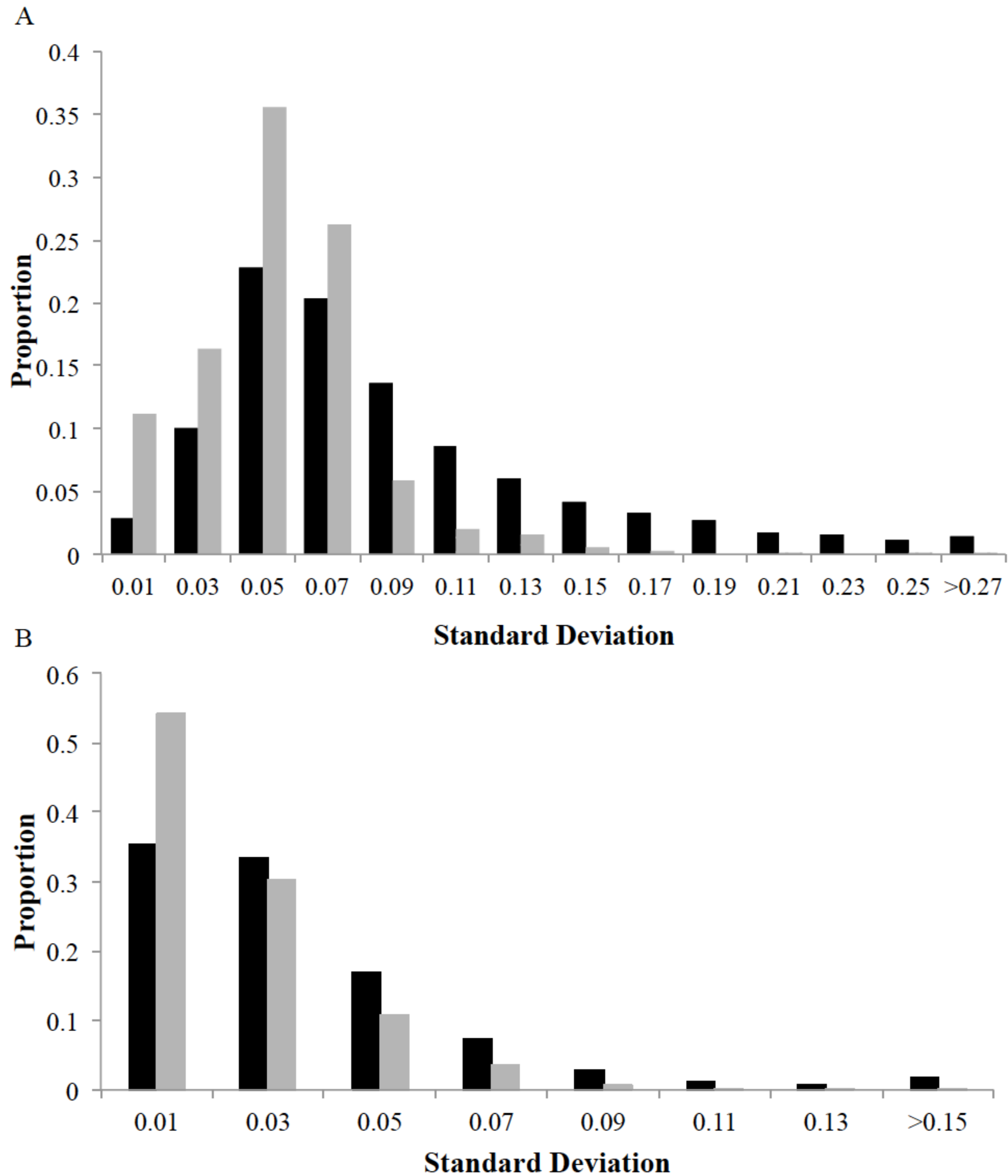


Figure 2-3: Correlated CpG sites and their enrichment in meQTLs. The x-axis is the $-\log_{10}$ of the statistical threshold (alpha level) used to define meQTLs. The y-axis shows odds ratios comparing the odds that a correlated CpG site (vs. an uncorrelated site) is also a meQTL and the vertical lines represent the confidence interval for each odds ratio.

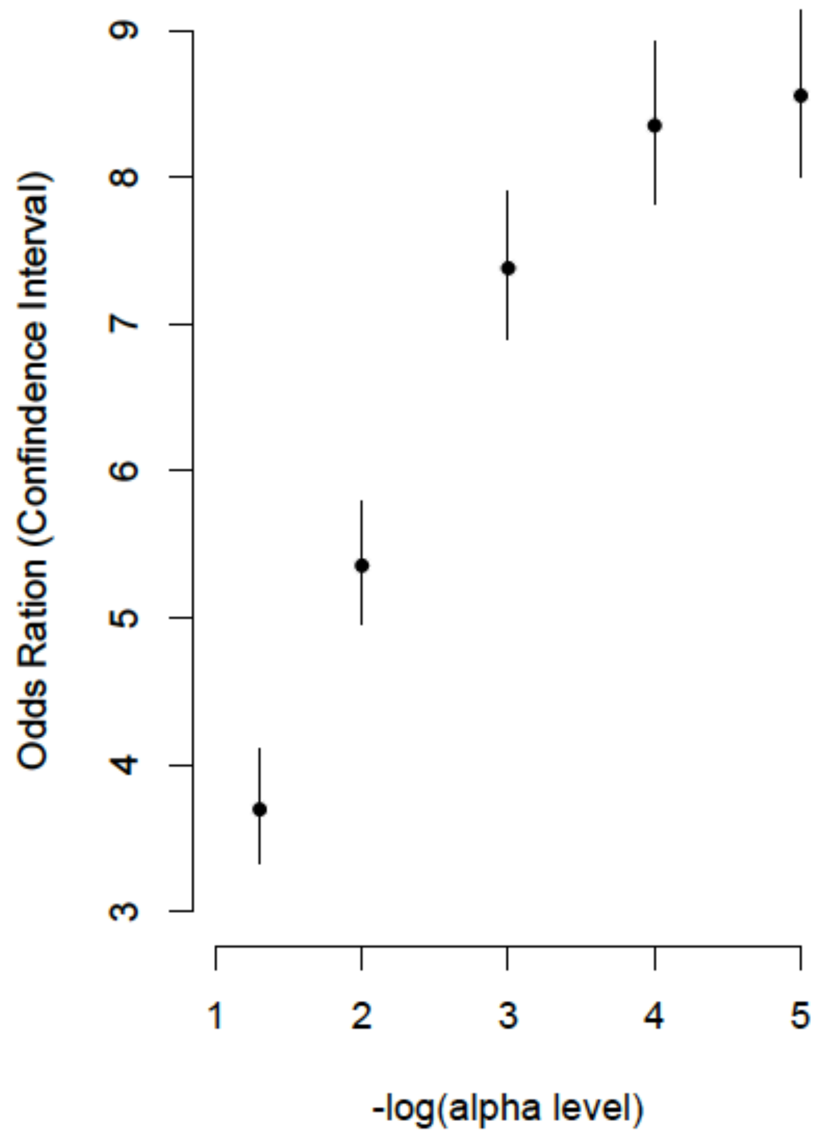
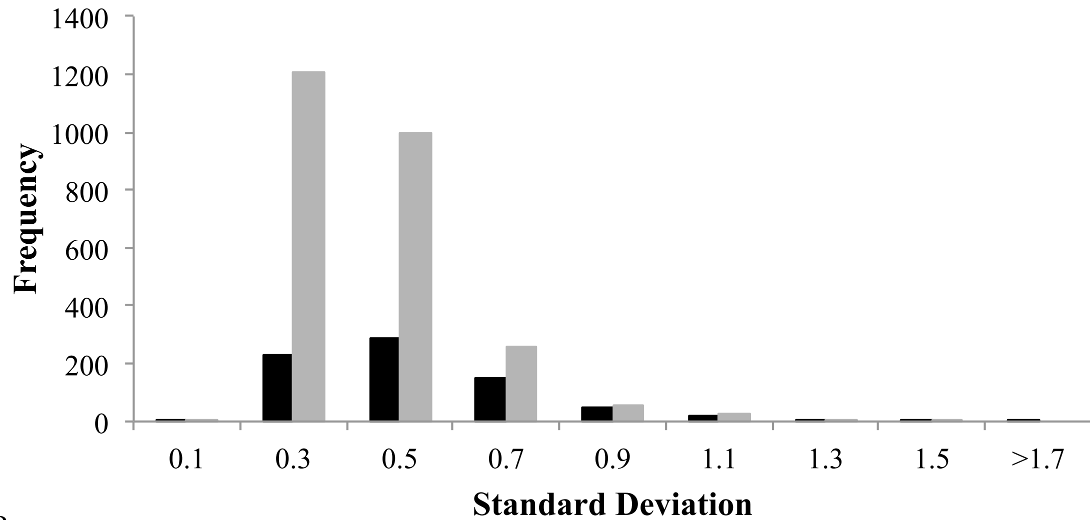


Figure 2-4: Distribution of maternal and cord blood expression. The x-axis is the standard deviation (SD) of the log2 transformation of each transcript in (A) cord blood or (B) maternal blood. The y-axis indicates the frequency of transcripts in each SD category. Black represents transcripts that were correlated while grey represents transcripts that are not correlated in maternal-neonatal pairs.

A



B

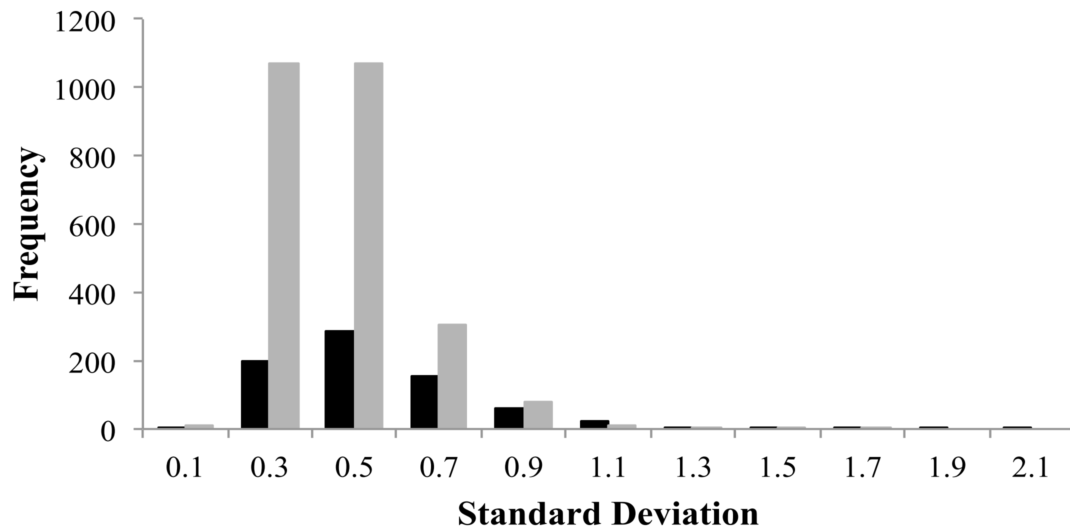
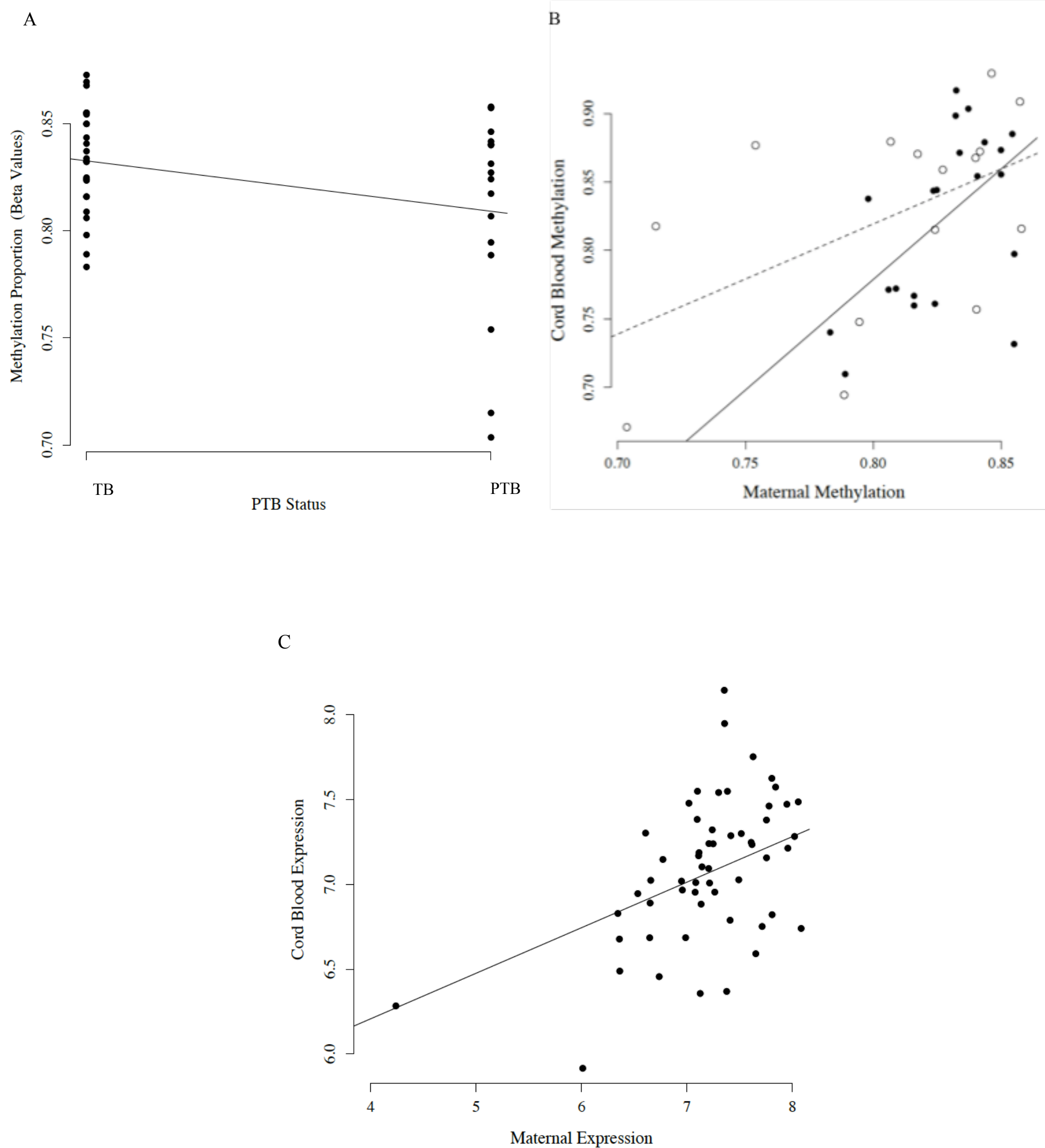


Figure 2-5: Association of *MICB* in PTB, and its correlation in expression and DNA methylation. Graph (A) is the association of cg06284756 maternal methylation with PTB with the x-axis representing PTB status where 0 denotes term birth and 1 denotes PTB and the y-axis representing the methylation (beta values). Graph (B) is the correlation between maternal methylation (x-axis) for cg0628476 with fetal methylation (y-axis) where open circles represent pairs that are preterm and closed circles represent pairs that are term. The dashed line represents correlation in PTB pairs, and the solid line represents correlation in term birth (TB) pairs. Graph (C) is the correlation between maternal *MICB* expression (x-axis) with fetal MICB expression (y-axis).



