Associations of Maternal Exposure and Characteristics with Placental Transcription Regulation and Consequences on Fetal Growth

Michael R. Hussey

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Daniel A. Enquobahrie, Chair

Sheela Sathyanarayana

Fasil Tekola-Ayele

Program Authorized to Offer Degree:

Epidemiology

©Copyright 2023

Michael R. Hussey

University of Washington

Abstract

Associations of Maternal Exposure and Characteristics with Placental Transcription Regulation and Consequences on Fetal Growth

Michael R. Hussey

Chair of the Supervisory Committee:

Daniel A. Enquobahrie

Department of Epidemiology

Background: The placenta is a key organ in pregnancy and its dysfunction can lead to pregnancy complications as well as adverse maternal and fetal outcomes. While evidence for adverse consequences of maternal characteristics (such as obesity) or environmental exposures (e.g. particulate matter under 2.5 μm in diameter, $PM_{2.5}$) on the course and outcomes of pregnancy has been described, specific mechanisms are largely unknown. The transcriptome plays important physiological roles in growth, development, and metabolism, and may help elucidate these mechanistic links. This dissertation project investigated non-coding elements of the placental transcriptome and transcriptomic regulation via epigenetic mechanisms (long non-coding RNAs, lncRNAs and DNA methylation, DNAme) in relation to maternal pre-pregnancy obesity and prenatal $PM_{2.5}$ exposure.

Objectives: This dissertation project had three aims: 1) investigate associations of maternal $PM_{2.5}$ exposure with placental lncRNA expression and subsequent effects on birthweight, 2) investigate associations of pre-pregnancy BMI on placental lncRNA expression and subsequent effects on birthweight, and 3) investigate associations of pre-pregnancy obesity with differential methylation of genome-wide placental CpG sites and differential transcription of placental mRNA and lncRNA, as well as explore pre-transcriptional regulation of RNAs by DNAme in the placenta.

Methods: Aims 1 and 2 were addressed using data and placental samples from the CANDLE (N = 776) and GAPPS (N = 205) cohorts within the ECHO PATHWAYS Consortium. CANDLE, the Conditions Affecting Neurocognitive Development in Early Childhood cohort, is a pregnancy cohort based in Shelby County, TN. GAPPS, the Global Alliance to Prevent Prematurity and Stillbirth, is a pregnancy biorepository of participants recruited in Seattle and Yakima, WA. Aim 3 utilized data and placental samples from the Fetal Growth Studies (FGS) cohort (N = 301), a pregnancy study on women recruited from twelve medical centers across the United States. Maternal prenatal exposure to $PM_{2.5}$ was estimated via residential address history using spatio-temporal air pollution models developed at the University of Washington, while maternal pre-pregnancy BMI (ppBMI) was calculated from self-reported height and pre-pregnancy weight at initial study visits. Placental transcriptome-wide RNA and epigenome-wide DNAme were assayed using Illumina HiSeq sequencers and Illumina Infinium Human Methylation 450 Beadchip, respectively. To address these study aims, linear regression models were used to examine associations of pre-pregnancy $PM_{2.5}$ exposure and maternal obesity with placental lncRNA (CANDLE, GAPPS, FGS), DNAme (FGS), and gene expression (FGS).

Contextually appropriate covariate and precision variables determined *a priori* were included in the models, and effect modification by infant-sex was examined using both sex-stratified analyses and additional inclusion of exposure–infant-sex interaction terms. All analyses used a false discovery rate calculation to control for multiple testing.

Results: In CANDLE, first-trimester $PM_{2.5}$ was positively associated expression of the lncRNA *LINC00702* (logFC=0.103, FDR p-value=0.027) and second-trimester $PM_{2.5}$ was negatively associated expression the lncRNA *AC105345.1* (logFC=-0.163, FDR p-value=0.028). In CANDLE, interaction of $PM_{2.5}$ with infant-sex was observed for the lncRNA, *LINC00339* (interaction FDR<0.001). Additionally, in CANDLE, 47 transcripts were significantly associated with second-trimester $PM_{2.5}$ among female infants. Further two transcripts, *LINC00467* and *AC023157.3*, were expressed at higher and lower levels, respectively, among male infants in GAPPS. In CANDLE, ppBMI was associated with increased expression of three lncRNAs (*ERVH48-1*, *AC139099.1*, *CEBPA-DT*) among males, and pre-pregnancy obesity (ppOB) was associated with decreased expression of one lncRNA, *ZNF225-AS1*, among females. In GAPPS, ppBMI was associated with decreased expression of two lncRNA transcripts (*AP000879.1* and *AL365203.2*) among male infants. In FGS, methylation at five CpG sites, cg11844079[*PIGM*], cg22591875[*MEDAG*], cg27278787[*DLGAP1*], and cg04069951[*CD81*], was associated with ppOB. Additionally, ppOB was associated with decreased methylation of cg09858237[*WDR16*; *STX8*], cg15606914[*H1FOO*], and cg00311799[*CTSF*] among male infants. We also observed significant interaction of ppOB and infant-sex on methylation of cg08956510[*NME1*]. Finally, ppOB was associated with increased expression of the *ACE2* gene (logFC=0.922, FDR=0.033)

and decreased expression of the lncRNA *RP4-55D20.2* (logFC=-1.699, FDR<0.001) among female infants.

Conclusions: Overall, in the current study, we found potential evidence supporting associations of maternal $PM_{2.5}$ exposures during pregnancy and pre-pregnancy BMI with alterations in the placental transcriptome and epigenome. Further, evidence indicated that some of these associations, including lncRNAs and DNAme related to cell proliferation, growth, and transcription control, were infant-sex dependent, suggesting mechanisms that may also influence the known differences in placental morphology and development between males and females. These results need to be replicated in other populations with similar characteristics and exposure distributions to further explore their implications. In addition, functional studies are needed to characterize downstream effects of reported transcriptome and epigenome findings. Our gained knowledge can add mechanistic insight that facilitates efforts to promote optimal health for mother and child over the course of pregnancy.

ACKNOWLEDGEMENTS

I would like to thank my committee members, Daniel Enquobahrie, Sheela Sathyanarayana, Fasil Tekola-Ayele, Kathleen Kerr and Lloyd Mancl for their critical support and guidance throughout the various stages of this dissertation project. In particular, I am exceedingly grateful for the work of my committee chair and advisor, Daniel, who has been present throughout my time at the University of Washington and a never-ending source of knowledge and well-timed encouragement to spur me on throughout this process. I cannot thank you all enough.

I would also like to thank all the PIs and researchers I have brushed shoulders with in the last three years as a part of the ECHO PATHWAYS group, particularly the placental workgroup. You all have been instrumental in shaping my experience throughout my degree, and I have learned so much from you. Also, thank you as well to the participants of and all those who have contributed to the CANDLE, GAPPS, FGS and Omega cohorts. It has been a privilege to work with these populations and everyone who has worked so hard to bring them together.

Thank you as well to Lianne Sheppard and all those that manage the BEBTEH Training Grant, which provided a great deal of my funding.

Finally, thank you to all my dear friends and family who have stood by me throughout the last three and a half years of struggles, triumphs, failures and milestones. You all have been a constant reminder of God's provision amidst some of the most challenging times in recent memory, and I would not have reached the end without your support.

TABLE OF CONTENTS

LIST OF TABLES vii

CHAPTER 1: Introduction 1

CHAPTER 2: Associations of placental long non-coding RNAs with prenatal $PM_{2.5}$ exposure and infant birthweight 8

- **2.1 Introduction** 10
- **2.2 Methods** 11
- **2.3 Results** 17
- **2.4 Discussion** 19
- **2.5 Tables** 26

CHAPTER 3: Associations of placental long non-coding RNAs with maternal pre-pregnancy obesity and infant birthweight 36

- **3.1 Introduction** 38
- **3.2 Methods** 39
- **3.3 Results** 45
- **3.4 Discussion** 46
- **3.5 Tables** 55

CHAPTER 4: Associations of maternal pre-pregnancy obesity with placental DNA methylation and the placental transcriptome 61

- **4.1 Introduction** 63
- **4.2 Methods** 64
- **4.3 Results** 70
- **4.4 Discussion** 72
- **4.5 Tables** 81

CHAPTER 5: Discussion 88

REFERENCES 94

SUPPLEMENT 114

LIST OF TABLES

CHAPTER 2: Associations of placental long non-coding RNAs with prenatal $PM_{2.5}$ exposure and birthweight

Table 2-1: Select Characteristics of the Study Populations ____ *26*
Table 2-2: Maternal $PM_{2.5}$ Exposure Distributions ____ *28*
Table 2-3: Significant LncRNA Transcripts associated with $PM_{2.5}$ Exposure ____ *29*
Table 2-4: $PM_{2.5}$–Infant-Sex Interactions on LncRNA transcripts ____ *30*
Table 2-5: Infant-sex Stratified $PM_{2.5}$-LncRNA Associations ____ *31*
Table 2-6: LncRNA Transcripts Associated with Birthweight in Infant-sex Stratified Models ____ *35*

CHAPTER 3: Associations of placental long non-coding RNAs with maternal pre-pregnancy obesity and infant birthweight

Table 3-1: Select Characteristics of the Study Populations ____ *55*
Table 3-2: Maternal Pre-pregnancy Body Mass Index (BMI) Distributions ____ *57*
Table 3-3: Top Sex-specific Associations of Maternal Pre-pregnancy BMI with LncRNA transcripts *58*
Table 3-4: Top Sex-specific Associations of Birthweight with LncRNA Transcripts ____ *59*

CHAPTER 4: Associations of maternal pre-pregnancy obesity with placental DNA methylation and the placental transcriptome

Table 4-1: Select Characteristics of the Study Population ____ *81*
Table 4-2: Maternal Pre-pregnancy Body Mass Index (BMI) Distributions ____ *82*
Table 4-3: Placental CpG Sites and RNA Transcripts Associated with Maternal Pre-pregnancy Obesity ____ *83*
Table 4-4: CpG Sites Differentially Methylated in relation to Maternal Pre-pregnancy BMI ____ *84*
Table 4-5: Infant-sex Specific Differentially Methylated CpG Sites Associated with Maternal Pre-pregnancy BMI ____ *85*

SUPPLEMENT

Supplemental Table 1: Comparison of Analytic Sample Characteristics with Respective Cohort Populations ____ *114*
Supplemental Table 2: Select Infant-sex Stratified Associations of $PM_{2.5}$ Exposure with LncRNAs (Infant-sex Interaction Transcripts) ____ *116*
Supplemental Table 3: Select Infant-sex Stratified Associations of Pre-pregnancy BMI with CpG methylation (Infant-sex Interaction CpGs) ____ *117*
Supplemental Table 4: Top mRNA transcripts proximal (within 200 kilobases) to BMI-associated CpG sites ____ *118*

CHAPTER 1: Introduction

The Perinatal Period and Placenta

The prenatal period represents one of the most vulnerable times in human development, and disturbances to its delicate processes can have lasting consequences on both neonatal and lifelong health[1]. Therefore, etiologic and mechanistic investigations of associations of maternal health and exposures during pregnancy and impaired fetal development during pregnancy, characterized by birthweight – an important marker of fetal growth and development – offer excellent insights on disease prevention measures over the life course.

