12-31-2010

DNA Methylation and its Association with Prenatal Exposures and Pregnancy Outcomes

Jennifer Straughen
University of South Florida

DNA Methylation and its Association with Prenatal Exposures and Pregnancy Outcomes

by

Jennifer Kornosky Straughen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Epidemiology and Biostatistics
College of Public Health
University of South Florida

Major Professor: Hamisu Salihu, M.D., Ph.D.
Kathleen O'Rourke, Ph.D.
Erin M. Siegel, Ph.D.
Getachew Dagne, Ph.D.
Terry Ashmeade, M.D.

Date of Approval:
August 11, 2010

Keywords: prenatal smoking, intrauterine growth restriction, epigenetics, small for gestational age, birthweight, cotinine

DEDICATION

To all those who provided love, support, perspective, and encouragement, especially my parents who taught me the importance of education and always encouraged me to chase my dreams and my husband, Rick whose patience and support helped me navigate my most difficult challenges.

ACKNOWLEDGEMENTS

Words cannot express my gratitude toward all those who have shaped my life and career. Along the way I learned from so many people, like those at the Tiedje Lab at Michigan State and at Los Alamos National Laboratory. I would really like to thank the Tiedje Lab for teaching me how to make lemonade out of lemons. It is a recipe I have frequently called upon. Then there was the incredible woman, Anne Sweeney, at Texas A&M who taught me how to write a 200 word abstract and had the patience to help me refine my writing skills. I am forever indebted.

Other influential people didn't enter my life until I came to USF, but they have been just as instrumental in helping me accomplish my goals. Jenny, first I must thank you for helping me drink my lemonade. I never could have succeeded without you and your sugar! I am so lucky to have you as a friend and confidant. Our weekly check-ins always managed to help me muster up just a little more motivation and somehow you helped me get through all the ups and downs unscathed. Kevin, thank you for teaching me to have faith. Having faith in myself, in god, and in those around me because I know it enabled others to have more faith in me.

My deepest gratitude is extended to my committee members whose guidance, support, and assistance helped this entire project come to fruition. I would also like to thank Phil Marty, Jane Carver, Morna Dorsey, Tina Pastoor, Maja Okuka, Celso Silva, Lewis Rubin, Xiaoming Gong, and Doris Wiener for all their support and assistance. Most importantly, I am grateful for Ileana's willingness to share her samples.

Finally, although many went unmentioned, your assistance has not been forgotten and your words and advice never went unheard.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Altered DNA methylation may lead to suboptimal fetal programming, increasing the risk of adverse pregnancy outcomes such as small for gestational age (SGA); however, few studies have examined the associations between DNA methylation, prenatal exposures, and fetal outcomes. Cross-sectional data from a larger, ongoing study were used to assess the impact of prenatal smoking on gene specific methylation of umbilical cord blood derived DNA and to investigate the association between gene-specific methylation and risk of SGA. The association between gene-specific DNA methylation and birthweight was also assessed. Maternal and infant covariates were abstracted from medical records, cigarette smoke exposure was determined by measuring cotinine in umbilical cord blood plasma, and the Illumina Infinium Methylation27 assay was used to assess CpG site specific methylation. Methylation was represented by a beta value ranging from 0 to 1. Gene-level methylation was calculated by averaging the methylation levels over the CpG sites interrogated in that gene. Logistic regression was used to generate adjusted odds ratios (OR) and 95% confidence intervals (CI) for the association between SGA and methylation of CYP1A1, HIF1A, GSTT1, and GSTM1 and the association between cotinine level and hypermethylation of CYP1A1, HIF1A, GSTT1, and GSTM1. DNA was considered hypermethylated if the beta value was greater than or equal to the 75th percentile. Univariate and multivariable linear regression were used to examine the association between birthweight and methylation of the IGF1 and IGF2 gene. The analyses included 90 singleton births. A 0.10 unit increase in methylation of GSTT1 increased the risk of SGA almost 3-fold (OR=2.69, 95%CI=1.34, 5.43). A 5ng/ml increase in cotinine level increased the risk of hypermethylation of

GSTT1 (OR=1.18, 95%CI=1.02, 1.37). Birthweight did not appear to be impacted by methylation of IGF2 (β=0.07, 95%CI=-2.91, 3.05), but a one standard deviation increase in methylation of IGF1 was associated with a 3.63% decrease in birthweight (95%CI= -6.49, -0.78). No differences in DNA methylation by prenatal vitamin intake were detected. These findings suggest that DNA methylation plays a critical role in fetal growth and may mediate the risk of SGA and low birthweight.

CHAPTER ONE

Introduction and theoretical framework

Specific aims

An increasing number of studies are linking prenatal exposures to adverse birth outcomes and adult onset disease. For example, heart disease, diabetes, and high blood pressure have all been associated with suboptimal pregnancy outcomes such as low birth weight and fetal growth restriction (1-4). Although it is posited that poor fetal growth is an in utero survival mechanism that enables the fetus to adapt to its environment, the mechanism underlying these adaptations are still poorly described.

Epigenetic modifications, heritable changes in gene expression that are not accompanied by changes in genotype, such as DNA methylation, have not been largely explored as potential mechanisms by which suboptimal uterine conditions leads to poor fetal growth in humans (5). Nonetheless, experimental evidence has demonstrated that DNA methylation is critical to normal development of mammals and that abnormal methylation can result in diseases such as Rett syndrome, neoplasias, and facial abnormalities (5-7). Accordingly, several genes critical for fetal and postnatal growth are epigenetically regulated (8-11). Furthermore, DNA methylation appears to be directly altered by exposures such as cigarette smoke and folic acid (12-15). Although aberrant DNA methylation has been linked to cigarette smoke, folic acid, and other potential in utero exposures, less is known about the association between DNA methylation and birth outcomes. Normal methylation of DNA, a process that is important for regulation of gene expression and DNA stability, may be disrupted in response to suboptimal uterine conditions, altering gene expression and subsequently preventing normal growth.

Given the lack of information about fetal exposures, DNA methylation, and pregnancy outcomes, this study sought to examine the association between DNA methylation of specific genes and indicators of fetal growth (small for gestational age and birthweight). In addition, it examined the impact of prenatal smoking on DNA methylation. Consequently, the analysis centered on the methylation of genes related to the metabolism of cigarette smoke carcinogens and genes important for fetal growth. The central hypothesis of this study is that suboptimal uterine conditions during pregnancy results in aberrant DNA methylation in umbilical cord blood-derived nucleated cells, which manifests itself as impaired fetal growth. The hypothesis is tested in a cross-sectional study including 92 infants recruited from a larger, on-going study at Tampa General Hospital (TGH). The study will help elucidate the biological mechanism by which cigarette smoke exerts its negative effects and, importantly, the role of gene-specific DNA methylation in mediating fetal growth.

The specific aims of this study include:

1. To investigate the association between small for gestational age (SGA) and the degree of DNA methylation in genes involved in xenobiotic metabolism and hypoxic response, specifically hypoxia inducible factor 1, cytochrome P450, glutathione s-transferase (GST) M1, and GSTT1
2. To determine if there is a dose-response relationship between DNA methylation of selected xenobiotic metabolism and hypoxic response genes and risk of SGA
3. To determine whether cigarette smoke exposure is associated with altered methylation levels in genes involved in metabolism of xenobiotics and hypoxic response
4. To determine whether suboptimal methylation of the Insulin-like growth factor (IGF) genes, IGF1 and IGF2 are associated with infant birthweight

Fetal growth and birthweight

Birthweight and intrauterine growth restriction (IUGR) are frequently used as indicators of fetal growth and survival potential. IUGR occurs when infants or fetuses fail to meet their growth potential and is typically operationalized based on birthweight and its appropriateness for a given race and gestational age (16, 17). Under this classification scheme, infants weighing less than the 10^{th} percentile for a given gestational age, race and gender are considered small for gestational age (SGA), those above the 90^{th} percentile are large for gestational age (LGA), and the remainder are appropriate for gestational age (AGA) (16). The percentage of SGA infants born in the United States ranges from 7.87% for non-Hispanic whites to 15.43% for non-Hispanic blacks (16). Hispanics have an intermediate rate of 9.30% (16). IUGR can also be categorized as symmetrical (proportional reductions in weight, length, and head circumference) or asymmetrical (reductions in weight are markedly higher than reductions in length or head circumference) (18, 19). Symmetrical growth restriction is characterized by smaller head dimensions and abdominal size that usually results from insults that occur early in gestation (birth defects, chromosomal anomalies, smoking, etc.) when growth occurs by cell division (17, 20). Asymmetrical growth restriction is marked by normal head dimensions and decreased abdominal size (17, 19). Asymmetrical growth restriction usually occurs in response to placental factors, diabetes mellitus, or inadequate nutrition in late pregnancy when cells are increasing in size (17, 20).

On the other hand, birthweight is reflective of fetal growth throughout the pregnancy. Birthweight is a product of gestational age and fetal growth, thus low birthweight, a common pregnancy outcome measure, can be attributed to poor fetal growth or preterm birth. Birthweight is frequently categorized as low birthweight (less

than 2500 grams), normal birthweight (2500 grams-4000 grams), or macrosomic (>4000 grams) (21). Birthweight is one of the best predictors of infant mortality, although the association may not be causal (22). The mean birthweight of singleton infants is 3,325 grams with about 6.2% of singleton infants and 7.9% of all infants being born low birthweight (21). However, the rates of low birthweight differ by race and ethnicity with non-Hispanic blacks having the highest rates (13.6%) followed by non-Hispanic whites (7.0%) and Hispanics (6.2%) (21). Birthweight has a very low heritability, thus it is thought that environmental factors play a critical role in birthweight determination (23).

The mechanism by which poor fetal growth manifests in response to environmental factors is unclear, but it may also depend on maternal morbidity and nutrition (16). Growth is dependent on adequate nutrition, oxygen, and hormones and growth factors, and suboptimal uterine conditions may impair availability of these factors, possibly by altering gene expression. Epigenetic modifications are a reversible mechanism that could account for the alterations in gene expression, and the ability of some fetuses to circumvent adverse outcomes attributed to adverse uterine environments.

Exposures associated with fetal growth

Maternal smoking

In the United States, about 21% of reproductive aged women smoke, and about 12% of women continue to smoke during pregnancy (24, 25). An even greater proportion of women are exposed to environmental tobacco smoke (33%) (26). Numerous large epidemiologic studies have provided evidence linking prenatal tobacco smoke exposure to a number of poor birth outcomes, including IUGR, stillbirth, and low birthweight, and now the relationship between IUGR and prenatal tobacco smoke exposure is presumed to be causal (27).

Reduced fetal growth attributed to prenatal tobacco smoke exposure may be due to nicotine exposure itself, hypoxia, placental changes, or direct effects of other chemicals, but the exact mechanism remains unknown. Nicotine, a known vasoconstrictor, can cross the placenta, exposing the developing fetus to higher nicotine concentrations than the mother (28, 29). The vasoconstrictive effect of nicotine may decrease the transfer of nutrients across the placenta and cause hypoxia, impairing normal processes of fetal growth and development (30). Conversely, animal studies suggest that carbon monoxide, not nicotine or other chemical exposures is responsible for the observed decreases in birth weight in fetuses prenatally exposed to tobacco smoke (31). High carbon monoxide concentrations can cause hypoxia, resulting in altered trophoblast expression, which are important for placental transfer, hormone production, and metabolism (32, 33).

Morphological changes in the placentas of smokers have been documented, but these changes have not been found to be associated with growth restriction (34). Placentas of smokers have thickened villous membranes, reduced capillary volumes, and decreased weight, but oxygen diffusion, a factor that may limit growth, does not seem to be impaired (34-36) Transfer of folate across the placenta may also be impaired by placental changes, a problem compounded by observations indicating that smokers have lower serum folate levels than nonsmokers (37). Studies examining maternal-fetal folate transfer in smokers and nonsmokers have not supported this hypothesis. Jauniaux et al. reported folate levels inside the first trimester gestational sac as measured in coelomic fluid did not differ between smokers and nonsmokers, but they were lower than those found in maternal serum, suggesting that smoking does not impair placental transfer of folate (37).

Several studies have examined the anthropometric measurements of infants born to smokers and nonsmokers, noting that tobacco smoke exposure usually causes

symmetrical growth restriction (38, 39). Higgins et al. demonstrated that maternal smoking decreases birth weight, crown-heel length, head circumference, and reduces brain:body weight ratio (BRR) (40). The same study reported that smoking cessation before 32 weeks' gestation eliminated the reductions in birth weight and head circumference, but deficits in crown-heel length, BRR, and ponderal index remained (40). Other studies have also linked prenatal smoking to reductions in linear growth, femur length, head circumference, and abdominal circumference (41-43). Numerous studies have demonstrated a dose-response inverse relationship between number of cigarettes smoked and birthweight (42, 44). In addition, fetal growth restriction correlates with nicotine concentration in fetal blood in a dose-response pattern (45). Both term and preterm infants are affected by maternal smoking (42).

Maternal smoking is associated with lower concentrations of amino acids in umbilical cord blood plasma and altered activity of trophoblasts and enzymatic activity (34, 46). Conversely, studies have shown that there are no differences in triglyceride, glucose, or albumin concentrations in umbilical cord blood between infants exposed to cigarette smoke prenatally and their unexposed counterparts (39). When compared to infants of nonsmokers, those born to smokers have lower insulin-like growth factor and IGF binding protein 3 in umbilical cord blood (39). In addition, smokers tend to gain less weight during pregnancy, but this does not explain the decreases in fetal size as an increase in nutrient intake does not prevent poor fetal growth in infants exposed to tobacco smoke prenatally (47, 48).

Folate

Folate is a water soluble B vitamin that plays a critical role in human reproduction (49). It occurs naturally in leafy green vegetables, beans, and liver, but it also exists in synthetic forms. Folic acid (naturally occurring or synthetic folate) is an important

substrate in one carbon metabolism, purine and amino acid synthesis, and methylation reactions (50-53). Folate is an essential component in reactions proceeding RNA and DNA synthesis which highlights the significance of folic acid for fetal growth and development (54).

Food fortification is one method that can help ensure adequate folic acid intake during the periconceptional period (55, 56). In 1998, the United States (US) Food and Drug Administration (FDA) began fortifying staple foods, including flour, rice, cereal, and pasta, with approximately 10% of the recommended daily intake (RDI) of folic acid (400μg) in an effort to increase folic acid intake (57, 58). Although successful, many reproductive aged women (aged 15-44) still do not have adequate folate intake (59). Non-Hispanic white women have the highest intake and the greatest proportion of women meeting the RDI (30.3%), followed by Mexican-Americans (17.1%) and non-Hispanic blacks (9.1%) (59). Dietary supplements including prenatal vitamins and other vitamins are an important source of folic acid for reproductive aged women. Women who took supplements containing folic acid were more than 10 times more likely to meet the RDI than non-users (95%CI=7.1, 14.7) (59).

The success of folate supplementation in decreasing the incidence of neural tube defects has lead to the investigation of the impact of folate on other fetal outcomes, such as low birthweight and IUGR (49). A large meta-analysis reported a decreased risk of low birthweight among infants born to women who used multimicronutrients (OR=0.81, 95%CI=0.73, 0.91) or iron-folic acid supplements (OR=0.83, 95%CI=0.74, 0.93) as compared to placebo, but there was no difference in risk for SGA (60). Similarly, Neggers et al found that folate supplementation increased birthweight by 48 grams (61). In addition, the Generation R Study found that periconceptional folic acid use increases birthweight by 68 grams (95%CI=37.2, 99.0 grams) (62). The same study also reported a reduced risk of SGA (OR=0.40, 95%CI=0.22, 0.72) (62). Although some studies

reported conflicting results, the importance of folic acid in one carbon metabolism support the observed associations and additional studies examining the mechanism underlying the association between folic acid and fetal growth are warranted (50-53, 63).

Biological mechanism linking exposures to birth outcomes and DNA methylation

Epigenetics and DNA methylation

Gene expression is mediated not only by DNA sequence, but also by epigenetic factors. Epigenetic modifications encompass three main processes: (1) DNA methylation; (2) histone acetylation; (3) micro-RNA molecules. Epigenetic modifications such as DNA methylation have been shown to affect disease susceptibility in human and animal studies as it regulates gene expression (5-7). DNA methylation occurs primarily at cytosine dinucleotides in the sequence cytosines and guanine (CpG) (64, 65). In most of the DNA sequence, CpG dinucleotides occur infrequently and a majority (~80%) are methylated. Alternatively, there are regions of DNA sequence that are very rich in CpG dinucleotides, termed CpG islands which primarily occur in gene promoter regions. Interestingly, during active gene transcription (open chromatin structure), CpG islands are unmethylated (64, 65). About 60% of genes have a CpG island at the 5' end of the promoter region which are important for transcriptional regulation (50, 52, 64). Expression of the promoter associated with a given CpG island is not dependent on methylation status (i.e. it may not be expressed even though it is not methylated), but methylation silences the promoter by promoting histone de-acetylation and a closed chromatin structure (64). Similar activity is seen in X-chromosome inactivation and imprinting (64). Consequently, disease can often be ascribed to failure to maintain normal DNA methylation, an epigenetic marker that can be assessed globally or at specific gene sites.

In normal fetal development, there is a wave of demethylation after fertilization that affects all methylated regions except the imprinted loci (66, 67). DNA methylation patterns are restored as development continues. De novo genome wide methylation occurs between formation of the blastocyst and gastrulation and then as cells differentiate, changes in gene-specific methylation occur (66). Imprinting occurs prior to fertilization and acts to silence either the maternal or paternal allele so that there is monoallelic expression of imprinted genes (66, 68).Disruption of normal imprinting is associated with several different syndromes, such as Beckwith Weidemann syndrome (66, 69).

DNA methylation requires S-adenosylmethionine (SAM), thus nutritional manipulation of the factors that are involved in the synthesis of SAM could impact DNA methylation reactions (50-52). However, SAM can be derived from a number of different dietary factors via different metabolic pathways and disruptions in one factor results in compensatory changes in other metabolic pathways (50, 51, 70). Figure 1.1 provides a schematic of some of the substrates and reactions that are involved in DNA methylation. Complex metabolic pathways involved in the methylation cycle may explain some authors' findings. Maloney et al. found that in rats, folate deficient diet intake during pregnancy did not impact DNA methylation in offspring (71). The authors did observe metabolic alterations, suggesting that alternative pathways were used to prevent aberrant methylation. Conversely, studies that looked at folate supplementation as opposed to folate deficiencies found an increase in DNA methylation, suggesting that diets deficient in methyl donors, such as folate do not directly influence the methylation of DNA, but diets rich in methylating factors can restore normal methylation (15, 52). Other studies have reported hypomethylated DNA in animals fed diets deficient in choline and methonine (50).

Maternal supplementation with methyl donors (ie; SAM, folic acid, choline, etc.) may be able to compensate for the negative effects of adverse intrauterine environmental exposures that impact DNA methylation. Dolinoy et al (2007) demonstrated that dietary supplementation of folic acid can prevent CpG site specific DNA hypomethylation caused by exposure to bisphenol A (BPA), a chemical used in the manufacturing of certain plastics (14). Restoration of normal methylation patterns prevents incorrect programming and could preclude subsequent diseases such as increased body weight, cancer and poor reproductive function (14). Similar results have been observed in animal studies (15, 72). Lillycrop et al. demonstrated that the offspring of rats fed protein-restricted diets had hypomethylated DNA in the glucocorticoid receptor and that supplementation with folic acid could prevent hypomethylation of the glucocorticoid receptor (15). Therefore, it seems that while high folate diets can prevent gene specific DNA hypomethylation, low folate diets do not necessary cause DNA hypomethylation, but it does cause changes in metabolism that may induce alterations in methylation (71).

Genetic alterations may also impact DNA methylation, such as single nucleotide polymorphisms (SNPs), especially when SNPs occur within genes that are integrally linked to the methylation cycle. Methionine synthase reductase, MTHFR, and methionine synthase are enzymes that have common polymorphisms and the latter two are influenced by folate deficiencies (51). Associations between these polymorphisms and diseases such as cancer, birth defects, and cardiovascular disease have been identified. However, their role in DNA methylation is not entirely clear as it may be modulated by gene polymorphisms as well as interactions between numerous dietary methyl donors (51, 73).

Alterations in DNA methylation may be the mechanism by which smoking causes poor fetal growth as methylation affects gene expression and in turn enzymatic activity

important for normal growth. Recent studies have found that several imprinted genes (genes that are turned on or off permanently via DNA methylation, histone acetylation, or chromatin modification) play a critical role in placental growth and nutrient transfer (33, 74). Furthermore, maternally expressed genes suppress fetal growth whereas paternally expressed genes enhance it, thus errors in imprinting can have various affects (68, 74). Environmental factors, such as smoking can have detrimental effects on methylation and thus gene imprinting; therefore, a better understanding of the smoking-induced changes in methylation may further delineate the mechanism by which prenatal smoke exposure causes IUGR (12, 74). Nonetheless, conflicting evidence has left the exact mechanism by which smoking exerts its negative effects elusive and to date, smoking cessation is the only strategy that is known to prevent IUGR and low birthweight.

Epigenetics, smoking, and birth outcomes

While DNA methylation has not been largely studied in relation to smoking during human pregnancies, it has been investigated in animal studies and studies of cancer. Reports of smoking-related cancers have demonstrated that tobacco smoke exposure may impact global and CpG site specific DNA methylation (12, 75). Consistent with these results, a study of bladder cancer cases suggests that cases had decreased DNA methylation and that tobacco smoke exposure modified the association between global DNA methylation and disease (13). The risk of bladder cancer was highest among current smokers with the highest levels of global DNA methylation (13). Interestingly, the study noted that global DNA methylation was not associated with genetic polymorphisms in 1-carbon metabolism such as MTHFR (13). Associations between global DNA methylation and smoking have been reported in other studies and some noted correlations between MTHFR and methylation only under conditions of low folate (76).