References

1. Goldenberg RL, Culhane JF, Iams JD, Romero R (2008) Epidemiology and causes of preterm birth. Lancet 371: 75-84.

2. Kistka ZA, Palomar L, Lee KA, Boslaugh SE, Wangler MF, et al. (2007) Racial disparity in the frequency of recurrence of preterm birth. Am J Obstet Gynecol 196: 131 e131-136.

3. Hamilton BE, Martin JA, Ventura SJ (2012) Births: preliminary data for 2011. Natl Vital Stat Rep 61: 1-18.

4. Cho G, Min KJ, Hong HR, Kim S, Hong JH, et al. (2013) High-risk human papillomavirus infection is associated with premature rupture of membranes. BMC Pregnancy Childbirth 13: 173.

5. Jeffcoat MK, Geurs NC, Reddy MS, Goldenberg RL, Hauth JC (2001) Current evidence regarding periodontal disease as a risk factor in preterm birth. Ann Periodontol 6: 183-188.

6. Kelly RH, Russo J, Holt VL, Danielsen BH, Zatzick DF, et al. (2002) Psychiatric and substance use disorders as risk factors for low birth weight and preterm delivery. Obstet Gynecol 100: 297-304.

7. Spiegler J, Stichtenoth G, Weichert J, Konig IR, Schlaud M, et al. (2013) Pregnancy risk factors for very premature delivery: what role do hypertension, obesity and diabetes play? Arch Gynecol Obstet 288: 57-64.

8. Marret S, Ancel PY, Marpeau L, Marchand L, Pierrat V, et al. (2007) Neonatal and 5-year outcomes after birth at 30-34 weeks of gestation. Obstet Gynecol 110: 72-80.

9. Goldenberg RL, Cliver SP, Mulvihill FX, Hickey CA, Hoffman HJ, et al. (1996) Medical, psychosocial, and behavioral risk factors do not explain the increased risk for low birth weight among black women. Am J Obstet Gynecol 175: 1317-1324.

10. McGrady GA, Sung JF, Rowley DL, Hogue CJ (1992) Preterm delivery and low birth weight among first-born infants of black and white college graduates. Am J Epidemiol 136: 266-276.

11. (2007) Preterm birth causes, consequences, and prevention. . In: Behrman RE BA, editor. Institute of Medicine (US) Understanding premature birth and assuring healthy outcomes. Washington (DC): National Academies Press (US).

12. Ward K, Argyle V, Meade M, Nelson L (2005) The heritability of preterm delivery. Obstet Gynecol 106: 1235-1239.

13. Haataja R, Karjalainen MK, Luukkonen A, Teramo K, Puttonen H, et al. (2011) Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. PLoS Genet 7: e1001293.

14. Treloar SA, Macones GA, Mitchell LE, Martin NG (2000) Genetic influences on premature parturition in an Australian twin sample. Twin Res 3: 80-82.

15. Clausson B, Lichtenstein P, Cnattingius S (2000) Genetic influence on birthweight and gestational length determined by studies in offspring of twins. BJOG 107: 375-381.

16. Committee on Practice Bulletins-Obstetrics TACoO, Gynecologists (2012) Practice bulletin no. 130: prediction and prevention of preterm birth. Obstet Gynecol 120: 964-973.

17. Parets SE, Bedient CE, Menon R, Smith AK (2014) Preterm birth and its long-term effects: methylation to mechanisms. Biology (Basel) 3: 498-513.

18. Myking S, Boyd HA, Myhre R, Feenstra B, Jugessur A, et al. (2013) X-chromosomal maternal and fetal SNPs and the risk of spontaneous preterm delivery in a Danish/Norwegian genome-wide association study. PLoS One 8: e61781.

19. Steffen KM, Cooper ME, Shi M, Caprau D, Simhan HN, et al. (2007) Maternal and fetal variation in genes of cholesterol metabolism is associated with preterm delivery. J Perinatol 27: 672-680.

20. Ehn NL, Cooper ME, Orr K, Shi M, Johnson MK, et al. (2007) Evaluation of fetal and maternal genetic variation in the progesterone receptor gene for contributions to preterm birth. Pediatr Res 62: 630-635.

21. Plunkett J, Muglia LJ (2008) Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. Ann Med 40: 167-195.

22. Kaprio J (2012) Twins and the mystery of missing heritability: the contribution of gene-environment interactions. J Intern Med 272: 440-448.

23. Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. Proc Natl Acad Sci U S A 109: 1193-1198.

24. Menon R, Conneely KN, Smith AK DNA methylation: an epigenetic risk factor in preterm birth. Reprod Sci 19: 6-13.

25. Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, et al. (2013) Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age. PLoS One 8: e67489.

26. Liu Y HC, Murphy S, Huang Z, Overcash F, Thompson J, Brown H, Murtha AP (2013) DNA methylation at imprint regulatory regions in preterm birth and infection. Am J Obst Gynecol pii S0002-9378: 00142–00147.

27. Tendl KA, Schulz SM, Mechtler TP, Bohn A, Metz T, et al. (2013) DNA methylation pattern of CALCA in preterm neonates with bacterial sepsis as a putative epigenetic biomarker. Epigenetics 8: 1261-1267.

28. Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, et al. (2013) Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. Genome Med 5: 96.

29. Yuan W, Basso O, Sorensen HT, Olsen J (2001) Indicators of fetal growth and infectious disease in childhood--a birth cohort with hospitalization as outcome. Eur J Epidemiol 17: 829-834.

30. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ (2002) Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA 288: 728-737.

31. Singhal A, Fewtrell M, Cole TJ, Lucas A (2003) Low nutrient intake and early growth for later insulin resistance in adolescents born preterm. Lancet 361: 1089-1097.

32. Gluckman PD, Hanson MA, Cooper C, Thornburg KL (2008) Effect of in utero and early-life conditions on adult health and disease. N Engl J Med 359: 61-73.

33. Barker DJ (2007) The origins of the developmental origins theory. J Intern Med 261: 412-417.

34. Kessous R, Shoham-Vardi I, Pariente G, Holcberg G, Sheiner E (2013) An association between preterm delivery and long-term maternal cardiovascular morbidity. Am J Obstet Gynecol 209: 368 e361-368.

35. Catov JM, Newman AB, Roberts JM, Kelsey SF, Sutton-Tyrrell K, et al. (2007) Preterm delivery and later maternal cardiovascular disease risk. Epidemiology 18: 733-739.

36. Hastie CE, Smith GC, Mackay DF, Pell JP (2011) Maternal risk of ischaemic heart disease following elective and spontaneous pre-term delivery: retrospective cohort study of 750 350 singleton pregnancies. Int J Epidemiol 40: 914-919.

37. Smith GC, Pell JP, Walsh D (2001) Pregnancy complications and maternal risk of ischaemic heart disease: a retrospective cohort study of 129,290 births. Lancet 357: 2002-2006.

38. James-Todd T, Wise L, Boggs D, Rich-Edwards J, Rosenberg L, et al. (2014) Preterm birth and subsequent risk of type 2 diabetes in black women. Epidemiology 25: 805-810.

39. Lykke JA, Paidas MJ, Damm P, Triche EW, Kuczynski E, et al. (2010) Preterm delivery and risk of subsequent cardiovascular morbidity and type-II diabetes in the mother. BJOG 117: 274-281.

40. Catov JM, Dodge R, Yamal JM, Roberts JM, Piller LB, et al. (2011) Prior preterm or small-for-gestational-age birth related to maternal metabolic syndrome. Obstet Gynecol 117: 225-232.

41. Robbins CL, Hutchings Y, Dietz PM, Kuklina EV, Callaghan WM (2014) History of preterm birth and subsequent cardiovascular disease: a systematic review. Am J Obstet Gynecol 210: 285-297.

42. Melbye M, Wohlfahrt J, Andersen AM, Westergaard T, Andersen PK (1999) Preterm delivery and risk of breast cancer. Br J Cancer 80: 609-613.

43. James-Todd TM, Karumanchi SA, Hibert EL, Mason SM, Vadnais MA, et al. (2013) Gestational age, infant birth weight, and subsequent risk of type 2 diabetes in mothers: Nurses' Health Study II. Prev Chronic Dis 10: E156.

44. Dunlop AL, Kramer MR, Hogue CJ, Menon R, Ramakrishan U (2011) Racial disparities in preterm birth: an overview of the potential role of nutrient deficiencies. Acta Obstet Gynecol Scand 90: 1332-1341.

45. Backdahl L, Bushell A, Beck S (2009) Inflammatory signalling as mediator of epigenetic modulation in tissue-specific chronic inflammation. Int J Biochem Cell Biol 41: 176-184.

46. Rodriguez-Cortez VC, Hernando H, de la Rica L, Vento R, Ballestar E (2011) Epigenomic deregulation in the immune system. Epigenomics 3: 697-713.

47. Sun YV, Lazarus A, Smith JA, Chuang YH, Zhao W, et al. (2013) Gene-specific DNA methylation association with serum levels of C-reactive protein in African Americans. PLoS One 8: e73480.

48. Barres R, Zierath JR (2011) DNA methylation in metabolic disorders. Am J Clin Nutr 93: 897S-900.

49. Gluckman PD, Hanson MA, Buklijas T, Low FM, Beedle AS (2009) Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. Nat Rev Endocrinol 5: 401-408.

50. Ordovas JM, Smith CE (2010) Epigenetics and cardiovascular disease. Nat Rev Cardiol 7: 510-519.

51. Kim J, Pitlick MM, Christine PJ, Schaefer AR, Saleme C, et al. (2013) Genome-Wide Analysis of DNA Methylation in Human Amnion. The Scientific World Journal 2013: 11.

52. Oken E, Kleinman KP, Rich-Edwards J, Gillman MW (2003) A nearly continuous measure of birth weight for gestational age using a United States national reference. BMC Pediatr 3: 6.

53. Barfield RT, Kilaru V, Smith AK, Conneely KN (2012) CpGassoc: an R function for analysis of DNA methylation microarray data. Bioinformatics 28: 1280-1281.

54. Kilaru V, Barfield RT, Schroeder JW, Smith AK, Conneely KN (2012) MethLAB: a graphical user interface package for the analysis of array-based DNA methylation data. Epigenetics 7: 225-229.

55. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, et al. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 13: 86.

56. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen SE, et al. (2012) Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One 7: e41361.

57. Ahmad QR, Allen RC, Andersen TC, J DA, Barton JC, et al. (2002) Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. Phys Rev Lett 89: 011301.

58. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57.

59. Huang da W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: 1-13.

60. Smith AK, Kilaru V, Kocak M, Almli LM, Mercer KB, et al. (2014) Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. BMC Genomics 15: 145.

61. Barfield RT, Almli LM, Kilaru V, Smith AK, Mercer KB, et al. (2014) Accounting for population stratification in DNA methylation studies. Genet Epidemiol 38: 231-241.

62. Mehta D, Klengel T, Conneely KN, Smith AK, Altmann A, et al. (2013) Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. Proc Natl Acad Sci U S A 110: 8302-8307.

63. Norrholm SD, Jovanovic T, Smith AK, Binder E, Klengel T, et al. (2013) Differential Genetic and Epigenetic Regulation of catechol-O-methyltransferase is Associated with Impaired Fear Inhibition in Posttraumatic Stress Disorder. Front Behav Neurosci 7: 30.

64. Votavova H, Dostalova Merkerova M, Fejglova K, Vasikova A, Krejcik Z, et al. (2011) Transcriptome alterations in maternal and fetal cells induced by tobacco smoke. Placenta 32: 763-770.

65. Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. Bioinformatics 24: 1547-1548.

66. McRae AF, Powell JE, Henders AK, Bowdler L, Hemani G, et al. (2014) Contribution of genetic variation to transgenerational inheritance of DNA methylation. Genome Biol 15: R73.

67. van Dongen J, Ehli EA, Slieker RC, Bartels M, Weber ZM, et al. (2014) Epigenetic variation in monozygotic twins: a genome-wide analysis of DNA methylation in buccal cells. Genes (Basel) 5: 347-365.

68. Huang SY, Chiang CH, Chen FP, Yu CL (2011) The alteration of placental-derived soluble MHC class I chain-related protein A and B during pregnancy. Acta Obstet Gynecol Scand 90: 802-807.

69. Geiss LS, Wang J, Cheng YJ, Thompson TJ, Barker L, et al. (2014) Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. JAMA 312: 1218-1226.

70. Gibbs RS, Romero R, Hillier SL, Eschenbach DA, Sweet RL (1992) A review of premature birth and subclinical infection. Am J Obstet Gynecol 166: 1515-1528.

71. Kerkhof GF, Willemsen RH, Leunissen RW, Breukhoven PE, Hokken-Koelega AC (2012) Health profile of young adults born preterm: negative effects of rapid weight gain in early life. J Clin Endocrinol Metab 97: 4498-4506.

72. Melville JM, Moss TJ (2013) The immune consequences of preterm birth. Front Neurosci 7: 79.

73. Corwin EJ, Guo Y, Pajer K, Lowe N, McCarthy D, et al. (2013) Immune dysregulation and glucocorticoid resistance in minority and low income pregnant women. Psychoneuroendocrinology 38: 1786-1796.

Chapter 3

Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age

This chapter has been adapted and was originally published in PLoS One under a CC-BY 3.0 license (http://creativecommons.org/licenses/by/3.0/). Readers can access the published article at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.006748 or from the citation below:

Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, G Saade, AK Smith, R Menon. (2013) Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age. PLoS One 8: e67489.