The placenta is an active organ with critical roles during pregnancy that shape the intrauterine environment, mediating effects of maternal exposures and characteristics on fetal growth and development. Its multiple roles in pregnancy include hormone metabolism, immune functions, protection from toxicants, transfer of nutrients and oxygen to and removal of waste from the fetus[2,3]. Therefore, placental growth, development, and function have the potential to illuminate underlying pathomechanisms. For instance, external stress placed upon the intrauterine environment, whether due to prior maternal health or exogenous exposures, has the potential to alter any number of placental processes, triggering inflammation, altered placental growth, insufficient vascularization, and changes to nutrient transport[4–6]. The placenta acts as a partial barrier to some toxicants (e.g. air pollutants, see below), and these compounds can remain in the placenta and exert their toxic effects locally with similar downstream consequences[7,8]. Given well-established differences in male and female placental size, growth rates, and sex-hormone specific functions, infant-sex must also be considered as a potential modifier of placenta-mediated health outcomes subsequent to maternal exposures and characterstics[9,10].

Air Pollution and Prenatal Complications

Elevated exposure to airborne ambient particulate matter (PM) – a common metric for overall air quality – is ubiquitous in many developed and developing nations. Defined based on its size, particles below 2.5 μm in diameter ($PM_{2.5}$) can be composed of numerous types of compounds depending on their local natural or anthropogenic sources, including black carbon, elemental carbon, inorganic carbon, sulfates or metals[11]. Exposure to $PM_{2.5}$ is associated with a wide variety of detrimental health effects among the general population, including impaired lung function, increased cardiovascular morbidity and mortality, type 2 diabetes, and childhood lung infection[12]. During pregnancy and the perinatal period, higher exposure to air pollutants (including $PM_{2.5}$) has been related to elevated rates of both maternal and infant pregnancy complications, including gestational diabetes, preeclampsia, miscarriage, spontaneous abortion, preterm birth, and fetal growth restriction (FGR, characterized by low birthweight, LBW)[13–19]. The consequences of $PM_{2.5}$ may be due to its effect on maternal health, the placenta (where it might accumulate[7]), or direct effect on the fetus (by passing through the placenta). A recent systematic review of *in-vitro* and *in-vivo* animal and human evidence found that particles around a micron (1,000 nm) in diameter or smaller could consistently reach and enter the placenta, while particles less than 500 nm are potentially able to cross it into fetal circulation[7]. Despite significant evidence supporting the multiple adverse consequences of $PM_{2.5}$ on maternal and fetal pregnancy outcomes, underlying mechanisms, particularly placenta-mediated genetic and epigenetic mechanisms, are largely unknown. In particular, there is a broad lack of investigations on roles of non-protein-coding placental RNA transcripts in mediating the

downstream consequences of $PM_{2.5}$ exposure on fetal growth and development, including whether these effects and mechanisms are infant-sex specific.

Maternal Pre-pregnancy Obesity and Adverse Pregnancy Complications

The prevalence of mothers who are overweight ($30 > BMI \geq 25$ kg/m^2) or obese (BMI $\geq$ 30 kg/m^2) prior to pregnancy has risen sharply over the course of the last few decades. In 2019, CDC estimates of pre-pregnancy obesity (ppOB) prevalence in the United States were around 29%, a 3% increase from an estimates of 26% in 2016[20]. In addition, around two thirds of mothers in the US are estimated to be overweight, highlighting the relevance of this adverse risk factor[21]. Maternal ppOB has been linked to higher rates of pregnancy complications such as gestational diabetes, pre-eclampsia, gestational hypertension, and preterm delivery[22,23]. It has also been linked to both macrosomia (high birthweight) and, in rarer cases, low birthweight (LBW), which have been related to life-course chronic diseases[24,25]. Beyond the observed impacts of maternal obesity on hormone (e.g. adipokine) production, hormone metabolism (e.g. insulin)[26], and inflammatory responses during pregnancy[27], ppOB may limit nutrient transfer to support fetal growth and development[27,28]. However, underlying mechanisms, particularly mechanisms that involve the transcriptome and epigenetic regulation of the placenta, are poorly described.

Placental Transcriptome Regulation Epigenetics

The transcriptome comprises the entirety of RNA transcripts, coding and non-coding, produced during transcription of cellular DNA sequences. Unlike the genome, which remains fixed across the life-course, the transcriptome can vary widely depending on environment. Recent advancements in sequencing technology have provided excellent opportunities for

investigations of the transcriptome and transcriptome regulation by epigenetic biomarkers. Importantly, the transcriptome and its regulation by epigenetic biomarkers provide excellent opportunities to assess gene-environment interactions by reflecting ongoing cellular responses to changing environmental conditions. Relatively few transcriptome-wide association studies (TWAS) have been performed relating the placental transcriptome to $PM_{2.5}$ exposure[17,29–33] or ppOB[28,34–39], and essentially none have examined placental non-coding RNA transcripts in relation to these exposures. An *in vitro* sequencing study that exposed placental JEG-3 cells to $PM_{2.5}$ [31] found that genes broadly related to immune response, apoptosis regulation, steroid hormone biosynthesis, and multiple types of signaling pathways were expressed in these cell lines. In a study in the PATHWAYS research consortium that examined $PM_{2.5}$ and gene expression associations, our group observed that first- and second-trimester $PM_{2.5}$ exposures were associated with expression of several placental genes related to mRNA splicing, molecular transport, GABAergic signaling and mitochondrial function[33]. Cox et al., in an investigation of maternal BMI and the placental transcriptome, found that genes for immune response, myeloid differentiation, organ/tissue development, blood vessel morphogenesis and extracellular matrix structure were significantly associated with pre-pregnancy BMI (ppBMI) and birthweight[39]. Relevant to this dissertation project, even fewer studies examined epigenetic regulation of the transcriptome in the context of perinatal exposure to $PM_{2.5}$ or ppBMI/ppOB[40–43].

Long non-coding RNAs (lncRNAs) are a relatively newly discovered class of RNA transcripts greater than 200 nucleotides in length that have no protein-coding functionality when transcribed. LncRNAs have now been characterized to have highly tissue-specific regulatory roles in a number of different processes, including pre-transcriptional and post-transcriptional regulation of mRNA[44,45]. Many previous analyses have examined LncRNAs in the placental,

though most studies focused on transcript candidates in the context of placental pathologies (e.g. preeclampsia[46–48]) or specific placental cell types[49–51]. Further, studies assessing more general associations of placental lncRNAs with maternal health conditions or exogenous exposures have broadly been limited in their transcriptional scope rather than conducting transcriptome-wide association studies (TWAS)[52,53]. One rare example of placental TWAS investigation was conducted by Long et al. (2016) that compared placental biopsies among women who had been diagnosed with early-onset preeclampsia (EOPE) compared to controls and found differences in levels of lncRNA transcripts involved in gene ontology groups for cell migration, cell adhesion, stimulus response and focal adhesion – among others[54]. Another such study examined the relationship between maternal phthalate exposures in pregnancy – as measured by urinary metabolites – and placental lncRNAs, finding that several transcripts, including the prominent *NEAT1*, were differentially for multiple phthalates measured during different trimesters of pregnancy[55]. Only two published studies have examined transcriptome-wide lncRNAs in relation to birthweight, one assessing placental cadmium-lncRNA-birthweight associations[56] and another comparing placentas of macrosomic and normal weight infants of non-diabetic mothers[57]. To our knowledge, no published study has examined the role of lncRNAs in placental response to $PM_{2.5}$ or ppOB.

Another epigenetic mechanism that has been previously investigated is DNA methylation (DNAme), the addition of a methyl group to the C5 position of cytosine nucleotides in DNA – a modification that can either up- or downregulate RNA transcripts[58]. Janssen et al. found that $PM_{2.5}$ exposure, particularly during the first trimester of pregnancy, was associated with decreased global placental DNAme[59]. Saenen et al. found hypomethylation of the *LEP* gene promoter related to $PM_{2.5}$ exposure during the second-trimester[60]. These studies, despite having

relatively large sample sizes, were limited in scope by examining either global methylation or only specific genomic regions. For ppOB, a prior study within the Fetal Growth Studies cohort investigated global and site-specific DNAme, and found that pre-pregnancy BMI (ppBMI) was associated with significant changes in DNAme at CpG sites for three genes (increases in methylation for cg14568196[related to *EGFL7* expression] and cg15339142[*VETZ* expression], and decreases in methylation at cg02301019[*AC092377.1*])[61]. While a few other studies have examined $PM_{2.5}$ or birthweight in relation to placental gene expression, DNAme, or both, these studies have been small in sample size or limited in sequencing scope[62–64].