Methylation changes in specific genes may mediate fetal response to uterine conditions ultimately impacting fetal growth. Several genes that are involved in fetal response to hypoxia and cigarette smoke exposure may also mediate IUGR risk. Glutathione S-transferases (GST) metabolize environmental pollutants such as insecticides and carcinogens as well as by-products of oxidative stress (77). GSTM1 is a major phase 2 enzyme that catalyzes the conversion of phase one metabolites into glutathione and impaired enzyme activity may alter the response to cigarette smoke sand other toxicants (77-79). Hypermethylation of GSTM1 or GSTT1 may decrease the ability to metabolize xenobiotics, prolonging exposure and increase the risk of IUGR. The cytochrome (CYP) P450 enzymes and GSTT1 are also important for the metabolism of polyaromatic hydrocarbons (PAHs). Previous studies suggest that altered expression of genes involved in xenobiotic metabolism may increase the risk of IUGR (79-81). Further, cancer studies suggest that expression of some of the CYP450 enzymes are down regulated by promoter hypermethylation and abnormal methylation is associated with some cancers (82). Hypoxia inducible factor 1 (HIF-1) also modulates gene expression in response to hypoxia and recent evidence indicates that it is mediated by methylation whereby hypermethylation decreases transcription (83, 84). A summary of the genes described above is provided in Table 1.1.

Follow-up data from the National Collaborative Perinatal Project examined DNA methylation in relation to exposures that occurred throughout the life course and reported an association between prenatal smoke exposure and higher levels of global DNA methylation in peripheral blood mononuclear cells; however, prenatal smoke exposure was based on maternal self-report (85). Additional studies are needed to examine the impact of prenatal smoking on gene-specific methylation as this may impact gene expression and subsequent fetal growth.

Epigenetics, folate, and birth outcomes

Both IUGR and low birthweight manifest in response to a number of different causes, including birth defects and other chromosomal anomalies, multiple gestation, high altitude, extreme malnutrition, dietary deficiencies, abnormal placenta, or maternal smoking (17, 86, 87). Most studies examining folic acid and prenatal vitamin use focused on prevention of birth defects, but DNA methylation is one possible mechanism by which folic acid may mediate fetal growth.

Folic acid deficiency may prevent normal methylation of epigenetically regulated genes such as insulin-like growth factor (IGF) 1 and IGF2. In addition to IGF1 and IGF2, the IGF system also includes insulin, four receptors, and six binding proteins (88). It regulates fetal and placental growth, promoting cell growth and differentiation and inhibiting apoptosis (8, 9). Both IGF1 and IGF2 are expressed early in fetal development, however IGF2 expression exceeds that of IGF1 (88). IGF2 is a paternally imprinted gene and since expression is regulated by DNA methylation, it may vulnerable to abnormal methylation during development. Although imprinting can be detected as early as the 8-cell stage, after birth, IGF2 expression becomes biallelic in most tissues (88, 89). Imprinting of IGF2 is regulated by H19, but loss of imprinting of IGF2 can occur regardless of whether imprinting is disrupted in H19 (90). The importance of methylation in the expression of IGF2 is underscored by the fact that fetal overgrowth is associated with imprinting disorders such as Beckwith-Wiedemann syndrome (69, 91). Knockout studies of IGF1 or IGF2 decreases fetal weight in mice and partial deletion of IGF1 in humans has similar effects (8, 11). For example, in mice, deletion of IGF2 results in a fetus that is only 60% of the normal weight (8, 92).

Although epigenetic control of IGF1 has received less attention, animal studies suggest that IGF1 methylation is altered in intrauterine growth restricted (IUGR) rats and that hypermethylation decreases IGF1 expression (93). IGF1 and birthweight are

positively associated and growth restricted infants have low umbilical cord blood levels of IGF1 compared to their counterparts with normal growth (11, 93, 94). Infant sex and concentrations of IGF1 and IGF binding protein 3 in umbilical cord blood plasma explain about 38% of the variability in birthweight after adjusting for gestational age, parity and maternal height (95). Together, these findings suggest that IGF1 and IGF2 are critical drivers of fetal growth and that sub-optimal methylation may impair fetal growth (see Table 1.2).

Although the mechanism by which folic acid impacts fetal growth and development, its universal methyl-donor status, may enable it to help prevent sub-optimal methylation of genes critical for fetal growth. Findings of recent studies such as that of Steegers-Theunissen et al support this theory as they reported that infants exposed perinatally to folic acid had higher methylation of the IGF2 DMR (differentially methylated region) than their unexposed counterparts (10). In addition, a methylation increase of 1.7% in the IGF2 DMR was associated with increased birthweight (10). However, it is possible that the time of sampling (about 17 months after delivery) impacted methylation of IGF2 as IGF2 expression changes after birth (8, 10, 88). Similarly, prenatal exposure to famine has also been associated with decreased methylation of IGF2 (96). Few studies have evaluated the impact of methylation of IGF1 or IGF2 on fetal growth in humans and additional studies are needed to explore the possible associations.

Summary

An increasing number of studies suggest that DNA methylation is a critical component of fetal development, yet it has not been largely explored as potential mechanism by which suboptimal uterine conditions leads to poor fetal growth in humans (5). This study seeks to address the lack of information by examining the relationship

between DNA methylation of genes involved in xenobiotic metabolism and 1) biochemically validated prenatal cigarette smoke exposure and 2) risk of SGA. This information may help elucidate the biological mechanism by which cigarette smoke exposure causes adverse pregnancy outcomes. In addition, this study examines the relationships between DNA methylation of IGF1 and IGF2, prenatal vitamin use, and birthweight. Previous studies have reported conflicting evidence for the association between birthweight and prenatal vitamin use, but few have examined this association in relation to a biological mechanism (54, 61, 63). Understanding the mechanisms by which adverse pregnancy outcomes manifest may lead to enhanced prevention strategies to reduce the morbidity and mortality associated with adverse pregnancy outcomes such as SGA and low birthweight.

Table 1.1 Overview of CYP1A1, HIF1A, GSTM1, and GSTT1 as they relate to fetal growth, DNA methylation, and response to prenatal cigarette smoke exposure

Gene	Description/function	Literature overview	References
CYP1A1	• Phase 1 enzyme • Important for detoxification and metabolism of xenobiotics • Involved in metabolic activation of PAHs from tobacco smoke	• Expression of CYP1A1 associated with cigarette smoke exposure • cancer studies suggest that expression of some of the CYP450 enzymes are down regulated by promoter hypermethylation and abnormal methylation is associated with some cancers • Placental CYP1A1 upregulated in rat model of smoking induced IUGR • Excess PAHs may lead to DNA adducts	(80-82, 97)
HIF1A	• Modulates gene expression in response to hypoxia • Tightly regulated by oxygen concentration and determines the level of HIF1 activity	• HIF1a expression is critical for downstream activation of a number of genes involved in cell growth and viability as well as in vascularization, factors critical for normal fetal growth • It has been shown that the expression of HIF1a is epigenetically regulated and DNA methylation suppresses expression in some cell types • Abnormal methylation of HIF1a may suppress HIF1a and lead to fetal growth inhibition.	(83, 84, 98, 99)
GSTT1 and GSTM1	• 2 types of glutathione S-transferases • Phase 2 enzymes involved in detoxification of phase 1 metabolites into compounds that can be easily excreted • Act on a wide range of epoxides, hyperperoxides, and other substrates	• Metabolize environmental pollutants such as insecticides and carcinogens as well as by-products of oxidative stress • Enzymes involved in the metabolism of cigarette smoke and polyaromatic hydrocarbons (PAHs), thus impaired enzymatic activity may alter detoxification ability • Loss of expression of these genes may impair clearance of PAHs or their metabolic by-products, interfering with DNA transcription and replication, or impairment of trophoblast proliferation, all of which may increase the risk of IUGR	(77-79)

Table 1.2 Summary of IGF1 and IGF2

Gene	Description/function	Literature overview	References
IGF1	• Regulates fetal and placental growth, promoting cell growth and differentiation and inhibiting apoptosis	• Knockout studies of IGF1 decreases fetal weight in mice and partial deletion of IGF1 in humans has similar effects • Animal studies suggest that IGF1 methylation is altered in intrauterine growth restricted (IUGR) rats and that hypermethylation decreases IGF1 expression • IGF1 and birthweight are positively associated and growth restricted infants have low umbilical cord blood levels of IGF1 compared to their counterparts with normal growth • Infant sex and concentrations of IGF1 and IGF binding protein 3 in umbilical cord blood plasma explain about 38% of the variability in birthweight after adjusting for gestational age, parity and maternal height	(11, 93-95)
IGF2	• Regulates fetal and placental growth, promoting cell growth and differentiation and inhibiting apoptosis • Paternally imprinted gene • Expression is greater than that of IGF1 • After birth, IGF2 expression becomes biallelic in most tissues	• Knockout studies of IGF2 decreases fetal weight in mice • Findings of recent studies reported that infants exposed perinatally to folic acid had higher methylation of the IGF2 DMR (differentially methylated region) than their unexposed counterparts. • A methylation increase of 1.7% in the IGF2 DMR was associated with increased birthweight • Prenatal exposure to famine has also been associated with decreased methylation of IGF2	(10, 88, 89, 96)

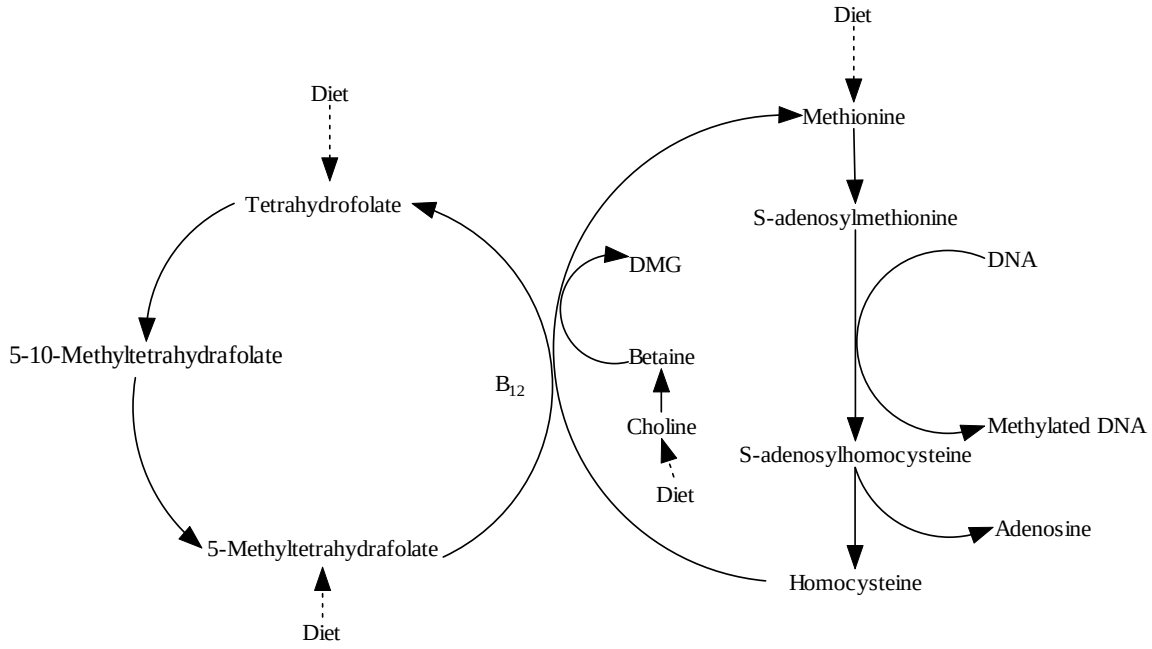

Figure 1.1 Schematic of some of the substrates (folate, choline, methionine) and reactions involved in the methylation of DNA (52, 73)

CHAPTER TWO

Manuscript 1: The association between fetal growth restriction, cotinine, and DNA methylation of detoxification and hypoxia related genes

Abstract

Objective: We assessed the impact of prenatal smoking on DNA methylation and the association between methylation and risk of small for gestational age (SGA).

Methods: Medical record data and biological samples from 90 singleton births were obtained from an ongoing, cross-sectional study. Cigarette smoke exposure was determined by measuring cotinine in plasma and CpG site-specific methylation in DNA extracted from umbilical cord blood was measured with the Illumina Infinium Methylation27 assay. Gene-level methylation was calculated by averaging the methylation levels over the CpG sites interrogated in that gene. Maternal and infant characteristics were compared by SGA status as well as by hypermethylation status using fisher's exact test and t-tests as appropriate. Logistic regression was used to generate adjusted odds ratios (OR) and 95% confidence intervals (CI) for the association between SGA and methylation of CYP1A1, HIF1A, GSTT1, and GSTM1 and the association between cotinine level and hypermethylation of the aforementioned genes.

Results: SGA infants were less likely to have adequate prenatal care and were more likely to be black and female. Infants with hypermethylation of GSTT1 were more likely to be black. A 0.10 unit increase in methylation of GSTT1 increased the risk of SGA

almost 3-fold (OR=2.69, 95%CI=1.34, 5.43) and the association appeared to be dose-dependent (p<0.001). The risk of hypermethylation of GSTT1 increased with increasing cotinine level (5ng/ml increase: OR=1.18, 95%CI=1.02, 1.37; 20ng/ml increase: OR=1.94, 95%CI=1.06, 3.53).

Conclusion: Methylation appears to play a critical role in fetal response to cigarette smoke and may influence the risk of SGA.

Introduction

Intrauterine growth restriction (IUGR) is a significant contributor to infant morbidity and mortality, thus its prevention has important public health implications (17, 100). The spectrum of evidence suggests that the relationship between prenatal smoking and IUGR is causal (27, 41, 101, 102). However, the mechanism by which prenatal tobacco smoke exposure causes adverse pregnancy outcomes remains poorly defined. Studies suggest a number of different modes of action, including vasoconstriction of the placenta, hypoxia, inhibited amino acid transport, and disrupted lipid metabolism (30, 103, 104). Although each of these mechanisms is biologically plausible and has supporting evidence, none has been able to explain why some fetuses, although exposed to tobacco smoke, are not growth restricted.

Previous studies have identified a number of genes involved in xenobiotic metabolism and these genes may mediate IUGR risk (79, 105). Glutathione S-transferases (GST) metabolize environmental pollutants such as insecticides and carcinogens as well as by-products of oxidative stress (77). GSTM1 is one of the enzymes involved in the metabolism of cigarette smoke and polyaromatic hydrocarbons (PAHs), thus impaired enzymatic activity may alter detoxification ability (77-79). Similarly, GSTT1 and the cytochrome (CYP) P450 enzymes are important for the

metabolism of PAHs and loss of expression of these genes may impair clearance of PAHs or their metabolic by-products, interfering with DNA transcription and replication, or impairment of trophoblast proliferation, all of which may increase the risk of IUGR (79, 80, 106, 107). Further, expression of some of the CYP450 enzymes is mediated by promoter methylation, with aberrant methylation occurring in some colorectal cancers, suggesting that changes in methylation may impact disease risk (82). In addition, hypoxia inducible factor 1 (HIF-1) modulates gene expression in response to hypoxia and recent evidence indicates that it is silenced by DNA methylation (83, 84).

Studies have shown that DNA methylation can be directly altered by exposure to cigarette smoke (12, 13). Recent studies have also found that in utero tobacco smoke exposure changes global and gene-specific methylation profiles in young children (108). However, the exposure and methylation changes in the aforementioned study were assessed about 5 to 6 years apart, thus, it is possible that postnatal exposures resulted in the observed DNA methylation changes (108). Other studies examining detrimental prenatal exposures also suggest that they may alter DNA methylation. For example, Pilsner et al found that maternal tibia lead burden was negatively associated with methylation of umbilical cord genomic DNA (109). Together, these observations suggest that epigenetic modifications, such as DNA methylation, may be a potential mechanism by which suboptimal uterine conditions caused by tobacco smoke exposure leads to IUGR, yet this potential pathway remains insufficiently explored. To examine this mechanism as a potential pathway, we undertake this study with the following hypotheses: 1) suboptimal uterine conditions may prevent normal fetal programming through altered DNA methylation of CYP1A1, HIF1A, GSTM1, and GSTT1, an event that is subsequently displayed phenotypically as SGA; 2) that there is a dose-response relationship between gene specific DNA methylation in DNA isolated from mononuclear cells and risk of SGA whereby infants with the highest methylation levels have the

highest risk of SGA; and 3) that cigarette smoke exposure is associated with altered methylation levels in genes involved in metabolism of xenobiotics.

Methods

Study sample and data collection

All study participants in this cross-sectional study were enrolled at Tampa General Hospital in Tampa, Florida as part of a larger ongoing study examining lymphocyte subpopulations and prematurity. All pregnant females delivering at Tampa General Hospital were eligible to participate in the Lymphocyte Study. However, infants born to women whose prenatal tests indicated that they were HIV positive or Hepatitis B positive were excluded. Maternal race and ethnicity are not factors for inclusion. For the present study, infants with birth defects were excluded and only singleton infants were eligible for inclusion.

De-identified demographic and clinical variables initially collected via medical record abstraction using standardized forms as part of the parent study were also obtained. The data elements collected include: gestational age, infant birth weight, infant sex, presence of infection, delivery complications, presence of birth defects, plurality, parity, gravidity, prenatal care usage, maternal age, and race.

Umbilical cord blood collection

In addition to medical record data, the Lymphocyte Study also collected umbilical cord blood samples. The umbilical cord blood samples were collected by venipuncture of the umbilical cord after delivery of the placenta into tubes containing EDTA and were processed within 24 hours of collection. Samples were processed at the University of South Florida. Plasma was removed and stored in 1ml tubes at -80°C for subsequent cotinine analysis. A ficoll gradient separation was used to isolate the mononuclear layer.

The amount of cord blood processed varied as the amount collected differed for each infant. However, it ranged from 0.5ml to 5.5ml. After separation, samples were suspended in freeze media (fetal bovine serum and 10% DMSO) and stored in liquid nitrogen. Studies have demonstrated that long term storage of cryopreserved cells does not impact cell viability or recovery with greater than 80% of nucleated cells recovered (110).

DNA isolation and DNA methylation assessment

DNA isolation and methylation assessment was done at Wayne State University Applied Genomics Technology Center. Laboratory personnel were blinded to birth outcome and other maternal and infant health indicators. DNA was isolated from the mononuclear fraction of umbilical cord blood using the Qiagen EZ1 DNA tissue kit according to Lum et al with the exception that PBS (phosphate buffered saline) was substituted for TE (tris ethylenediaminetetraacetic acid) buffer (111). The mononuclear fraction is largely comprised of monocytes and lymphocytes, but also contains hematopoietic stem cells (112). Changes in cells derived from umbilical cord blood should more directly reflect changes that occurred in relation to suboptimal fetal environment, leading to IUGR. After extraction, DNA was quantified by loading 3µl of the DNA suspension in the Trinean Dropsense96.

Bisulfite modification of 0.5 µg of DNA was then done with the EZ-96 DNA Methylation Kit™ per the manufacturer's instructions (Zymo Research Corp., USA). Quantitative, loci-specific methylation of the bisulfite modified DNA was assessed using the Infinium HumanMethylation27 BeadArray™ (Illumina Inc, San Diego, CA) according to the manufacturer's guidelines. The array interrogates 27,578 CpG loci located in more than 14,000 genes. For each CpG site, two different probes are hybridized with the bisulfite modified DNA (one against the methylated site and one against the

unmethylated site). Next, a single-base extension adds one of two possible fluorescent probes (one for methylated (C) and one for unmethylated (T) alleles). Methylation status is then represented by a beta value which is calculated from the ratio of fluorescent signals from methylated to the sum of methylated and unmethylated probes and ranges from 0 (unmethylated) to 1 (methylated). Background normalization was done according to the guidelines recommended by Illumina using the GenomeStudio Methylation module. In short, this method subtracts the average signal of the negative control bead-types from the probe signals. Normalized beta values were then output and used in subsequent analyses. Heat maps were generated with the MultiExperiment Viewer (113, 114).

A subset of samples was run in duplicate in order to assess inter-chip variability. In addition, CpGenome Universal Methylated DNA (Millipore, Temecula, CA) was bisulfite treated and run with the methylation assay as a positive control. Inter-chip variability was found to be highly reproducible. Pearson correlation coefficients were greater than 0.99 for each set of replicates ($p<0.0001$). In addition to running a positive control to ensure bisulfite conversion and accuracy of methylation, internal validity was assessed by examining gender specific methylation of 6 x-linked housekeeping genes (EFNB1, ELK1, FMR1, G6PD, GPC3, GLA) (115, 116).Overall, methylation of these 6 genes was as expected in that females exhibited hemimethylation and males had very little methylation at the loci in these genes ($p<0.0001$ for each gene). Figure 2.1 depicts the gender specific methylation patterns of these 6 housekeeping genes.

Cotinine assessment

In utero exposure to tobacco smoke (through either passive or active smoking) was evaluated by measuring cotinine, a metabolite of nicotine, in plasma from umbilical cord blood. Cotinine has a long half-life and has been previously validated as a

biomarker of tobacco exposure; therefore, it is the gold standard measure of tobacco smoke exposure (117-119). A solid phase competitive ELISA was used to assess cotinine level (Calbiotech, California). All samples were run in duplicate with controls and standards per the manufacturer's instructions. In short, 10 µl of plasma sample was combined with 100 µl of enzyme conjugate, mixed, and incubated for 60 minutes at room temperature in the dark. Samples were then washed with distilled water and residual moisture removed. After adding 100 µl of substrate reagent, the samples were incubated for 30 minutes in the dark and 100 µl of stop solution was added. Absorbance was read on a plate reader at 450nm.