Introduction

Despite advances in health care, the rate of preterm birth (PTB; birth before 37 weeks of gestation) has been increasing for the last 25 years [1]. Specifically, children born preterm are more likely be hospitalized and have diminished cognitive performance and develop behavioral problems such as ADHD during childhood [2,3]. Along these lines, many adult onset diseases have been linked to adverse intrauterine conditions or adverse pregnancy outcomes [4,5]. Thus, PTB not only imparts a difficult start but also considerable challenges throughout life [1,6]. Spontaneous preterm birth (PTB), which occurs without indications, is common and contributes to significant neonatal morbidity and mortality over time [7].

Several epidemiologic, behavioral and biological factors (i.e. race, socioeconomic status, malnutrition, smoking, and infection) have been associated with PTB, but the mechanistic pathways that underlie the association of the risk factors to PTB are still unclear [8-10]. The field of epigenetics has the potential to provide a greater understanding of the pathways that contribute to or result from PTB [11]. Indeed specific risk factors may promote epigenetic changes that result in PTB or that predisposes a neonate to adult-onset diseases. Although epigenetic differences associate with many prenatal exposures and complex traits, published studies that evaluate maternal and fetal epigenetic changes during pregnancy, influence on pregnancy outcome, and fetal programming of adult-onset diseases are limited [12,13]. The study of epigenetic patterns during early development is likely to provide more information about environmental and behavioral influences on long-term outcomes than the study of individuals later in life. In time, such studies may suggest biomarkers for developmental outcomes.

DNA methylation is an epigenetic modification required for proper gene regulation and cellular differentiation during fetal development [14,15]. Over the first years of life, DNA methylation of many genes appears to be relatively stable [16,17]. Therefore, DNA methylation patterns of certain genes established at birth may result in a developmental trajectory with long-term consequences. We have previously shown that DNA methylation of certain genes associates with gestational age (GA) in term deliveries [18], and evidence suggests that DNA methylation differences in key genes may provide insight into biological pathways that underlie PTB. The primary objective of this study is to interrogate methylation patterns across the genome in DNA derived from umbilical cord blood leukocytes of a high risk African American cohort and to evaluate the association of each CpG site with PTB and GA.

Methods

This study was approved by the Institutional Review Boards of Centennial Women's Hospital, Western Institutional Review Board and the University of Texas Medical Branch.

Subjects and sample collection

The Nashville Birth Cohort (NBC) was established to examine genetic risk factors and changes in the biochemical pathways that distinguish spontaneous preterm from term labor. All subjects were recruited at Centennial Women's Hospital and the Perinatal Research Center in Nashville, TN beginning in 2003. Pregnant women were enrolled during their first clinical visit after obtaining informed consent. Maternal demographic and clinical data were recorded from medical records or by interviews during the consenting process. Demographic and clinical data specific to the fetus was collected

from clinical records. Gestational age of the neonate was determined by maternal reporting of the last menstrual period and corroboration by ultrasound dating. Race was identified by self-reporting that traced back to three generations from maternal and paternal side of the fetus. Only African Americans of non-Hispanic ethnicity were included in this study.

Subjects were included in this study if they had contractions (rate of 2 contractions/10 minutes) leading to delivery either at preterm or term. Cases were delivered preterm with intact membranes between $24^{1/7}$ weeks and $34^{0/7}$ weeks. Controls were delivered (> $39^{0/7}$ weeks) with spontaneous term labor and delivery and no current or history of pregnancy-related complications including PTB and preterm or prelabor rupture of the membranes (pPROM). Subjects who had multiple gestations, preeclampsia, placenta-previa, fetal anomalies, and/or medical or surgical complications during pregnancy were excluded from the study. Subjects with any surgical procedures during pregnancy were treated for preterm labor or for suspected intra-amniotic infection and delivered at term were excluded from the control group. Maternal demographic and clinical data were collected from medical records or thorough self-report at the time of consent.

Race, socioeconomic (education, household income, marital status, and insurance status), behavioral (cigarette smoking) factors were documented by maternal self-report. Intraamniotic infection was determined by amniotic fluid culture or by PCR for 16s ribosomal RNA. In cases where culture or PCR data were not available, infection was assessed with four of the following clinical or histologic symptoms: high fever (> 102°C),

high CRP (> 0.8 U/ml), abdominal tenderness, fetal tachycardia, mucopurulent vaginal discharge or histologic chorioamnionitis, funisitis.

Biological sample collection and DNA extraction

Umbilical cord blood samples were collected in EDTA tubes soon after placental delivery. Blood samples were centrifuged at 3,000 RPM to separate plasma, and buffy coats were aliquoted and stored at -80°C. DNA was extracted using the Autopure automated system (Gentra Systems, Minneapolis, MN).

DNA methylation analysis

For each subject, > 485,000 CpG sites across the genome were interrogated using the HumanMethylation450 BeadChip (Illumina, San Diego, CA)[19,20]. Briefly, 1 ug of DNA was converted with sodium bisulfite, amplified, fragmented, and hybridized on the HumanMethylation450 BeadChip (Illumina, San Diego, CA) according to the manufacturer's instructions. CpGassoc [21] was used to perform quality control and calculate ß values. Data points with probe detection p-values >.001 were set to missing, and CpG sites with missing data for >10% of samples were excluded from analysis; 483,830 CpG sites passed the above criteria. Samples with probe detection call rates <90% and those with an average intensity value of either <50% of the experiment-wide sample mean or <2,000 arbitrary units (AU) were excluded from further analysis. One sample of male DNA was included on each BeadChip as a technical control throughout the experiment and assessed for reproducibility using the Pearson correlation coefficient, to ensure that Pearson correlation coefficient >0.99 for all pairwise comparisons of technical replicates. For each individual sample and CpG site, the signals from

methylated (M) and unmethylated (U) bead types were used to calculate a beta value as ß = M/(U + M).

Statistical Analysis

We used MethLAB [22] to test for association with PTB via linear regressions that modeled β-values as the outcome and PTB as the independent variable, adjusting for GA, gender, chip, and row on the chip. Based on previous reports and the potential contribution to PTB we examined the association of birth weight percentile, gravidity, parity, infection and smoking as confounding factors in our analysis; these factors did not associate with methylation of any CpG site after adjustment for multiple testing (FDR<.05; data not shown). Birth weight percentile was based on estimated gestational age (GA) in accordance with the United States national registry [23]. We subsequently used MethLAB to fit similar linear regressions that modeled GA as the independent variable, adjusting for gender, chip, and row on the chip. Because it has been suggested that logit-transformed β values (a.k.a. M values) may perform better in statistical analyses [24], we also examined associations with M values using the strategy described above. Because there was no significant difference between the results, we present results based on untransformed β to ease biological interpretation.

The location of each CpG site was determined using the Illumina array annotation for the HumanMethylation450 BeadChip based on build 37 of the human genome. We tested for enrichment among GA-associated sites by comparing the number of GA-associated CpG sites that did or did not occur in a particular gene region (e.g. promoter, 5'UTR, Body, 1st exon, 3'UTR, or intragenic regions) to the number of non-GA-associated sites that did or did not occur in that gene region, using Fisher's exact test. We

then performed similar tests of enrichment for CpG-rich regions defined as islands or CpG poor regions defined as shores [25,26]. CpG sites with 1000 Genomes Project variants physically contained within the Illumina probe were noted in the analyses but not excluded a priori. In addition we examined whether significant GA-associated CpG sites were enriched or depleted on the X chromosome using Fisher's exact test.

We used GSEAPrerank [27,28] to evaluate whether GA-associated CpG sites were located in genes that were enriched for specific biological processes and cellular components. Significance of the gene ontology enrichment was corrected for an FDR<.05 following 1000 permutations.

Results

The cohort, described in Table 3-1, consists of African American preterm (GA range 24.1-34.0 weeks) and term (39.0-40.9 weeks) births. Though the groups differed by GA and birthweight, they did not differ significantly in demographic or clinical factors.

Preterm Birth (PTB)

After accounting for multiple comparisons (FDR <.05) and confounding factors (gender, gestational age, and chip effects), 29 CpG sites associate with PTB independently of GA (Figure 3-1A; Table 3-2; $5.7 \times 10^{-10} < p < 2.9 \times 10^{-6}$; $-.17 < \Delta\beta < .26$). Based on annotation with data from the 1000 Genomes Project, 5 of these 29 CpG probes (17.2%) do contain a SNP (estimated average minor allele frequency of 15.5%), suggesting that we could be observing a genetic rather than an epigenetic association for these 5 CpG sites; the methylated and unmethylated signals for these five sites are shown in Figure 3-2. In some cases, the pattern appears consistent with SNP-induced methylation differences, while in other cases there is no strong pattern of

clustering. Results were not significantly altered by adjustment for maternal smoking, or infection, birth weight percentile, and gravidity (data not shown) nor were they altered by logit-transformation of the beta values. Among the CpG sites associated with PTB, we observed increased DNA methylation of a site (cg13250001) in *GSK3B* (glycogen synthase kinase 3 beta; $p=1.7 \times 10^{-6}$; $\Delta\beta=-.06$) and decreased methylation of a CpG site (cg25376491) in *MAML1* (mastermind-like 1; $p=1.8 \times 10^{-6}$; $\Delta\beta=.14$) in fetuses with PTB. In addition, 3 other CpG sites in GSK3B and 4 in MAML1 were nominally associated with PTB ($p<.05$).

Gestational Age

Our above analyses of PTB all included GA as a covariate because PTB and GA are by definition correlated ($r=.93$), and there is overwhelming agreement in the association of DNA methylation with PTB unadjusted for GA, or GA itself (Figure 3-3). In fact, 9637 CpG sites associated with GA independent of gender and chip effects (FDR <.05; $9.5\times10^{-16}<p<1.0\times10^{-3}$; $-.024<\Delta\beta$ per week$<.023$; Figure 3-1B). GA-associated CpG sites were depleted in the promoter, first exon and 3'UTR regions and enriched in the 5'UTR, gene body and intragenic regions ($2.2\times10^{-16}<p<2.6\times10^{-3}$; Table 3-3) when compared to CpG sites that were not associated with GA via Fisher's exact test. Associated CpG sites were also depleted in CpG islands (14.9% vs. 31.3%; $p<2.2\times10^{-16}$) and enriched in CpG shores (34.1% vs. 22.8%; $p<2.2\times10^{-16}$). Examining the directionality of GA-associated CpG sites, 61.8% (5958 CpG sites) had lower methylation in subjects with lower GA; these CpG sites were twice as likely to be located in CpG islands ($p<2.2\times10^{-16}$; Table 3-3) and less likely to occur in the gene body ($p<2.2\times10^{-16}$;) and 3'UTR ($p=1.5\times10^{-9}$). While the sample size was not sufficient to look

for sex-specific differences (i.e. interactions between age and sex), we did note a depletion of GA-associated CpG sites on the X chromosome (.5% vs. 2.4 %; $p<2.2x10^{-16}$); both the depletion of GA-associated variants on CpG islands and the X chromosome are consistent with a previous report of age-associated methylation in children [29].

Gene set enrichment analysis (GSEA) was used to gain further insight into the functional context of GA-associated CpG sites (FDR<.05; Table 3-4). Prominent biological processes that were enriched in GA-associated CpG sites were related to embryonic development. For example, 9 sites in the 5'UTR and body of histone deacetylase 4 (*HDAC4*, $1.3.x10^{-11}<p<9.8x10^{-4}$; $-.0023<\Delta\beta$ per week<-.01) have higher methylation levels in fetuses with lower GA. *HDAC4* is involved in numerous identified pathways including system development and multicellular organismal development, anatomical structure development, organ development, and nervous system development. Several other CpG sites involved in epigenetic regulation during development were also identified. Specifically, CpG sites in the gene body of *DNMT1* (DNA methyltransferase 1; $p=3.4x10^{-5}$; t=-4.7; $\Delta\beta$ per week=-.0034), the gene body of *DNMT3A (*$p=6.7x10^{-4}$; t=-3.7; $\Delta\beta$ per week=-.0042*),* the 5'UTR of *DNMT3B* ($4.5x10^{-6}<p<8.4x10^{-4}$; 3.6<t<5.3; $.0040<\Delta\beta$ per week<.0053) and the 5'UTR of *TET1* (tet methylcytosine dioxygenase 1; $1.5x10^{-7}<p<2.7x10^{-4}$; 4.0<t<6.4; $.0046<\Delta\beta$ per week<.01) also associate with GA.

Among the enriched cellular components are several groups that relate to extracellular regions. Remodeling of the extracellular matrix is required to support pregnancy and parturition [30] and increased attention has recently been focused on the role of matrix metallopeptidases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) in preterm birth [31]. In this study, 4 CpG sites in the promoter of *MMP9*

($5.6x10^{-7}<p<3.2x10^{-4}$; $4.0<t<6.0$; $.0021<\Delta\beta$ per week<.0033) had higher methylation with increasing gestational age. *MMP9* is involved in the breakdown of the extracellular matrix in the process of cervical ripening, and increased expression has been seen in pPROM compared to preterm birth with intact membranes [32]. Furthermore, one CpG site in the gene body of the MMP9 inhibitor, *TIMP2* also associates with GA ($p=1.4x10^{-5}$; $t=-5.0$; $\Delta\beta$ per week=-.0053).

To complement our discovery approach, we evaluated the association between CpG sites in genes that had been associated with GA in a previous study that used a less dense array with 27,578 CpG sites [18] (Table 3-5). 21 of 26 CpG sites (80.8%) significantly associated with GA in the previous study replicate in the present cohort ($7.5 x10^{-11}<p<.05$; $-8.1<t<8.9$; $-.01<\Delta\beta$ per week<.01). Notably, methylation of a CpG site in corticotrophin-releasing hormone binding protein (*CRHBP*) increased with decreasing GA ($t=-4.49$; $p=6.5x10^{-5}$; $\Delta\beta$ per week=.01). CRHBP regulates corticotrophin-releasing hormone (CRH), a principal regulator of the hypothalamic-pituitary-adrenal (HPA) axis. In addition, methylation increased in a CpG site in the promoter of *PIK3CD* (phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit delta) with decreasing GA ($p=2.4x10^{-8}$; $t=-7.0$; $\Delta\beta$ per week=-.0062). The therapeutic value of PIK3CD inhibitors is currently being explored as anti-inflammatory drugs [33].