Overview of the Dissertation

This dissertation project addresses gaps in knowledge relating perinatal maternal $PM_{2.5}$ exposure and pre-pregnancy BMI to the placental transcriptome (mRNA) and epigenome (DNAme and lncRNAs) by leveraging well-characterized cohorts, in-depth sequencing efforts, and state-of-the-art statistical approaches. The study was conducted using study populations and data collected as part of three separate cohorts. The Conditions Affecting Neurocognitive Development and Learning in Early Childhood (CANDLE) and Global Alliance to Prevent Prematurity and Stillbirth (GAPPS) studies are birth cohorts – based in Memphis, TN and Seattle/Yakima, WA, respectively – with high resolution estimates of residential particulate matter air pollution exposure, characterization of pre-pregnancy BMI and full placental RNA sequencing at birth[65]. Fetal Growth Studies (FGS) is an NICHD birth cohort that included participants across cities in the US to investigate standardization of fetal growth metrics, and features maternal BMI characterization, full RNA sequencing, genomic sequencing and DNAme measurement[66]. The next three chapters (2-4) of this report summarize the investigations that

were conducted to address each of the three aims of this project. To address Aim 1 (Chapter 2), we sought to determine the associations (overall and infant-sex specific) of maternal $PM_{2.5}$ exposure with placental lncRNA and the relationships of $PM_{2.5}$-associated lncRNA levels with birthweight in the CANDLE and GAPPS cohorts. To address Aim 2 (Chapter 3), we examined the associations (overall and infant-sex specific) of pre-pregnancy BMI phenotypes with placental lncRNA in the CANDLE and GAPPS cohorts, and similarly examined the role of lncRNAs in mediating associations of ppOB with birthweight. Finally, to address Aim 3 (Chapter 4), we examined associations (overall and infant-sex specific) of ppOB or pre-pregnancy overweight status with placental epigenome-wide DNAme and transcriptome-wide mRNA/lncRNA levels. In addition to examining transcriptomic regulation, we explored pre-transcriptional regulation of RNA transcripts whose genomic locations were proximal to differentially methylated CpG sites. Finally, Chapter 5 describes an overall summary of study findings, study strengths and limitations, and future research recommendations.

CHAPTER 2: Associations of placental long non-coding RNAs with prenatal $PM_{2.5}$ exposure and infant birthweight

Abstract

Background: Long non-coding RNAs (lncRNA) play essential roles in regulation of cell processes in multiple human tissues. However, few studies have examined placental lncRNA expression in relation to environmental exposures and birth outcomes. We addressed this gap by investigating associations (overall and infant-sex specific) of genome-wide placental expression of lncRNA transcripts with prenatal maternal $PM_{2.5}$ exposure and infant birthweight in two birth cohorts.

Methods: Placental samples were collected from mother-child dyads recruited to the CANDLE (Memphis, TN) (N = 763) and GAPPS (Seattle and Yakima, WA) (N = 142) cohorts. Average exposures to $PM_{2.5}$ were estimated with spatiotemporal models for six timespans: full-pregnancy, each trimester, and the first and last months of pregnancy. RNA was extracted from fetal-side and full-thickness placental samples in CANDLE and GAPPS, respectively, and sequenced using an Illumina HiSeq 4000. Cohort- and lncRNA-specific weighted linear models were fit for lncRNA transcripts (1,077 and 1,033 in CANDLE and GAPPS, respectively) and $PM_{2.5}$ exposure windows or birthweight, adjusting for potential confounders and experimental variables. Effect modification of associations by infant-sex was examined using stratified models and models containing relevant interaction terms. A false discovery rate (FDR<0.10) was used in all models to control for multiple testing.

Results: In CANDLE and GAPPS, mean maternal age was 27 and 31 years, respectively, and full pregnancy mean $PM_{2.5}$ was 10.5 and 6.2 μg/m^3, in CANDLE and GAPPS, respectively. In CANDLE, first-trimester $PM_{2.5}$ was associated with *LINC00702* expression (logFC = 0.103, FDR p-value = 0.027), while second-trimester $PM_{2.5}$ was associated with *AC105345* expression (logFC = -0.163, FDR p-value = 0.028). We did not find main-effect associations in the GAPPS cohort or main-effect associations of birthweight with lncRNA in either cohort. Full-pregnancy $PM_{2.5}$ interacted with infant-sex on expression of *LINC00339* in CANDLE (interaction FDR<0.001), though stratified associations for *LINC00339* were not significant for either sex. In CANDLE, second-trimester $PM_{2.5}$ was associated with 47 transcripts among female infants and first-month $PM_{2.5}$ was associated with two transcripts, *LINC00656* and *LINC00342*, among males. Late-pregnancy $PM_{2.5}$ exposures were associated with *LINC00467* and *AC023157.3* among male infants in GAPPS.

Conclusions: We observed cohort-specific associations of $PM_{2.5}$ exposure with several placental lncRNA transcripts. We also observed potential cohort- and infant-sex specific $PM_{2.5}$-lncRNA associations for multiple lncRNA transcripts. Identified lncRNA transcripts have functions related to cell proliferation, growth control and stress response. Findings need further validation to foster additional mechanistic insights on the consequences of particulate exposure.

2.1 Introduction

The placenta plays an essential role in pregnancy given its array of functions supporting the developing fetus, including nutrient transport, gas exchange, and waste expulsion[67]. Therefore, exogenous maternal exposures that alter placental development and function have both maternal and fetal consequences. Exposure to $PM_{2.5}$ (particulate matter under 2.5 μm in diameter) is common worldwide and is associated with pregnancy complications and poor neonatal outcomes (such as low birthweight)[14–16,18,19,68], however, placental mechanisms for these outcomes have not been fully characterized.

Placental inflammation[69,70], abnormal vascularization (related to angiogenic proteins)[71–73], as well as alterations in trophoblast migration/invasion[74,75], MAPK-signaling, progesterone production, *COX-2* expression, amino acid transport, and cellular respiration are potential consequences of $PM_{2.5}$ exposure observed *in vitro* and *in vivo* investigations, including in studies we reported before[17,31–33]. Regulatory mechanisms of these underlying pathologic processes are poorly understood.

Long non-coding RNA transcripts (lncRNA) have been shown to play prominent epigenetic regulatory roles in disease pathogenesis, including those involving placental function[49,54,76–82]. LncRNAs are >200 base pairs in length untranslated tissue-specific RNAs that carry out their regulatory roles at all stages of transcription or translation through physical interactions with DNA, other RNAs and proteins[45]. To our knowledge, no prior study has investigated human placental lncRNA transcription in relation to maternal $PM_{2.5}$ exposure across pregnancy, while only one investigated placental lncRNA in relation to infant birthweight[83]. Therefore, we investigated overall and infant-sex specific associations of prenatal exposure to

$PM_{2.5}$ during various time windows in pregnancy with genome-wide placental levels of lncRNAs in two birth cohorts. We also examined overall and infant-sex specific associations of placental lncRNA expression with birthweight.

2.2 Methods

Study Setting and Study Participants

This study utilized two populations involved in the ECHO-PATHWAYS Consortium: the Conditions Affecting Neurocognitive Development in Early Childhood (CANDLE) cohort[84] and the Global Alliance to Prevent Prematurity and Stillbirth (GAPPS) study. In-depth information on these study populations and data collected is described elsewhere[65,84]. Briefly, CANDLE is a prospective pregnancy cohort based in Shelby County, TN that recruited pregnant women – largely from the metro Memphis, TN – for longitudinal monitoring of maternal and infant health during pregnancy, at birth and, for infants, in the years following birth. To be eligible, pregnant women between 16 and 40 years of age needed to be residents of Shelby County, TN and plan to deliver at a participating study hospital, and have a singleton pregnancy without complications at enrollment. A total of N = 1,503 women and their infants were, upon informed consent, recruited into the population between 2006 and 2011. All study protocols and data collection efforts were sanctioned by the University of Tennessee Health Sciences Center IRB.

GAPPS started in 2007 at Seattle Children's Hospital as a pregnancy biorepository established to study adverse birth outcomes and highlight targets for intervention[65]. Women at regional WA state hospitals provided data and various tissue samples during pregnancy to the

biorepository, and were eligible for re-recruitment as part of the ECHO-PATHWAYS study. Additional re-recruitment criteria included providing a placenta sample and at least one urine sample, having delivered at either the Seattle, WA or Yakima, WA hospitals, having a GAPPS child currently 4-7 years of age, and completion of the original enrollment questionnaire. While this re-recruitment process remains in progress, at the time of this study N = 1,271 were potentially eligible and N = 392 women enrolled in the study. GAPPS protocols were approved by IRBs at both Seattle Children's Hospital and the University of Washington.

Given our specific interest in placental transcriptomics, this investigation centered on participants from these cohorts who provided placental samples at birth, N = 794 and N = 392 for CANDLE and GAPPS, respectively, prior to additional exclusion criteria. Mothers whose pregnancies resulted in stillbirth, who encountered major complications during pregnancy (such as fetal chromosomal abnormalities, placental infarction, chorioamnionitis, oligohydramnios, placental abruption, or placenta previa) or who did not give a residential address were excluded from analysis. After these exclusions, the resulting transcriptomic sub-cohorts contained 776 women from CANDLE and 205 women from GAPPS.

Measurement of Particulate Matter Air Pollution under 2.5 µm ($PM_{2.5}$)

$PM_{2.5}$ was estimated using spatiotemporal models formulated at the University of Washington. As described and utilized previously[33,85,86], these models produce biweekly $PM_{2.5}$ average exposure levels for a given maternal residential address using a combination of cohort specific monitors and regulatory air-quality data; this estimation is done in a way that decomposes the space-time field of PM concentrations to account for pure spatial variation and combined spatial and seasonal variation in long-term trends. Averages were calculated from bi-

weekly estimates produced by these models for six periods during pregnancy: each trimester of pregnancy (first, second, and third trimesters), the initial month of pregnancy (the time spanning implantation and placental formation), the final month of pregnancy, and the full pregnancy duration.

Placental Sample Collection

As described previously[33,55], placental samples in CANDLE included one 2 cm x 0.5 cm x 0.5 cm rectangular piece of villous tissue cut from the fetal placental parenchyma. These samples were taken no more than 15 minutes after delivery, sliced into four cubes with an edge length of 0.5 cm, and each placed in 50 ml tubes with 20 ml of RNAlater buffer and stored at 4°C for between 8 and 24 hours. Cubes were then placed in separate 1.8 ml cryovials with new RNAlater and stored at -80˚C. Manual dissection was then performed to separate residual maternal decidua from fetal villous tissue, the latter retained and stored at -80°C in fresh RNAlater.