Statistical analyses

Demographic and clinical variables were classified as follows: parity (nulliparous or multiparous), and race (black or non-black), prenatal care (adequate or not), labor and delivery complications (yes or no), and infant sex (male or female). Gestational age was assessed using both the clinical estimate and date of last menstrual period. Small for gestational age (SGA) was used as a surrogate indicator of intrauterine growth restriction. The birth weight percentiles for gestational age created by Alexander et al. were used to classify infants as SGA (<10th percentile for a given gestational age), appropriate for gestational age (AGA) (10-90th percentile), or large for gestational age (LGA) (>90th percentile) (16). Demographic and clinical variables of were compared between SGA and non-SGA infants (AGA and LGA infants) using fisher's exact test.

Methylation level was measured at multiple CpG sites for CYP1A1 and HIF1A; therefore, DNA methylation level for each CpG site was averaged over the gene of interest. In analyses examining the risk of SGA, DNA methylation was treated as a continuous variable so that we could assess the impact of several methylation levels on SGA risk. However, in analyses examining the impact of cigarette smoke exposure on

level of DNA methylation, DNA methylation was dichotomized as hypermethylated (≥75th percentile) or unmethylated (<75th percentile). There is currently no consensus on what level of methylation is indicative of hypermethylation and the 75th percentile was chosen so that we could adequately differentiate infants with the highest levels of methylation (120, 121). Although infants are considered exposed to cigarette smoke when the umbilical cord cotinine level is greater than 1ng/ml, 5 ng/ml and 14 ng/ml increases in cotinine levels are more indicative of active cigarette smoking (119, 122, 123). Therefore, these cut-points were used in our analyses examining cigarette smoke exposure and its association with DNA methylation.

A t-test was used to compare the mean methylation levels of each gene of interest between SGA and non-SGA infants as well as between infants exposed and unexposed to cigarette smoke. Logistic regression was used to compute adjusted odds ratios and 95% confidence intervals for the association between 1) SGA and the methylation level and 2) methylation level and cigarette smoke exposure. In logistic regression models, covariates and potential confounders were treated as continuous variables whenever possible (e.g.cotinine level, maternal age). The contribution of each differentially methylated gene of interest was assessed independently. Adjusted models controlled for covariates and potential confounders. Confounders were identified by comparing the crude and adjusted odds ratios. If the estimate changed by more than 10%, the variable was adjusted for. Dose-response trends were assessed with the Cochran-Armitage trend test and by generating effect estimates for different levels of exposure. All tests of hypothesis were two-tailed with a type 1 error rate fixed at 5 percent. SAS version 9.1 (SAS Institute, Cary, NC) was used to perform all analyses. This study was approved by the University of South Florida Institutional Review Board.

Results

After excluding infants with missing data (n=1) and birth defects (n=1), 90 infants were included in the analysis, of which 11.11% (n=10) were SGA and 88.89% (n=80) were non-SGA. Table 2.1 presents maternal and infant covariates of SGA and non-SGA infants. Almost 30% of black infants as opposed to 5.80% of non-black infants were SGA (p=0.01) When compared to non-SGA infants, SGA infants were more likely to be female (p=0.01). The cotinine levels of the 11 exposed infants ranged from 1.00ng/ml to 100ng/ml, with an overall mean of 36.43ng/ml. However, the mean cotinine level was higher among SGA infants (11.80 ng/ml) than non-SGA infants (3.53 ng/ml).

A heat map representing the methylation level for the CpG sites used to determine the gene-specific methylation level for each sample is depicted in Figure 2.2. This figure also depicts the positive control DNA for which the sample was almost completely methylated as expected. Mean methylation of the 4 genes of interest (CYP1A1, HIF1A, GSTM1, GSTT1) ranged from 0.04 (standard deviation (SD)=0.02) for HIF1A to 0.10 (SD=0.10) for GSTM1. Mean methylation levels by SGA and smoking status are presented in Figure 2.3. Visual inspection suggested that there was little to no difference in methylation level by SGA or smoking status in the CYP1A1 or HIF1A genes. There appeared to be marginal differences by SGA status in the GSTM1 gene whereas GSTT1 had the greatest variation in methylation level by for both SGA and smoking status. Furthermore, when data were compared by methylation level few differences in maternal and infant characteristics were found (data not shown). However, of the non-black infants, 20.29% had hypermethylation of GSTT1 whereas of the black infants 42.86% had hypermethylation of GSTT1 (p=0.05). In addition, of the infants with complications, 40.74% had hypermethylation of CYP1A1 whereas only 17.46% of infants without complications had hypermethylation of CYP1A1 (p=0.03).

In both crude and adjusted logistic regression models, only 1 of the 4 genes was significantly associated with SGA (Table 2.2). A 0.01 unit increase in methylation of the GSTT1 gene was associated with an increased risk of SGA (OR=1.10, 95%CI=1.03, 1.18) and the association appeared to be dose dependant. The risk of SGA was most pronounced among infants with the highest methylation levels of GSTT1 as the risk of SGA increased with increasing methylation ($p<0.001$). A 0.05 unit increase in methylation level was associated with a 22% increased risk of SGA (OR= 1.22, 95%CI=1.06, 1.40), and a 0.10 unit increase in methylation increased the risk of SGA almost 3-fold (OR=2.69, 95%CI=1.34, 5.43). Increased methylation of the HIF1A gene appeared to increase the risk of SGA, but the association did not reach statistical significance in either crude (OR=1.23, 95%CI=0.88, 1.72) or adjusted analyses (OR=1.59, 95%CI=0.99, 2.56).

Table 2.3 presents the crude and adjusted odds ratios and 95% confidence intervals for the association between hypermethylation (methylation level greater than or equal to the 75^{th} percentile) and prenatal smoking. Prenatal smoking as measured by a 5 ng/ml increase in cotinine level was not associated with methylation level of CYP1A1 (OR=0.99, 95%CI=0.85, 1.16), HIF1A (OR=1.08, 95%CI=0.94, 1.24), or GSTM1 (OR=0.94, 95%CI=0.77, 1.16). Similarly, a 14 ng/ml increase in cotinine was not associated with hypermethylation of the same three genes (CYP1A1, HIF1A, and GSTM1). However, methylation of GSTT1 was associated with smoking. A 5 ng/ml increase in cotinine level was associated with an 18% increased risk of hypermethylation (OR=1.18, 95%CI=1.02, 1.37) and a 14 ng/ml increase in cotinine level was associated with a 60% increased risk of hypermethylation of GSTT1 (OR=1.59, 95%CI=1.04, 2.42). Further, the risk of hypermethylation increased with increased cotinine level ($p=0.02$) whereby a 20 ng/ml increase in cotinine increased the risk of hypermethylation 2-fold (OR=1.94, 95%CI=1.06, 3.53) and a 30.0 ng/ml increase in cotinine level was

associated with a 2.7 fold increased risk of hypermethylation (OR=2.70, 95%CI=1.10, 6.64).

Discussion

This study found that the risk of SGA increased with increasing methylation of the GSTT1 gene. GSTT1 is important in fetal response to hypoxia and metabolism of environmental contaminants, thus it is interesting to note that methylation of GSTT1 was also associated with cigarette smoke exposure.

Few studies have examined the impact of methylation changes in DNA isolated from umbilical cord blood in relation to fetal outcome or fetal exposures. Two previous studies reported associations between prenatal tobacco smoke exposure and DNA methylation. Terry et al found higher levels of global methylation in exposed infants whereas Breton et al found lower levels of methylation in AluYb8 and higher methylation 8 other genes, though none of the genes were the same as those examined in the present study (85, 108). Although the study by Terry et al examined global methylation changes rather than gene-specific changes as presented here, the conclusions were similar to ours as the authors observed that cigarette smoke exposure increases methylation in mononuclear cells (85). However, while promoter specific methylation is associated with transcriptional silencing, global methylation is more representative of DNA stability and cancer studies indicate that global and CpG site specific methylation can be quite different (i.e. global hypomethylation and CpG site specific hypermethylation can coexist) (124).

Although previous studies have identified methylation changes in other genes associated with IUGR, to our knowledge previous studies have not examined the risk of SGA in relation to methylation of CYP1A1, HIF1A, GSTT1, or GSTM1 in DNA from umbilical cord blood-derived nucleated cells (125, 126). Although the association

between HIF1A and SGA did not reach statistical significance, this study may have been under-powered to detect an association. Loss of expression of HIF1A may impact the ability of the developing fetus to respond to hypoxia, thus it is biologically plausible that methylation of HIF1a may contribute to IUGR risk. HIF1a expression is critical for downstream activation of a number of genes involved in cell growth and viability as well as in vascularization, factors critical for normal fetal growth (98). Abnormal methylation of HIF1a may suppress HIF1a and lead to fetal growth inhibition. Further, it has been shown that the expression of HIF1a is epigenetically regulated (84). Additional studies with larger samples sizes are needed to assess the association between methylation of HIF1A and SGA.

Loss of GSTT1 expression could inhibit detoxification of xenobiotics, increasing the risk of IUGR. Hypermethylation of GSTT1 may decrease GSTT1 expression, causing an excess of phase 1 metabolites and increased oxidative stress (127, 128). In turn, this may lead to the formation of DNA adducts, cellular damage, or altered cell signaling (127). Although these findings have not yet been replicated, it is conceivable that SGA manifests in response to adverse uterine conditions via a methylation-mediated mechanism. The reported association between cigarette smoke exposure and hypermethylation of GSTT1 further supports the proposed mechanism as hypermethylation may alter gene expression and allow the accumulation of reactive oxygen species. Unfortunately, our small sample size prevented us from conducting a mediation analysis. In addition, our study did not examine the effects of differential methylation on gene expression, thus additional studies are needed to examine the impact of methylation changes on gene expression.

This study has several strengths including its use of cotinine to assess cigarette smoke exposure. Due to societal stigmas associated with maternal smoking during pregnancy, many women underreport prenatal smoking. Our study overcomes this bias

by evaluating smoking status with a previously validated biomarker, cotinine, a metabolite of nicotine (119). It has been successfully measured in umbilical cord blood by several studies and enables investigators to quantify active and passive exposure to tobacco smoke (119). Cotinine in umbilical cord blood only represents exposures during the end of pregnancy, thus the exposure estimates may underestimate the true exposure level as some women may have quit smoking early in pregnancy. As a result, our risk estimates may be biased toward the null.

In spite of this study's strengths, several limitations merit mention. We used a state-of-the-art, high-throughput methylation array to assess over 27,000 CpG loci. It has been reported that some CpG sites assessed in this array may fall within polymorphic sites, which may interfere with our method used to assess methylation (129, 130). The net impact of SNPs on methylation assessments that rely on bisulfite modified DNA remains unclear, thus the implications of SNPs corresponding to CpG sites included from the four genes of interest in this study is unknown. However, SNPs are rare and should not significantly alter study results. We were also unable to control for gene polymorphisms that may impact the risk of SGA or response to cigarette smoke. However, several recent studies suggest that polymorphisms in CYP1A1, GSTT1, and the GSTM1 genes have little or no impact on the association between IUGR and cigarette smoke exposure, although the literature is somewhat inconsistent (79, 80, 128).

This study relies on derivations from birthweight and gestational age to classify infants as SGA. Inaccuracies in gestational age measurement can affect how infants are classified. The clinical estimate can be calculated in several different ways, some of which are more accurate than others (131). Wingate et al. noted that clinical estimates and LMP differ in about 50% of the population studied and that the discordance varied by race and ethnicity, thus there may be some degree of non-differential

misclassification (131). However, the authors also noted that LMP is an imperfect measure as well due to recall errors and bleeding early in pregnancy (131). Although these indicators have faults, they are commonly employed, and in the absence of serial ultrasound measurements (the gold standard for IUGR classification), they are the best measure available (17). To ensure that factors associated with extreme prematurity did not influence our results, we re-ran the analyses excluding infants less than 32 weeks of gestation. However, the measures of association remained relatively unchanged when excluding them from the analysis (data not shown). However, it is important to note that the association between methylation of HIF1A and SGA was significant after excluding extremely preterm infants (OR=1.62, 95%CI=1.05, 2.60).

Overall, this is the first study to evaluate the methylation status of 4 detoxifying genes in umbilical cord blood and assess the association with SGA. The associations between SGA and DNA methylation and smoking and DNA methylation found here should be investigated further using larger samples. A better understanding of the impact of methylation change on gene expression and risk of SGA may lead to more targeted intervention methods. For example, investigation and development of methods to prevent adverse epigenetic changes may decrease the risk of SGA among infants born to smokers.

Acknowledgements

This study was supported by grant 1R36DP001898-01 from the CDC. Its content is solely the responsibility of the authors and does not necessarily represent the official views of the CDC. The study was also supported by funds from the Flight Attendant Medical Research Institute and a Research Initiative Grant from the Office of the Dean at the College of Public Health and the Office of the Associate Vice-President for Research of USF Health at the University of South Florida.

Table 2.1 Demographic and clinical variables by SGA status

	SGA[1] (n=10)		**Non-SGA (n=80)**		**p-value**
	n	**%**	**n**	**%**	
Parity					
Nulliparous	5	17.24	24	82.76	0.28
Multiparous	5	8.20	56	91.80	
Race					0.01
Black	6	28.57	15	71.43	
Non-black	4	5.80	65	94.20	
Pregnancy complications[2]					0.72
Yes	2	7.41	25	92.59	
No	8	12.70	55	87.30	
Adequate prenatal care[3]					0.21
Yes	6	8.45	65	91.55	
No	4	21.05	15	78.95	
Prenatal vitamin use					1.00
Yes	8	10.96	65	89.04	
No	2	11.76	15	88.24	
Infant sex					<.01
Male	1	1.96	50	98.04	
Female	9	23.08	30	76.92	
	Mean	**SD[4] (±)**	**Mean**	**SD (±)**	
Cotinine level (ng/ml)	11.80	25.24	3.53	14.57	0.33
Maternal age	24.90	4.33	28.99	6.58	0.06
Gestational age	38.10	1.52	38.18	2.30	0.92

[1]SGA=small for gestational age

[2] This includes the presence of diabetes, gestational diabetes, preeclampsia, placental previa, or hypertension

[3] As reported in the medical record

[4]SD=standard deviation

Table 2.2 Risk of SGA associated with a 0.01 unit increase in methylation level

	OR[1]	95%CI[2]	OR	95%CI
CYP1A1[3]	0.87	0.49, 1.55	0.97	0.41, 2.29
HIF1A[3]	1.23	0.88, 1.72	1.59	0.99, 2.56
GSTM1[4]	0.97	0.90, 1.05	1.01	0.93, 1.09
GSTT1[4]	1.08	1.03, 1.13	1.10	1.03, 1.18

[1]OR=odds ratio

[2]CI=confidence interval

[3] The adjusted odds ratio controls for gender, smoking, prenatal care, race, maternal age, and parity

[4] The adjusted odds ratio controls for smoking, prenatal care, race, maternal age, and parity

Table 2.3 Impact of a 5ng/ml and a 14ng/ml increase in cotinine level on the risk of hypermethylation

Cotinine level:	5ng/ml				14 ng/ml	
	Crude		Adjusted		Adjusted	
	OR[1]	95%CI[2]	OR[3]	95%CI	OR	95%CI
CYP1A1	0.99	0.85, 1.16	0.99	0.85, 1.16	0.98	0.63, 1.53
HIF1A	1.05	0.92, 1.20	1.08	0.94, 1.24	1.25	0.85, 1.84
GSTM1	0.93	0.76, 1.13	0.94	0.77, 1.16	0.85	0.48, 1.50
GSTT1	1.16	1.00, 1.35	1.18	1.02, 1.37	1.59	1.04, 2.42

[1]OR=odds ratio

[2]CI=confidence interval

[3] The adjusted odds ratio controls for maternal age and race

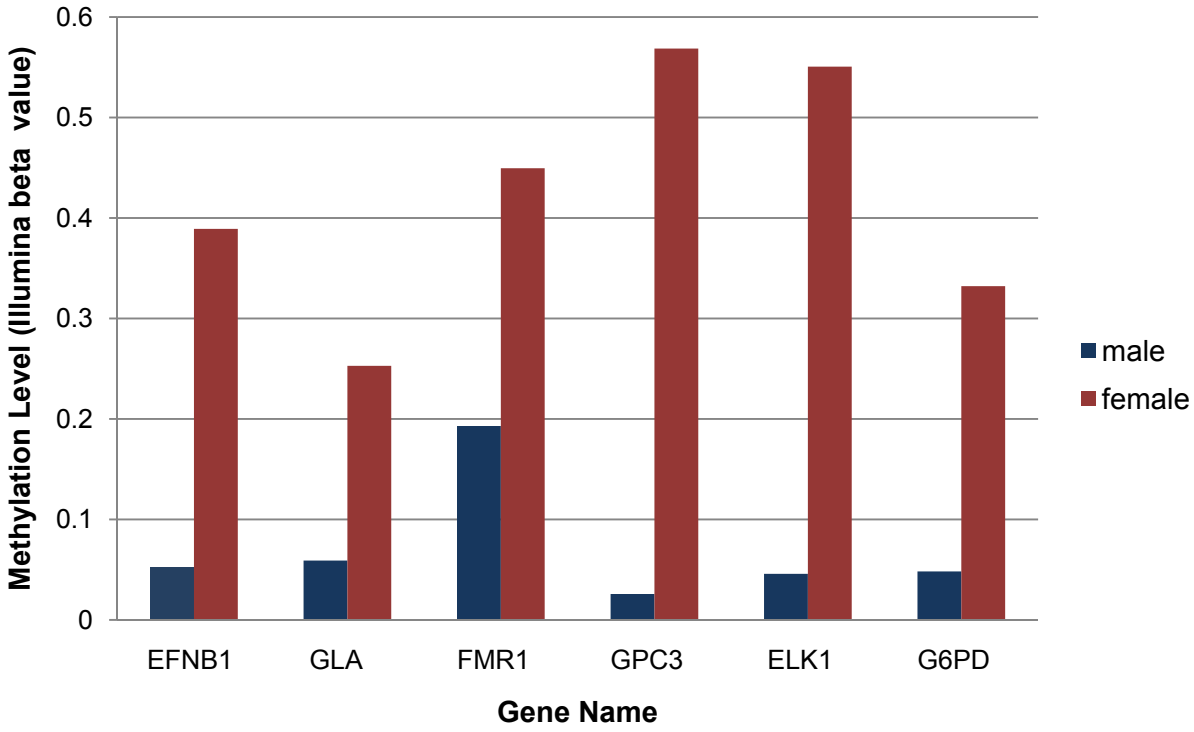

Figure 2.1 Methylation of selected housekeeping genes by gender

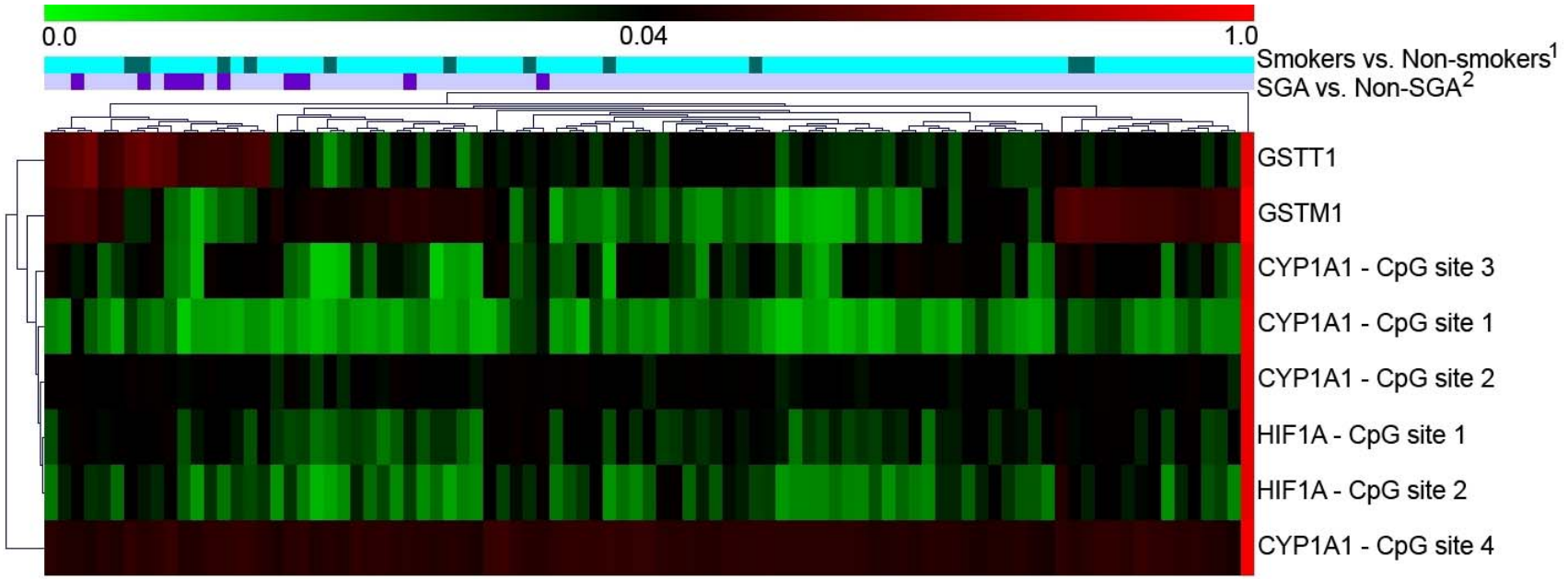

Figure 2.2 Heat map depicting the methylation level of all loci used to calculate the gene-specific methylation level for each sample and the control DNA

[1]The smokers are represented by the darker shade of teal

[2]The SGA (small for gestational age) infants are represented by the darker shade of purple

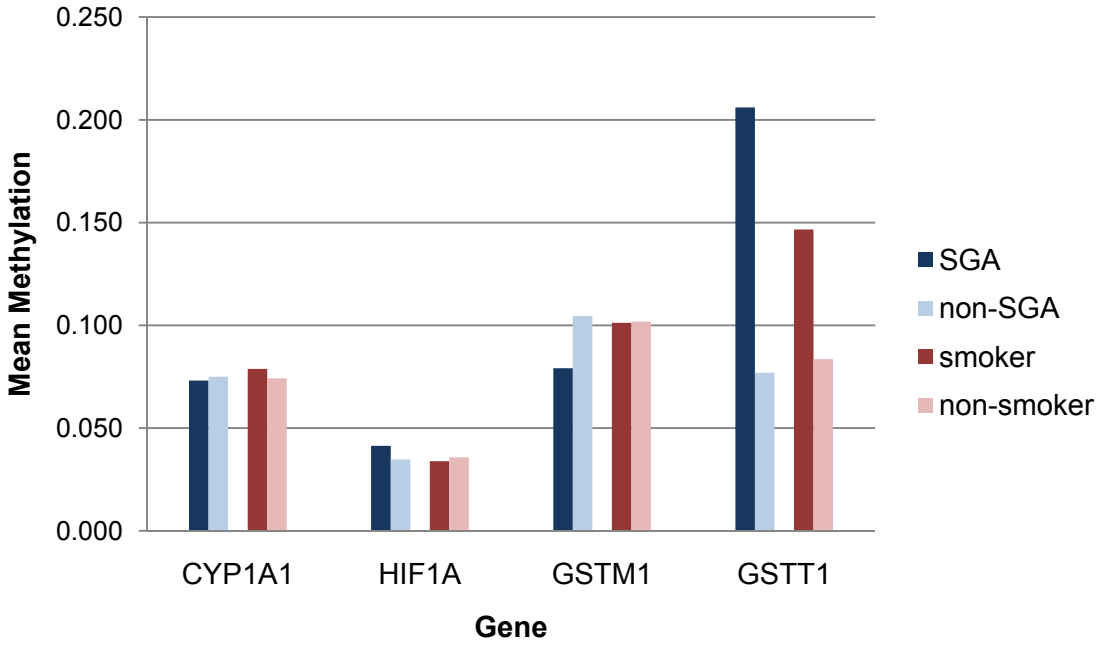

Figure 2.3 Mean methylation of selected genes by SGA and smoking status

CHAPTER THREE

Manuscript 2: Is there an association between birthweight and DNA methylation of IGF1 and IGF2?