One limitation of this strategy is that GA and PTB represent correlated but etiologically distinct phenotypes. Thus, replicating associations observed with GA may not capture the same breadth of candidate genes that could be explored in a study focused on PTB. For example, *IGFBP1* has been considered as a marker for preterm in vaginal infection and leaking amniotic fluid [34]. We observed associations between GA and 6

CpG sites in insulin-like growth factor 2 mRNA binding protein 1 *(IGF2BP1;* $2.1\times10^{-12}<p<1.9\times10^{-4}$; $-4.1<t<10.2$; $-.0087<\Delta\beta$ per week$<.02$) located in the gene body though the direction of the association changed based on proximity of the CpG site to the CpG island.

Discussion

By examining DNA methylation across the genome, we identified 29 CpG sites that associated with PTB independently of GA in leukocyte DNA from high-risk African American fetuses. Among these are CpG sites in *GSK3B* (glycogen synthase kinase 3 beta), which is involved in neuronal migration, development, and polarization, particularly during early embryonic development [35,36]. Interestingly, *GSK3B* is a negative regulator of *MAML1* (mastermind-like 1) [37], a component of the Notch pathway [38,39], and a CpG site in *MAML1* also associated with PTB. *GSK3B* decreases transcription in the notch pathway through inhibition of *MAML1* [37]. Consistent with the role of *GSK3B* in regulating *MAML1*, there was an inverse relationship in the associations for the CpG sites in these genes. During development, the Notch pathway is integral to several developmental processes including neurogenesis, cardiovascular function, angiogenesis as well as intestinal and bone development [40].

Additionally, 9637 CpG sites associated with GA when it was modeled separately from PTB. Our analyses suggest enrichment of GA-associated CpG sites in biological processes involved not only in embryonic and organ development but also in neurogenesis, nervous system development and neuron development. These processes involve extensive epigenetic regulation so it is not surprising that we observed associations with CpG sites in genes related to shaping epigenetic patterns during

development: *HDAC4, DNMT1, DNMT3A, DNMT3B,* and *TET1*. For example, CpG sites in *TET1* and *DNMT3B* have lower DNA methylation in subjects with shorter GA. *TET1* functions to hydroxylate 5'methylcyctosine(mC) into 5'hydroxymethyl cytosine (hmC) [41]. *TET1* has been implicated is normal embryogenesis, and the depletion of *TET1* leads to low birth weight (LBW) in mouse pups [42]. *TET1* promotes active demethylation while *DNMT3B* promotes *de novo* methylation; these two processes are highly involved in the establishment of tissue-specific DNA methylation patterns during development [41,43]. Though these results are indicative of the developmental time sampled (i.e. 32 versus 38 weeks), they may also support the hypothesis of epigenetic programming during fetal development [44].

The cellular components most enriched for genes with GA-associated CpG sites were primarily related to the extracellular region. Genes such as *MMP9* and *TIMP2* are integral to the process of parturition [45]. MMP9 has previously been considered as a biomarker for preterm birth [46] and has been thought to play a role in premature rupture of the membranes (PROM) because of its role in the degradation of the amniochorion basement membranes [47]. MMP9 levels are higher following PROM when compared to term deliveries, while TIMP2 levels decrease. DNA methylation differences in these and other genes related to extracellular matrix function support further study of the role of the fetal extracellular matrix throughout pregnancy and during parturition.

Many studies of fetal programming or prenatal exposures focus on fetuses with intrauterine growth restrictions or that were small for gestational age. Recent studies in the field support associations between GA and both DNA methylation and gene expression differences, but note lesser or no associations with birth weight [18,48].

Similarly, in this study we identified numerous associations between DNA methylation and PTB, which is measured by GA, but no associations with percentile birth weight. Based on this, Stunkel and colleagues hypothesize that birth weight may be a less appropriate measure of adverse outcomes than GA [48]. Along these lines, we identified associations between GA and DNA methylation of CpG sites in insulin-like growth factor 2 mRNA binding protein 1 *(IGF2BP1),* a developmentally regulated gene that binds IGF2 and has been a focus of the fetal programming literature [49]. DNA methylation in IGF2 has been linked to various pregnancy-related conditions including birth weight [50]. IGFBP proteins are secreted from the placenta, decidua and fetal membranes in increasing amounts across gestation and are abundant in amniotic fluid [51]. Detection of IGFBP-1 in cervical–vaginal secretions is reliably used to detect preterm premature rupture of the membranes, which precedes 40% of spontaneous PTB cases [52,53]. However, we were not able identify PTB-associated DNA methylation differences.

Our results were consistent with previous studies of DNA methylation in gestational age. Despite differences between cohorts and study design, we replicated >80% of CpG sites associated with GA in a previous study [18] further supporting the role of these genes in embryonic development and parturition. For example, CpG sites in *CRHBP* associated with GA. CHRBP binds CRH limiting its activity, and changes in the relative ratios of CRH to CRHBP associate with timing of birth [54,55]. Prior to parturition, CRHBP levels decrease while CRH levels increase facilitating labor in both term and preterm deliveries [56]. In women who deliver preterm there is a decrease in plasma levels of CRHBP compared to women who deliver term [57].

The goal of this study was to identify associations between DNA methylation and PTB. However, PTB is defined by GA at birth; thus, the differences observed may correspond to differences in the developmental stage versus the causes or consequences of PTB. In this study, the correlation between association tests for PTB and GA is strong (r=.93; Figure 3-3), and delineation of these factors is complex, particularly in a study with a relatively small sample size. Thus, larger studies will be required to identify DNA methylation differences exclusive to PTB. Future studies of methylation as a risk factor for PTB should also focus on maternal methylation during pregnancy; a prospective study design could avoid confounding due to differences in GA by sampling at standardized time points, and could allow comparisons between maternal and fetal methylation changes. However, even with our relatively small sample of fetal cord blood DNA, we were able to identify robust associations using a stringent phenotype definition that compared samples from early preterm and later term deliveries in a high-risk cohort; in general, African American women are 3-4 times more likely than Caucasian women to deliver in the early preterm period [7]. Another limitation is the use of whole umbilical cord blood. While an ideal design would examine DNA methylation in a single cell type, this approach and our results were consistent with previous studies [18,58]. Still, our results support the idea that epigenetic differences exist in fetuses born at different gestational ages. Recent studies suggest that DNA methylation patterns in many genes may be relatively stable over the first two years of life [16,17], and further studies will be necessary to determine whether persisting differences in DNA methylation may underlie the physiological correlates of PTB.

Table 3-1: Clinical and demographic characteristics of the cohort. All subjects are African American.

Phenotype	**PTB (N=22)** Mean ± SD	**Term Birth (N=28)** Mean +/- SD	**p-value**[*]
Male, %	(14) 63.6%	(11) 39.3%	NS
Gestational age, weeks	30.8± 3.3	39.8±0.4	<.0001
Birthweight, grams	1524.1 ± 638.1	3304.9± 333.4	<.0001
Birthweight percentile	32± 27.7	47± 25.3	NS
Gravidity	2.2± 1.5	2.4± 2.7	NS
Maternal Age	25.5±5.2	21.0±4.6	NS
Employed	(5) 22.7%	(8) 36.3%	NS
Married	(7) 35.0%	(7) 25.0%	NS
Maternal Smoking	(5) 22.7%	(5) 17.9%	NS
Income			
<15K	(12) 54.5%	(12) 42.9%	
15-30K	(5) 22.7%	(8) 28.5%	NS
>30K	(5) 22.7%	(8) 28.5%	

[*] NS indicates the p-value is not significant (p>.05)

Table 3-2: 29 CpG sites that associated with PTB (adjusted for GA). Δβ represents the average methylation difference between early PTB and term birth after adjustment for covariates

Probe ID	Gene	Δβ	t-statistic	p-value
cg04549583		-0.133	-8.30	5.70×10^{-10}
cg13290254		0.122	7.52	5.77×10^{-9}
cg03254336		0.178	6.87	4.24×10^{-8}
cg16447680	*KIAA0748*	-0.131	-6.60	9.84×10^{-8}
cg03272932		0.175	6.54	1.19×10^{-7}
cg03152187	*SEPT9*	-0.112	-6.15	3.91×10^{-7}
cg18721397	*SUB1*	0.259	6.12	4.31×10^{-7}
cg03706951		-0.142	-6.04	5.57×10^{-7}
cg13514049		0.129	6.01	6.13×10^{-7}
cg01142526	*N4BP3*	-0.049	-5.92	8.00×10^{-7}
cg03901454		0.174	5.86	9.65×10^{-7}
cg19787650	*FAM49B*	0.181	5.85	9.88×10^{-7}
cg20253872	*AMPD2*	-0.174	-5.85	1.00×10^{-6}
cg06320380	*TNS1*	-0.111	-5.79	1.20×10^{-6}
cg23471393		0.185	5.79	1.23×10^{-6}
cg26501007		0.135	5.75	1.38×10^{-6}
cg13250001	*GSK3B*	-0.062	-5.68	1.70×10^{-6}
cg20519581		0.224	5.68	1.73×10^{-6}
cg04212285	*PTPRN2*	0.155	5.67	1.75×10^{-6}
cg25376491	*MAML1*	0.14	5.66	1.82×10^{-6}
cg00101629	*KAZN*	0.147	5.65	1.86×10^{-6}
cg19921917	*PALLD*	0.177	5.61	2.15×10^{-6}
cg12207930	*MED12L*	0.198	5.59	2.24×10^{-6}
cg10131972	*TXNRD2*	-0.077	-5.59	2.25×10^{-6}
cg09964921		0.162	5.57	2.44×10^{-6}
cg01476222	*TRAF6*	0.144	5.56	2.49×10^{-6}
cg03318906	*RAB11FIP1*	0.128	5.54	2.60×10^{-6}
cg01621943		0.185	5.54	2.65×10^{-6}
cg13749927	*DDB2*	-0.070	-5.51	2.91×10^{-6}

Table 3-3: Enrichment analysis. To examine whether there is an enrichment in certain regions that associated with GA, or whether there is an enrichment of a certain direction of a t-statistic for associated CpG sites. Note that each row in the table represents a Fisher's exact test that tests for enrichment of associated or unassociated sites for the relevant category (e.g. CpG islands).

	GA-associated	Not GA-associated	p-value	(+) GA-associated	(-) GA-associated	p-value*
CpG Islands	14.9%	31.3%	$< 2.2 \times 10^{-16}$	18.4%	9.4%	$< 2.2 \times 10^{-16}$
CpG Shores	34.1%	22.8%	$< 2.2 \times 10^{-16}$	33.4%	35.2%	NS
Promoter	22.9%	25.0%	2.5×10^{-6}	25.3%	18.7%	4.7×10^{-14}
5' UTR	10.1%	8.8%	6.0×10^{-6}	8.6%	12.5%	1.7×10^{-9}
1st Exon	3.1%	4.7%	4.7×10^{-15}	3.0%	3.3%	NS
Gene Body	35.0%	33.3%	4.4×10^{-4}	29.6%	43.7%	$< 2.2 \times 10^{-16}$
3' UTR	2.9%	3.6%	$< 2.2 \times 10^{-16}$	2.1%	4.3%	1.5×10^{-9}
Intragenic	26.0%	24.6%	.003	31.4%	17.1%	$< 2.2 \times 10^{-16}$

[*] NS indicates the p-value is not significant ($p>.05$)

Table 3-4: Gene enrichment analysis of CpG sites that associated with GA (FDR<.05). NES is the normalized enrichment score.

GO Term: Biological Processes	**Size**	**NES**	**FDR q-value**
Neuron differentiation	29	2.69	0.001
Generation of neurons	33	2.53	0.003
Multicellular organismal development	307	2.39	0.006
Anatomical structure development	309	2.32	0.007
System development	258	2.24	0.011
Neurogenesis	34	2.23	0.011
Cellular morphogensis during differentiation	16	2.17	0.014
Neuron development	20	2.09	0.023
Embryonic development	24	2.08	0.022
Organ development	171	2.04	0.026
Nervous system development	133	2	0.033
Skeletal development	34	1.95	0.042
Anatomical structure morphogenesis	122	1.9	0.048
Negative regulation of biological process	183	-2.18	0.043
Apoptosis GO	127	-2.17	0.036
Positive regulation of I-kappaB kinase NF-kappaB cascade	23	-2.13	0.042
GO Term: Cellular Components			
Extracellular region part	75	3.13	<.001
Extracellular region	97	3.02	<.001
Extracellular space	48	2.66	<.001
Proteinaceous extracellular matrix	26	2.24	0.009
Extracellular matrix	26	2.2	0.009
Extracellular matrix part	20	2.05	0.017
Intracellular organelle part	219	-2.12	0.042
Organelle part	219	-2.12	0.021
Nuclear part	102	-2.08	0.019
Microtubule cytoskeleton	37	-2.03	0.023
Nucleus	313	-2.02	0.02
Ribonucleoprotein complex	17	-1.97	0.026
Membrane enclosed lumen	73	-1.93	0.029
Cytoskeletal part	53	-1.93	0.025
Nuclear lumen	66	-1.92	0.023
Organelle lumen	73	-1.91	0.023
Cytoskeleton	105	-1.9	0.022
Nucleoplasm	56	-1.87	0.025
Nucleoplasm part	37	-1.79	0.038
Cell cortex	16	-1.74	0.048
Non membrane bound organelle	138	-1.73	0.047
Intracellular non membrane bound organelle	138	-1.73	0.046
Macromolecular complex	166	-1.71	0.049

Table 3-5: Replication of CpG sites previously associated with GA [18]. Δβ per week represents the average increase in β-value associated with each additional week of GA.