Placental samples for the GAPPS biorepository were collected within 30 minutes of delivery. Vertical tissue punches (8 mm full-thickness) vertical tissue punches from the placental disc were collected and immediately placed in5 ml tubes containing 3 ml of RNAlater and stored at -20°C. The tissue specimens were shipped to the GAPPS facility and stored at −80 °C until the fetal villous tissue was manually dissected and cleared of maternal decidua using standard protocols developed by the GAPPS placental biorepository. Samples were then divided into 1-3 pieces with 10 mg-30 mg in mass. Each sample was immersed in 1 ml RNAlater and stored at -20°C until RNA was isolated.

Placental Sample Processing and RNA Sequencing

Approximately 30 mg of placental tissue was used for RNA isolation. The tissue was homogenized in tubes with 600 µl Buffer RLT Plus with mercaptoethanol using a TissueLyser LT instrument (Qiagen, Germantown, MD). RNA was then isolated using Qiagen's AllPrep DNA/RNA/miRNA Universal Kits (Qiagen, Germantown, MD) according to the manufacturer's recommended protocol. RNA purity was assessed by measuring $OD_{260/230}$ and $OD_{260/260}$ ratios using a NanoDrop 8000 Spectrophotometer (Thermo Fischer Scientific, Waltham MA). RNA integrity was determined with RNA 6000 Nanochips (Agilent, Santa Clara, CA) on a Bioanalyzer 2100. Only samples with an RNA integrity number (RIN) >7 were used for RNAseq analysis.

All RNA sequencing was conducted at the University of Washington Northwest Genomics Center1 μg of total RNA was poly-A selected and cDNA libraries were prepared using the TruSeq Stranded mRNA kit (Illumina, San Diego, CA). Each library was uniquely barcoded and subsequently amplified with 13 cycles of PCR and quantified using the Qubit Quant-it dsDNA High Sensitivity fluorometric assay (Life Technologies, Carlsbad, CA). Average fragment size and overall quality were assessed with a DNA1000 assay on an Agilent 2100 Bioanalyzer. An Illumina HiSeq 4000 sequencer was used to sequence all libraries to a depth of approximately 30 million reads.

The quality of sequencing reads was assessed with the FASTX-toolkit (v0.0.13) and FastQC (v0.11.2)[87]. We estimated transcript abundances by aligning reads with the GRCh38 transcriptome (Gencode v33) in Kallisto[88]. The Bioconductor *tximport* package in R was then used to summarize these values at the RNA transcript level and scale them to the average transcript length[89]. The Bioconductor *edgeR* package's Trimmed Mean of M-values (TMM)

function[90] was then applied to compute library normalization factors before filtering to only LncRNA transcripts. Based on visual assessment of count distributions, transcripts with unreliably low expression (in this case mean log counts/million (logCPM) ≤ 0) were excluded, reducing our transcripts from ~16,000 to 1,077 and 1,033 for analysis in CANDLE and GAPPS, respectively. All modeling was conducted using the Bioconductor *limma* package in R and its limma-voom pipeline[91], which converts the counts per transcript to logCPM via TMM normalization factors and estimates observation-level weights for inclusion in weighted linear regression from the mean-variance relationship of the transcripts being analyzed.

Statistical Analyses

The transcriptome sub-populations for CANDLE and GAPPS were first compared against one another and against their respective overall cohorts using descriptive statistics. This included t-tests for continuous covariates, as well as X^2 tests for categorical variables. Following these initial comparisons, weighted linear regression models were constructed using the Bioconductor *limma* package. First round models included transcript-specific lncRNA expression (in logCPM) as the dependent variable and individual $PM_{2.5}$ exposure parameters (the six estimates outlined above) as the independent variable, using the limma-voom pipeline[91] which utilizes observation-level weights to control for heteroscedasticity of the logCPM values. Further, limma also computes empirical Bayes adjusted t-statistics to assess significance of changes in expression of a specific transcript using a variance estimate prior based on all transcripts[89].

All models were cohort-specific (CANDLE or GAPPS) due to differences in demographics and $PM_{2.5}$ exposures between the cohorts. All models also adjusted for additional

demographic covariates and precision variables collected from enrollment questionnaires, abstraction from medical records, or as elements of the experiment (e.g. sequencing batch). These *a priori*–selected covariates included continuous maternal age, self-reported race (White, Black, Asian, other or multiple race), infant-sex, calendar year of birth, continuous gestational age (in weeks), family income (continuous), mode of delivery (vaginal or cesarean section), labor type (labor present or absent), self-reported smoking during pregnancy (yes or no), maternal urine cotinine (for CANDLE only, binary variable using 200ng/mL cut off [92]), and maternal education (less than high school, high school, college or technical school, and graduate or professional degree), RNA sequencing batch, and hospital location (for GAPPS only, adjusting for differences between the Seattle and Yakima, WA locations). Mothers who reported American Indian/Alaska Native for self-reported race were included as a part of "other" in GAPPS due to low sample size. Next, to examine exposure–infant-sex interactions, we fit similarly-constructed models with an additional exposure–infant-sex interaction term, followed by infant-sex stratified analyses that excluded infant-sex as a covariate. In all model types, surrogate variable analysis (SVA)[93], a technique controlling for cell heterogeneity and unmeasured sources of variation in high-throughput analyses, was applied to any models whose output p-value distributions for the coefficient of interest deviated from the expected uniform distribution. Models using SVA were then rerun, all prior to examining results. A false discovery rate (FDR) < 0.10 [94] was used in all models to identify statistically significant finding while taking multiple testing into account.

We then constructed models that assessed the relationship between lncRNA expression and birthweight. Models were specified with birthweight as the independent variable and transcript expression as the dependent variable, allowing for the continued use of the limma-

voom pipeline. Models included a similar set of covariates to those described above. Further, exposure–infant-sex interaction was examined using interaction terms and sex-stratified models as described above. Finally, in the case that individual transcripts were associated with both $PM_{2.5}$ and birthweight, we planned to conduct mediation analysis using the *mediation* R package. We conducted complete-case analyses when fitting all models, with final sample sizes of N = 763 in CANDLE and between N = 138 and N = 142 in GAPPS for various models due to varied levels of missing covariate and exposure information. All analyses were completed in R (v 4.1.2).

2.3 Results

Mothers in CANDLE and GAPPS were on average 27.3 (SD = 5.5) and 30.5 (5.7) years old, respectively (**Table 2-1**). The majority of mothers recruited to CANDLE were Black (56%) while in GAPPS, 80% of mothers were White. The transcriptome sub-cohorts had comparable characteristics, generally, to their respective parent cohorts (**Supplemental Table 1**).

Exposure to $PM_{2.5}$ was significantly different between CANDLE and GAPPS for all exposure windows. Broadly, average exposures were much higher in CANDLE, ranging between 10.5 and 10.7 $\mu g/m^3$, than GAPPS, where average exposures ranged between 6.0 and 6.3 $\mu g/m^3$ (**Table 2-2**). However, larger exposure variability was observed in GAPPS (SDs between 1.8 and 3.7 $\mu g/m^3$) compared with CANDLE (SDs between 0.9 and 2.1 $\mu g/m^3$).

We found two lncRNA transcripts differentially expressed in CANDLE in relation to $PM_{2.5}$ exposure: *LINC00702* with first trimester $PM_{2.5}$ exposure (logFC = 0.109, FDR = 0.027)

and *AC105345.1* with $PM_{2.5}$ exposure in the second trimester (logFC = 0.159, FDR = 0.028) (**Table 2-3**). No overall $PM_{2.5}$ and lncRNA association was observed in GAPPS and no overall association of $PM_{2.5}$ with birthweight was observed in either cohort.

In CANDLE infant-sex and $PM_{2.5}$ interaction models, the interaction term for infant-sex and full-pregnancy $PM_{2.5}$ exposure was statistically significant for the *LINC00339* transcript (interaction FDR < 0.001) (**Table 2-4**). In infant-sex stratified models, full-pregnancy $PM_{2.5}$ exposure and *LINC00339* transcript associations were positive among both male and female infants (stronger among female infants, logFCs = 0.003 and 0.041 in males and females, respectively), though neither association was statistically significant (**Supplemental Table 2**).

In infant-sex stratified models in CANDLE, we found that $PM_{2.5}$ exposure in the second trimester was associated with expression of 45 lncRNA transcripts among female infants (**Table 2-5**). This list included well known lncRNAs such as X inactive specific transcript (*XIST*), nuclear paraspeckle assembly transcript 1 (*NEAT1*) and metastasis associated lung adenocarcinoma transcript 1 (*MALAT1*), yet 69% were novel, and only half of that set had known anti-sense or sense intronic genes they were associated with. Additionally, first-month $PM_{2.5}$ was associated with two transcripts among males in CANDLE: *LINC00656* and *LINC00342* (logFC = 0.069 and 0.051, FDR = 0.048 and 0.075, respectively). Further, $PM_{2.5}$ exposures were associated with two transcripts among males in GAPPS, *LINC00467* with third-trimester $PM_{2.5}$ (logFC = 0.128, FDR = 0.034) and *AC023157.3* with last month $PM_{2.5}$ exposure (logFC = -0.179, FDR = 0.038) (**Table 2-5**).

While we did not find significant infant-sex and lncRNA transcript interactions on birthweight in CANDLE or GAPPS, in stratified models, among male infants in CANDLE, we found association of birthweight with expression of *AC104083.1* (logFC = 4.4×10^{-4}, FDR =

0.032) **(Table 2-6)**. This association was not observed among female infants in CANDLE (logFC = 2.2×10^{-4}, FDR = 0.407) or among male or female infants in GAPPS. Lacking any overlapping findings between models for $PM_{2.5}$ and birthweight, mediation analysis was not performed.

2.4 Discussion

In the current study, we observed few overall, cohort-specific associations of $PM_{2.5}$ with placental lncRNA transcripts. First- and second-trimester $PM_{2.5}$ exposures were associated with two transcripts, *LINC00702* and *AC105345.1*, respectively, in CANDLE. Infant-sex interacted with full-pregnancy $PM_{2.5}$ on *LINC00339*. Further, second-trimester $PM_{2.5}$ exposure was associated with numerous transcripts among female infants in CANDLE, including prominent transcripts known to be related to the placenta or pregnancy complications. Among male infants, first-month $PM_{2.5}$ was associated with two transcripts in CANDLE, *LINC00656* and *LINC00342*, and third-trimester and last-month $PM_{2.5}$ were associated with the transcripts *LINC00467* and *ACO23157.3*, respectively, in GAPPS. No main-effect lncRNA and infant birthweight associations were observed in CANDLE or GAPPS, while only one male-infant specific association was observed between *AC104083.1* and birthweight in CANDLE.