Abstract

Objective: To evaluate the association between DNA methylation of IGF1 and IGF2 and birthweight.

Study Design: Medical record data and biological samples from 92 singletons were obtained from an ongoing, cross-sectional study. Methylation of DNA extracted from umbilical cord blood was measured with the Illumina Infinium Methylation27 assay. Univariate and multivariable linear regression were used to assess the impact of methylation on percent change in birthweight.

Results: The 90 infants included in the study had a mean birthweight of 3242 grams and a mean gestational age of 38 weeks. After adjusting for gender, maternal age, parity, and pregnancy complications, a one standard deviation increase in methylation of IGF1 decreased birthweight by 3.63% (95%CI= -6.49, -0.78). Birthweight was not associated with increased methylation of IGF2, even after adjusting (β= 0.07, 95%CI= -2.91, 3.05).

Conclusion: Methylation of IGF1, but not IGF2 influences birthweight and may be an important target for interventions aimed at preventing low birthweight.

Introduction

Birthweight is an important indicator of fetal growth and is associated with infant mortality, though the association is unlikely to be causal (22). Low birthweight infants (less than 2500 grams) have a higher infant mortality rate than their normal weight counterparts with the infant mortality rate of low birthweight infants being more than 110 times that of infants with normal birthweights (greater than or equal to 2500 grams) (132). Further, low birthweight infants have a higher risk of several adult-onset diseases. For example, birthweight is associated with elevated blood pressure, diabetes, and heart disease later in life (3, 133-135). As a result, it is important to understand the factors that mediate fetal growth and birthweight and currently the mechanism by which poor fetal growth manifests in response to environmental factors is unclear, but the insulin like growth factor (IGF) system appears to play an important role.

The IGF system regulates fetal and placental growth, promoting cell growth and differentiation and inhibiting apoptosis (8, 9). Although the IGF system includes insulin, several binding proteins, and multiple receptors, studies suggest that IGF1 and IGF2 are critical drivers of fetal growth. IGF2 is an imprinted gene that is expressed only from the paternal allele in most fetal tissues (88). Knockout studies of IGF1 or IGF2 suggest that it decreases fetal weight in mice and partial deletion of IGF1 in humans has similar effects (8, 11). However, prenatal IGF2 expression is greater than IGF1 and expression of IGF2 changes postnatally as it becomes biallelic (8, 88). IGF1 and birthweight are positively associated and growth restricted infants have low umbilical cord blood levels of IGF1 compared to their counterparts with normal growth (11, 93, 94). Given this, it is conceivable that altered expression of IGF1 or IGF2 may be associated with altered fetal growth.

Epigenetic modifications, namely DNA methylation, are a biologically plausible mechanism by which environmental and nutritional factors mediate gene expression to

impact phenotype, including birthweight. Findings of recent studies such as that of Steegers-Theunissen et al support this theory as they reported that infants exposed perinatally to folic acid had higher methylation of the IGF2 DMR (differentially methylated region) than their unexposed counterparts (10). In addition, a methylation increase of 1.7% in the IGF2 DMR was associated with decreased birthweight (10). However, it is possible that the time of sampling (about 17 months after delivery) impacted methylation of IGF2 as IGF2 expression changes after birth (8, 10, 88). Similarly, prenatal exposure to famine has also been associated with decreased methylation of IGF2 (96). Although epigenetic control of IGF1 has received less attention, animal studies suggest that IGF1 methylation is altered in intrauterine growth restricted (IUGR) rats and that hypermethylation decreases IGF1 expression (93).

The association of birthweight with lifelong health consequences such as diabetes and heart disease underscores the importance of understanding the mechanisms that are the foundation of fetal programming (136, 137). Therefore, we sought to examine the relationship between birthweight and DNA methylation of IGF1 and IGF2.

Methods

Study sample and data collection

All study participants in this cross-sectional study were enrolled at Tampa General Hospital in Tampa, Florida as part of a larger ongoing study examining lymphocyte subpopulations and prematurity. All pregnant females delivering at Tampa General Hospital were eligible to participate in the Lymphocyte Study. However, infants born to women whose prenatal tests indicated that they were HIV positive or Hepatitis B positive were excluded. Maternal race and ethnicity are not factors for inclusion. For the

present study, infants with birth defects were excluded and only singleton infants were eligible for inclusion.

De-identified demographic and clinical variables initially collected via medical record abstraction using standardized forms as part of the parent study were also obtained. The data elements collected include: gestational age, infant birthweight, infant sex, presence of infection, delivery complications, presence of birth defects, plurality, parity, gravidity, prenatal care usage, maternal age, and race.

Umbilical cord blood collection

In addition to medical record data, the Lymphocyte study also collected umbilical cord blood samples. After delivery of the placenta, umbilical cord blood was collected in tubes containing EDTA. During cord blood collection, no contact with the mother or infant occurred. Samples were processed at the University of South Florida within 24 hours of collection. Plasma was removed and stored at -80°C for subsequent cotinine analysis. The mononuclear layer was isolated using a ficoll gradient. After separation, samples were suspended in fetal bovine serum and 10% DMSO and stored in liquid nitrogen. Long term storage of cryopreserved cells does not impact cell viability or recovery as previous studies have indicated that more than 80% of nucleated cells can be recovered (110).

DNA isolation and methylation assessment

DNA isolation and methylation assessment was performed at Wayne State University Applied Genomics Technology Center. Laboratory personnel were blinded to birth outcome and other maternal and infant health indicators. DNA was isolated from the mononuclear fraction of umbilical cord blood using the Qiagen EZ1 DNA tissue kit according to Lum et al with the exception that PBS (phosphate buffered saline) was

substituted for TE (tris ethylenediaminetetraacetic acid) buffer (111). The mononuclear fraction is largely comprised of monocytes and lymphocytes, but also contains hematopoietic stem cells (112). Methylation changes in DNA from nucleated cells derived from umbilical cord blood should reflect changes that occurred in relation to the fetal environment. The Trinean Dropsense96 was used to quantify DNA after extraction.

Bisulfite modified DNA was prepared using the EZ-96 DNA Methylation Kit™ according to the manufacturer's instructions (Zymo Research Corp., USA). Quantitative, loci-specific methylation was assessed using the Infinium HumanMethylation27 BeadArray™ (Illumina Inc, San Diego, CA) per the manufacturer's instructions. The array interrogates 27,578 loci located in more than 14,000 genes. For each CpG site, two different probes (one against the methylated site and one against the unmethylated site) are hybridized with the bisulfite modified DNA. Next, a single-base extension adds one of two possible fluorescent probes (one for methylated (C) and one for unmethylated (T) alleles). Methylation status is then represented by a beta value which is calculated from the ratio of fluorescent signals from methylated to the sum of methylated and unmethylated probes and ranges from 0 (unmethylated) to 1 (methylated).

Background normalization was done according to the guidelines recommended by Illumina using the GenomeStudio Methylation module. In short, this method subtracts the average signal of the negative control bead-types from the probe signals. Normalized beta values were then output for use in subsequent analyses. Heat maps were generated with the MultiExperiment Viewer (113, 114). In this study, we are focused on the methylation status of IGF1 and IGF2; therefore we are only analyzed 6 CpG loci. The nucleotide positions of 5 of the 6 the CpG sites included in this analysis were located in the CpG islands at chr11: 2,110,452-2,111,041 and chr11:2,115,427-2,119,259 and the other was in chr12:101,398,416 in NCBI build 36.1.

A subset of samples was run in duplicate in order to assess inter-chip variability. In addition, CpGenome Universal Methylated DNA was used as a positive control (Millipore, Temecula, CA) and was bisulfite treated and run with the methylation assay. The positive control was used to ensure bisulfite conversion and accuracy of methylation measurement. The positive control DNA was almost completely methylated as expected. Inter-chip variability was assessed and was found to be highly reproducible. Pearson correlation coefficients were greater than 0.99 for each set of replicates ($p<0.0001$). Internal validity was assessed by examining gender specific methylation of 6 x-linked housekeeping genes (EFNB1, ELK1, FMR1, G6PD, GPC3, GLA) (115, 116). Overall, methylation of the 6 aforementioned housekeeping genes was as expected in that females exhibited hemimethylation and males had very little methylation at the loci in these genes ($p<0.0001$ for each gene).

Statistical analyses

Demographic and clinical variables were classified as follows: parity (nulliparous or multiparous), race (black or non-black), prenatal care (adequate or not as recorded in the medical record), pregnancy complications (yes or no), and infant sex (male or female). Prenatal vitamin use was dichotomized as yes or no as recorded in the medical record. Gestational age assessment was based on clinical estimate and the date of last menstrual period. Tobacco smoke exposure was assessed by measuring cotinine, a metabolite of nicotine. Gestational age, cotinine level and maternal age were kept as continuous variables.

Methylation level was measured at one CpG site in the IGF1 gene and multiple CpG sites in the IGF2 gene. The association between methylation of each gene as well as each CpG site in the IGF2 gene and infant birthweight was assessed independently. The association between methylation of the complete IGF2 gene and birthweight was

assessed by averaging the methylation levels (i.e. the illumina beta values) from each CpG site across the gene. In order to ease interpretation, Illumina beta values were converted to z-scores.

Univariate linear regression was used to assess the unadjusted association between birthweight and DNA methylation of the IGF1 gene, the IGF2 gene, and each CpG site measured in the IGF2 gene. Multivariable linear regression models were then constructed to control for primary predictors and potential confounders. The response variable, birthweight, did not initially meet all the assumptions of linear regression, therefore birthweight was log transformed to achieve normality. As a result, the reported effect estimates indicate the percent change in mean birthweight per standard deviation change in methylation. Confounders were identified by examining the significance of the covariate upon addition to the unadjusted model. If the p-value was less than 0.05, the variable was adjusted for. The final models were adjusted for gender, maternal age, parity, gestational age, and pregnancy complications. All hypothesis tests were two-tailed with a type 1 error rate fixed at 5 percent. SAS version 9.1 (SAS Institute, Cary, NC) was used to perform all analyses. This study was approved by the University of South Florida Institutional Review Board.

Results

In total, 2 infants were excluded from the analysis because of missing data or presence of birth defects. The mean birthweight of 90 infants included in the analysis was 3242.27 grams (Standard Deviation (SD)=654.18 grams) (Table 3.1). The mean gestational age was 38.17 weeks (SD=2.22). A majority of the women (>80%) used prenatal vitamins and about 32% were nulliparous.

Methylation was measured at 5 CpG sites in the IGF2 gene. The average methylation of IGF1, IGF2, and the 5 CpG sites measured in the IGF2 gene are

presented in Table 3.2. The methylation levels of the CpG sites range from a minimum of 0.02 to a maximum of 0.52. The overall methylation of the IGF2 gene was 0.24 (SD=0.02), which was slightly lower than that of IGF1 (mean=0.26, SD=0.05). The methylation level was variable in the IGF2 gene, with CpG site 3 having the lowest methylation (mean=0.02) and CpG site 2 having the highest level of methylation (mean=0.52). The heat map in Figure 3.1 depicts the differential methylation of all the CpG sites for which methylation was measured (includes sites in both IGF1 and IGF2).

In univariate linear regression models, all CpG sites except CpG site 5 suggested that birthweight decreased with increasing methylation (Table 3.3). Methylation of the entire IGF2 gene reflected this trend, but the association was not significant (β= -0.62, 95%CI= -5.75, 4.51). Similar results were found for IGF1 (β= -4.05, 95%CI=-9.11, 1.01). CpG site 3, the site with the lowest mean methylation, was the only site that was significantly associated with birthweight. Every one standard deviation increase in methylation was associated with a 5.10% decrease in birthweight. However, after adjusting for gender, maternal age, parity, gestational age, and pregnancy complications, the association between birthweight and methylation of CpG site 3 was no longer significant (β=-1.82, 95%CI= -4.81, 1.16). The association between percent change in birthweight and one standard deviation increase in methylation of the IGF2 gene remained non-significant (β= 0.07, 95%CI= -2.91, 3.05) as did the associations between methylation of each of the other CpG sites in the IGF2 gene and birthweight. After adjusting for confounders, methylation of IGF1 was significantly associated with birthweight and one standard deviation increase in methylation of IGF1 decreased birthweight by 3.63% (95%CI= -6.49, -0.78).

Initial assessments indicated that race was not a confounder and therefore it was not included in the multivariable models. However, we sought to explore this further and re-ran the analysis including race as well as the other confounders (gender, maternal

age, parity, gestational age, and pregnancy complications). This did not appreciably alter the results or the conclusions; therefore race was not included in the final multivariable model (data not shown).

Discussion

We found that birthweight is associated with methylation of the IGF1 gene, but not the IGF2 gene. A one standard deviation increase (SD=0.05) in methylation of the IGF1 gene decreased birthweight by 3.63%. Partial deletion or knockout of IGF1 has been shown to decrease birthweight in animal models, thus if IGF1 was epigenetically silenced, similar findings would be expected (8, 11). Although we did not measure IGF1 expression, the methylation site that was assessed was near the transcription start site, so it is likely that methylation would impact expression, but additional studies are needed to confirm this.

Previous studies have reported an association between birthweight and folic acid use and folic acid has been shown to increase gene-specific methylation, thus it is plausible that a methylation mediated mechanism controls fetal growth (10, 15, 61, 62, 138). Interestingly, prenatal vitamin use, a surrogate indicator of folic acid intake, did not significantly impact the association between birthweight and methylation of the IGF1 or IGF2 gene in our regression models. Our small sample size precluded a detailed investigation of this, but results from analyses examining methylation of IGF1 and IGF2 stratified by prenatal vitamin use were similar to that of the unstratified analysis, but in most cases did not reach statistical significance (data not shown). Further, 3 women in this study used high dose folic acid supplements during pregnancy and the mean methylation levels of these infants did not differ from those of regular prenatal vitamin users or non-users (data not shown). Nonetheless, our assessment of folic acid intake was limited to information abstracted from the medical record. We did not have any

information regarding the trimester that prenatal vitamin use began, the actual dose received, or dietary folic acid intake, thus these results must be interpreted with caution.

Few previous studies have examined methylation of IGF2 and birthweight. Our results conflict with those of Steegers-Theunissen et al as they found that increased methylation of IGF2 decreased birthweight after controlling for periconceptional folic acid use and gestational age whereas we did not find an association (10). In contrast, a study by Tabano et al examined methylation in DNA from umbilical cord blood of 60 normal and 66 IUGR infants (139). Although the study examined a different indicator of fetal growth (IUGR as opposed to birthweight) the results were similar to those found here in that infants had similar methylation levels in the IGF2/H19 imprinted region regardless of whether the infant was growth restricted (139). Other factors may interact with IGF2 to modulate fetal growth. A study by Ong et al reported a relationship between umbilical cord blood levels of IGF2 and the IGF2 receptor and that when considered together, these factors were significantly associated with birthweight (140). Alternatively, one previous study suggested that methylation of the IGF2 gene is highly conserved, thus it is possible that more extreme changes in methylation of IGF2 may only be associated with other outcomes not captured in this study such as Beckwith-Wiedermann syndrome or miscarriage (139).

Although this unique study provides much needed information on how methylation of two important genes in the IGF system influence birthweight, several limitations merit mention. We used a state-of-the-art, high-throughput methylation array to assess over 27,000 CpG loci. It has been reported that some CpG sites fall within polymorphic sites and may interfere with our method used to assess methylation (129, 130). The net impact of SNPs on methylation assessments that rely on bisulfite modified DNA remains unclear, but in this study, there were no known SNPs in 4 of the 6 CpG sites assessed. The implications of SNPs corresponding to the remaining 2 CpG sites

included from the two genes of interest in this study is unknown. However, SNPs are rare and should not significantly alter study results. The small sample size may have affected study results and prevented some associations from reaching statistical significance, thus additional studies are needed to examine the role of epigenetics in fetal growth.

An increasing number of studies are reporting an association between birthweight and a number of adverse health outcomes such as diabetes and heart disease (136, 137). A better understanding of the mechanisms that curtail normal fetal growth, may lead to enhanced strategies that are able to prevent suboptimal fetal growth and low birthweight. The findings of the current study highlight the need for additional investigations into the role of epigenetic modifications in the IGF system and their interactions with folic acid and other methyl donors.

Acknowledgements

This study was supported by grant 1R36DP001898-01 from the CDC. Its content is solely the responsibility of the authors and does not necessarily represent the official views of the CDC. The study was also supported by funds from the Flight Attendant Medical Research Institute and a Research Initiative Grant from the Office of the Dean at the College of Public Health and the Office of the Associate Vice-President for Research of USF Health at the University of South Florida.