Probe ID	Gene	Δβ per week	t-statistic	p-value
cg09523691	*ATG12*	-.0030	-3.33	.0020
cg16536918	*AVP*	-.0065	-3.08	.0039
cg25551168	*AVP*	-.0067	-3.61	$8.7x10^{-4}$
cg01143454	*C20orf141*	-.0034	-1.66	.10
cg26799474	*CASP8*	-.0087	-4.03	$2.6x10^{-4}$
cg13813391	*CMTM2*	-.0067	-3.99	$2.9x10^{-4}$
cg21842274	*CRHBP*	-.014	-4.49	$6.5x10^{-5}$
cg11540997	*DUOX2*	.0032	3.27	.0023
cg14409083	*EMP1*	-.0052	-3.00	.0047
cg15626350	*ESR1*	.011	3.83	$4.6x10^{-4}$
cg20291222	*GLIPR1L2*	-.0065	-3.63	$8.4x10^{-4}$
cg16098726	*GP9*	.0018	1.30	.20
cg14423778	*MBNL1*	-.0064	-2.96	.0052
cg05294455	*MYL4*	.010	4.83	$2.3x10^{-5}$
cg26267561	*OXT*	-.0029	-1.43	.16
cg20994801	*PIK3CD*	-.0061	-7.01	$2.4 x10^{-8}$
cg15561986	*POMT2*	-.0029	-3.55	.0010
cg00594952	*RIMS3*	.0035	3.21	.0027
cg22417398	*SCYL1*	.-.0033	-4.48	$6.7x10^{-5}$
cg10652277	*SLC30A9*	$1.4x10^{-4}$	.58	.57
cg16301617	*TMC6*	.-.011	-6.86	$3.8x10^{-8}$
cg26385222	*TMEM176B*	.0057	3.42	.0015
cg00411097	*TMEM184A*	-.0073	-5.63	$1.8x10^{-6}$
cg27210390	*TOM1L1*	.0086	8.92	$7.5 x10^{-11}$
cg06051311	*TRIM15*	-.011	-8.08	$8.8 x10^{-10}$
cg09244244	*TTC37*	-.0025	-1.53	.13

Figure 3-1: Manhattan plots depicting the association of all CpG sites with (A) PTB and with (B) GA. The y-axis is the negative log10 of the p-value for the association while the x-axis is position on each chromosome. The dashed line indicates experiment-wide significance.

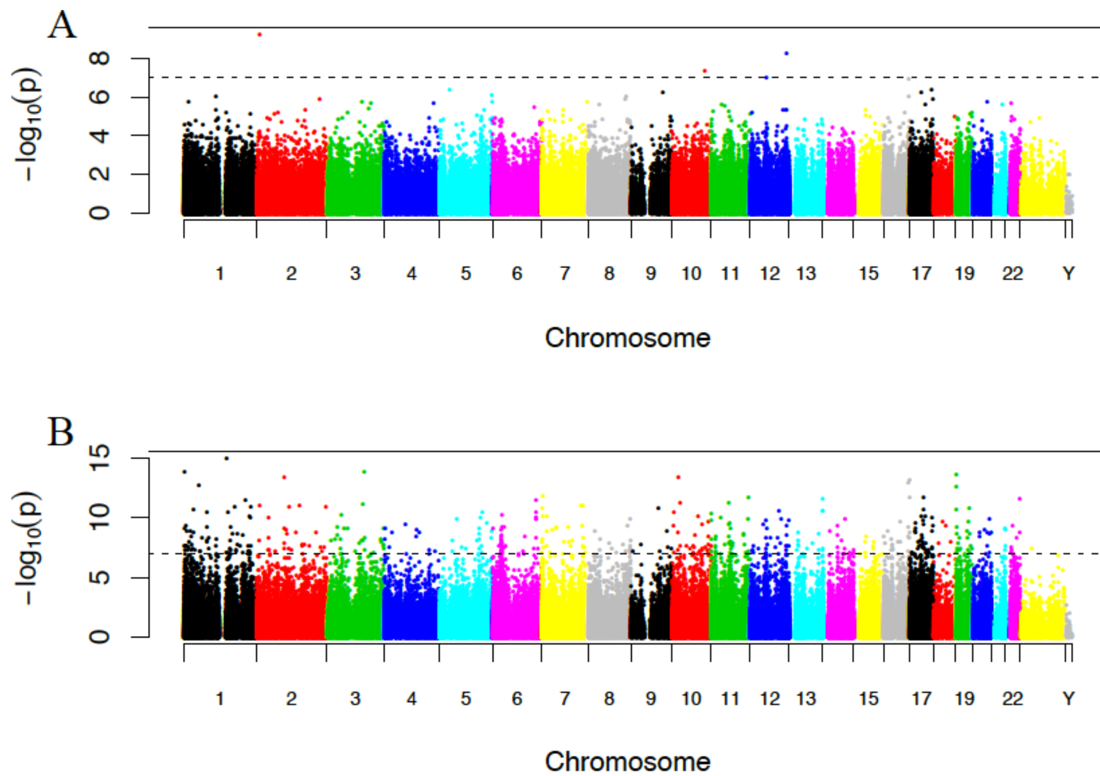


Figure 3-2: Scatter plots of the unmethylated vs. methylated signals (A versus B) for the five PTB-associated CpG sites that have 1000 Genomes SNPs within the probe**.**

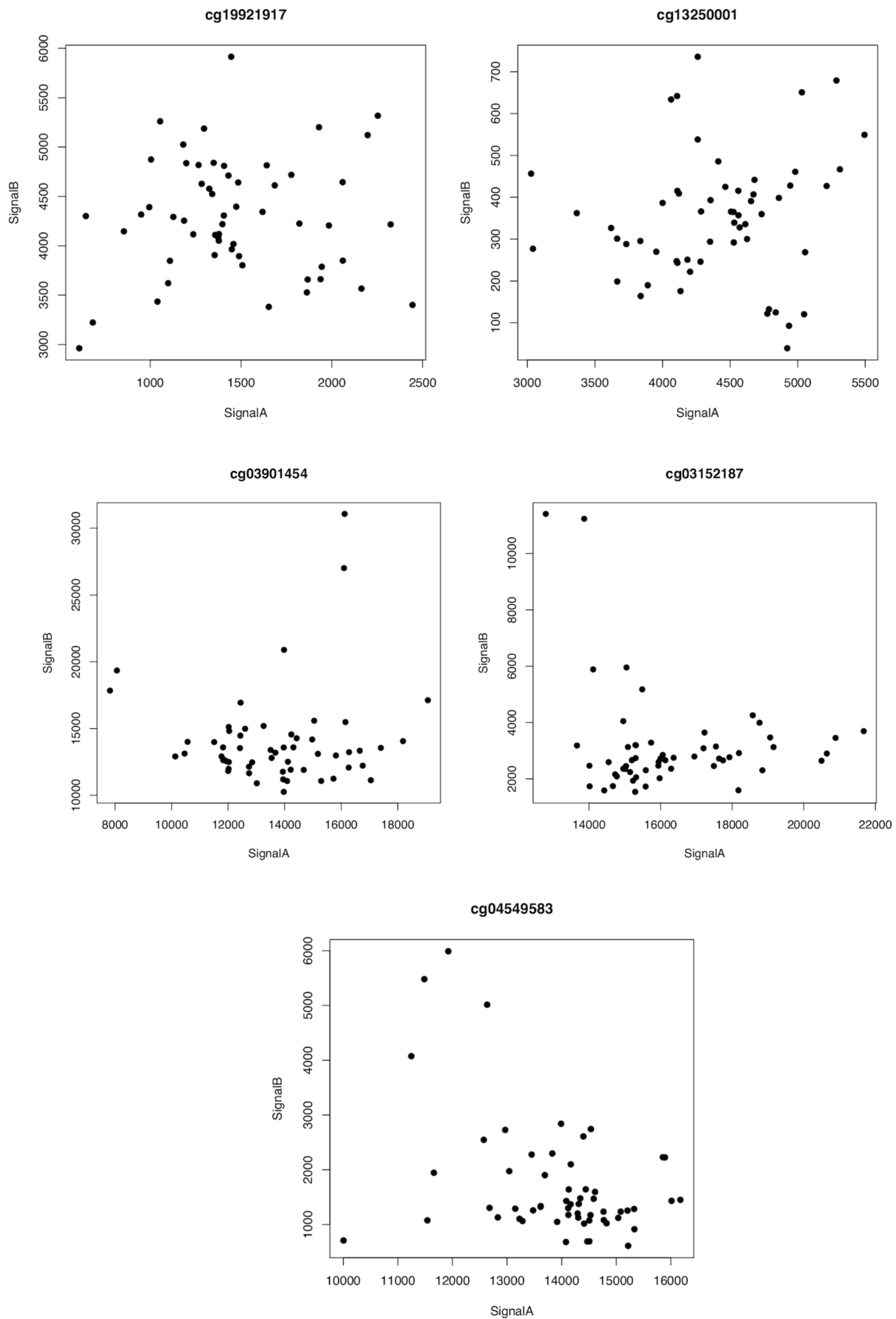
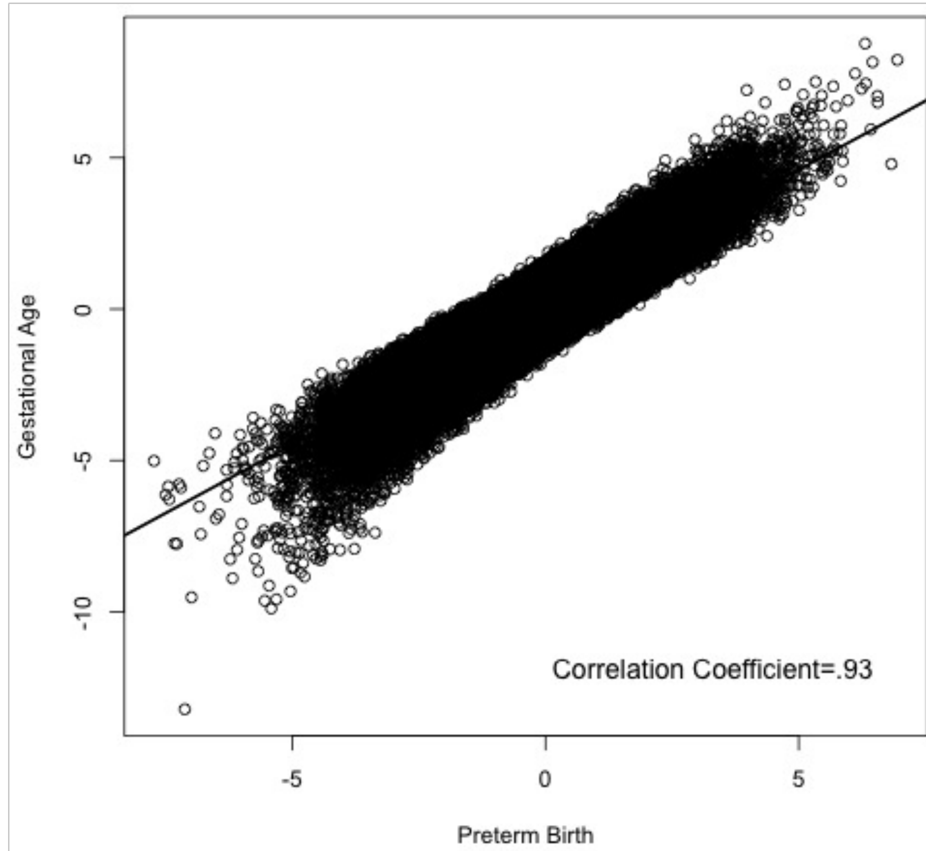


Figure 3-3: Correlation between the t-statistics depicting association analysis of CpG sites with PTB (x-axis) compared to GA (y-axis). All CpG sites are depicted whether or not they were associated with the outcome. In order to compare more directly compare the results from analyses of PTB and GA, we reversed the sign of the t-statistics for PTB in this plot.



References

1. Beck S, Wojdyla D, Say L, Betran AP, Merialdi M, et al. (2010) The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. Bull World Health Organ 88: 31-38.

2. Yuan W, Basso O, Sorensen HT, Olsen J (2001) Indicators of fetal growth and infectious disease in childhood--a birth cohort with hospitalization as outcome. Eur J Epidemiol 17: 829-834.

3. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ (2002) Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA 288: 728-737.

4. Loftin RW, Habli M, Snyder CC, Cormier CM, Lewis DF, et al. (2010) Late preterm birth. Rev Obstet Gynecol 3: 10-19.

5. Nosarti C, Reichenberg A, Murray RM, Cnattingius S, Lambe MP, et al. (2012) Preterm birth and psychiatric disorders in young adult life. Arch Gen Psychiatry 69: E1-8.

6. Behrman REB, A.S. (2007) Preterm Birth: Causes, Consequences, and Prevention; Outcomes CoUPBaAH, editor. Washington, D.C.: The National Academies Press. 792 p.

7. Goldenberg RL, Culhane JF, Iams JD, Romero R (2008) Epidemiology and causes of preterm birth. Lancet 371: 75-84.

8. Menon R, Conneely KN, Smith AK DNA methylation: an epigenetic risk factor in preterm birth. Reprod Sci 19: 6-13.

9. Kyrklund-Blomberg NB, Granath F, Cnattingius S (2005) Maternal smoking and causes of very preterm birth. Acta Obstet Gynecol Scand 84: 572-577.

10. Jeffcoat MK, Geurs NC, Reddy MS, Goldenberg RL, Hauth JC (2001) Current evidence regarding periodontal disease as a risk factor in preterm birth. Ann Periodontol 6: 183-188.

11. Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: a landscape takes shape. Cell 128: 635-638.

12. Barker DJ (1990) The fetal and infant origins of adult disease. BMJ 301: 1111.

13. Barker DJ, Gelow J, Thornburg K, Osmond C, Kajantie E, et al. (2010) The early origins of chronic heart failure: impaired placental growth and initiation of insulin resistance in childhood. Eur J Heart Fail 12: 819-825.

14. Paulsen M, Ferguson-Smith AC (2001) DNA methylation in genomic imprinting, development, and disease. J Pathol 195: 97-110.

15. Ginder GD, Gnanapragasam MN, Mian OY (2008) The role of the epigenetic signal, DNA methylation, in gene regulation during erythroid development. Curr Top Dev Biol 82: 85-116.

16. Wang D, Liu X, Zhou Y, Xie H, Hong X, et al. (2012) Individual variation and longitudinal pattern of genome-wide DNA methylation from birth to the first two years of life. Epigenetics 7: 594-605.

17. Beyan H, Down TA, Ramagopalan SV, Uvebrant K, Nilsson A, et al. (2012) Guthrie card methylomics identifies temporally stable epialleles that are present at birth in humans. Genome Res 22: 2138-2145.

18. Schroeder JW, Conneely KN, Cubells JC, Kilaru V, Newport DJ, et al. Neonatal DNA methylation patterns associate with gestational age. Epigenetics 6: 1498-1504.

19. Pan H, Chen L, Dogra S, Teh AL, Tan JH, et al. (2012) Measuring the methylome in clinical samples: improved processing of the Infinium Human Methylation450 BeadChip Array. Epigenetics 7: 1173-1187.

20. Roessler J, Ammerpohl O, Gutwein J, Hasemeier B, Anwar SL, et al. (2012) Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. BMC Res Notes 5: 210.