To our knowledge, this is the first transcriptome-wide study of placental lncRNAs in relation to maternal $PM_{2.5}$ in pregnancy, and only the third transcriptome-wide association study examining associations of lncRNAs with birthweight. Prior to this work, a number of studies have examined lncRNAs and $PM_{2.5}$ exposure as a part of *in vitro* investigations[95–99], however,

while some of these studies were transcriptome-wide[95,98], most investigated candidate lncRNAs[96,97,99], including *MALAT1*[95] and *NEAT1*[99]. Of the 45 transcripts associated with $PM_{2.5}$ exposure among female infants in CANDLE (**Table 2-5**), two transcripts with positive associations, *LINC00470* and *NEAT1*, were previously reported in placental transcriptome studies. Placental expression of *LINC00470*, which was increased by around 8.3% in relation to second-trimester $PM_{2.5}$ in our study, was associated with increased birthweight (ß = 230.8 g, 95% CI: 88.9, 372.7) in a study by Hussey and colleagues[83]. Similarly, *NEAT1*, upregulated by around 18.6% among female infants (in relation to second-trimester $PM_{2.5}$) in our study, was positively associated with urine phthalate urine metabolites in a study by Paquette et al. in CANDLE[55]. Another study comparing the placentas of normal and high birthweight infants found that 2,929 and 2,127 lncRNAs were up- or down-regulated, respectively, among macrosomic infants compared to controls[57]. However, only the top 10 lncRNA associations (which did not overlap with lncRNAs identified in our study) were presented, so we could not compare our findings.

Our primary models linked *LINC00702* and *AC105345.1* to maternal $PM_{2.5}$ exposures in early-to-mid pregnancy. While *AC105345.1* has not yet been characterized, *LINC00702* (long intergenic non-coding RNA 702) has previously been associated with a breadth of different cancers. Cancer may be considered to have some parallels with placental growth given the placenta's tumor-like proliferative character[100]. The transcript was noted by Chen and colleagues to be a heavily downregulated in cancerous endometrial tissue when compared to normal tissue (logFC = -2.953), a property that was specific to endometrial cancer as opposed to ovarian and cervical cancers[101]. *LINC00702's* cancer connections are not limited to female reproductive cancers, however. For example, *LINC00702* expression was noted to be

significantly lower in bladder tumor samples compared to normal controls, with markedly lower expression serving as an indicator of increased tumor grade and predictor of lower patient survival over observation[102]. The authors also observed that overexpression of *LINC00702* suppressed cancer cell proliferation *in vitro*, and linked part of its anti-cancer mechanism to upregulation of dual specificity phosphatase 1 (*DUSP1*), an inhibitor of the tumor inflammatory microenvironment[102]. A similar trend was observed in colorectal CaCo2 cell lines by Yu et al., who found that *LINC00702* expression was lower in these cells and inversely associated with proliferation, migration and metastasis[103]. The opposite effect, *LINC00702* promoting cell proliferation, has also been observed in models of ovarian cancer, indicating the roles of this transcript are potentially highly tissue-specific[104]. Additional studies indicate that some of these mechanisms may require interaction of *LINC00702* with various miRNAs to alternately promote or suppress cancer proliferation, progression, and invasion[105,106]. These studies suggest that the positive association we observed for $PM_{2.5}$ and *LINC00702* may be consistent with broad opposition to localized tissue inflammation during the first trimester, a period of rapid growth and trophoblast proliferation, but further placental studies are needed to characterize the impacts of *LINC00702* and *AC105345.1* in this context.

In contrast to the sparse overall findings, in stratified models, among female infants in CANDLE, we observed associations of $PM_{2.5}$ with 45 different lncRNAs; 31 of these transcripts were novel, with the remaining lncRNAs varying widely in their levels of characterization. Two transcripts were found in each cohort (*LINC00656*, *LINC00342* in CANDLE, *LINC00467*, *AC023157.3* in GAPPS) to be associated with $PM_{2.5}$ among male infants; of these transcripts, two were previously linked to cancer etiology[107–110]. In CANDLE, we found a significant infant-sex interaction with full-pregnancy $PM_{2.5}$ on *LINC00339*, an lncRNA explicitly linked to cell

proliferation, migration and invasion in cancer[111] and endometriosis, a condition defined by endometrial tissue growing outside the uterus[112–115]. *LINC00339* is selectively transcribed in endometrial lesions and associated with localized expression of genes related to immune defense and inflammation[115]. Female infants in CANDLE were observed to have stronger associations of *LINC00339* with $PM_{2.5}$ (logFC = 0.041) compared to male infants (logFC = 0.003), yet neither association was statistically significant (**Supplemental Table 2**). Our findings regarding lncRNA-birthweight associations were limited to one uncharacterized transcript, *AC104083.1*, which was found to be upregulated among male infants in GAPPS.

Using online iPathwayGuide software (Advaita Corporation, Ann Arbor, MI), we performed an exploratory pathway analysis to assess functions of the 45 lncRNAs that were associated with second-trimester $PM_{2.5}$ among female infants in CANDLE. Given 31 of the 45 transcripts available for analysis lacked any functional characterization, 14 transcripts had available annotations for gene-ontology comparisons. Ultimately, upregulation of two lncRNAs, *NEAT1* (18.6%) and *MALAT1* (13.8%) (**Table 2-5**), was significantly associated with five biological pathways with clear relevance to the placenta: cell population proliferation, regulation of cell population proliferation, positive regulation of cell population proliferation, response to stimulus, and response to stress (p = 0.027 for all five pathways, though FDR = 0.236 for all). *XIST* was also the sole-representative in the dosage compensation and dosage compensation by inactivation of X chromosome pathways, fitting its role as the primary transcript involved in condensing the extra X-chromosome into the Barr body in female cells[116]; however, both dosage compensation pathways were not statistically significant based on *XIST*'s transcription (p = 0.111). These pathways are broadly consistent with those highlighted in other studies of lncRNAs in other cell types, including elements of transcriptional control (e.g. regulation of

DNA-methyl-transferase 1)[96], G2/M cell cycle arrest, apoptosis induction (via ROS damage to the mitochondria)[97], and vascular toxicity[98] as a part of cellular responses to $PM_{2.5}$ exposure.

Numerous recent investigations have contributed strong evidence that lncRNAs play a role in normal[50,78,116] and pathologic placental growth, development and function[48,54,117–120], yet the impact of prenatal particulate matter exposure on placental lncRNA transcription is an area of research that is highly understudied. We believe this full-transcriptome studies to be the first of its kind (of any size) examining impacts of spatio-temporally estimated prenatal $PM_{2.5}$ exposure on humans placental lncRNAs and one of the largest of its kind to attempt to link these lncRNAs to birthweight. We anticipated that replication of findings between cohorts might allow for mediation analysis of transcripts, yet no overlapping associations were observed between birthweight and $PM_{2.5}$ exposures. Information in the current study was comprised of data collected through extensive surveys and medical record abstraction, allowing for more rigorous covariate control than many other investigations. Additionally, the CANDLE cohort, given 56.4% of its population was Black women and their infants, also represents the largest analysis to date of placental lncRNAs in this demographic group, which is broadly underrepresented in TWAS. Our large sample sizes provided enough power to examine sex-specific associations with lncRNAs, something not possible in many earlier analyses of this type; our findings in these sex-specific analyses were generally in line with potential impacts of $PM_{2.5}$ exposure and biologically plausible sex-specific differences in placental function[9,121,122] and other sex-specific disease etiologies known to involve lncRNAs[123,124]. While well-known and characterized transcripts should certainly be further examined in the context of our findings, not all of the uncharacterized transcripts differentially expressed in the current study are likely to be true associations.

Despite the numerous strengths of this investigation, several limitations must also be addressed. The possibility of replication motivated the use of two cohorts, yet the cohorts utilized in the current analyses differed fundamentally in several ways, including population demographics and exposures. For instance, mothers in GAPPS and CANDLE differed significantly across maternal age, family income, maternal race/ethnicity, smoking, education and the presence or absence of labor. The combination of higher study power in CANDLE analyses (N = 763) relative to GAPPS (max N = 142) and significantly more extreme $PM_{2.5}$ average exposures in CANDLE may have contributed to our observation of more associations in that cohort. It is also possible that differences in $PM_{2.5}$ composition between the cohorts influenced associations, yet this aspect of regional exposures was not assessed. Another limitation of the current analyses, similar to most placental lncRNA analyses, is that the analyses are cross-sectional given samples can only be taken at birth. Consequently, despite prior observations that the placental transcriptome changes over time in pregnancy as fetal and maternal needs change[125], we were restricted to assessing only the late-pregnancy placenta. It was not feasible in this investigation to include methods that address placental cell heterogeneity of lncRNA transcription (e.g. single-cell sequencing), control of which would have rigorously adjusted for potential differences in cell mixtures of samples that might occur based on differences in placental sampling, storage and processing. Our exploratory FDR < 0.1 also increased the odds of false positive findings and, therefore, our findings must be interpreted cautiously. Numerous transcripts we identified are well known growth and proliferation regulators (including under stress conditions), and while this lends credence to our findings the inconsistency of findings across cohorts – especially in light of the known high inter-individual

variability of lncRNA expression[126] – reinforces that these results are likely only generalizable to populations with relatively similar distributions of $PM_{2.5}$ exposure and other characteristics.

In conclusion, we observed cohort-specific associations of placental *LINC00702* transcription with maternal $PM_{2.5}$ exposure in CANDLE as well as numerous potential infant-sex specific associations – particularly specific to female infants in CANDLE – of $PM_{2.5}$ across pregnancy with placental lncRNA. We also observed potential sex-specific association of *AC104083.1* with infant birthweight among male infants in CANDLE. Several of the identified lncRNAs, including those identified in infant-sex specific associations, have hypothesized roles in placental growth and development. These findings warrant future confirmatory and functional studies. Findings, if confirmed will enhance mechanistic understanding of adverse prenatal complications of particulate matter exposure.