Table 3.1 Summary of maternal and infant covariates

	n	**%**
Race		
Black	69	76.67
Non-black	21	23.33
Adequate prenatal care[1]		
Yes	71	78.89
No	19	21.11
Pregnancy Complications[2]		
Yes	27	30.00
No	63	70.00
Nulliparous		
Yes	29	32.22
No	61	67.78
Prenatal vitamin use[1]		
Yes	73	81.11
No	17	18.89
Gender		
Female	39	43.33
Male	51	56.67
	Mean	**SD (±)[3]**
Cotinine level (ng/ml)	4.45	16.12
Maternal age	28.53	6.48
Gestational age	38.17	2.22

[1] As reported in the medical record

[2] This includes the presence of diabetes, gestational diabetes, preeclampsia, placental previa, or hypertension

[3] SD=standard deviation

Table 3.2 Mean methylation level of the IGF1 gene, the IGF2 gene, and the 5 CpG sites in the IGF2 gene used to assess the overall methylation level

	Mean methylation	Standard Deviation (+/-)
IGF1	0.26	0.05
IGF2	0.24	0.02
CpG site 1	0.23	0.03
CpG site 2	0.52	0.04
CpG site 3	0.02	0.01
CpG site 4	0.11	0.02
CpG site 5	0.33	0.04

Table 3.3 Association between birthweight and DNA methylation in the IGF1 and IGF2 genes

	Crude		Adjusted	
	B[1]	95%CI[3]	B[1,2]	95%CI
IGF1	-4.05	-9.11, 1.01	-3.63	-6.49, -0.78
IGF2	-0.62	-5.75, 4.51	0.07	-2.91, 3.05
CpG site 1	-0.31	-5.44, 4.83	-0.11	-3.09, 2.86
CpG site 2	-0.24	-5.37, 4.90	0.15	-2.84, 3.14
CpG site 3	-5.10	-10.11, -0.08	-1.82	-4.81, 1.16
CpG site 4	-0.54	-5.67, 4.59	-1.72	-4.67, 1.24
CpG site 5	0.39	-4.74, 5.53	1.25	-1.72, 4.22

[1]Percent change in birthweight per standard deviation change in methylation level

[2]Adjusted for gender, maternal age, parity, gestational age, and pregnancy complications

[3] CI=confidence interval

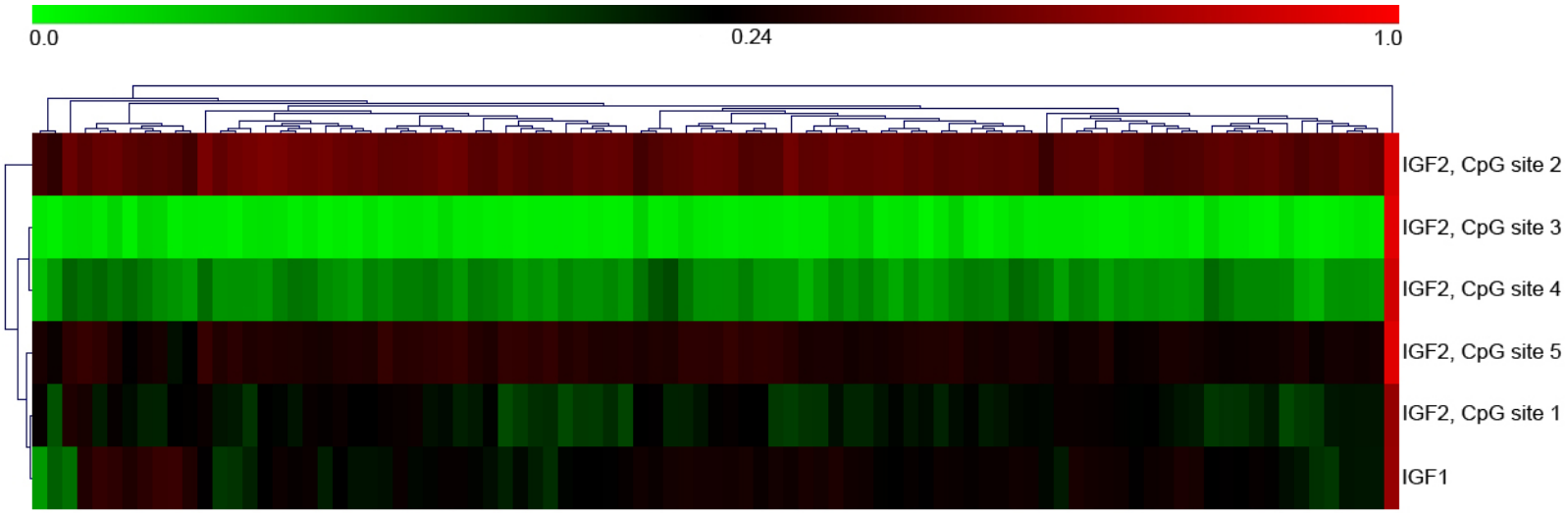

Figure 3.1 Heat map of methylated CpG sites in the IGF1 and IGF2 genes

CHAPTER FOUR

Conclusions and Recommendations

Although previous studies have found that prenatal exposures such as cigarette smoke and folic acid influence birthweight and SGA risk, the mechanism underlying these associations remains unclear (41, 61, 62, 102). One possible mechanism, DNA methylation, has not yet been fully explored and there is a lack of information about the influence of prenatal exposures on methylation patterns in DNA isolated from umbilical cord blood and subsequent pregnancy outcomes. This study used a candiate gene approach to address this gap by first examining whether methylation of CYP1A1, HIF1A, GSTM1, or GSTT1 modulates SGA risk and then examining the relationship between birthweight and methylation of IGF1 and IGF2. The study findings suggest that methylation may control some of the genes critical for normal fetal growth. Moreover, we found the risk of SGA increased with increasing methylation of GSTT1, a gene important for metabolism of cigarette smoke and other xenobiotics (77). Similarly, hypermethylation of a critical fetal growth gene, IGF1, was associated with birthweight decrements (8, 11).

We also examined whether abnormal methylation could be the mechanism underlying previously described associations between prenatal smoking and suboptimal fetal growth (41, 102). Additional analyses demonstrated that smoking impacts methylation of GSTT1, but not CYP1A1, GSTM1, or HIF1A. This finding suggests that abnormal methylation may be caused by in utero exposures and supports our hypothesis that methylation is the mechanism underlying the association between

prenatal smoking and SGA. Additional studies are needed to determine if hypermethylation alters the expression of the GSTT gene.

Although this unique study provides valuable information that helps fill a critical gap in our knowledge of the mechanisms that influences birthweight and SGA risk, it seemingly raises more questions than it answers. First, there is a need to further examine the role of folate in preventing abnormal methylation. Folate appears to have a beneficial role in pregnancy, decreasing the risk of neural tube defects and preventing low birthweight, and its role in one carbon metabolism, methylation reactions, and amino acid synthesis suggests that the mechanism underlying these outcomes may be tied to abnormal methylation (50, 52, 53). This study found that hypermethylation of IGF1 decreases birthweight, but prenatal vitamin use did not appear to be important in the association. However, this study used prenatal vitamin use as a surrogate indicator of folic acid intake and did not consider the duration, dosage, or additional dietary sources of folic acid which may have masked the true association. Additional studies with more accurate measures of folic acid intake are needed to determine whether folic acid can prevent abnormal methylation and associated adverse pregnancy outcomes.

It is also important to acknowledge the possibility that our study results may not translate to other populations. This study may over-represent high risk pregnancies since the biological and clinical data was initially collected for a different study. Some of the discrepancies may be attributed to the design of the parent study as the principle investigator of that study was based in the neonatal intensive care unit and had greater access to high risk pregnancies. When several key indicators for our study were compared to that of the county in which the hospital is located (Hillsborough) and that of the state, some differences were noted. Our sample had a higher proportion of low birthweight infants (10.00%) than that of the Hillsborough County (7.61%) or the state (7.04%) (141). In addition, when compared to the county, the present study had a higher

proportion of black infants (24.14% versus 21.00%) and a lower proportion of births to women under 35 years of age (80.00% versus 86.08%) (141).

Perhaps some of the main shortcomings of this study are its sample size and exploratory design. We were unable to examine how genes in the same system interact together to produce an adverse outcome. We hope to expand the sample size and obtain better indicators of folate intake. Although this study used a targeted gene approach to examine methylation and fetal growth indicators, we also plan to do a more epigenome wide approach and identify all CpG sites with a high degree of variation in methylation level. Although this complex, the data gathered for this study are conducive to such a design as this study obtained data on methylation of more than 27,000 CpG sites in about 14,000 different genes.

Nonetheless, the mechanisms underlying SGA and low birthweight remain unclear and additional efforts are needed to increase our knowledge of the role of DNA methylation in mediating fetal growth as it may lead to the development of methods to circumvent adverse pregnancy outcomes.

REFERENCES

1. Godfrey KM, Barker DJ. Fetal nutrition and adult disease. Am J Clin Nutr 2000;71(5 Suppl):1344S-52S.

2. Adair LS, Cole TJ. Rapid child growth raises blood pressure in adolescent boys who were thin at birth. Hypertension 2003;41(3):451-6.

3. Barker DJ, Osmond C, Golding J, Kuh D, Wadsworth ME. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. BMJ 1989;298(6673):564-7.

4. Eriksson JG, Forsen T, Tuomilehto J, Jaddoe VW, Osmond C, Barker DJ. Effects of size at birth and childhood growth on the insulin resistance syndrome in elderly individuals. Diabetologia 2002;45(3):342-8.

5. Li X, Zhao X. Epigenetic regulation of mammalian stem cells. Stem Cells Dev 2008;17(6):1043-52.

6. Kremenskoy M, Kremenska Y, Ohgane J, Hattori N, Tanaka S, Hashizume K, et al. Genome-wide analysis of DNA methylation status of CpG islands in embryoid bodies, teratomas, and fetuses. Biochem Biophys Res Commun 2003;311(4):884-90.

7. Santos KF, Mazzola TN, Carvalho HF. The prima donna of epigenetics: the regulation of gene expression by DNA methylation. Braz J Med Biol Res 2005;38(10):1531-41.

8. Fowden AL. The insulin-like growth factors and feto-placental growth. Placenta 2003;24(8-9):803-12.

9. Morison IM, Reeve AE. Insulin-like growth factor 2 and overgrowth: molecular biology and clinical implications. Mol Med Today 1998;4(3):110-5.

10. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, Lindemans J, Siebel C, Steegers EA, et al. Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. PLoS One 2009;4(11):e7845.

11. Woods KA, Camacho-Hubner C, Savage MO, Clark AJ. Intrauterine growth retardation and postnatal growth failure associated with deletion of the insulin-like growth factor I gene. N Engl J Med 1996;335(18):1363-7.

12. Watson RE, Curtin GM, Doolittle DJ, Goodman JI. Progressive alterations in global and GC-rich DNA methylation during tumorigenesis. Toxicol Sci 2003;75(2):289-99.

13. Moore LE, Pfeiffer RM, Poscablo C, Real FX, Kogevinas M, Silverman D, et al. Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. Lancet Oncol 2008;9(4):359-66.

14. Dolinoy DC, Huang D, Jirtle RL. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci U S A 2007;104(32):13056-61.

15. Lillycrop KA, Phillips ES, Jackson AA, Hanson MA, Burdge GC. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. J Nutr 2005;135(6):1382-6.

16. Alexander GR, Kogan MD, Himes JH. 1994-1996 U.S. singleton birth weight percentiles for gestational age by race, Hispanic origin, and gender. Matern Child Health J 1999;3(4):225-31.

17. Resnik R. Intrauterine growth restriction. Obstet Gynecol 2002;99(3):490-6.

18. Sifianou P. Small and growth-restricted babies: drawing the distinction. Acta Paediatr 2006;95(12):1620-4.

19. Peleg D, Kennedy CM, Hunter SK. Intrauterine growth restriction: identification and management. Am Fam Physician 1998;58(2):453-60, 66-7.

20. Vrachnis N, Botsis D, Iliodromiti Z. The fetus that is small for gestational age. Ann N Y Acad Sci 2006;1092:304-9.

21. Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Munson ML. Births: final data for 2003. Natl Vital Stat Rep 2005;54(2):1-116.

22. Wilcox AJ. On the importance--and the unimportance--of birthweight. Int J Epidemiol 2001;30(6):1233-41.

23. Lunde A, Melve KK, Gjessing HK, Skjaerven R, Irgens LM. Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data. Am J Epidemiol 2007;165(7):734-41.

24. Cnattingius S. The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. Nicotine Tob Res 2004;6 Suppl 2:S125-40.

25. U.S. Department of Health and Human Services. Health, United States, 2002, with chartbooks on trends in the health of Americans (DHHS Publication No.

1232). Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; 2002.

26. Pirkle JL, Flegal KM, Bernert JT, Brody DJ, Etzel RA, Maurer KR. Exposure of the US population to environmental tobacco smoke: the Third National Health and Nutrition Examination Survey, 1988 to 1991. JAMA 1996;275(16):1233-40.

27. Salihu HM, Wilson RE. Epidemiology of prenatal smoking and perinatal outcomes. Early Hum Dev 2007;83(11):713-20.

28. Shea AK, Steiner M. Cigarette smoking during pregnancy. Nicotine Tob Res 2008;10(2):267-78.

29. Luck W, Nau H, Hansen R, Steldinger R. Extent of nicotine and cotinine transfer to the human fetus, placenta and amniotic fluid of smoking mothers. Dev Pharmacol Ther 1985;8(6):384-95.

30. Pastrakuljic A, Derewlany LO, Koren G. Maternal cocaine use and cigarette smoking in pregnancy in relation to amino acid transport and fetal growth. Placenta 1999;20(7):499-512.

31. Carmines EL, Rajendran N. Evidence for carbon monoxide as the major factor contributing to lower fetal weights in rats exposed to cigarette smoke. Toxicol Sci 2008;102(2):383-91.

32. Regnault TR, Friedman JE, Wilkening RB, Anthony RV, Hay WW, Jr. Fetoplacental transport and utilization of amino acids in IUGR--a review. Placenta 2005;26 Suppl A:S52-62.

33. Myatt L. Placental adaptive responses and fetal programming. J Physiol 2006;572(Pt 1):25-30.

34. Jauniaux E, Burton GJ. Morphological and biological effects of maternal exposure to tobacco smoke on the feto-placental unit. Early Hum Dev 2007;83(11):699-706.

35. Bush PG, Mayhew TM, Abramovich DR, Aggett PJ, Burke MD, Page KR. Maternal cigarette smoking and oxygen diffusion across the placenta. Placenta 2000;21(8):824-33.

36. Demir R, Demir AY, Yinanc M. Structural changes in placental barrier of smoking mother. A quantitative and ultrastructural study. Pathol Res Pract 1994;190(7):656-67.

37. Jauniaux E, Johns J, Gulbis B, Spasic-Boskovic O, Burton GJ. Transfer of folic acid inside the first-trimester gestational sac and the effect of maternal smoking. Am J Obstet Gynecol 2007;197(1):58 e1-6.

38. Salihu HM, Aliyu MH, Pierre-Louis BJ, Alexander GR. Levels of excess infant deaths attributable to maternal smoking during pregnancy in the United States. Matern Child Health J 2003;7(4):219-27.

39. Ingvarsson RF, Bjarnason AO, Dagbjartsson A, Hardardottir H, Haraldsson A, Thorkelsson T. The effects of smoking in pregnancy on factors influencing fetal growth. Acta Paediatr 2007;96(3):383-6.

40. Higgins S. Smoking in pregnancy. Curr Opin Obstet Gynecol 2002;14(2):145-51.

41. Jaddoe VW, Verburg BO, de Ridder MA, Hofman A, Mackenbach JP, Moll HA, et al. Maternal smoking and fetal growth characteristics in different periods of pregnancy: the generation R study. Am J Epidemiol 2007;165(10):1207-15.

42. Ohmi H, Hirooka K, Mochizuki Y. Fetal growth and the timing of exposure to maternal smoking. Pediatr Int 2002;44(1):55-9.

43. Vielwerth SE, Jensen RB, Larsen T, Greisen G. The impact of maternal smoking on fetal and infant growth. Early Hum Dev 2007;83(8):491-5.

44. Bernstein IM, Mongeon JA, Badger GJ, Solomon L, Heil SH, Higgins ST. Maternal smoking and its association with birth weight. Obstet Gynecol 2005;106(5 Pt 1):986-91.

45. Bardy AH, Seppala T, Lillsunde P, Kataja JM, Koskela P, Pikkarainen J, et al. Objectively measured tobacco exposure during pregnancy: neonatal effects and relation to maternal smoking. Br J Obstet Gynaecol 1993;100(8):721-6.

46. Jauniaux E, Biernaux V, Gerlo E, Gulbis B. Chronic maternal smoking and cord blood amino acid and enzyme levels at term. Obstet Gynecol 2001;97(1):57-61.

47. Muscati SK, Koski KG, Gray-Donald K. Increased energy intake in pregnant smokers does not prevent human fetal growth retardation. J Nutr 1996;126(12):2984-9.

48. Ellard GA, Johnstone FD, Prescott RJ, Ji-Xian W, Jian-Hua M. Smoking during pregnancy: the dose dependence of birthweight deficits. Br J Obstet Gynaecol 1996;103(8):806-13.

49. Obican SG, Finnell RH, Mills JL, Shaw GM, Scialli AR. Folic acid in early pregnancy: a public health success story. Faseb J 2010.

50. Niculescu MD, Zeisel SH. Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. J Nutr 2002;132(8 Suppl):2333S-5S.

51. Van den Veyver IB. Genetic effects of methylation diets. Annu Rev Nutr 2002;22:255-82.

52. Waterland RA. Assessing the effects of high methionine intake on DNA methylation. J Nutr 2006;136(6 Suppl):1706S-10S.

53. Beaudin AE, Stover PJ. Folate-mediated one-carbon metabolism and neural tube defects: balancing genome synthesis and gene expression. Birth Defects Res C Embryo Today 2007;81(3):183-203.

54. Tamura T, Picciano MF. Folate and human reproduction. Am J Clin Nutr 2006;83(5):993-1016.

55. Folate status in women of childbearing age, by race/ethnicity--United States, 1999-2000. MMWR Morb Mortal Wkly Rep 2002;51(36):808-10.

56. Canfield MA, Annegers JF, Brender JD, Cooper SP, Greenberg F. Hispanic origin and neural tube defects in Houston/Harris County, Texas. II. Risk factors. Am J Epidemiol 1996;143(1):12-24.

57. Honein MA, Paulozzi LJ, Mathews TJ, Erickson JD, Wong LY. Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. Jama 2001;285(23):2981-6.

58. Oakley GP, Jr., Adams MJ, Dickinson CM. More folic acid for everyone, now. J Nutr 1996;126(3):751S-5S.

59. Tinker SC, Cogswell ME, Devine O, Berry RJ. Folic acid intake among U.S. women aged 15-44 years, National Health and Nutrition Examination Survey, 2003-2006. Am J Prev Med 2010;38(5):534-42.

60. Shah PS, Ohlsson A. Effects of prenatal multimicronutrient supplementation on pregnancy outcomes: a meta-analysis. CMAJ 2009;180(12):E99-108.

61. Neggers YH, Goldenberg RL, Tamura T, Cliver SP, Hoffman HJ. The relationship between maternal dietary intake and infant birthweight. Acta Obstet Gynecol Scand Suppl 1997;165:71-5.

62. Timmermans S, Jaddoe VW, Hofman A, Steegers-Theunissen RP, Steegers EA. Periconception folic acid supplementation, fetal growth and the risks of low birth weight and preterm birth: the Generation R Study. Br J Nutr 2009;102(5):777-85.

63. Nilsen RM, Vollset SE, Monsen AL, Ulvik A, Haugen M, Meltzer HM, et al. Infant birth size is not associated with maternal intake and status of folate during the second trimester in Norwegian pregnant women. J Nutr 2010;140(3):572-9.

64. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev 2002;16(1):6-21.

65. Nafee TM, Farrell WE, Carroll WD, Fryer AA, Ismail KM. Epigenetic control of fetal gene expression. Bjog 2008;115(2):158-68.

66. Burdge GC, Lillycrop KA. Nutrition, epigenetics, and developmental plasticity: implications for understanding human disease. Annu Rev Nutr 2010;30:315-39.

67. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. Nature 2000;403(6769):501-2.

68. Reik W, Walter J. Genomic imprinting: parental influence on the genome. Nat Rev Genet 2001;2(1):21-32.

69. Gluckman PD, Pinal CS. Regulation of fetal growth by the somatotrophic axis. J Nutr 2003;133(5 Suppl 2):1741S-6S.

70. Allegrucci C, Denning CN, Burridge P, Steele W, Sinclair KD, Young LE. Human embryonic stem cells as a model for nutritional programming: an evaluation. Reprod Toxicol 2005;20(3):353-67.

71. Maloney CA, Hay SM, Rees WD. Folate deficiency during pregnancy impacts on methyl metabolism without affecting global DNA methylation in the rat fetus. Br J Nutr 2007;97(6):1090-8.

72. Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. J Nutr 2002;132(8 Suppl):2393S-400S.

73. Davis CD, Uthus EO. DNA methylation, cancer susceptibility, and nutrient interactions. Exp Biol Med (Maywood) 2004;229(10):988-95.

74. Reik W, Constancia M, Fowden A, Anderson N, Dean W, Ferguson-Smith A, et al. Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. J Physiol 2003;547(Pt 1):35-44.

75. Piyathilake CJ, Johanning GL, Macaluso M, Whiteside M, Oelschlager DK, Heimburger DC, et al. Localized folate and vitamin B-12 deficiency in squamous cell lung cancer is associated with global DNA hypomethylation. Nutr Cancer 2000;37(1):99-107.

76. Hsiung DT, Marsit CJ, Houseman EA, Eddy K, Furniss CS, McClean MD, et al. Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev 2007;16(1):108-14.

77. Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. Annu Rev Pharmacol Toxicol 2005;45:51-88.

78. Chen X, Woodcroft KJ. Polymorphisms in metabolic genes CYP1A1 and GSTM1 and changes in maternal smoking during pregnancy. Nicotine Tob Res 2009;11(3):225-33.

79. Grazuleviciene R, Danileviciute A, Nadisauskiene R, Vencloviene J. Maternal smoking, GSTM1 and GSTT1 polymorphism and susceptibility to adverse pregnancy outcomes. Int J Environ Res Public Health 2009;6(3):1282-97.

80. Delpisheh A, Brabin L, Topping J, Reyad M, Tang AW, Brabin BJ. A case-control study of CYP1A1, GSTT1 and GSTM1 gene polymorphisms, pregnancy smoking and fetal growth restriction. Eur J Obstet Gynecol Reprod Biol 2009;143(1):38-42.

81. Yan YE, Wang H, Feng YH. Alterations of placental cytochrome P450 1A1 and P-glycoprotein in tobacco-induced intrauterine growth retardation in rats. Acta Pharmacol Sin 2005;26(11):1387-94.

82. Habano W, Gamo T, Sugai T, Otsuka K, Wakabayashi G, Ozawa S. CYP1B1, but not CYP1A1, is downregulated by promoter methylation in colorectal cancers. Int J Oncol 2009;34(4):1085-91.

83. Semenza GL. Hypoxia-inducible factor 1: control of oxygen homeostasis in health and disease. Pediatr Res 2001;49(5):614-7.

84. Walczak-Drzewiecka A, Ratajewski M, Pulaski L, Dastych J. DNA methylation-dependent suppression of HIF1A in an immature hematopoietic cell line HMC-1. Biochem Biophys Res Commun 2010;391(1):1028-32.

85. Terry MB, Ferris JS, Pilsner R, Flom JD, Tehranifar P, Santella RM, et al. Genomic DNA methylation among women in a multiethnic New York City birth cohort. Cancer Epidemiol Biomarkers Prev 2008;17(9):2306-10.

86. Sheridan C. Intrauterine growth restriction--diagnosis and management. Aust Fam Physician 2005;34(9):717-23.

87. Kramer MS. Intrauterine growth and gestational duration determinants. Pediatrics 1987;80(4):502-11.

88. Randhawa R, Cohen P. The role of the insulin-like growth factor system in prenatal growth. Mol Genet Metab 2005;86(1-2):84-90.