21. Barfield RT, Kilaru V, Smith AK, Conneely KN (2012) CpGassoc: an R function for analysis of DNA methylation microarray data. Bioinformatics 28: 1280-1281.

22. Kilaru V, Barfield RT, Schroeder JW, Smith AK, Conneely KN (2012) MethLAB: a graphical user interface package for the analysis of array-based DNA methylation data. Epigenetics 7: 225-229.

23. Oken E, Kleinman KP, Rich-Edwards J, Gillman MW (2003) A nearly continuous measure of birth weight for gestational age using a United States national reference. BMC Pediatr 3: 6.

24. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, et al. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 11: 587.

25. Feinberg AP (2007) Phenotypic plasticity and the epigenetics of human disease. Nature 447: 433-440.

26. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, et al. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41: 178-186.

27. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102: 15545-15550.

28. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, et al. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34: 267-273.

29. Alisch RS, Barwick BG, Chopra P, Myrick LK, Satten GA, et al. (2012) Age-associated DNA methylation in pediatric populations. Genome Res 22: 623-632.

30. Weiss A, Goldman S, Shalev E (2007) The matrix metalloproteinases (MMPS) in the decidua and fetal membranes. Front Biosci 12: 649-659.

31. Tency I, Verstraelen H, Kroes I, Holtappels G, Verhasselt B, et al. (2012) Imbalances between matrix metalloproteinases (MMPs) and tissue inhibitor of metalloproteinases (TIMPs) in maternal serum during preterm labor. PLoS One 7: e49042.

32. Romero R, Chaiworapongsa T, Espinoza J, Gomez R, Yoon BH, et al. (2002) Fetal plasma MMP-9 concentrations are elevated in preterm premature rupture of the membranes. Am J Obstet Gynecol 187: 1125-1130.

33. Harris SJ, Foster JG, Ward SG (2009) PI3K isoforms as drug targets in inflammatory diseases: lessons from pharmacological and genetic strategies. Curr Opin Investig Drugs 10: 1151-1162.

34. Vogel I, Gronbaek H, Thorsen P, Flyvbjerg A (2004) Insulin-like growth factor binding protein 1 (IGFBP-1) in vaginal fluid in pregnancy. In Vivo 18: 37-41.

35. Aparicio IM, Garcia-Herreros M, Fair T, Lonergan P (2010) Identification and regulation of glycogen synthase kinase-3 during bovine embryo development. Reproduction 140: 83-92.

36. Hur EM, Zhou FQ (2010) GSK3 signalling in neural development. Nat Rev Neurosci 11: 539-551.

37. Saint Just Ribeiro M, Hansson ML, Lindberg MJ, Popko-Scibor AE, Wallberg AE (2009) GSK3beta is a negative regulator of the transcriptional coactivator MAML1. Nucleic Acids Res 37: 6691-6700.

38. Watanabe T, Oyama T, Asada M, Harada D, Ito Y, et al. (2013) MAML1 Enhances the Transcriptional Activity of Runx2 and Plays a Role in Bone Development. PLoS Genet 9: e1003132.

39. Shen H, McElhinny AS, Cao Y, Gao P, Liu J, et al. (2006) The Notch coactivator, MAML1, functions as a novel coactivator for MEF2C-mediated transcription and is required for normal myogenesis. Genes Dev 20: 675-688.

40. Alva JA, Iruela-Arispe ML (2004) Notch signaling in vascular morphogenesis. Curr Opin Hematol 11: 278-283.

41. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, et al. (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473: 343-348.

42. Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, et al. (2011) Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell 9: 166-175.

43. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99: 247-257.

44. Hogg K, Price EM, Hanna CW, Robinson WP (2012) Prenatal and perinatal environmental influences on the human fetal and placental epigenome. Clin Pharmacol Ther 92: 716-726.

45. Maymon E, Romero R, Pacora P, Gervasi MT, Gomez R, et al. (2000) Evidence of in vivo differential bioavailability of the active forms of matrix metalloproteinases 9 and 2 in parturition, spontaneous rupture of membranes, and intra-amniotic infection. Am J Obstet Gynecol 183: 887-894.

46. Botsis D, Makrakis E, Papagianni V, Kouskouni E, Grigoriou O, et al. (2006) The value of cervical length and plasma proMMP-9 levels for the prediction of preterm delivery in pregnant women presenting with threatened preterm labor. Eur J Obstet Gynecol Reprod Biol 128: 108-112.

47. Fortunato SJ, Menon R, Lombardi SJ (1999) MMP/TIMP imbalance in amniotic fluid during PROM: an indirect support for endogenous pathway to membrane rupture. J Perinat Med 27: 362-368.

48. Stunkel W, Pan H, Chew SB, Tng E, Tan JH, et al. (2012) Transcriptome changes affecting Hedgehog and cytokine signalling in the umbilical cord: implications for disease risk. PLoS One 7: e39744.

49. Perkins E, Murphy SK, Murtha AP, Schildkraut J, Jirtle RL, et al. (2012) Insulin-like growth factor 2/H19 methylation at birth and risk of overweight and obesity in children. J Pediatr 161: 31-39.

50. Hoyo C, Fortner K, Murtha AP, Schildkraut JM, Soubry A, et al. (2012) Association of cord blood methylation fractions at imprinted insulin-like growth factor 2 (IGF2), plasma IGF2, and birth weight. Cancer Causes Control 23: 635-645.

51. Martina NA, Kim E, Chitkara U, Wathen NC, Chard T, et al. (1997) Gestational age-dependent expression of insulin-like growth factor-binding protein-1 (IGFBP-1) phosphoisoforms in human extraembryonic cavities, maternal serum, and decidua suggests decidua as the primary source of IGFBP-1 in these fluids during early pregnancy. J Clin Endocrinol Metab 82: 1894-1898.

52. Lockwood CJ, Wein R, Chien D, Ghidini A, Alvarez M, et al. (1994) Fetal membrane rupture is associated with the presence of insulin-like growth factor-binding protein-1 in vaginal secretions. Am J Obstet Gynecol 171: 146-150.

53. Rutanen EM, Karkkainen TH, Lehtovirta J, Uotila JT, Hinkula MK, et al. (1996) Evaluation of a rapid strip test for insulin-like growth factor binding protein-1 in the diagnosis of ruptured fetal membranes. Clin Chim Acta 253: 91-101.

54. Hillhouse EW, Grammatopoulos DK (2002) Role of stress peptides during human pregnancy and labour. Reproduction 124: 323-329.

55. McLean M, Bisits A, Davies J, Woods R, Lowry P, et al. (1995) A placental clock controlling the length of human pregnancy. Nat Med 1: 460-463.

56. Hobel CJ, Arora CP, Korst LM (1999) Corticotrophin-releasing hormone and CRH-binding protein. Differences between patients at risk for preterm birth and hypertension. Ann N Y Acad Sci 897: 54-65.

57. Perkins AV, Eben F, Wolfe CD, Schulte HM, Linton EA (1993) Plasma measurements of corticotrophin-releasing hormone-binding protein in normal and abnormal human pregnancy. J Endocrinol 138: 149-157.

58. Adkins RM, Krushkal J, Tylavsky FA, Thomas F Racial differences in gene-specific DNA methylation levels are present at birth. Birth Defects Res A Clin Mol Teratol 91: 728-736.

Chapter 4

Fetal DNA Methylation of Calcitonin (*CALCA*) Does Not Associate with Behavioral and Cognitive Outcomes in Childhood

Introduction

Maternal depression during pregnancy is a risk factor for preterm delivery [1], but maternal depression associates with lower gestational age even among term deliveries (GA) [2]. Furthermore, maternal depression during pregnancy increases the risk of behavioral problems in children [3]. A prominent hypothesis, known as the developmental origins of health and disease (DOHaD), suggests that the intrauterine and early life environments can increase risk for the development of chronic diseases over the lifespan [4]. DOHaD was first proposed based on the Dutch Hunger Winter famine studies, which first connected the gestational environment to increased rates of cardiovascular disease [5], psychiatric disorders [6] and metabolic dysregulation [7,8] in adults. Later studies showed epigenetic differences in this group [9,10], prompting the hypothesis that epigenetics may mediate the relationship between intrauterine and early life environment and outcomes later in life.

Recently, this concept has been demonstrated in an animal model. Jiao and colleagues examined genome-wide expression of the hippocampal region of offspring of stress pregnant mice who developed depression during adulthood [11]. They found 3 transcripts (*NDN*, *USP29*, and *αCGRP*) that were differentially expressed according to whether the mother was stressed or not during pregnancy. They found that *CALCA* methylation mediated the relationship between this gestational environment and the development of depression. Furthermore, when they administered a protein encoded by *CALCA*, there was an increase in depressive symptoms in the adult mice. Conversely, when they administered its antagonist, there was a decrease in depressive symptoms. Depressive symptoms were measured using well-accepted behavioral measures such as

the forced swim test [12], during which a mouse with depressive symptoms would spend more time floating then swimming. This comprehensive study not only showed the importance of the gestational environment in programming long-term behavioral effects, but it also suggests that *CALCA* plays a role in mediating the relationship between maternal depressive symptoms and offspring behavior. Based on this study, we prioritized calcitonin (*CALCA*) in order to examine whether *CALCA* methylation in umbilical cord blood mediates the relationship between maternal depressive symptoms during pregnancy and the development of internalizing behaviors during early childhood.

In the literature *CALCA* is implicated in the timing of delivery as well as regulation of the intrauterine environment [13-15]. For example, a protein encoded by *CALCA* is involved in implantation and angiogenesis of the placenta, which is necessary for a normal pregnancy [15]. DNA methylation of *CALCA* has also been proposed as a biomarker for bacterial sepsis in preterm neonates [14]. In data generated in our lab, CpG sites in *CALCA* associated with GA in cord blood [13], preterm birth (PTB) in maternal blood, and correlate between a mother and her fetus. In addition, *CALCA* has previously associated with psychiatric outcomes [11,16]. SNPs within *CALCA* associate with schizophrenia, Parkinson's disease and depression [16]. Proteins encoded by *CALCA* have been used to treat osteoporosis [17], mania [18] and have been proposed as a treatment for bipolar disease [19]. Antagonists to the proteins have been used to treat migraines [20-22], suggesting the potential for treatment if *CALCA* regulation at birth is predictive of childhood behavioral problems.

CALCA has six exons that produce two polypeptides through alternative splicing of exon 4 [23]. One of these polypeptides is calcitonin and is comprised of exons 1-4 and

is involved in calcium homeostasis through parathyroid regulation [24,25]. The other protein is calcitonin gene-related peptide (CGRP) and is made up of exons 1-3, 5 and 6 and is primarily expressed in the nervous system [26]. CGRP is involved in multiple functions such as vasodilation in the peripheral and cerebral blood vessels [27,28]. It is involved in placental implantation, trophoblast invasion, fetal organ development, and hormone signaling through the placenta [29]. Furthermore, injection of CGRP into the paraventricular nucleus stimulates the hypothalamic-pituitary-adrenal (HPA) axis, which is involved in stress response [30,31]. HPA axis dysfunction associates with depression [32]. Depressed women and their neonates have higher levels of plasma cortisol, suggesting that the maternal depression can signal across the maternal and fetal HPA axes [33]. It is therefore, important to evaluate the regulation and expression of *CALCA* and its potential role in the relationship between maternal depressive symptoms, PTB, and childhood behavioral outcomes. It is with all this is mind that we hypothesized that maternal depression during pregnancy associates with childhood internalizing behavioral outcomes through *CALCA* methylation in humans. Findings in this study can bring insight into the mechanism of the association between maternal depression with length of GA and childhood outcomes.

Methods

Subject Selection

To our knowledge, there is no spontaneous PTB cohort that has child behavioral data or psychiatric measure from the mothers available. Therefore, we used subjects from the Women's Mental Health Program (WMHP) at Emory University, a cohort comprised of primarily term deliveries. Women with psychiatric illnesses were prospectively

assessed to evaluate the perinatal course of their illness and the impact that illness has on their offspring. Mothers completed an intake questionnaire for demographic, socioeconomic, medical and psychiatric history, and were administered the Beck Depressive Inventory (BDI) to assess depression symptoms [34]. Women were evaluated at four to six week intervals throughout pregnancy to assess maternal depression and stress. Depressive symptoms scores at each time point were used to calculate the area under the curve (AUC) across pregnancy. To calculate the AUC each patient must have been seen at least three times during pregnancy. In total, data throughout pregnancy were available for 148 Caucasian and 132 African-American women who also provided umbilical cord blood for methylation.

A subset of the children born to these women was evaluated for behavior problems between 3 and 5 years of age (N=155). The children were administered the Childhood Behavioral Checklist (CBCL) [35] to assess behavioral outcomes. The CBCL is given to the parents and an alternative caregiver that assesses the child in syndrome scales and in DSM scales. CBCL syndrome scale is comprised of one total summary score and then grouped into externalizing and internalizing behaviors, which are also summary scores. In this study we focus on internalizing behaviors, because of the previous association in mice between maternal stress during gestational environment and depressive behavior. [11] Internalizing behaviors include withdrawal behaviors, anxiety/depression, and somatic symptoms.

Biological Sample Collection and DNA Extraction

Umbilical cord blood samples were collected at birth, and processed within 2 h of delivery. Blood was centrifuged at 4°C to separate out plasma and then frozen, at −80°C

until processing. DNA was extracted from the cellular fraction at the Emory Biomarker Service Center using a Qiagen Biorobot M48.

DNA methylation

The HumanMethylation27 BeadChip was used to interrogate > 27,000 independent CpG sites across the genome. In brief, 1ug of genomic DNA was bisulfite-converted, and analyzed using the HumanMethylation27 BeadChip procedure according to the manufacturer's instructions (Illumina). For a technical replicate we used a single female genomic DNA sample to assess chip quality and chip-to-chip variability in signal. Samples with probe detection call rates <90% or with an average intensity value of either <50% of the experiment-wide sample mean or <2,000 arbitrary units (AU) were excluded from the analysis. For each sample, the signals from methylated (*M*) and unmethylated (*U*) bead types were used to determine a β value for each queried locus. Each β value is calculated as $\beta = M/(U + M)$.

MethLAB [36] was used to test linear models where the β is logit transformed and modeled as the independent variable and the phenotype of interest (i.e. CBCL scores or BDI) is the response variable and adjusted for cell type proportions, fetal sex, and chip effects. The HumanMethylation27 array has 8 probes (Figure 4-1) in the *CALCA* gene. To examine the relationship between maternal BDI and gestational age (GA), we modeled BDI as the independent variable and GA as the outcome. For all test we used a Bonferroni correction for the 8 *CALCA* probes. In order to evaluate whether *CALCA* DNA methylation mediated the relationship between maternal depressive symptoms or GA and internalizing behaviors, we used a Sobel test [37], which tests for indirect effects of the casual variable on an outcome.