2.5 Tables

Table 2-1: Select Characteristics of the Study Populations

Population Covariates	**CANDLE** (N=776)	**GAPPS** (N=205)
Continuous	**Mean (SD)**	**Mean (SD)**
Maternal Age (years)	27.3 (5.5)*	30.5 (5.7)*
Gestational Age (weeks)	39.0 (1.5)	38.3 (3.0)*
Pre-Pregnancy BMI (kg/m^3)	27.7 (7.4)*	26.3 (9.1)*
Family Income (dollars)	41400 (27600)*	58900 (27600)*
Birthweight (grams)	3310 (516)	3270 (806)
Categorical	**N (%)**	**N (%)**
Maternal Race/Ethnicity		
White	292 (37.6)	163 (79.5)†
Black/African American	438 (56.4)	4 (2.0)†
Asian	6 (0.8)	7 (3.4)†
Native Hawaiian Pacific Islander	0 (0)	0 (0)
American Indian/Alaska Native	0 (0)	2 (1.0)†
Other	37 (4.8)	12 (5.9)†
Multiple Race	3 (0.4)	3 (1.5)†
Offspring Sex		
Male	381 (49.1)	108 (52.7)
Female	395 (50.9)	97 (47.3)
Season of Birth		
Spring	166 (21.4)†	53 (25.9)
Summer	219 (28.2)†	50 (24.4)
Fall	217 (28.0)†	52 (25.4)
Winter	170 (21.9)†	50 (24.4)
Mode of Delivery		
Normal/Vaginal	470 (60.6)	134 (65.4)
C-section	306 (39.4)	71 (34.6)
Presence/Absence of Labor		
Spontaneous, spontaneous augmented, or induced labor	627 (80.8)†	180 (87.8)†
No labor	148 (19.1)†	16 (7.8)†
Maternal Smoking		
Yes/Cotinine positive (>200 ng/mL)	37 (4.8)†	0 (0)†
No/Cotinine negative (<200 ng/mL)	736 (94.8)†	2 (1.0)†

Maternal Smoking (Self-report)		
Yes	65 (8.4)	4 (2.0)†
No	711 (91.6)	199 (97.1)†
Maternal Education		
Less than high school	66 (8.5)	10 (4.9)†
High school completion	346 (44.6)	57 (27.8)†
Graduated college or technical school	260 (33.5)	85 (41.5)†
Some graduate work or professional degree	104 (13.4)	52 (25.4)†

* Mean and standard deviation values calculated after removal of missing values

† Percentages do not sum to 100% due to missing values of the covariates

Table 2-2: Maternal $PM_{2.5}$ Exposure Distributions

Exposure Variable	Distribution Characteristics						
	Mean	SD	Min	25%	Median	75%	Max
CANDLE							
$PM_{2.5}$ - Full Pregnancy (ug/m^3)	10.5	1.4	7.9	9.5	10.3	11	16.8
$PM_{2.5}$ - 1st Trimester (ug/m3)	10.6	1.4	7.9	9.7	10.4	11.2	16.6
$PM_{2.5}$ - 2nd Trimester (ug/m^3)	10.7	1.7	7.5	9.6	10.6	11.5	17.2
$PM_{2.5}$ - 3rd Trimester (ug/m^3)	10.7	2.1	5.5	9.5	10.5	11.7	19.2
$PM_{2.5}$ - First Month (ug/m^3)	10.5	2.1	5.8	9.4	10.3	11.3	19.0
$PM_{2.5}$ - Last Month (ug/m^3)	10.6	0.9	8.7	9.9	10.7	11.2	13.8
GAPPS							
$PM_{2.5}$ - Full Pregnancy (ug/m^3)	6.2	3.1	1.8	4.1	5.3	7.3	21.3
$PM_{2.5}$ - 1st Trimester (ug/m^3)	6.0	3.0	2.1	3.7	5.1	7.2	16.8
$PM_{2.5}$ - 2nd Trimester (ug/m^3)	6.0	3.1	2.3	3.9	5.2	7.0	20.3
$PM_{2.5}$ - 3rd Trimester (ug/m^3)	6.0	3.5	2.1	3.8	5.1	6.7	23.3
$PM_{2.5}$ - First Month (ug/m^3)	6.3	3.7	1.8	3.8	5.5	7.6	22.8
$PM_{2.5}$ - Last Month (ug/m^3)	6.1	1.8	2.1	4.8	5.7	7.2	12.0

* All air pollution estimates were generated based on home address using the R Spatiotemporal package; distance to nearest roadway was estimated using GIS

Table 2-3: Significant LncRNA Transcripts associated with $PM_{2.5}$ Exposure

ENSEMBL ID	Transcript Symbol*	Description	logFC	Fold Change	P-value	FDR† P-value
First-Trimester $PM_{2.5}$ (CANDLE)						
ENSG00000233117	LINC00702	Long intergenic non-protein coding RNA 702 (Chr. 10)	0.103	1.074	<0.001	0.027
Second-Trimester $PM_{2.5}$ (CANDLE)						
ENSG00000205959	AC105345.1	Novel transcript (Chr. 4)	0.159	1.116	<0.001	0.028

* Transcript symbols and descriptions from HGNC database

† The FDR cutoff for significance was <0.10

Table 2-4: $PM_{2.5}$–Infant-Sex Interactions on LncRNA transcripts

ENSEMBL ID	Transcript Symbol*	Description	Interaction P-value	Interaction┼ FDR P-value
Full-Pregnancy $PM_{2.5}$ (CANDLE)				
ENSG00000218510	LINC00339	Long intergenic non-protein coding RNA 339 (Chr. 1)	<0.001	<0.001

* Transcript symbols and descriptions from HGNC database

┼ The FDR cutoff for significance was <0.10

Table 2-5: Infant-sex Stratified $PM_{2.5}$-LncRNA Associations

ENSEMBL ID	Transcript Symbol*	Description	Infant-sex Strata	logFC	Fold Change	P-value	FDR† P-value
Second-Trimester $PM_{2.5}$ (CANDLE)							
ENSG00000286358	AC096647.1	Novel transcript (Chr. 2)	Male	0.035	1.024	0.171	0.755
			Female	**0.111**	**1.080**	**<0.001**	**0.023**
ENSG00000132204	LINC00470	long intergenic non-protein coding RNA 470 (Chr. 18)	Male	0.050	1.035	0.072	0.735
			Female	**0.115**	**1.083**	**<0.001**	**0.072**
ENSG00000229152	ANKRD10-IT1	ANKRD10 intronic transcript 1 (Chr. 13)	Male	0.129	1.094	0.043	0.735
			Female	**0.249**	**1.189**	**<0.001**	**0.072**
ENSG00000229807	XIST	X inactive specific transcript (Chr. X)	Male	0.090	1.065	0.494	0.852
			Female	**0.222**	**1.166**	**<0.001**	**0.072**
ENSG00000248445	SEMA6A-AS1	SEMA6A antisense RNA 1 (Chr. 5)	Male	0.083	1.059	0.130	0.745
			Female	**0.194**	**1.144**	**<0.001**	**0.072**
ENSG00000273373	AL355488.1	Novel transcript, antisense to SLC16A4 (Chr. 1)	Male	0.054	1.038	0.320	0.789
			Female	**0.195**	**1.144**	**<0.001**	**0.072**
ENSG00000284669	AC092053.4	Novel transcript, antisense CSRNP1 (Chr. 3)	Male	0.063	1.045	0.212	0.766
			Female	**0.177**	**1.130**	**<0.001**	**0.072**
ENSG00000287189	AL121956.6	Novel transcript (Chr. 6)	Male	0.035	1.025	0.397	0.801
			Female	**0.169**	**1.124**	**<0.001**	**0.072**
ENSG00000228434	AC004951.1	Novel transcript (Chr. 7)	Male	0.045	1.031	0.349	0.789
			Female	**0.142**	**1.104**	**<0.001**	**0.080**
ENSG00000232682	AL592430.2	Novel transcript, antisense to ANK3 (Chr. 10)	Male	0.118	1.085	0.107	0.745
			Female	**0.212**	**1.159**	**<0.001**	**0.080**
ENSG00000182873	PRKCZ-AS1	PRKCZ antisense RNA 1 (Chr. 1)	Male	0.130	1.094	0.033	0.735
			Female	**0.180**	**1.133**	**0.001**	**0.085**
ENSG00000231365	WARS2-AS1	WARS2 antisense RNA 1 (Chr. 1)	Male	0.019	1.013	0.445	0.825
			Female	**0.082**	**1.058**	**0.001**	**0.085**
ENSG00000267943	AC010328.1	Novel transcript (Chr. 19)	Male	0.001	1.001	0.968	0.989

			Female	**0.166**	**1.122**	**0.001**	**0.085**
ENSG00000273151	AC073957.3	Novel transcript, antisense to GET4 (Chr. 7)	Male	0.146	1.107	0.016	0.735
			Female	**0.189**	**1.140**	**<0.001**	**0.085**
ENSG00000286403	AC010378.2	Novel transcript (Chr. 5)	Male	0.031	1.022	0.509	0.856
			Female	**0.177**	**1.131**	**0.001**	**0.085**
ENSG00000177822	TENM3-AS1	TENM3 antisense RNA 1 (Chr. 4)	Male	0.001	1.001	0.978	0.989
			Female	**0.111**	**1.080**	**0.002**	**0.087**
ENSG00000184441	AP001062.1	Novel transcript, antisense to C21orf2 (Chr. 21)	Male	0.075	1.054	0.133	0.745
			Female	**0.143**	**1.104**	**0.002**	**0.087**
ENSG00000227706	AL713998.1	Novel transcript (Chr. 6)	Male	0.068	1.048	0.089	0.735
			Female	**0.129**	**1.094**	**0.001**	**0.087**
ENSG00000230551	AC021078.1	Novel transcript (Chr. 5)	Male	0.094	1.067	0.020	0.735
			Female	**0.128**	**1.093**	**0.002**	**0.087**
ENSG00000233006	MIR3936HG	MIR3936 host gene (Chr. 5)	Male	0.056	1.040	0.218	0.766
			Female	**0.137**	**1.099**	**0.002**	**0.087**
ENSG00000234072	AC074117.1	Novel transcript, antisense to GTF3C2 and EIF2B4 (Chr. 2)	Male	-0.018	0.988	0.512	0.861
			Female	**0.085**	**1.061**	**0.002**	**0.087**
ENSG00000234684	SDCBP2-AS1	SDCBP2 antisense RNA 1 (Chr. 20)	Male	0.019	1.013	0.536	0.873
			Female	**0.088**	**1.063**	**0.002**	**0.087**
ENSG00000253671	AC027117.1	Novel transcript (Chr. 8)	Male	0.029	1.020	0.549	0.875
			Female	**0.144**	**1.105**	**0.002**	**0.087**
ENSG00000259306	AC020891.2	Novel transcript (Chr. 15)	Male	0.031	1.021	0.642	0.911
			Female	**0.211**	**1.158**	**0.002**	**0.087**
ENSG00000260034	LCMT1-AS2	LCMT1 antisense RNA 2 (Chr. 16)	Male	-0.001	0.999	0.978	0.989
			Female	**0.150**	**1.109**	**0.002**	**0.087**
ENSG00000269825	AC022150.4	Novel transcript, sense intronic to ZNF83 (Chr. 19)	Male	0.054	1.038	0.109	0.745
			Female	**0.123**	**1.089**	**0.001**	**0.087**
ENSG00000272455	AL391244.2	Novel transcript (Chr. 1)	Male	0.066	1.047	0.150	0.745