89. Lighten AD, Hardy K, Winston RM, Moore GE. IGF2 is parentally imprinted in human preimplantation embryos. Nat Genet 1997;15(2):122-3.

90. Constancia M, Dean W, Lopes S, Moore T, Kelsey G, Reik W. Deletion of a silencer element in Igf2 results in loss of imprinting independent of H19. Nat Genet 2000;26(2):203-6.

91. Weksberg R, Shen DR, Fei YL, Song QL, Squire J. Disruption of insulin-like growth factor 2 imprinting in Beckwith-Wiedemann syndrome. Nat Genet 1993;5(2):143-50.

92. DeChiara TM, Efstratiadis A, Robertson EJ. A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. Nature 1990;345(6270):78-80.

93. Fu Q, Yu X, Callaway CW, Lane RH, McKnight RA. Epigenetics: intrauterine growth retardation (IUGR) modifies the histone code along the rat hepatic IGF-1 gene. FASEB J 2009;23(8):2438-49.

94. Martos-Moreno GA, Barrios V, Saenz de Pipaon M, Pozo J, Dorronsoro I, Martinez-Biarge M, et al. Influence of prematurity and growth restriction on the

adipokine profile, IGF1, and ghrelin levels in cord blood: relationship with glucose metabolism. Eur J Endocrinol 2009;161(3):381-9.

95. Geary MP, Pringle PJ, Rodeck CH, Kingdom JC, Hindmarsh PC. Sexual dimorphism in the growth hormone and insulin-like growth factor axis at birth. J Clin Endocrinol Metab 2003;88(8):3708-14.

96. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A 2008;105(44):17046-9.

97. Whyatt RM, Bell DA, Jedrychowski W, Santella RM, Garte SJ, Cosma G, et al. Polycyclic aromatic hydrocarbon-DNA adducts in human placenta and modulation by CYP1A1 induction and genotype. Carcinogenesis 1998;19(8):1389-92.

98. Semenza GL. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. J Appl Physiol 2000;88(4):1474-80.

99. Semenza GL. Expression of hypoxia-inducible factor 1: mechanisms and consequences. Biochem Pharmacol 2000;59(1):47-53.

100. Garite TJ, Clark R, Thorp JA. Intrauterine growth restriction increases morbidity and mortality among premature neonates. Am J Obstet Gynecol 2004;191(2):481-7.

101. Olsen J. Cigarette smoking in pregnancy and fetal growth. Does the type of tobacco play a role? Int J Epidemiol 1992;21(2):279-84.

102. Cnattingius S, Haglund B. Decreasing smoking prevalence during pregnancy in Sweden: the effect on small-for-gestational-age births. Am J Public Health 1997;87(3):410-3.

103. Lambers DS, Clark KE. The maternal and fetal physiologic effects of nicotine. Semin Perinatol 1996;20(2):115-26.

104. Mosier HD, Jr., Capodanno CC, Li IO, Magruder CS, Jansons RA. Resistance of rat fetuses to nicotine-induced lipolysis. Teratology 1974;9(2):239-45.

105. Infante-Rivard C, Weinberg CR, Guiguet M. Xenobiotic-metabolizing genes and small-for-gestational-age births: interaction with maternal smoking. Epidemiology 2006;17(1):38-46.

106. Bostrom CE, Gerde P, Hanberg A, Jernstrom B, Johansson C, Kyrklund T, et al. Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. Environ Health Perspect 2002;110 Suppl 3:451-88.

107. Zhang L, Connor EE, Chegini N, Shiverick KT. Modulation by benzo[a]pyrene of epidermal growth factor receptors, cell proliferation, and secretion of human

chorionic gonadotropin in human placental cell lines. Biochem Pharmacol 1995;50(8):1171-80.

108. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med 2009;180(5):462-7.

109. Pilsner JR, Hu H, Ettinger A, Sanchez BN, Wright RO, Cantonwine D, et al. Influence of prenatal lead exposure on genomic methylation of cord blood DNA. Environ Health Perspect 2009;117(9):1466-71.

110. Navarrete C, Contreras M. Cord blood banking: a historical perspective. Br J Haematol 2009;147(2):236-45.

111. Lum A, Le Marchand L. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. Cancer Epidemiol Biomarkers Prev 1998;7(8):719-24.

112. Newcomb JD, Sanberg PR, Klasko SK, Willing AE. Umbilical cord blood research: current and future perspectives. Cell Transplant 2007;16(2):151-8.

113. Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, et al. TM4 microarray software suite. Methods Enzymol 2006;411:134-93.

114. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 2003;34(2):374-8.

115. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res 2006;16(3):383-93.

116. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 2005;434(7031):400-4.

117. Jarvis MJ, Russell MA, Benowitz NL, Feyerabend C. Elimination of cotinine from body fluids: implications for noninvasive measurement of tobacco smoke exposure. Am J Public Health 1988;78(6):696-8.

118. Jauniaux E, Gulbis B, Acharya G, Thiry P, Rodeck C. Maternal tobacco exposure and cotinine levels in fetal fluids in the first half of pregnancy. Obstet Gynecol 1999;93(1):25-9.

119. Pichini S, Basagana XB, Pacifici R, Garcia O, Puig C, Vall O, et al. Cord serum cotinine as a biomarker of fetal exposure to cigarette smoke at the end of pregnancy. Environ Health Perspect 2000;108(11):1079-83.

120. Dietrich D, Krispin M, Dietrich J, Fassbender A, Lewin J, Harbeck N, et al. CDO1 promoter methylation is a biomarker for outcome prediction of anthracycline treated, estrogen receptor-positive, lymph node-positive breast cancer patients. BMC Cancer 2010;10:247.

121. Zhu W, Qin W, Hewett JE, Sauter ER. Quantitative evaluation of DNA hypermethylation in malignant and benign breast tissue and fluids. Int J Cancer 2010;126(2):474-82.

122. Bearer C, Emerson RK, O'Riordan MA, Roitman E, Shackleton C. Maternal tobacco smoke exposure and persistent pulmonary hypertension of the newborn. Environ Health Perspect 1997;105(2):202-6.

123. Nafstad P, Kongerud J, Botten G, Urdal P, Silsand T, Pedersen BS, et al. Fetal exposure to tobacco smoke products: a comparison between self-reported maternal smoking and concentrations of cotinine and thiocyanate in cord serum. Acta Obstet Gynecol Scand 1996;75(10):902-7.

124. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. Carcinogenesis 2010;31(1):27-36.

125. Einstein F, Thompson RF, Bhagat TD, Fazzari MJ, Verma A, Barzilai N, et al. Cytosine methylation dysregulation in neonates following intrauterine growth restriction. PLoS One 2010;5(1):e8887.

126. Guo L, Choufani S, Ferreira J, Smith A, Chitayat D, Shuman C, et al. Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. Dev Biol 2008;320(1):79-91.

127. Luo ZC, Fraser WD, Julien P, Deal CL, Audibert F, Smith GN, et al. Tracing the origins of "fetal origins" of adult diseases: programming by oxidative stress? Med Hypotheses 2006;66(1):38-44.

128. Aagaard-Tillery K, Spong CY, Thom E, Sibai B, Wendel G, Jr., Wenstrom K, et al. Pharmacogenomics of maternal tobacco use: metabolic gene polymorphisms and risk of adverse pregnancy outcomes. Obstet Gynecol 2010;115(3):568-77.

129. Laird PW. Principles and challenges of genome-wide DNA methylation analysis. Nat Rev Genet 2010;11(3):191-203.

130. Byun HM, Siegmund KD, Pan F, Weisenberger DJ, Kanel G, Laird PW, et al. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Hum Mol Genet 2009;18(24):4808-17.

131. Wingate MS, Alexander GR, Buekens P, Vahratian A. Comparison of gestational age classifications: date of last menstrual period vs. clinical estimate. Ann Epidemiol 2007;17(6):425-30.

132. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2003 period linked birth/infant death data set. Natl Vital Stat Rep 2006;54(16):1-29.

133. Hardy R, Wadsworth ME, Langenberg C, Kuh D. Birthweight, childhood growth, and blood pressure at 43 years in a British birth cohort. Int J Epidemiol 2004;33(1):121-9.

134. Barker DJ, Winter PD, Osmond C, Margetts B, Simmonds SJ. Weight in infancy and death from ischaemic heart disease. Lancet 1989;2(8663):577-80.

135. Hales CN, Barker DJ, Clark PM, Cox LJ, Fall C, Osmond C, et al. Fetal and infant growth and impaired glucose tolerance at age 64. Bmj 1991;303(6809):1019-22.

136. Iliadou A, Cnattingius S, Lichtenstein P. Low birthweight and Type 2 diabetes: a study on 11 162 Swedish twins. Int J Epidemiol 2004;33(5):948-53; discussion 53-4.

137. Barker DJ, Clark PM. Fetal undernutrition and disease in later life. Rev Reprod 1997;2(2):105-12.

138. Pogribny IP, Miller BJ, James SJ. Alterations in hepatic p53 gene methylation patterns during tumor progression with folate/methyl deficiency in the rat. Cancer Lett 1997;115(1):31-8.

139. Tabano S, Colapietro P, Cetin I, Grati FR, Zanutto S, Mando C, et al. Epigenetic modulation of the IGF2/H19 imprinted domain in human embryonic and extra-embryonic compartments and its possible role in fetal growth restriction. Epigenetics 2010;5(4).

140. Ong K, Kratzsch J, Kiess W, Costello M, Scott C, Dunger D. Size at birth and cord blood levels of insulin, insulin-like growth factor I (IGF-I), IGF-II, IGF-binding protein-1 (IGFBP-1), IGFBP-3, and the soluble IGF-II/mannose-6-phosphate receptor in term human infants. The ALSPAC Study Team. Avon Longitudinal Study of Pregnancy and Childhood. J Clin Endocrinol Metab 2000;85(11):4266-9.

141. Florida Department of Health Office of Health Statistics and Assessment. Florida CHARTS. 2010.

APPENDIX A: STUDY METHODS

Participant recruitment and data collection

The Lymphocyte Study

All participants were recruited from a larger, on-going study at Tampa General Hospital (TGH) which seeks to examine prematurity in relation to lymphocyte subpopulations in umbilical cord blood samples from infants born at TGH. Participants in the lymphocyte project are identified at admission to TGH for delivery. The principal investigator of this study, a neonatology fellow, was based in the neonatal intensive care unit, thus it is possible that the infants included in this study disproportionately reflect high risk pregnancies and deliveries. Maternal age, race, and ethnicity are not factors for inclusion. So as to prevent unnecessary risk, women whose prenatal tests indicated that they were HIV or hepatitis B positive were not eligible for this study.

Umbilical cord blood collection occurs after delivery of the infant and afterbirth and involves no direct contact with the mother or infant. After delivery of the placenta, a transport nurse or physician wiped the umbilical cord with 70% alcohol and betadine as this is the needle insertion (collection) site. The needle is inserted into the umbilical cord and held in place while the blood from the cord flows into the attached collection tube containing EDTA. Approximately 3-5 mL of blood is collected, but the amount collected depended on the amount available. After collection, the tube was gently inverted to mix the cord blood with the anticoagulant. The collection tube was then labeled with the date of collection and the medical record number and placed into a plastic bag. This was then sealed and sent to the USF lab for storage and preliminary processing. Samples were processed within 24 hours of collection. At the time of processing, the sample was assigned a study ID number and the medical record number was transcribed into the study log so that the medical record information could be abstracted by the principal investigator at a later date. Thereafter, the samples were labeled only with the study ID and no personal identifiers.

APPENDIX A (CONTINUED)

Samples were transferred to conical tubes and spun. Then plasma was removed and stored in 1mL tubes at -80°C for subsequent cotinine analysis. A ficoll gradient separation was then used to isolate the mononuclear layer. After separation, samples were suspended in freeze media (fetal bovine serum and 10% DMSO) and stored in liquid nitrogen.

The principal investigator of the Lymphocyte study abstracted maternal and infant covariates from electronic medical records (birth record) onto a standardized form (see Appendix B). This form includes maternal age, the infant's gestational age, birth weight, presence of congenital anomalies, type of delivery, placental infection, and plurality (singleton vs. multiple). The form also collected information on prenatal care. Although this was initially planned to collected as a yes/no response, the medical record listed prenatal care as adequate or inadequate if there was prenatal care provided; therefore this information was also collected on the form. Blood samples are linked to the mother-infant dyad information with a unique ID.

<u>The current project</u>

In this study, no contact or participation of the mother/infant dyad occurred as all covariates and blood samples are collected by the Lymphocyte Study. The Lymphocyte Study provided paper copies of the forms used to abstract medical records as well as frozen plasma and nucleated cell samples. The paper forms were then entered into an Access database. After the data was entered into the database, paper forms were randomly selected compared to the electronic database in order to verify that there were no transcription errors. The data were also inspected for out of range values. None were detected. In total, 92 records were entered. One infant's medical record could not be located (per the principal investigator's note) and one infant had a birth defect, thus

these 2 infants were excluded from the analysis as they did not meet the study criteria. Although the Lymphocyte Study included some multiple births, the medical record information for these infants was not provided. Therefore, the total number of infants in the present study is less than that of the parent study.

Cotinine assays were done at the University of South Florida using a solid phase competitive ELISA (Calbiotech, California). All cotinine assays were performed by the principal investigator in duplicate per the manufacturer's instructions. Absorbance was read on a Biotek Synergy 2 plate reader using Gen5 software.

DNA isolation and methylation assessment was done at Wayne State University Applied Genomics Technology Center using 300ul to 500ul of cell suspension. The amount of DNA obtained from the cells varied greatly as the number of nucleated cells per microliter of freeze media varied. The variation in the number of cells in the freeze media is attributed to the fact that the number of cells collected was not quantified for a majority of the samples prior to aliquoting and freezing. The concentration of DNA extracted from each sample is presented in Table A.1.

The Infinium HumanMethylation27 BeadArray™ (Illumina Inc, San Diego, CA) is designed so that 24 samples can be run per chip, therefore, 4 chips were required to run the 92 samples. Samples were run in duplicate to assess inter-chip variability, so 1 set of duplicates was run on chips 1 and 2, another on chips 2 and 3, and the last one on chips 3 and 4. The 3 samples run in duplicate were: JS020, JS052, and JS079. The samples were randomly selected from those samples with enough DNA for more than one methylation assay.

The methylation assay requires bisulfite modified DNA which was prepared using the EZ-96 DNA Methylation Kit™ according to the manufacturer's instructions (Zymo Research Corp., USA). The bisulfite modification step converts unmethylated cytosines

to uracil whereas methylated cytosines are protected by the covalently bound methyl group. The DNA is then amplified and applied to a chip. Quantitative, loci-specific methylation is assessed using the Infinium HumanMethylation27 BeadArray™ (Illumina Inc, San Diego, CA). The array interrogates 27,578 loci located in more than 14,000 genes. Two different probes (one against the methylated site and one against the unmethylated site) are hybridized with the bisulfite modified DNA. Next, a single-base extension adds one of two possible fluorescent probes (one for methylated (C) and one for unmethylated (T) alleles). The samples are stained and scanned. Methylation status is then represented by a beta value which is calculated from the ratio of fluorescent signals from methylated to the sum of methylated and unmethylated probes and ranges from 0 (unmethylated) to 1 (methylated).

Background normalization was done according to the guidelines recommended by Illumina using the GenomeStudio Methylation module. This method subtracts the average signal of the negative control bead-types from the probe signals. The median absolute deviation method is used to remove outliers. Step by step instructions for normalizing the data are detailed in the Illumina user manual. Normalized beta values were then output into text files and used in subsequent analyses. Heat maps were generated with the MultiExperiment Viewer (113, 114).

Assessment of methylation analysis success

First the methylation level of the control DNA was examined in order to ensure bisulfite conversion. Since this DNA was almost completely methylated as expected, I then assessed internal validity by examining gender specific methylation of 6 x-linked housekeeping genes (EFNB1, ELK1, FMR1, G6PD, GPC3, GLA) (115, 116). Overall,

methylation of these 6 genes was as expected in that females exhibited hemimethylation and males had very little methylation at the loci in these genes ($p<0.0001$ for each gene). The mean methylation level of each of the housekeeping genes by gender is presented in Table A.2. The Illumina Infinium assay required 4 chips, thus a subset of samples were run in duplicate to assess inter-chip variability. Pearson correlation coefficients were greater than 0.99 for each set of replicates ($p<0.0001$). Table A.3 presents the correlation coefficients and p-values for each set of duplicates.

Analysis of manuscript 1

SGA and methylation level

A total of 90 infants were included in this analysis. Maternal and infant characteristics of SGA and non-SGA infants were compared using fisher's exact test and t-tests as appropriate. SGA was defined as having a birthweight that is less than the 10th percentile for a given gestational age based on the birthweight percentiles for gestational age created by Alexander et al. A t-test was used to compare the mean methylation levels of each gene of interest between SGA and non-SGA infants. Logistic regression was used to compute adjusted odds ratios and 95% confidence intervals for the association between SGA and the methylation level of each gene. The beta values obtained from the methylation assay were treated as continuous variables. The methylation of the CYP1A1 gene was determined by averaging the beta-values of 4 CpG sites whereas the methylation level of the HIF1A gene was determined by averaging the beta values of 2 CpG sites. The methylation of GSTM1 and the GSTT1 genes were each based on the methylation level of one CpG site. Table A.4 presents the mean, standard deviation, and the minimum and maximum beta values of each CpG site assessed.

APPENDIX A (CONTINUED)

Adjusted models controlled for covariates and potential confounders. Variables used in this analysis were classified as follows:

1. Gender: Male or female
2. Parity: nulliparous or multiparous
3. Cigarette smoke exposure: Two indicators of cigarette smoke exposure were considered. The first (smoker) was a dichotomous variable (yes or no). Smokers included everyone with a cotinine level ≥1.0. Cigarette smoke exposure was also evaluated as a continuous variable (cotinine) by including the cotinine measurement in the model.
4. Prenatal vitamin use: Yes or no as indicated in the medical record.
5. Pregnancy complications: yes or no. Pregnancy complications included the presence of diabetes, gestational diabetes, preeclampsia, placental previa, or hypertension.
6. Maternal age: This was left as a continuous variable.
7. Prenatal care: This was originally classified as adequate, inadequate, or none. However, due to the small sample size, this was reclassified as adequate or not.
8. Race: Race was originally classified as black, white, Hispanic, or other. However, the small sample size required that some of these categories be combined. Therefore it was reclassified as black or non-black. In addition, we examined race classified as black, white, or other.

In this analysis, SGA was the main outcome of interest and DNA methylation of each gene was the primary predictor. Although DNA methylation is a possible mediator in the SGA-smoking relationship, we were unable to assess this because the study sample

had few smoke exposed SGA infants. Gender, parity, race, prenatal vitamin use, pregnancy complications, prenatal care, maternal age, and smoking were examined as potential confounders. Confounders were identified by comparing the crude and adjusted odds ratios. The models for the crude analysis are:

$$\text{logit } \{Pr(SGA=1|X)\}=\beta_0+\beta_1(CYP1A1)$$

$$\text{logit } \{Pr(SGA=1|X)\}=\beta_0+\beta_1(HIF1A)$$

$$\text{logit } \{Pr(SGA=1|X)\}=\beta_0+\beta_1(GSTM1)$$

$$\text{logit } \{Pr(SGA=1|X)\}=\beta_0+\beta_1(GSTT1)$$

Potential confounders were added to the models one at a time and if the estimate changed by more than 10%, the variable was adjusted for. Table A.5 below summarizes the models and confounders assessed as well as the different classifications that were considered.

The final models used race as a dichotomous variable (black or non-black) and kept variables continuous whenever possible (maternal age and cotinine level as the indicator of cigarette smoke exposure). The final models are:

$$\text{logit } \{Pr(SGA=1|X)\}=\beta_0+\beta_1(CYP1A1)+\beta_2(gender)+\beta_3(cotinine)+\beta_4(prenatal\ care)+\beta_5(race)+\beta_6(maternal\ age)+\beta_7(parity)$$

APPENDIX A (CONTINUED)

$$\text{logit}\{Pr(SGA=1|X)\}=\beta_0+\beta_1(HIF1A)+\beta_2(gender)+\beta_3(cotinine)+\beta_4(prenatal\ care)+\beta_5(race)+\beta_6(maternal\ age)+\beta_7(parity)$$

$$\text{logit}\{Pr(SGA=1|X)\}=\beta_0+\beta_1(GSTM1)+\beta_2(cotinine)+\beta_3(prenatal\ care)+\beta_4(race)+\beta_5(maternal\ age)+\beta_6(parity)$$

$$\text{logit}\{Pr(SGA=1|X)\}=\beta_0+\beta_1(GSTT1)+\beta_2(cotinine)+\beta_3(prenatal\ care)+\beta_4(race)+\beta_5(maternal\ age)+\beta_6(parity)$$

The small sample size precluded the use of additional variable classifications. Both crude and adjusted estimates are reported in the final manuscript and the adjusted estimates controlled for confounders and covariates that previous authors found to be strongly associated with DNA methylation or SGA risk. Gender was the only confounder identified and it was included in models examining methylation of CYP1A1 and HIF1A. In addition, models were adjusted for smoking, prenatal care, race, maternal age, and parity. These variables were selected because smoking, race, and age have all be found to be associated with DNA methylation. Parity is frequently tied to maternal age and it as well as each of the other covariates has been reported to be associated with SGA risk. A summary of the parameter estimates and model fit statistics are presented in Table A.6. Dose-response trends were assessed with the Cochran-Armitage trend test and by generating effect estimates for different levels of exposure. This statistic is the same as the score test statistics in the testing global null hypothesis section of the proc logistic output.