Regulation of CALCA expression

In order to examine the association between *CALCA* DNA methylation and expression in tissues relevant to birth outcomes and behavior (umbilical cord blood, placenta and brain regions), we used publically available data. To understand the role of *CALCA* in pregnancy we first evaluated the relationship of DNA methylation and expression in cord blood and placenta. In the GEO datasets GSE36831, expression of umbilical cord blood (N=48) and placenta (N=48) was examined using the HumanHT-12 V3.0 (Illumina) that contains 3 *CALCA* transcripts (ILMN_1702566, ILMN_2327244, ILMN_1807034; Figure 1). We next assessed whether brain regions play an important role in the regulation of *CALCA* because of its previous association with maternal depression during pregnancy. In GSE15745 [38] expression of the frontal cortex (N=130), pons (N=119), cerebellum (N=118), and temporal cortex (N=124) was examined using the HumanRef-8 v2.0 (Illumina) that contains 2 *CALCA* transcripts (ILMN_1702566, ILMN_2327244; Figure1). DNA methylation was evaluated using the HumanMethylation27 BeadChip (Illumina) in both datasets. We used linear regression to compare *CALCA* CpG sites to *CALCA* transcripts. All expression data was log2 transformed.

Results

The cohort is primarily comprised of Caucasian infants born in the late preterm and term range (Table 4-1). Maternal depressive symptoms associate with lower GA ($p=7.55 \times 10^{-4}$, $t=-3.41$) but do not predict any internalizing CBCL outcomes ($p>.05$). Similarly, GA does not predict any of the CBCL outcomes assessed ($p>.05$).

Association between fetal CALCA methylation and prenatal exposures

First, we examined the association between maternal depressive symptoms throughout pregnancy and fetal *CALCA* methylation. Both cg09188980 (t=-2.32, p=.021) and cg14348532 (t=-2.11 p=.036) associated with maternal depressive symptoms, but neither CpG site overcame the significance threshold for multiple test correction (p>.0063). Similarly, we evaluated the relationship between fetal *CALCA* methylation and GA. Only one CpG site (cg09188980; t=3.02,p=.0028; Figure 4-2) associated after multiple test correction; however, methylation of cg14348532 (t=2.51p=.013) was also nominally associated with GA.

Association of fetal CALCA methylation and internalizing behaviors

We next examined whether *CALCA* DNA methylation at birth would predict internalizing behavior between the ages of 3-5 years. We found one CpG site (cg22183706) that associated with internalizing behaviors (t=-2.21, p=.029), anxiety/depression (t=-2.00, p=.048), and withdrawal behavior (t=-2.66, p=.0088). However, cg22183706 did not remain associated after multiple test correction. No other CpG site associated with any of the 4 behavioral outcomes examined (data not shown).

We also tested whether fetal *CALCA* methylation mediated the relationship between maternal depression during pregnancy and the development of childhood behavioral problems and found no association. Similarly, when we evaluated whether fetal *CALCA* methylation mediated the relationship between GA and childhood internalizing behaviors, there was no association.

Regulation of CALCA expression

We evaluated the relationship between *CALCA* methylation and *CALCA* expression in tissues whose expression related to birth outcomes and behavior. Although there was no association after multiple test correction in any of the tissues examined, some findings are nominally significant and may be important for follow-up. For perinatal tissues we evaluated the regulation of *CALCA* expression in umbilical cord blood and in placenta. In cord blood, methylation of cg09068492 associated with increased exon 3 expression (ILMN_2327244; t=2.07; p=.043). In placenta, cg01971122 associated with decreased expression in 2 probes located in exon 2 and exon 3 (ILMN_1702566, t=-2.71, p=.0095; ILMN_2327244, t=-2.55, p=.014). For evaluating behavior we assessed the relationship between *CALCA* methylation and expression in different brain regions. In the frontal cortex, cg09068492 associated with decreased expression in exon 6 (ILMN_1807034, t=-2.13, p=.035). In the pons, cg10467022 associated with increased expression in exon 2 (ILMN_1702566, t=2.11, p=.036). There was no association of *CALCA* DNA methylation and expression in cerebellum or the temporal cortex.

Discussion

In this study, we examined the association of *CALCA* DNA methylation with perinatal maternal depression, gestational age, and childhood internalizing behaviors. However, after multiple test correction we found no evidence that *CALCA* methylation associated directly with maternal depressive symptoms or childhood behaviors. Furthermore, there was no association between GA and CBCL. GA may not predict CBCL in these data because our cohort is primarily born at term. We next evaluated whether *CALCA* methylation mediates these relationships. There was no evidence that

CALCA methylation mediates the relationship between maternal depressive symptoms and childhood internalizing behaviors. We also evaluated whether maternal depressive symptoms during pregnancy and GA associate and whether that association was mediated by *CALCA* methylation. Maternal depressive symptoms associated with lower GA, which is consistent with prior studies that reported the association of maternal depression during pregnancy with preterm birth [1,39]. Furthermore, *CALCA* methylation associated with lower GA. However, *CALCA* methylation did not mediate the relationship between maternal depressive symptoms during pregnancy and GA, though we recognize that this study is not well powered to test for mediation.

When we examined the regulation of *CALCA* expression, we found that different CpG sites predicted *CALCA* expression in different tissues. Methylation did not predict expression after multiple test correction, but that may be because we lacked a number of variables to adjust for technical and biological artifacts that influence power. Interestingly, in the frontal cortex, a *CALCA* CpG site associated with expression of exon 6, which would be exclusively included in the polypeptide CGRP that is primarily found in the nervous system.

Overall, this study provides no evidence that *CALCA* methylation mediates the relationship between maternal depressive symptoms, gestational age and childhood internalizing behaviors in our cohort. We might not have had the power to detect the relationship between DNA methylation in *CALCA* and maternal depressive symptoms, GA, or internalizing behavior. However, other studies that evaluate other tissues or a wider range of CpG sites may be capable of detecting more subtle effects. The original study was performed in hippocampus of mice and therefore may not be reflected in cord

blood or in humans. Finally, the results of this study do not preclude other genes from mediating the relationship between the intrauterine environment and development of subsequent childhood behavioral outcomes.

Table 4-1: Demographic table of the WMHP umbilical cord blood samples

Phenotype	Prenatal N=280
Maternal Age, years, Mean ± SD	33.74 ± 4.7
Weeks Gestation, Mean ± SD	38.87 ±1.1
Granulocytes, proportion, Mean ± SD	55.54 ± 8.6
Lymphocytes, proportion, Mean ± SD	31.12± 7.6
Fetal Sex, N,% male	(80) 51.6%
White, N, %	(148) 95.5%
BDI, area under the curve, Mean ± SD	376.78±311.13
3-5 year follow up; N=155	
Anxiety/depression, t score, Mean ± SD	52.35±4.41
Internalizing, t score, Mean ± SD	46.99±10.38
Withdrawn, t score, Mean ± SD	53.31±5.14

Figure 4-1: Schematic of *CALCA*. Green horizontal lines represent 2 CpG islands in this region. Verticles lines represent mRNA probe locations on the HumanHT-12 V3.0. Light blue verticle lines are CpG sites assessed on the HumanMethylation27 BeadChip.

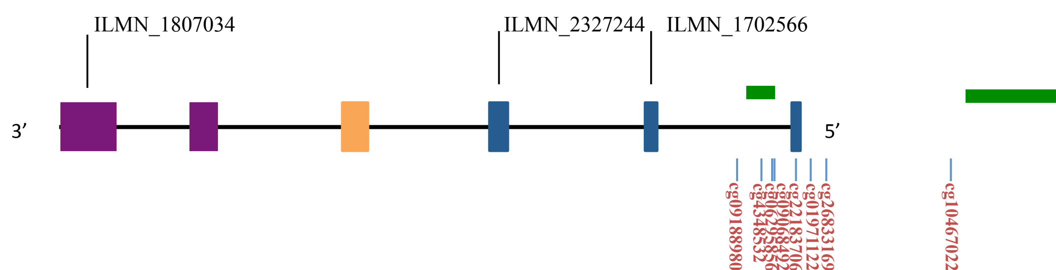
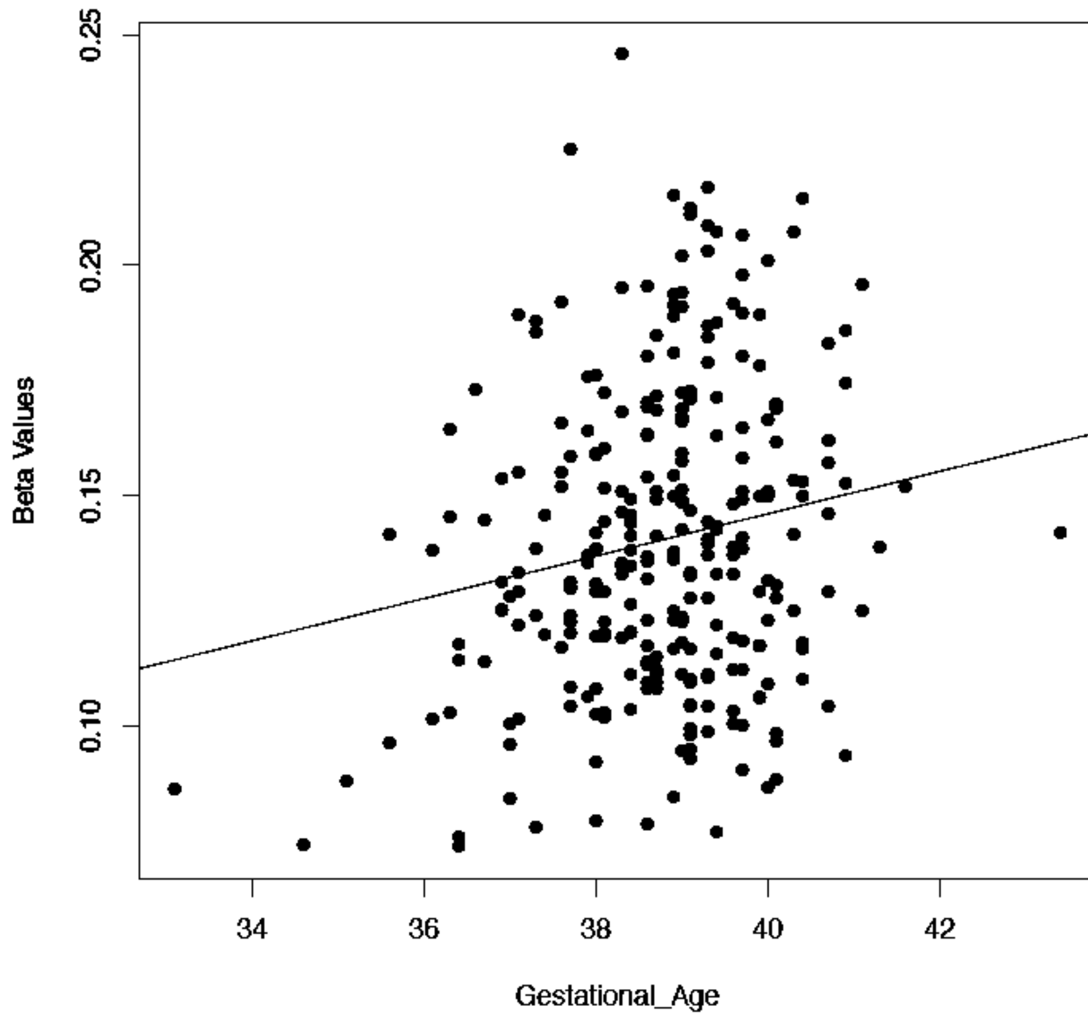


Figure 4-2: Association between DNA methylation in *CALCA* (cg09188980) with GA. The x-axis representing GA measured in weeks and the y-axis representing the methylation (logit transformed beta values).



References

1. Orr ST, James SA, Blackmore Prince C (2002) Maternal prenatal depressive symptoms and spontaneous preterm births among African-American women in Baltimore, Maryland. Am J Epidemiol 156: 797-802.

2. Field T, Diego M, Hernandez-Reif M (2010) Prenatal depression effects and interventions: a review. Infant Behav Dev 33: 409-418.

3. Beck CT (1999) Maternal depression and child behaviour problems: a meta-analysis. J Adv Nurs 29: 623-629.

4. Wadhwa PD, Buss C, Entringer S, Swanson JM (2009) Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med 27: 358-368.

5. Roseboom TJ, van der Meulen JH, Osmond C, Barker DJ, Ravelli AC, et al. (2000) Coronary heart disease after prenatal exposure to the Dutch famine, 1944-45. Heart 84: 595-598.

6. Susser ES, Lin SP (1992) Schizophrenia after prenatal exposure to the Dutch Hunger Winter of 1944-1945. Arch Gen Psychiatry 49: 983-988.

7. Ravelli AC, van der Meulen JH, Michels RP, Osmond C, Barker DJ, et al. (1998) Glucose tolerance in adults after prenatal exposure to famine. Lancet 351: 173-177.

8. Ravelli AC, van Der Meulen JH, Osmond C, Barker DJ, Bleker OP (1999) Obesity at the age of 50 y in men and women exposed to famine prenatally. Am J Clin Nutr 70: 811-816.

9. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, et al. (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum Mol Genet 18: 4046-4053.

10. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, et al. (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A 105: 17046-17049.

11. Jiao J, Opal MD, Dulawa SC (2013) Gestational environment programs adult depression-like behavior through methylation of the calcitonin gene-related peptide gene. Mol Psychiatry 18: 1273-1280.

12. Can A, Dao DT, Arad M, Terrillion CE, Piantadosi SC, et al. (2012) The mouse forced swim test. J Vis Exp: e3638.

13. Parets SE, Conneely KN, Kilaru V, Fortunato SJ, Syed TA, et al. (2013) Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age. PLoS One 8: e67489.

14. Tendl KA, Schulz SM, Mechtler TP, Bohn A, Metz T, et al. (2013) DNA methylation pattern of CALCA in preterm neonates with bacterial sepsis as a putative epigenetic biomarker. Epigenetics 8: 1261-1267.