			Female	**0.149**	**1.108**	**0.002**	**0.087**
ENSG00000278600	AC015871.3	Novel transcript, sense intronic to ST20 (Chr. 15)	Male	0.064	1.045	0.169	0.749
			Female	**0.156**	**1.114**	**0.002**	**0.087**
ENSG00000245532	NEAT1	nuclear paraspeckle assembly transcript 1 (Chr. 11)	Male	0.164	1.121	0.048	0.735
			Female	**0.246**	**1.186**	**0.002**	**0.091**
ENSG00000253200	AC037459.2	Novel transcript, antisense to KIAA1967 (Chr. 8)	Male	-0.006	0.996	0.850	0.966
			Female	**0.092**	**1.066**	**0.003**	**0.091**
ENSG00000254988	AP002498.1	Novel transcript (Chr. 11)	Male	0.064	1.046	0.314	0.789
			Female	**0.189**	**1.140**	**0.003**	**0.091**
ENSG00000272941	AC083862.2	Novel transcript, antisense to C7orf49 (Chr. 7)	Male	0.009	1.007	0.844	0.966
			Female	**0.127**	**1.092**	**0.003**	**0.091**
ENSG00000279467	AP000350.7	Novel transcript, antisense to SLC2A11 (Chr. 22)	Male	0.106	1.076	0.082	0.735
			Female	**0.172**	**1.127**	**0.003**	**0.091**
ENSG00000285979	AC009090.6	Novel transcript, antisense to FAM192A (Chr. 16)	Male	0.104	1.075	0.065	0.735
			Female	**0.175**	**1.129**	**0.003**	**0.091**
ENSG00000224536	AC096677.1	Novel transcript (Chr. 1)	Male	0.055	1.039	0.120	0.745
			Female	**0.098**	**1.070**	**0.003**	**0.094**
ENSG00000233928	AL591501.1	Novel transcript (Chr. X)	Male	0.040	1.028	0.367	0.797
			Female	**0.132**	**1.096**	**0.003**	**0.094**
ENSG00000249348	UGDH-AS1	UGDH antisense RNA 1 (Chr. 4)	Male	0.016	1.011	0.502	0.853
			Female	**0.080**	**1.057**	**0.003**	**0.094**
ENSG00000251562	MALAT1	metastasis associated lung adenocarcinoma transcript 1 (Chr. 11)	Male	0.086	1.062	0.160	0.745
			Female	**0.186**	**1.138**	**0.003**	**0.094**
ENSG00000269514	AC024257.3	Novel transcript, antisense to OR10AD1 (Chr. 12)	Male	0.014	1.009	0.634	0.911
			Female	**0.086**	**1.061**	**0.003**	**0.094**
ENSG00000258744	AL132800.1	Novel transcript, antisense to CMA1 (Chr. 14)	Male	0.018	1.012	0.539	0.873
			Female	**0.087**	**1.063**	**0.004**	**0.097**
ENSG00000286828	AC008883.3	Novel transcript (Chr. 5)	Male	0.058	1.041	0.177	0.766

			Female	**0.124**	**1.089**	**0.004**	**0.097**
ENSG00000213062	Z99572.1	Novel transcript (Chr. 1)	Male	0.042	1.030	0.270	0.789
			Female	**0.114**	**1.082**	**0.004**	**0.098**
ENSG00000235703	LINC00894	long intergenic non-protein coding RNA 894 (Chr. X)	Male	0.200	1.149	0.031	0.735
			Female	**0.243**	**1.184**	**0.004**	**0.098**
ENSG00000288569	AC008394.1	Novel transcript (Chr. 5)	Male	0.103	1.074	0.059	0.735
			Female	**0.167**	**1.123**	**0.004**	**0.098**
ENSG00000261505	AL031714.1	Novel transcript, antisense to UBE2I (Chr. 16)	Male	0.094	1.068	0.148	0.745
			Female	**0.177**	**1.131**	**0.004**	**0.098**
First-Month $PM_{2.5}$ (CANDLE)							
ENSG00000267040	LINC00656	long intergenic non-protein coding RNA 656 (Chr. 18)	**Male**	**0.069**	**1.049**	**<0.001**	**0.048**
			Female	-0.018	0.988	0.408	0.988
ENSG00000232931	LINC00342	long intergenic non-protein coding RNA 342 (Chr. 2)	**Male**	**0.051**	**1.036**	**<0.001**	**0.075**
			Female	-0.008	0.995	0.639	0.988
Third-Trimester $PM_{2.5}$ (GAPPS)							
ENSG00000153363	LINC00467	long intergenic non-protein coding RNA 467 (Chr. 1)	**Male**	**0.128**	**4.553**	**<0.001**	**0.034**
			Female	-0.050	-1.562	0.126	0.937
Last-Month $PM_{2.5}$ (GAPPS)							
ENSG00000276900	AC023157.3	Novel transcript, antisense to AMN1 (Chr. 12)	**Male**	**-0.179**	**0.883**	**<0.001**	**0.038**
			Female	0.044	1.031	0.128	0.981

* Transcript symbols and descriptions from HGNC database
† The FDR cutoff for significance was <0.10

Table 2-6: LncRNA Transcripts Associated with Birthweight in Infant-sex Stratified Models

ENSEMBL ID	Transcript Symbol*	Description	Infant-sex Strata	logFC	Fold Change	P-value	FDR† P-value
Birthweight (GAPPS)							
ENSG00000260244	AC104083.1	Novel transcript, overlapping GUCY1A3 (Chr. 4)	**Male**	**4.4E-04**	**1.000**	**<0.001**	**0.032**
			Female	2.2E-04	1.000	0.021	0.407

* Transcript symbols and descriptions from HGNC database

† The FDR cutoff for significance was <0.10

CHAPTER 3: Associations of placental long non-coding RNAs with maternal pre-pregnancy obesity and infant birthweight

Abstract

Background: Pre-pregnancy obesity (ppOB) has been linked to both pregnancy complications and abnormal fetal growth through mechanisms that involve the placenta. Long non-coding RNAs (lncRNAs) play potentially critical epigenetic regulatory roles in cellular processes and known roles in pregnancy complications, yet to our knowledge, no prior study has examined the impact of maternal pre-pregnancy BMI (ppBMI) on human placental lncRNAs and subsequent consequences on infant birthweight. We investigated overall and sex-specific associations of ppBMI and ppOB with placental lncRNA transcripts in two birth cohorts. In addition, we examined associations of these lncRNAs with birthweight.

Methods: Study participants were mother-child dyads recruited to the CANDLE (Memphis, TN) (N = 776) and GAPPS (Seattle and Yakima, WA) (N = 205) cohorts. Maternal ppBMI was assessed at enrollment using interviewer-administered questionnaires. LncRNAs (1,077 and 1,033 transcripts for CANDLE and GAPPS, respectively) were measured using paired-end RNA sequencing of placental samples collected at birth. Cohort-specific linear models were fit to evaluate associations of placental lncRNA with ppBMI, ppOB (BMI≥30kg/m^2). Models included *a priori*-specified confounders and experimental variables. To explore potential effect modification by infant-sex, we conducted sex-stratified analyses and fitted models that additionally included BMI–infant-sex interaction terms. Associations of continuous birthweight

with lncRNAs were modelled similarly. All analyses used FDR<0.10 to determine statistical significance.

Results: Mean CANDLE and GAPPS ppBMIs were 27.7kg/m^3 and 27.4kg/m^3 (both SDs=7.4), respectively, while corresponding ppOB proportions were 32% and 25%. Overall, no lncRNA transcript was significantly associated with ppBMI, ppOB, or infant birthweight. In CANDLE, ppBMI was associated with three lncRNA transcripts (*ERVH48-1*, *AC139099.1*, *CEBPA-DT*) among male infants (logFCs = 0.012, 0.029, 0.016, respectively, all FDR p-values = 0.085) while ppOB was associated with one lncRNA transcript (*ZNF225-AS1*) among female infants (logFC = -1.175, FDR p-value = 0.020). In addition, in CANDLE, one lncRNA transcript (*AC104083.1*) was associated with birthweight among male infants (logFC = 4.3e-4, FDR p-value = 0.039). In GAPPS, ppBMI was associated with two lncRNA transcripts (*AP000879.1* and *AL365203.2*) among male infants (logFCs = -0.077, -0.053, respectively, all FDR p-values = 0.085), while 17 lncRNA transcripts (including *LINC02709*, *KANSL1-AS1*, *DANCR*, *EPB41L4A-AS1*, and *GABPB1-AS1*) were associated with birthweight among male infants. However, no BMI-infant-sex interaction terms were statistically significant.

Conclusions: We observed potential infant-sex specific associations of several lncRNA transcripts with ppBMI, ppOB, and infant birthweight. Though many of these transcripts are uncharacterized, several of the identified lncRNAs (e.g. *ERVH48-1* and *CEBPA-DT*) have previously been linked to pathways controlling cancer or placental growth, trophoblast differentiation and gene expression. These associations must be validated in future studies.