In sub-analyses related to the discussion portion of the manuscript, the impact of gender on the results was assessed by excluding male infants. This exclusion was done

because there was only 1 SGA male and gender was identified as a confounder in some of the analyses. The results were very similar. As in our original analysis, only methylation of GSTT1 was associated with SGA risk in both crude (OR=1.11, 95%CI=1.04, 1.20) and adjusted analyses (OR=1.13, 95%CI=1.01, 1.26). Similarly, excluding extremely preterm infants (<32 weeks of gestation) did not markedly alter the interpretation of the results. The adjusted odds ratios for a 0.01 increase in methylation were similar for CYP1A1 (OR=1.20, 95%CI=0.48, 3.03), HIF1A (OR=1.62, 95%CI=1.05, 2.60), GSTM1 (OR=1.00, 95%CI=0.92, 1.09), or GSTT1 (OR=1.10, 95%CI=1.03, 1.18). However, it is important to note that the association between methylation of HIF1A and SGA was significant after excluding extremely preterm infants.

Hypermethylation and cotinine level

For this analysis, the impact of cigarette smoke exposure on DNA methylation was assessed using the same 90 infants in the previous analysis. In order to assess this, the outcome, DNA methylation, was dichotomized. The current literature on DNA methylation classified hypermethylation in different ways and there does not appear to be a consensus on what level of methylation is indicative of hypermethylation. For example, Dietrich et al used the median and Zhu et al used greater than the 90th percentile (120, 121). Although I considered using a cut-point at the mean, it was thought that this may not adequately differentiate infants with much higher levels of methylation. Therefore, infants with methylation level at or above the 75th percentile were compared to those with a methylation level falling below the 75th percentile. The cut points (based on Illumina beta values) used to differentiate between hypermethylation and normal methylation are: 0.0823591 (CYP1A1), 0.0409117 (HIF1A), 0.0569107 (GSTT1), and 0.1739638 (GSTM1).

APPENDIX A (CONTINUED)

Although a cotinine level greater than 1ng/ml is indicative of cigarette smoke exposure, we left cotinine as a continuous variable so that we could look at the impact of several different levels of exposure in relation to methylation. Maternal and infant characteristics of infants with hypermethylation were compared to those with normal methylation using fisher's exact test and t-tests as appropriate. Each gene was considered separately. Confounders and covariates considered for inclusion in the adjusted models were classified as follows:

1. Gender: Male or female
2. Parity: nulliparous or multiparous
3. Prenatal vitamin use: Yes or no as indicated in the medical record.
4. Pregnancy complications: yes or no. Pregnancy complications included the presence of diabetes, gestational diabetes, preeclampsia, placental previa, or hypertension.
5. Maternal age: This was left as a continuous variable.
6. Prenatal care: This was originally classified as adequate, inadequate, or none. However, due to the small sample size, this was reclassified as adequate or not.
7. Race: Race was originally classified as black, white, Hispanic, or other. However, the small sample size required that some of these categories be combined; therefore it was reclassified as black or non-black. In addition, we examined race classified as black, white, or other.
8. Gestational age: This was left as a continuous variable.

Logistic regression was used to compute crude and adjusted odds ratios and 95% confidence intervals for the risk of hypermethylation of each gene for a given cotinine

level. The main outcome of interest was hypermethylation of each gene and cigarette smoke exposure acted as the primary predictor in this analysis. The crude models that were examined in this analysis are:

$$\text{logit}\{\Pr(\text{Hypermethylation of CYP1A1} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine})$$

$$\text{logit}\{\Pr(\text{Hypermethylation of HIF1A} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine})$$

$$\text{logit}\{\Pr(\text{Hypermethylation of GSTM1} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine})$$

$$\text{logit}\{\Pr(\text{Hypermethylation of GSTT1} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine})$$

Gender, parity, prenatal vitamin use, pregnancy complications, maternal age, prenatal care, race, and gestational age were examined as potential confounders. Adjusted models controlled for covariates and potential confounders. Confounders were identified by comparing the crude and adjusted odds ratios. If the estimate changed by more than 10%, the variable was adjusted for. Table A.7 below summarizes the models and confounders assessed as well as the different variable classifications that were considered. None of the variables that we assessed were considered confounders under the definition described above. However, the adjusted odds ratios control for maternal age and race because these are important variables in continued smoking during pregnancy and have been associated with DNA methylation. In the final model cotinine was kept as a continuous variable, race was dichotomized (black or non-black) and maternal age was kept as a continuous variable. Crude and adjusted odds ratios are reported in the final manuscript.

APPENDIX A (CONTINUED)

The final adjusted models are:

$$\text{logit}\{\Pr(\text{Hypermethylation of CYP1A1} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine}) + \beta_2(\text{race}) + \beta_3(\text{maternal age})$$

$$\text{logit}\{\Pr(\text{Hypermethylation of HIF1A} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine}) + \beta_2(\text{race}) + \beta_3(\text{maternal age})$$

$$\text{logit}\{\Pr(\text{Hypermethylation of GSTM1} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine}) + \beta_2(\text{race}) + \beta_3(\text{maternal age})$$

$$\text{logit}\{\Pr(\text{Hypermethylation of GSTT1} = 1|X)\} = \beta_0 + \beta_1(\text{cotinine}) + \beta_2(\text{race}) + \beta_3(\text{maternal age})$$

A summary of the parameter estimates and fit statistics for the crude and adjusted models are presented in Table A.8. The risk of hypermethylation was examined for a 5ng/ml increase in cotinine level and a 14ng/ml increase, levels which were selected based on previous research. Infants are generally considered to be exposed to cotinine if the cord blood cotinine level is greater than 1 ng/ml (119, 122, 123). Nafstad et al found that 14ng/ml differentiates active and passive smokers well, but it does not capture occasional smokers (123). Further, the study found that an increase in one cigarette per day increased cotinine levels by almost 5ng/ml (4.4 ng/ml; 95% CI: 1.1-7.6) (123). Dose-response trends were assessed with the Cochran-Armitage trend test and by generating effect estimates for different levels of exposure. This statistic is the same as the score test statistics in the testing global null hypothesis section of the proc logistic output.

APPENDIX A (CONTINUED)

Analysis of manuscript 2

For this manuscript we sought to examine the association between birthweight and methylation of the IGF2 gene, each CpG site assessed in the IGF2 gene, and the IGF1 gene. Birthweight, the dependant variable, was treated as continuous. Methylation of the IGF1 gene was based on the methylation level of one CpG site in that gene whereas the methylation level of the IGF2 gene was determined by averaging the methylation of 5 CpG sites in the IGF2 gene. Covariates were kept as continuous variables whenever possible. Variables considered for inclusion in the adjusted models were classified as follows:

1. Gender: Male or female
2. Parity: nulliparous or multiparous
3. Prenatal vitamin use: Yes or no as indicated in the medical record.
4. Pregnancy complications: yes or no. Pregnancy complications included the presence of diabetes, gestational diabetes, preeclampsia, placental previa, or hypertension.
5. Maternal age: This was kept as a continuous variable.
6. Prenatal care: This was originally classified as adequate, inadequate, or none. However, due to the small sample size, this was reclassified as adequate or not.
7. Race: Race was originally classified as black, white, Hispanic, or other. However, the small sample size required that some of these categories be combined; therefore it was reclassified as black or non-black. In addition, we examined race classified as black, white, or other.
8. Gestational age: This was kept as a continuous variable.

APPENDIX A (CONTINUED)

9. Preterm birth: yes or no. Infants were considered preterm if they were less than 37 weeks of gestation.
10. Cigarette smoke exposure: Two indicators of cigarette smoke exposure were considered. The first (smoker) was a dichotomous variable (yes or no). Smokers included everyone with a cotinine level ≥1.0. Cigarette smoke exposure was also evaluated as a continuous variable (cotinine) by including the cotinine measurement in the model.

In order to ease interpretation and ensure linearity between the outcome and methylation level, Illumina beta values were converted to z-scores. The z-scores were then used in the regression models. The validity of this transformation was assessed by plotting birthweight verses the transformed methylation values. The plots for IGF1 and IGF2 are depicted in Figures A.1 and A.2.

First, univariate linear regression was used to assess the crude association between birthweight (outcome) and DNA methylation of the IGF1 and IGF2 genes (primary predictors). The crude models assessed were:

$$E(\text{Birthweight}|X) = \beta_0+\beta_1(\text{IGF1})$$

$$E(\text{Birthweight}|X) = \beta_0+\beta_1(\text{IGF2})$$

I then checked to ensure that the normality assumption was not violated. Residuals were generated for the models and then plotted verses the predicted values. According to the Shapiro-Wilk statistic, the data were not normal for IGF2 ($p= 0.0306$), but IGF1 appeared to be normal ($p= 0.0742$). As a result, the outcome, birthweight, was log transformed.

APPENDIX A (CONTINUED)

This resulted in normally distributed data for IGF1 (p= 0.6260) and IGF2 (p=0.5548) based on the Shapiro-Wilk statistic. As a result of these transformations, the reported effect estimates indicate the percent change in the average birthweight (100*parameter estimate) per one standard deviation change in methylation. Crude models were then re-run using the transformed variable.

Gender, parity, prenatal vitamin use, cigarette smoke exposure, pregnancy complications, maternal age, prenatal care, race, and gestational age were examined as potential confounders. In addition, prenatal vitamin use was assessed as a possible moderator. I then identified potential confounders by adding each covariate to the unadjusted model. Each covariate was assessed independently and if the p-value associated with that variable was less than 0.05, the variable was adjusted for. A summary of the confounders assessed and their associated p-values are presented in Table A.9. Multivariable linear regression models were then constructed to control for potential confounders. The final models were adjusted for gender, maternal age, parity, gestational age, and pregnancy complications.

$$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF1})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$$

$$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF2})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$$

$$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF2, CPG site1})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$$

APPENDIX A (CONTINUED)

$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF2, CPG site 2})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$

$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF2, CPG site 3})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$

$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF2, CPG site 4})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$

$E\{(\log(\text{Birthweight}))|X\} = \beta_0+\beta_1(\text{IGF2, CPG site 5})+\beta_2(\text{gender})+\beta_3(\text{parity})+\beta_4(\text{gestational age})+\beta_5(\text{complication})$

A summary of the parameter estimates and fit statistics for the crude and final models are presented in Table A.10. Initial assessments indicated that race was not a confounder and therefore it was not included in the multivariable models. However, I sought to explore this further and re-ran the analysis including race as well as the other confounders (gender, maternal age, parity, gestational age, and pregnancy complications) in the model. This did not appreciably alter the results or the conclusions; therefore race was not included in the final multivariable model (Table A.11).

Interestingly, prenatal vitamin use, a surrogate indicator of folic acid intake, did not significantly impact the association between birthweight and methylation of the IGF1 or IGF2 gene in our regression models. Our small sample size precluded a detailed investigation of this, but we did stratify the data by prenatal vitamin use to do an exploratory examination of the data. Stratified analysis suggested that the associations between methylation of IGF2 and birthweight differed by maternal prenatal vitamin use,

but the associations were not significant even after adjusting for confounders. Among infants exposed to prenatal vitamins, one standard deviation increase in methylation of IGF2 was associated with a non-significant 0.78% decrease in birthweight (95%CI-4.17, 2.61). On the other hand, among infants born to women who did not use prenatal vitamins, a one standard deviation increase in methylation of IGF2 was associated with a 3.72% increase in birthweight (95%CI= -5.32, 12.76). Results from analyses examining IGF1 stratified by prenatal vitamin were similar to that of the unstratified analysis. Among prenatal vitamin users, a one standard deviation increase in methylation of IGF1 was associated with a 4.21% decrease in birthweight (95%CI= -7.44, -0.98). Comparable results were found for women who did not use prenatal vitamins, though they were not statistically significant (β= -3.71, 95%CI= -11.07, 3.64). In an attempt to increase power in the examination of prenatal vitamin use as a moderator of the DNA hypermethylation–birthweight association, prenatal vitamin use and an interaction term were added to the final model for IGF1 and then IGF2. In both cases, the interaction terms were non-significant [(IGF1=-3.93, 95%CI=-12.72, 4.87); (IGF2=-7.32, 95%CI=-16.95, 2.31)].

APPENDIX A (CONTINUED)

Table A.1 Concentration (ng/µl) of DNA extracted from each sample of umbilical cord blood derived nucleated cells

Sample ID	Concentration (ng/µl)	Sample ID (continued)	Concentration (ng/µl)	Sample ID (continued)	Concentration (ng/µl)
207	36.69	JS032	77.02	JS066	96.78
210	42.95	JS033	99.86	JS067	113.88
213	41.98	JS034	0.41	JS068	21.27
214	42.2	JS035	5.86	JS069	112.04
215	50.17	JS036	19.11	JS070	3.52
216	53.87	JS037	19.85	JS071	8.1
JS007	74.6	JS040	95.79	JS074	9.06
JS008	46.78	JS041	7.8	JS075	65.26
JS009	82.7	JS042	9.59	JS076	3.14
JS010	119.36	JS043	94.05	JS077	81.52
JS011	21.06	JS044	16.68	JS078	1.58
JS012	14.39	JS045	82.07	JS079	30.85
JS013	13.54	JS046	11.1	JS080	1.47
JS014	12.92	JS047	9.12	JS081	6.66
JS015	16.75	JS048	53.25	JS082	53.71
JS016	8.19	JS050	11.53	JS083	69.04
JS017	66.31	JS051	20.94	JS084	126.35
JS018	12.75	JS052	26.69	JS085	43.28
JS019	31.16	JS053	10.42	JS086	3.37
JS020	92.52	JS054	85.04	JS087	85.1
JS021	98.21	JS055	3.79	JS088	74.16
JS022	16.33	JS056	142.1	JS089	66.55
JS023	24.59	JS057	120.47	JS090	88.74
JS024	83.12	JS058	102.84	JS091	154.26
JS025	17.75	JS059	19.36	JS092	118.69
JS026	39	JS060	40.06	JS093	38.98
JS027	149.81	JS061	4.33	JS094	3.5
JS028	118.61	JS062	45.33	JS099	140.47
JS029	36.54	JS063	114.35	JS100	85.14
JS030	38.49	JS064	79.45	JS101	107.97
JS031	8.58	JS065	165.39	blank__93	0

APPENDIX A (CONTINUED)

Table A.2 Distribution of the mean methylation level of each housekeeping gene by gender

	Mean Methylation Level	
Gene	**Male**	**Female**
EFNB1	0.05	0.39
GLA	0.06	0.25
FMR1	0.19	0.45
GPC3	0.03	0.57
ELK1	0.05	0.55
G6PD	0.05	0.33

APPENDIX A (CONTINUED)

Table A.3 Pearson correlation coefficients for assessment of interchip variability

Sample ID	R	p-value
JS020	0.998	<0.0001
JS052	0.998	<0.0001
JS079	0.997	<0.0001

APPENDIX A (CONTINUED)

Table A.4 Distribution of the beta values for the CpG sites assessed in manuscript 1

Gene	Mean	Standard Deviation	Minimum	Maximum
CYP1A1, CpG site 1	0.02	0.01	0.01	0.04
CYP1A1, CpG site 2	0.05	0.01	0.03	0.08
CYP1A1, CpG site 3	0.04	0.03	0.01	0.14
CYP1A1, CpG site 4	0.19	0.03	0.13	0.25
HIF1A, CpG site 1	0.04	0.01	0.02	0.10
HIF1A, CpG site 2	0.03	0.02	0.01	0.17
GSTM1	0.10	0.10	0.01	0.35
GSTT1	0.09	0.11	0.02	0.45

APPENDIX A (CONTINUED)

Table A.5 Summary of the confounders assessed for the association between SGA[1] and methylation of selected genes

		CYP1A1		HIF1A		GSTM1		GSTT1	
Variable	**Variable type**	**OR[2,3]**	**% Change[4]**	**OR**	**% change**	**OR**	**% change**	**OR**	**% change**
crude association		0.868		1.232		0.972		1.077	
Gender	categorical	0.985	13.48	1.428	15.91	0.984	1.23	1.087	0.93
Parity	categorical	0.894	3.00	1.206	-2.11	0.964	-0.82	1.091	1.30
Smoker	categorical	0.836	-3.69	1.247	1.22	0.973	0.10	1.076	-0.09
Smoke exposure (cotinine level)	continuous	0.814	-6.22	1.241	0.73	0.976	0.41	1.069	-0.74
Prenatal vitamin use	categorical	0.870	0.23	1.241	0.73	0.972	0.00	1.079	0.19
Pregnancy complications	categorical	0.893	2.88	1.223	-0.73	0.973	0.10	1.077	0.00
Maternal age	continuous	0.889	2.42	1.231	-0.08	0.975	0.31	1.102	2.32
Prenatal care	categorical	0.877	1.04	1.184	-3.90	0.980	0.82	1.086	0.84
Race-white, black, other	categorical	0.873	0.58	1.213	-1.54	0.992	2.06	1.071	-0.56
Race-black, non-black	categorical	0.868	0.00	1.219	-1.06	0.994	2.26	1.073	-0.37

[1]SGA=small for gestational age

[2]OR=Odds ratio

[3]Odds ratios for a 0.01 increase in methylation level

[4]the percent change in the odds ratio from the crude odds ratio attributed to the addition of the selected variable

APPENDIX A (CONTINUED)

Table A.6 Parameter estimates and model fit statistics for the crude and adjusted models assessing the association between methylation of selected genes and SGA

						Hosmer and Lemeshow Goodness-of-fit test		Model fit statistics for the intercept and covariates	
	Variable	**Parameter Estimate**	**Standard Error**	**Wald Chi-Square**	**pr> Chi-square**	**Chi-square**	**p-value**	**AIC**	**-2 Log L**
CYP1A1						4.50	0.81	66.56	62.56
	Intercept	-1.03	2.18	0.22	0.64				
	CYP1A1	-14.10	29.43	0.23	0.63				
CYP1A1						2.43	0.97	55.76	39.76
	Intercept	-0.46	3.63	0.02	0.90				
	CYP1A1	-3.29	43.96	0.01	0.94				
	Gender	-3.06	1.31	5.47	0.02				
	Maternal Age	-0.06	0.09	0.45	0.50				
	Parity	1.07	0.94	1.30	0.25				
	Cotinine	0.03	0.02	2.35	0.13				
	Prenatal Care	-0.56	0.94	0.35	0.55				
	Race	2.12	0.97	4.79	0.03				
HIF1A						13.70	0.09	65.46	61.46
	Intercept	-2.86	0.75	14.48	<0.01				
	HIF1A	20.85	17.12	1.48	0.22				

APPENDIX A (CONTINUED)

Table A.6 (Continued) Parameter estimates and model fit statistics for the crude and adjusted models assessing the association between methylation of selected genes and SGA

						Hosmer and Lemeshow Goodness-of-fit test		Model fit statistics for the intercept and covariates	
	Variable	**Parameter Estimate**	**Standard Error**	**Wald Chi-Square**	**pr> Chi-square**	**Chi-square**	**p-value**	**AIC**	**-2 Log L**
HIF1A						4.77	0.78	52.23	36.23
	Intercept	-1.22	2.85	0.18	0.67				
	HIF1A	46.50	24.18	3.70	0.05				
	Gender	-4.02	1.67	5.83	0.02				
	Maternal Age	-0.10	0.10	1.02	0.31				
	Parity	0.99	1.01	0.96	0.33				
	Cotinine	0.04	0.02	2.67	0.10				
	Prenatal Care	-0.44	0.93	0.22	0.64				
	Race	2.10	1.00	4.39	0.04				
GSTM1						1.99	0.98	66.18	62.18
	Intercept	-1.82	0.46	15.96	<0.01				
	GSTM1	-2.83	3.78	0.56	0.46				
GSTM1						3.83	0.87	63.63	49.63
	Intercept	-1.40	2.33	0.36	0.55				
	GSTM1	0.48	4.24	0.01	0.91				
	Maternal Age	-0.05	0.08	0.44	0.51				
	Parity	0.97	0.89	1.19	0.27				
	Cotinine	0.02	0.02	1.17	0.28				
	Prenatal Care	-0.97	0.79	1.50	0.22				
	Race	2.02	0.84	5.73	0.02				

APPENDIX A (CONTINUED)

Table A.6 (Continued) Parameter estimates and model fit statistics for the crude and adjusted models assessing the association between methylation of selected genes and SGA

						Hosmer and Lemeshow Goodness-of-fit test		Model fit statistics for the intercept and covariates	
	Variable	**Parameter Estimate**	**Standard Error**	**Wald Chi-Square**	**pr> Chi-square**	**Chi-square**	**p-value**	**AIC**	**-2 Log L**
GSTT1						15.78	0.05	57.71	53.71
	Intercept	-3.05	0.55	30.61	<0.01				
	GSTM1	7.41	2.44	9.20	<0.01				
GSTT1						8.22	0.41	54.34	40.34
	Intercept	-0.50	2.89	0.03	0.86				
	GSTM1	9.91	3.58	7.66	0.01				
	Maternal Age	-0.12	0.10	1.46	0.23				
	Parity	1.38	1.05	1.73	0.19				
	Cotinine	0.00	0.02	0.02	0.88				
	Prenatal Care	-1.28	0.92	1.94	0.16				
	Race	1.64	0.91	3.23	0.07				

APPENDIX A (CONTINUED)

Table A.7 Summary of the confounders assessed for the association between methylation of selected genes and cigarette smoke exposure