15. Dong YL, Reddy DM, Green KE, Chauhan MS, Wang HQ, et al. (2007) Calcitonin gene-related peptide (CALCA) is a proangiogenic growth factor in the human placental development. Biol Reprod 76: 892-899.

16. Buervenich S, Xiang F, Sydow O, Jonsson EG, Sedvall GC, et al. (2001) Identification of four novel polymorphisms in the calcitonin/alpha-CGRP

(CALCA) gene and an investigation of their possible associations with Parkinson disease, schizophrenia, and manic depression. Hum Mutat 17: 435-436.

17. Chesnut CH, 3rd, Silverman S, Andriano K, Genant H, Gimona A, et al. (2000) A randomized trial of nasal spray salmon calcitonin in postmenopausal women with established osteoporosis: the prevent recurrence of osteoporotic fractures study. PROOF Study Group. Am J Med 109: 267-276.
18. Carman JS, Wyatt RJ (1979) Use of calcitonin in psychotic agitation or mania. Arch Gen Psychiatry 36: 72-75.
19. Vik A, Yatham LN (1998) Calcitonin and bipolar disorder: a hypothesis revisited. J Psychiatry Neurosci 23: 109-117.
20. Tepper SJ, Cleves C (2009) Telcagepant, a calcitonin gene-related peptide antagonist for the treatment of migraine. Curr Opin Investig Drugs 10: 711-720.
21. Durham PL, Vause CV (2010) Calcitonin gene-related peptide (CGRP) receptor antagonists in the treatment of migraine. CNS Drugs 24: 539-548.
22. Olesen J, Diener HC, Husstedt IW, Goadsby PJ, Hall D, et al. (2004) Calcitonin gene-related peptide receptor antagonist BIBN 4096 BS for the acute treatment of migraine. N Engl J Med 350: 1104-1110.
23. Lou H, Gagel RF (1998) Alternative RNA processing--its role in regulating expression of calcitonin/calcitonin gene-related peptide. J Endocrinol 156: 401-405.
24. Zaidi M, Moonga BS, Abe E (2002) Calcitonin and bone formation: a knockout full of surprises. J Clin Invest 110: 1769-1771.

25. Huebner AK, Keller J, Catala-Lehnen P, Perkovic S, Streichert T, et al. (2008) The role of calcitonin and alpha-calcitonin gene-related peptide in bone formation. Arch Biochem Biophys 473: 210-217.

26. Ishida-Yamamoto A, Tohyama M (1989) Calcitonin gene-related peptide in the nervous tissue. Prog Neurobiol 33: 335-386.

27. Girgis SI, Macdonald DW, Stevenson JC, Bevis PJ, Lynch C, et al. (1985) Calcitonin gene-related peptide: potent vasodilator and major product of calcitonin gene. Lancet 2: 14-16.

28. Brain SD, Grant AD (2004) Vascular actions of calcitonin gene-related peptide and adrenomedullin. Physiol Rev 84: 903-934.

29. Yallampalli C, Chauhan M, Endsley J, Sathishkumar K (2014) Calcitonin gene related family peptides: importance in normal placental and fetal development. Adv Exp Med Biol 814: 229-240.

30. Dhillo WS, Small CJ, Jethwa PH, Russell SH, Gardiner JV, et al. (2003) Paraventricular nucleus administration of calcitonin gene-related peptide inhibits food intake and stimulates the hypothalamo-pituitary-adrenal axis. Endocrinology 144: 1420-1425.

31. Smith SM, Vale WW (2006) The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. Dialogues Clin Neurosci 8: 383-395.

32. Mokrani MC, Duval F, Crocq MA, Bailey P, Macher JP (1997) HPA axis dysfunction in depression: correlation with monoamine system abnormalities. Psychoneuroendocrinology 22 Suppl 1: S63-68.

33. Smith AK, Newport DJ, Ashe MP, Brennan PA, Laprairie JL, et al. (2011) Predictors of neonatal hypothalamic-pituitary-adrenal axis activity at delivery. Clin Endocrinol (Oxf) 75: 90-95.

34. Beck AT, Steer RA, Ball R, Ranieri W (1996) Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. J Pers Assess 67: 588-597.

35. Achenbach TMR, L. A. (2000) Manual for the ASEBA Preschool Forms & Profiles. Burlington, VT University of Vermont, Research Center for Children, Youth, & Families.

36. Kilaru V, Barfield RT, Schroeder JW, Smith AK, Conneely KN (2012) MethLAB: a graphical user interface package for the analysis of array-based DNA methylation data. Epigenetics 7: 225-229.

37. Sobel ME (1982) Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. Sociological Methodology 13: 290-312.

38. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, et al. (2010) Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS Genet 6: e1000952.

39. Sanchez SE, Puente GC, Atencio G, Qiu C, Yanez D, et al. (2013) Risk of spontaneous preterm birth in relation to maternal depressive, anxiety, and stress symptoms. J Reprod Med 58: 25-33.

Chapter 5

A Pilot Study of DNA Methylation in Spontaneous Preterm Birth: Conclusions & Recommendations for Future Studies

Despite research and the advancement of healthcare systems in the US, the rate of preterm birth (PTB) remains high. Prevention strategies have not been successful at reducing the PTB rate, and interventions for PTB after initiation of labor are only intermittently successful. Advancement has been primarily in the management of the acute clinical needs of neonates that are born preterm [1]. However, PTB is still the primary cause of mortality in the first year [2], and the morbidities of being born preterm can last throughout life as children born preterm are more likely to develop chronic diseases [3,4]. Though the literature reports numerous risk factors that increase risk for PTB, the dire consequences of this condition cannot be fully addressed until the mechanisms that contribute to PTB are identified.

The objective of this pilot study was to examine DNA methylation in women and their fetuses to identify differences specific to PTB versus term birth (TB). Characterization of the genes being regulated in PTB may provide a greater understanding of the mechanisms that contribute to or result from PTB. We focused on DNA methylation in African Americans, a group with the highest risk for PTB, independent of socioeconomic status or other social factors [5,6]. Race-specific DNA methylation patterns distinguish African Americans from Caucasians [7]. DNA methylation patterns associate with race, and examining DNA methylation in the group with high risk for PTB may provide increased power to detect associations.

First we evaluated DNA methylation in maternal blood taken at delivery and found no association of large effect size between methylation of any CpG site and PTB. One limitation in this analysis is that maternal peripheral blood may not be the best tissue for studying PTB. The identity of the appropriate tissues is obscured by our lack of

understanding of the biological mechanisms underlying PTB. Blood could be the right tissue if a key mechanism that contributes to PTB is a systemic problem such as inflammation or disruption of the neuroendocrine system. However, if PTB is primarily triggered by localized factors such as intrauterine infection, blood is unlikely to be informative. Other tissues such as the myometrium may be appropriate because of its role in labor progression. Though DNA methylation patterns are tissue specific, some are consistent across tissues [8-10]. Therefore, future studies should establish epigenetic reference panels in multiple tissues throughout pregnancy and early development. This will serve as an important resource for interpreting the functional significance of PTB-associated DNA methylation differences that have already been identified. Future studies should also include prospective sample collection that could be used to assess DNA methylation patterns throughout pregnancy in multiple tissues. Prospectively collected tissues can be used to infer whether genes with PTB-associated regulation patterns cause PTB or are simply a consequence of it. Similarly, longitudinal studies may identify CpG sites that are informative for identifying and monitoring those at risk for delivering preterm or other adverse pregnancy outcomes. It should also be considered that DNA methylation is unlikely to be the only mechanism that is regulating gene expression. Therefore, a comprehensive approach should be taken that includes other epigenetic modifications, gene expression and proteomics.

Second, correlation between maternal and fetal DNA methylation was characterized to determine if DNA methylation patterns were shared based on PTB status. More than 5000 CpG sites associate between a mother and her fetus; correlated CpG sites could primarily be attributed to nearby genetic variation. In this analysis we

also identified CpG sites that correlate in paired PTB but not term birth samples. Similar to the results of the overall analysis, the majority of the CpG sites could be attributed to nearby genetic variation. This suggests that a complex relationship between genetic and epigenetic variation may contribute to the intergenerational transmission of PTB.

Third, the relationship between DNA methylation and gestational age (GA) was examined to identify CpG sites that distinguish early PTB from TB. We identified ~10,000 CpG sites that associated with GA in cord blood. We hypothesized that differences in methylation would reflect mechanisms underlying PTB risk. If this hypothesis were correct, a significant overlap between CpG sites that associated with GA in cord blood and with PTB in maternal blood would be observed. However, no CpG site associated with PTB in maternal samples after multiple test correction, and only 5.2% overlapped between cord blood and maternal blood at a nominal significance level. This comparison does not support the hypothesis that DNA methylation differences in these samples reveal common mechanisms underlying PTB. In umbilical cord blood, examination of the relationship between DNA methylation and PTB while adjusting for GA identified only 29 associated CpG sites. Interestingly, the majority of those sites could be attributed to a nearby genetic variant or a SNP directly in the probe sequence. This result is consistent with the observations from Chapter 2 that support a complex relationship between genetic and epigenetic factors in those with PTB. However, a significant limitation of this analysis is that we did not specifically evaluate sequence variation in this cohort. Future studies should include genotyping in order to better understand the relationship between genetics and epigenetics in PTB. It is very possible that the majority of the CpG sites that we identified primarily reflect developmental

differences. If this is true, they should replicate in cord blood from another PTB cohort. We were able to replicate 42% of the CpG sites that associated with PTB in a Caucasian cohort conducted by Cruickshank and colleagues [11]. However, we may not have been powered to detect PTB-specific CpG sites in this pilot study. Furthermore, comparing PTB samples with term birth samples may have identified CpG sites that represent developmental differences.

Finally, recognizing that the majority of GA-associated CpG sites are likely to reflect developmental differences, it remains unclear whether those developmental differences increase risk for long-term outcomes in children born preterm. DNA methylation of CpG sites in calcitonin (*CALCA)* associates with GA in cord blood, PTB in maternal blood and has been reported to mediate the relationship between gestational environment and later life behavior [12]. Thus, we evaluated the relationship between *CALCA* DNA, the intrauterine environment and childhood behavioral problems. In this study, *CALCA* methylation did not associate with maternal depressive symptoms during pregnancy. Furthermore, *CALCA* methylation did not associate with the development of internalizing, withdrawn or somatic behaviors or anxiety/depression in early childhood. *CALCA* methylation also did not mediate the relationship between maternal depressive symptoms and childhood behavioral outcomes. These negative results could have occurred for various reasons. Firstly, the relationship of *CALCA* methylation and the development of internalizing behaviors might not have manifested itself yet because the mouse study evaluated depression like behaviors in adult mice and we are evaluating it in children. Secondly, future studies should provide a focused examination in postmortem brain tissue from children who were born preterm. However, in such studies follow-up

cognitive and behavioral outcomes could not be assessed. Translating animals studies into human cohort studies is complicated by the availability of tissues in human cohorts specifically hippocampus in this study. Animal studies could begin to incorporate peripheral tissues, such as blood or saliva, into ongoing studies that focus on brain. This would better enable translation of those findings to human cohorts. However, a candidate gene approach may not be the most efficient way to address the overall question; genome-wide studies would be most informative.

Overall, this study had several limitations that are inherent to a pilot study using the HumanMethylation450 BeadChip. For example, the small sample size limited our power to only detect associations of large effect size. The addition of more samples will allow us to detect associations between CpG sites and PTB that have smaller effects. One limitation of using the HumanMethylation450 BeadChip is the limited coverage of the genome. Furthermore, by bisulfite converting the DNA we are unable to differentiate between 5' methylcytosine (5mC) and 5'hydroxymethylation (5hmC). 5hmC is an intermediate step from 5mC to demethylated CpG sites, though it is also a stable methylation state in its own right [13]. Future studies should incorporate evaluating 5hmc and 5mc to better understand the dynamic changes between the two in relation to PTB.

The results of this study provide insight into the patterns of DNA methylation in PTB. Our data suggest that it may not be simply genetic or epigenetic, but a complex interplay between two that contributes to PTB risk and its long-term consequences. In order to better understand this, future studies should use multi-generation cohorts with longitudinal collection of comprehensive phenotypic and biological measures to enable

hypothesis testing. This dissertation lays the foundation for future work examining the relationship between DNA methylation and the prevention, causes, and consequences of PTB.

References

1. Simmons LE, Rubens CE, Darmstadt GL, Gravett MG (2010) Preventing preterm birth and neonatal mortality: exploring the epidemiology, causes, and interventions. Semin Perinatol 34: 408-415.

2. Callaghan WM, MacDorman MF, Rasmussen SA, Qin C, Lackritz EM (2006) The contribution of preterm birth to infant mortality rates in the United States. Pediatrics 118: 1566-1573.

3. Goldenberg RL, Culhane JF, Iams JD, Romero R (2008) Epidemiology and causes of preterm birth. Lancet 371: 75-84.

4. (2007) Preterm birth causes, consequences, and prevention. . In: Behrman RE BA, editor. Institute of Medicine (US) Understanding premature birth and assuring healthy outcomes. Washington (DC): National Academies Press (US).

5. Goldenberg RL, Cliver SP, Mulvihill FX, Hickey CA, Hoffman HJ, et al. (1996) Medical, psychosocial, and behavioral risk factors do not explain the increased risk for low birth weight among black women. Am J Obstet Gynecol 175: 1317-1324.

6. McGrady GA, Sung JF, Rowley DL, Hogue CJ (1992) Preterm delivery and low birth weight among first-born infants of black and white college graduates. Am J Epidemiol 136: 266-276.

7. Adkins RM, Krushkal J, Tylavsky FA, Thomas F (2011) Racial differences in gene-specific DNA methylation levels are present at birth. Birth Defects Res A Clin Mol Teratol 91: 728-736.

8. Liang P, Song F, Ghosh S, Morien E, Qin M, et al. (2011) Genome-wide survey reveals dynamic widespread tissue-specific changes in DNA methylation during development. BMC Genomics 12: 231.

9. Smith AK, Kilaru V, Kocak M, Almli LM, Mercer KB, et al. (2014) Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. BMC Genomics 15: 145.

10. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, et al. (2012) Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol 13: R43.

11. Cruickshank MN, Oshlack A, Theda C, Davis PG, Martino D, et al. (2013) Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. Genome Med 5: 96.

12. Jiao J, Opal MD, Dulawa SC (2013) Gestational environment programs adult depression-like behavior through methylation of the calcitonin gene-related peptide gene. Mol Psychiatry 18: 1273-1280.

13. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, et al. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324: 930-935.