		GSTT1		GSTM1		HIF1A		CYP1A1	
Variable	**variable type**	**OR[1,2]**	**% change[3]**	**OR**	**% change**	**OR**	**% change**	**OR**	**% change**
crude		1.030		0.985		1.010		0.998	
Gender	categorical	1.031	0.10	0.984	-0.10	1.011	0.10	0.997	-0.10
Parity	categorical	1.030	0.00	0.984	-0.10	1.011	0.10	1.000	0.20
Prenatal vitamin use	categorical	1.029	-0.10	0.982	-0.30	1.007	-0.30	0.995	-0.30
Pregnancy complications	categorical	1.030	0.00	0.985	0.00	1.011	0.10	0.998	0.00
Maternal age	continuous	1.031	0.10	0.989	0.41	1.015	0.50	0.999	0.10
prenatal care	categorical	1.032	0.19	0.986	0.10	1.011	0.10	0.998	0.00
Race-white, black, other	categorical	1.032	0.19	0.983	-0.20	1.005	-0.50	0.992	-0.60
Race-black, non-black	categorical	1.031	0.10	0.986	0.10	1.011	0.10	0.998	0.00
Gestational age	continuous	1.032	0.19	0.987	0.20	1.010	0.00	0.997	-0.10

[1]OR=odds ratio

[2]The odds ratio is for a 1 ng/ml increase in cotinine level

[3]the percent change in the odds ratio from the crude odds ratio attributed to the addition of the selected variable

APPENDIX A (CONTINUED)

Table A.8 Parameter estimates and model fit statistics for the crude and adjusted models assessing the association between smoking and hypermethylation of selected genes

						Hosmer and Lemeshow Goodness-of-fit test		Model fit statistics for the intercept and covariates	
	Variable	**Parameter Estimate**	**Standard Error**	**Wald Chi-Square**	**pr> Chi-square**	**Chi-square**	**p-value**	**AIC**	**-2 Log L**
Hypermethylation of CYP1A1						3.13	0.08	104.09	100.09
	Intercept	-1.12	0.25	19.41	<0.01				
	Cotinine	-0.002	0.02	0.02	0.89				
Hypermethylation of CYP1A1						7.78	0.46	107.60	99.60
	Intercept	-1.28	1.22	1.11	0.29				
	Cotinine	0.00	0.02	0.01	0.94				
	Maternal Age	0.01	0.04	0.04	0.83				
	Race	-0.37	0.64	0.34	0.56				
Hypermethylation of HIF1A						0.08	.	105.73	101.73
	Intercept	-1.12	0.25	19.55	<0.01				
	Cotinine	0.01	0.01	0.60	0.44				
Hypermethylation of HIF1A						8.84	0.26	106.96	98.96
	Intercept	-3.15	1.29	5.93	0.01				
	Cotinine	0.02	0.01	1.23	0.27				
	Maternal Age	0.07	0.04	2.58	0.11				
	Race	0.46	0.60	0.61	0.44				

APPENDIX A (CONTINUED)

Table A.8 (Continued) Parameter estimates and model fit statistics for the crude and adjusted models assessing the association between smoking and hypermethylation of selected genes

						Hosmer and Lemeshow Goodness-of-fit test		Model fit statistics for the intercept and covariates	
	Variable	**Parameter Estimate**	**Standard Error**	**Wald Chi-Square**	**pr> Chi-square**	**Chi-square**	**p-value**	**AIC**	**-2 Log L**
Hypermethylation of GSTM1						0.90	0.34	105.65	101.65
	Intercept	-1.02	0.25	16.77	<0.01				
	Cotinine	-0.01	0.02	0.52	0.47				
Hypermethylation of GSTM1						8.81	0.29	106.75	98.75
	Intercept	-1.98	1.22	2.64	0.10				
	Cotinine	-0.01	0.02	0.31	0.58				
	Maternal Age	0.04	0.04	0.93	0.34				
	Race	-0.74	0.70	1.11	0.29				
Hypermethylation of GSTT1						0.01	.	101.78	97.78
	Intercept	-1.23	0.26	22.17	<0.01				
	Cotinine	0.03	0.02	3.89	0.05				
Hypermethylation of GSTT1						10.20	0.18	100.88	92.88
	Intercept	-2.72	1.35	4.02	0.05				
	Cotinine	0.03	0.02	4.67	0.03				
	Maternal Age	0.04	0.04	0.82	0.37				
	Race	1.29	0.59	4.77	0.03				

APPENDIX A (CONTINUED)

Table A.9 Summary of the confounders assessed for the association between methylation of IGF1 and IGF2 and birthweight

		IGF1		IGF2	
Variable	**Variable type**	**β[1]**	**p-value[2]**	**β**	**p-value**
Crude		-0.041		-0.006	0.81
Gender	categorical	-0.049	<0.01	-0.019	<0.01
Parity	categorical	-0.043	<0.01	-0.004	<0.01
Smoker	categorical	-0.041	0.54	-0.006	0.58
Smoke exposure (cotinine level)	continuous	-0.041	0.37	-0.004	0.39
Prenatal vitamin use	categorical	-0.040	0.89	-0.006	0.79
Pregnancy complications	categorical	-0.041	0.02	-0.008	0.03
Maternal age	continuous	-0.045	0.01	-0.009	0.02
Gestational age	continuous	-0.029	<0.01	0.007	<0.01
Preterm	categorical	-0.032	<0.01	-0.415	<0.01
Prenatal care	categorical	-0.038	0.17	-0.006	0.14
Race-white, black, other	categorical	-0.039	0.59, 0.70	-0.005	0.46, 0.79
Race-black, non-black	categorical	-0.038	0.70	-0.006	0.47

[1] Betas (β) are for the association of interest and represent the change in log(birthweight) for a one standard deviation increase in methylation

[2]p-value represents the significance of the covariate added to the univariate model

APPENDIX A (CONTINUED)

Table A.10 Summary of the parameter estimates and fit statistics for the crude and adjusted models examining the association between methylation of IGF1, IGF2, and the CpG sites in IGF2 and birthweight

	Variable	Parameter Estimate	Standard Error	t-value	pr>|t|	R sq
IGF1						
	Intercept	8.059	0.025	318.200	<0.0001	0.028
	Z-score of methylation of IGF1	-0.041	0.025	-1.590	0.115	
IGF1						0.716
	Intercept	5.044	0.278	18.130	<.0001	
	Z-score of methylation of IGF1	-0.036	0.014	-2.530	0.013	
	Gender	0.088	0.030	2.950	0.004	
	Maternal Age	0.003	0.002	1.430	0.158	
	Parity	-0.084	0.032	-2.600	0.011	
	Gestational Age	0.076	0.007	10.540	<.0001	
	Complications	0.008	0.033	0.230	0.820	
IGF2						0.001
	Intercept	8.059	0.026	313.830	<.0001	
	Z-score of methylation of IGF2	-0.006	0.026	-0.240	0.811	
IGF2						0.694
	Intercept	4.984	0.289	17.240	<.0001	
	Z-score of methylation of IGF2	0.001	0.015	0.050	0.964	
	Gender	0.080	0.031	2.580	0.012	
	Maternal Age	0.003	0.002	1.190	0.239	
	Parity	-0.083	0.033	-2.470	0.016	
	Gestational Age	0.078	0.007	10.410	<.0001	
	Complications	0.012	0.034	0.340	0.738	
IGF2, CpG site 1						0.0002
	Intercept	8.059	0.026	313.750	<.0001	
	Z-score of methylation of IGF2, CpG site 1	-0.003	0.026	0.905	-0.054	

APPENDIX A (CONTINUED)

Table A.10 (Continued) Summary of the parameter estimates and fit statistics for the crude and adjusted models examining the association between methylation of IGF1, IGF2, and the CpG sites in IGF2 and birthweight

	Variable	Parameter Estimate	Standard Error	t-value	pr>|t|	R sq
IGF2, CpG site 1						0.694
	Intercept	4.984	0.288	17.310	<.0001	
	Z-score of methylation of IGF2, CpG site 1	-0.001	0.015	-0.080	0.940	
	Gender	0.081	0.031	2.620	0.010	
	Maternal Age	0.003	0.002	1.190	0.236	
	Parity	-0.083	0.034	-2.470	0.016	
	Gestational Age	0.078	0.007	10.480	<.0001	
	Complications	0.012	0.035	0.340	0.735	
IGF2, CpG site 2						0.0001
	Intercept	8.059	0.026	313.740	<0.0001	
	Z-score of methylation of IGF2, CpG site 2	-0.002	0.026	-0.090	0.928	
IGF2, CpG site 2						0.694
	Intercept	4.982	0.290	17.210	<.0001	
	Z-score of methylation of IGF2, CpG site 2	0.001	0.015	0.100	0.922	
	Gender	0.080	0.031	2.570	0.012	
	Maternal Age	0.003	0.002	1.190	0.238	
	Parity	-0.083	0.033	-2.470	0.015	
	Gestational Age	0.078	0.007	10.410	<.0001	
	Complications	0.012	0.035	0.340	0.733	
IGF2, CpG site 3						0.044
	Intercept	8.059	0.025	320.900	<.0001	
	Z-score of methylation of IGF2, CpG site 3	-0.051	0.025	-2.020	0.047	

APPENDIX A (CONTINUED)

Table A.10 (Continued) Summary of the parameter estimates and fit statistics for the crude and adjusted models examining the association between methylation of IGF1, IGF2, and the CpG sites in IGF2 and birthweight

	Variable	Parameter Estimate	Standard Error	t-value	pr>|t|	R sq
IGF2, CpG site 3						0.699
	Intercept	5.037	0.288	17.460	<.0001	
	Z-score of methylation of IGF2, CpG site 3	-0.018	0.015	-1.210	0.228	
	Gender	0.083	0.031	2.730	0.008	
	Maternal Age	0.003	0.002	1.190	0.238	
	Parity	-0.077	0.033	-2.310	0.024	
	Gestational Age	0.076	0.007	10.260	<.0001	
	Complications	0.008	0.034	0.240	0.807	
IGF2, CpG site 4						0.001
	Intercept	8.059	0.026	313.810	<.0001	
	Z-score of methylation of IGF2, CpG site 4	-0.005	0.026	-0.210	0.835	
IGF2, CpG site 4						0.699
	Intercept	4.982	0.285	17.450	<.0001	
	Z-score of methylation of IGF2, CpG site 4	-0.017	0.015	-1.150	0.252	
	Gender	0.079	0.031	2.580	0.012	
	Maternal Age	0.003	0.002	1.340	0.183	
	Parity	-0.086	0.033	-2.580	0.012	
	Gestational Age	0.078	0.007	10.540	<.0001	
	Complications	0.009	0.034	0.260	0.795	
IGF2, CpG site 5						0.0003
	Intercept	8.059	0.026	313.770	<.0001	
	Z-score of methylation of IGF2, CpG site 5	0.004	0.026	0.150	0.880	

APPENDIX A (CONTINUED)

Table A.10 (Continued) Summary of the parameter estimates and fit statistics for the crude and adjusted models examining the association between methylation of IGF1, IGF2, and the CpG sites in IGF2 and birthweight

	Variable	Parameter Estimate	Standard Error	t-value	pr>|t|	R sq
IGF2, CpG site 5						0.697
	Intercept	4.962	0.288	17.230	<.0001	
	Z-score of methylation of IGF2, CpG site 5	0.012	0.015	0.840	0.406	
	Gender	0.077	0.031	2.470	0.016	
	Maternal Age	0.003	0.002	1.230	0.224	
	Parity	-0.084	0.033	-2.520	0.014	
	Gestational Age	0.078	0.007	10.550	<.0001	
	Complications	0.014	0.034	0.400	0.691	

APPENDIX A (CONTINUED)

Table A.11 Association between birthweight and DNA methylation of IGF1, IGF2, and the CpG sites in IGF2 after adding race as a covariate

	B[1, 2]	**95% CI[3]**
IGF1	-3.12	-6.09, -0.15
IGF2	0.26	-2.69, 3.21
CpG site 1	-0.04	-2.98, 2.89
CpG site 2	0.16	-2.79, 3.11
CpG site 3	-1.45	-4.44, 1.53
CpG site 4	-1.88	-4.79, 1.03
CpG site 5	1.62	-1.33, 4.56

[1]Percent change in birthweight per standard deviation change in methylation level

[2]Adjusted for gender, maternal age, parity, gestational age, pregnancy complications, and race

[3]CI=confidence interval

APPENDIX A (CONTINUED)

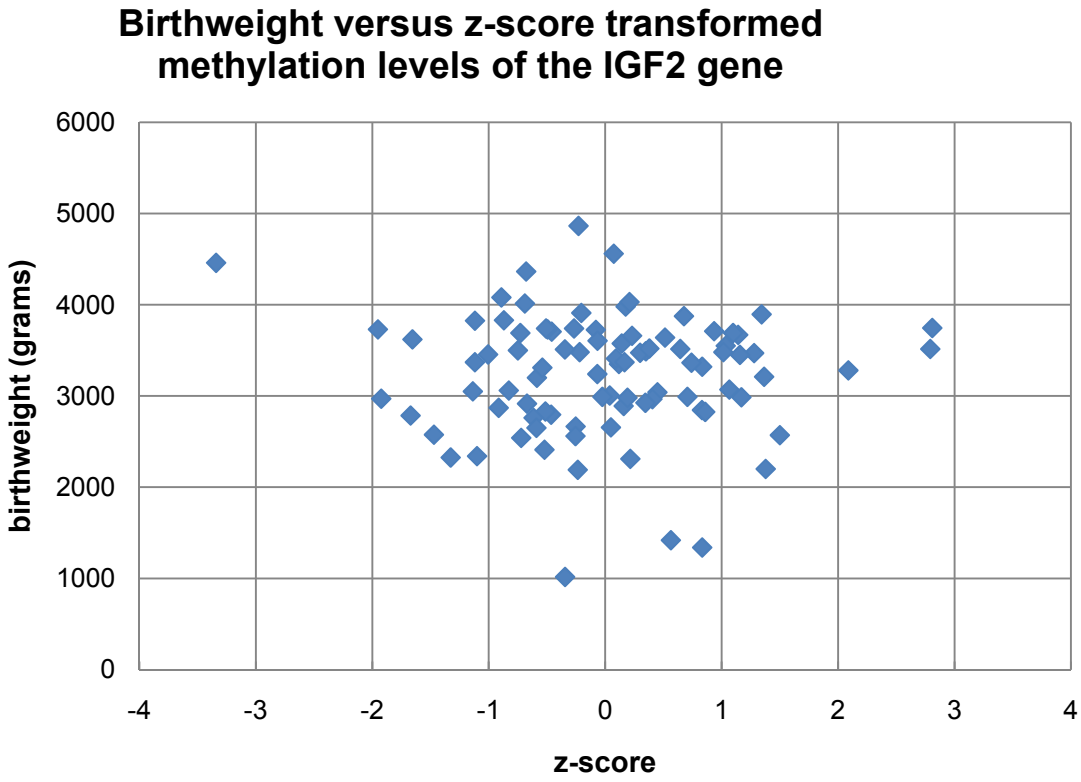

Figure A.1 Plot of birthweight verses the z-score transformed methylation values for the IGF2 gene

APPENDIX A (CONTINUED)

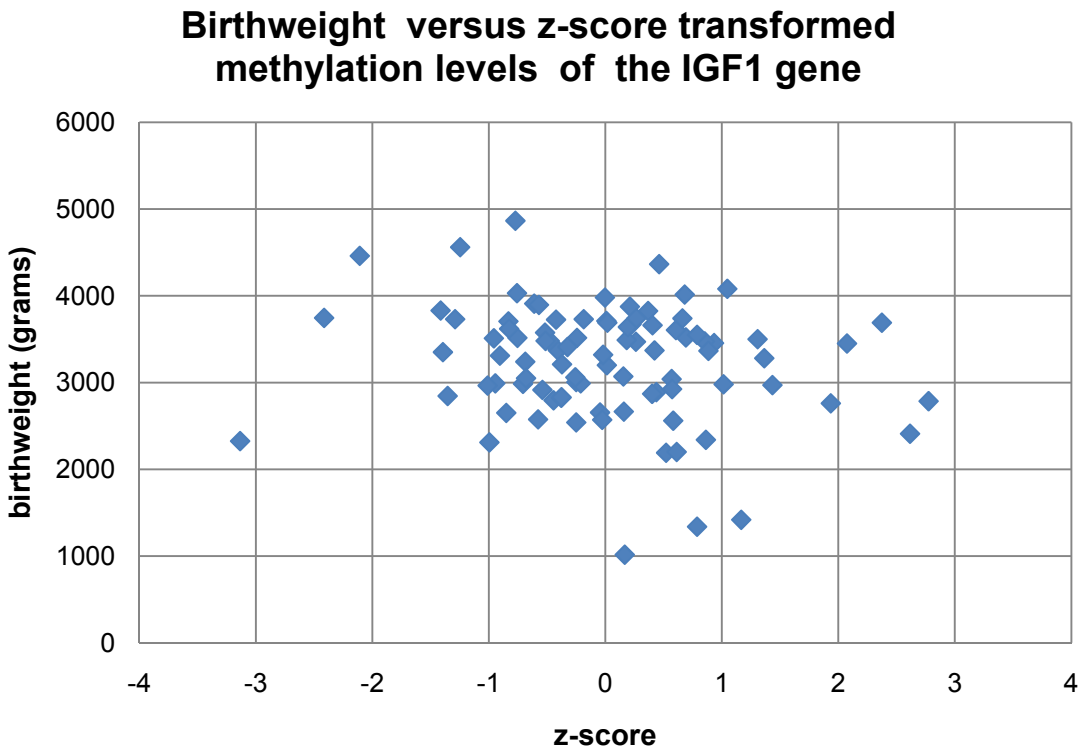

Figure A.2 Plot of birthweight verses the z-score transformed methylation values for the IGF1 gene

APPENDIX B: DATA COLLECTION FORM

Analysis of Lymphocyte Subpopulations in Cord Blood of Term and Preterm Infants

Data Collection Form

UPIN:

SUBJECT IDENTIFIER

Gender: ____ M ____ F

GA: ____ weeks BW: _____ grams

Race/Ethnicity: ____White ____Black/AA ____Native American ____Asian

____Pacific Islander ____Hispanic/Latino

PRENATAL HISTORY

Maternal Age: ____ years G ____ P____ A____ Prenatal Care: ____Yes ____No

Past Medical History: 1. _____ 2. _____

Pregnancy Complications: 1. _____ 2. _____

Pregnancy Medications: 1. _____ 2. _____

Substance Use: ____Smoking ____Alcohol ____Marijuana ____Cocaine ____Heroin

____ Other: _____

Prenatal Labs: Blood Type: ____ Antibody: ____ Rubella: ____ VDRL: ____ GC: ____

Chlamydia: ____ Hep B SAg:____ GBS:____ HIV:____

Pap Smear:_____

LABOR AND DELIVERY HISTORY

ROM: ____Spontaneous ____Artificial ROM prior to delivery: ____ hours

Type: ____NSVD ____Elective C/S Emergency C/S____

APGAR: ____ at 1min ____ at 5min ____ at 10 min

Complication during labor: 1. _____ 2. _____

Complications during delivery: 1. _____ 2. _____

Collected by: _____

Date Collected (mm/dd/yyyy): ____/____/_____

APPENDIX C: POWER ANALYSIS

For manuscript 1, the logistic statement in proc power (SAS 9.2) was used to perform the power analysis. For the analyses examining the association between SGA and DNA methylation of CYP1A1, HIF1a, GSTM1, or GSTT1, the distribution of the predictor of interest, DNA methylation, was specified for each gene, alpha was kept at 0.05, and the sample size was fixed at 90. Power was calculated for several different odds ratios (1.1, 1.5, 2.0, and 2.5) using a response probability of 0.11. Table C.1 provides an overview of the results from the power calculations. GSTT1 and GSTM1 had greater than 80% power for all the odds ratios examined. The study was underpowered to detect small changes in the odds ratio for CYP1A1 and HIF1A, but had greater than 90% power for larger odds ratios (i.e. odds ratios of 2.5). Manuscript one also sought to examine the association between cotinine level and DNA hypermethylation (outcome). For this power analysis, alpha was set at 0.05, response probability was 0.25, the sample size was fixed at 90, and the distribution of the cotinine variable was specified. Power was the same for all the genes (>99%) and is summarized in Table C.1.

For manuscript 2, the power analysis was done using the multreg statement in proc power (SAS 9.2). Alpha was fixed at 0.05, the sample size was 90 and the R-square of the full model as well as the change in R-square were specified for both IGF1 and IGF2. While there was sufficient power for IGF1 (86.6%), analyses of IGF2 lacked power (7.3%).

APPENDIX C (CONTINUED)

Table C.1. Summary of power analysis for Manuscript 1

	Power			
SGA and methylation of each gene	OR=1.1	OR=1.5	OR=2.0	OR=2.5
CYP1A1	6.2%	28.3%	68.7%	91.3%
HIF1A	6.8%	41.7%	87.7%	98.8%
GSTM1	81.0%	>99.9%	>99.9%	>99.9%
GSTT1	88.3%	>99.9%	>99.9%	>99.9%
Hypermethylation of each gene and cotinine level	>99.9%	>99.9%	>99.9%	>99.9%

ABOUT THE AUTHOR

Jennifer Kornosky Straughen received a Bachelor's Degree in Biochemistry from the Lyman Briggs School at Michigan State University in 2002 and then began working as a researcher at Los Alamos National Laboratory in Los Alamos, New Mexico. She then began pursuing a Master's degree and was awarded an M.S.P.H. in Epidemiology from the Texas A&M School of Rural Public Health in 2005. She continued her educational pursuits and entered the doctoral program at the University of South Florida the same year.

While in the Ph.D. program at the University of South Florida, Jennifer gained experiencing as a teaching assistant and later as a research assistant for the Florida Birth Defects Surveillance program and her major professor, Dr. Hamisu Salihu. She has co-authored numerous papers, presented research at local and national conferences, was the recipient of several merit scholarships and awards which included the Maternal and Child Health Epidemiology Training Grant and the U.S. Public Health Traineeship, among others. In addition, she was awarded a CDC dissertation grant.