3.1 Introduction

The intra-uterine environment plays a critical role in fetal growth and development during pregnancy, and adverse maternal conditions can set the stage for lifelong offspring health problems[3,127]. Prior research has linked maternal pre-pregnancy obesity (ppOB; defined as body mass index [BMI] $\geq$ 30 kg/m^2) to high birthweight (macrosomia) and, to a lesser degree, low birthweight (LBW), as well as higher risk for pregnancy complications including gestational diabetes, pre-eclampsia, gestational hypertension, and preterm delivery[22,23,27]. The placenta functions as a key mediator of these relationships. ppOB exerts distinct effects on the placenta[128], resulting in decreased ratios of placental length/breadth to birthweight, increased placental thickness[129], increased placental weight[130], placental vascular lesions and villous abnormalities[131–133], as well as other reductions in placental proliferation and apoptosis[134]. Some of these adverse impacts on placental growth and function may be due to alterations in expression of placental genes, including those involved in angiogenesis[34,35,37], or changes in their epigenetic regulators, such as long non-coding RNAs (lncRNAs)[49,54,77–79,81].

LncRNAs are RNA transcripts longer than 200 base pairs that are not translated into proteins[135]. They are now understood to be new class of epigenetic biomarkers known to have transcription regulation roles through physical interactions with mRNA, DNA, proteins and miRNAs[45] at transcriptional, post-transcriptional, translational, and post-translational stages. LncRNAs are less abundant than mRNAs and show considerable tissue-specificity and inter-individual variation[45]. This tissue-specificity has been observed in placenta[136]. Further, infant-sex differences in placental lncRNAs have been reported before[137,138], mirroring other differences between male and female infants in their response to adverse perinatal exposures and

maternal characteristics[30,139–141]. However, to our knowledge, no prior study examined human placental lncRNA transcript levels in relation to maternal ppOB and downstream effects on infant birthweight.

Therefore, this investigation sought to address these knowledge gaps by investigating overall and infant-sex specific associations of ppBMI and ppOB with placental lncRNA transcripts, as well as associations of placental lncRNAs with infant birthweight. To allow for potential replication of findings, we conducted this investigation separately in two birth cohorts.

3.2 Methods

Study Setting and Study Participants

Participants in this study came from two cohorts of the ECHO PATHWAYS Consortium. The first was the Conditions Affecting Neurocognitive Development in Early Childhood (CANDLE) cohort, described in detail elsewhere[84]. Briefly, CANDLE is a prospective pregnancy cohort study that enrolled urban, pregnant, women between 2006 and 2011 (N = 1,503). To be eligible, women were required to be Shelby County, TN residents – most were from the Memphis metro area – be 16 to 40 years of age, and plan to deliver at one of the participating study hospitals. Additionally, participants were verified at the time of enrollment to have uncomplicated singleton pregnancies. Participants who opted to provide placental samples at birth were available for transcriptomic analyses, with 794 of the 1,503 initial participants doing so. The University of Tennessee Health Sciences Center IRB approved all research protocol and CANDLE participants gave informed consent at enrollment.

The second cohort consisted of participants of a pregnancy biorepository based at Seattle Children's Hospital in Seattle. This repository was a part of the Global Alliance to Prevent Prematurity and Stillbirth (GAPPS)[65], an initiative to investigate risk factors for adverse birth outcomes and interventions addressing them. Women were eligible to be re-recruited from this biorepository into the ECHO PATHWAYS study if they delivered in an affiliated hospital (in Seattle, WA or Yakima, WA), provided one or more pregnancy urine samples and a placenta sample at delivery, completed an enrollment questionnaire during pregnancy, and had a GAPPS child between 4-7 years of age at time of recontact. A total of N = 1,271 women were eligible at the start of re-recruitment in 2017 and 392 met inclusion criteria. The IRBs of Seattle Children's Hospital and University of Washington approved all GAPPS study protocols. Participants gave informed consent.

A total of 794 and 392 women who provided placental samples at birth were ultimately available for inclusion in this ECHO-PATHWAYS study from CANDLE and GAPPS, respectively. In both cohorts, demographic covariates (e.g., maternal age, maternal education, self-reported race, family income, and smoking behavior) and self-reported maternal pre-pregnancy weight and height were collected as a part of interviewer-assisted questionnaires at enrollment. Maternal ppBMI was calculated as self-reported pre-pregnancy mass (in kg) divided by height (in meters) squared. Information on course of pregnancy and birth outcomes were abstracted from medical records. Based on these records, additional exclusions were applied to women from both cohorts. Exclusion criteria included lack of residential address data, multiple birth, or stillbirth. After final exclusions, 776 and 205 women from CANDLE and GAPPS, respectively, comprised the analytic samples.

Placental Sample Collection

Similar procedures were used to isolate placental samples for sequencing in CANDLE and GAPPS, as described in prior reports[33,55]. Briefly, in CANDLE, villous tissue samples were taken from the placental parenchyma within 15 minutes following delivery. Each sample was rectangular in shape with a length of approximately 2 cm, a width of 0.5 cm and a depth of 0.5 cm, and each was subsequently cut into four cubes with 0.5 cm edges and stored overnight in four separate 50 ml test tubes with 20 ml of RNAlater at -20˚C (for no more than 24 hours). Next, tissue cubes were moved to 1.8 ml cryovials of RNAlater and frozen at -80°C. Samples were later allowed to thaw just enough to remove any remaining maternal decidual tissue and then returned to -80˚ C storage with new RNAlater.

In GAPPS, vertical tissue punches (around 8 mm diameter) were cut from the full thickness of the placental disc no more than 30 minutes after delivery. These punches were placed in 3 ml of RNAlater in 5 ml tubes and stored at -20°C prior to being moved to the GAPPS labs. Once they had arrived, the samples were temporarily placed at -80°C prior to being manually dissected to remove decidual contamination. Following dissection, the placental tissue was then divided into one to three 10 – 30 mg pieces, each placed in 1ml RNAlater, and placed at -20°C until shipment for additional pre-sequencing processing steps.

Sample Processing and RNA Sequencing

RNA was separated for sequencing from 30 mg pieces of tissue. Tissue samples were removed from storage and allowed to warm to room temperature. Next, each sample was removed from its storage tube, gently dried with a Kimwipe to remove RNAlater buffer solution, and placed in new test tube containing 600 µl Buffer RLT Plus with mercaptoethanol. A 5 mm

ball bearing was then placed in each tube to aid in homogenization by a TissueLyser LT (Qiagen, Germantown, MD). We then used AllPrep DNA/RNA/miRNA Universal Kits (Qiagen, Germantown, MD) to isolate the RNA according to the manufacturer's instructions. RNA purity was then characterized by $OD_{260/230}$ and $OD_{260/260}$ ratios measured using a NanoDrop 8000 Spectrophotometer (Thermo Fischer Scientific, Waltham MA), and RNA integrity was further tested using a Bioanalyzer 2100 and RNA 6000 Nanochips (Agilent, Santa Clara, CA). Any RNA samples that did not meet an RNA Integrity Number (RIN) > 7 were not included in the RNA-Seq analysis.

Transcripts were isolated from 1µg of total RNA, starting with poly-A selection and followed by cDNA synthesis. These procedures were completed with TruSeq Stranded mRNA kits (Illumina, San Diego, CA) on a Sciclone NGSx Workstation (Perkin Elmer, Waltham, MA) prior to each cDNA library being marked with a unique barcode. Libraries were then amplified over 13 PCR cycles before being quantified with Qubit Quant-it dsDNA High Sensitivity fluorometric assays (Life Technologies, Carlsbad, CA). Further, a DNA1000 assay on an Agilent 2100 Bioanalyzer was used to assess overall quality and average fragment size before final sequencing of libraries on an Illumina HiSeq 4000 to a depth of around 30 million. Quality assessments of reads were performed using the FASTX-toolkit (v0.0.13) and FastQC (v0.11.2)[87]. All isolation and sequencing operations were performed at the University of Washington Northwest Genomics Center.

RNA transcripts were aligned to the GRCh38 transcriptome (Gencode v33) in Kallisto[88] to estimate their abundances (summarized at the transcript level), which were then scaled to the average transcript length using the R Bioconductor *tximport* package[89]. Library normalization factors were then estimated using the Bioconductor *edgeR* package's[90] Trimmed Mean of M-

values (TMM) function, allowing for lncRNA transcripts assessed to have unreliably low expression (mean log counts/million counts (logCPM) ≤ 0) to be excluded from analysis. Of the 16,000 lncRNA transcripts available, 1,077 met this threshold value in CANDLE and 1,033 in GAPPS. The limma-voom pipeline[91] of the Bioconductor *limma* package was used to convert the counts/transcript to logCPM values via the TMM normalization factors before computing observation-level weights based on the transcripts' mean-variance relationships. These weights were then applied in the linear modeling functions of the same package for analysis.

Statistical Analyses

We used descriptive statistics to characterize the analytic samples and the parent cohorts. As with prior transcriptomic studies in these populations[33], all analyses were cohort-specific to CANDLE or GAPPS. Weighted linear regression models were fit using the *limma* lmfit function[89], and regressed genome-wide placental expression of lncRNAs (as logCPM values, described above) on ppBMI (continuous) or ppOB (categorical, BMI ≥ 30 kg/m^2 vs. no). Models were fit for each BMI variable separately, and both ppBMI and ppOB models adjusted for *a priori*-identified confounding variables and experimental variables, including RNA sequencing batch, hospital site (for GAPPS only, controlling for differences between the Seattle and Yakima, WA locations), maternal age (continuous), self-reported race (White/Black/Asian/other/multiple race, with American Indian/Alaska Native counted as a part of "other" in GAPPS), infant-sex (male/female), calendar year of birth – as control for other year-linked exposures or conditions – gestational age at birth (continuous), family income (continuous), mode of delivery (cesarean section/vaginal), labor (presence/absence), smoking (yes/no for smoking during pregnancy), maternal urine cotinine (binary variable using 200 ng/mL cut off [92]), and maternal education

(less than high school/high school/college or technical school/graduate or professional degree). Prior to viewing results, the subset of models with p-value distributions deviating from the expected uniform distribution were rerun with added surrogate variable analysis (SVA)[93] to control excess unmeasured sources of variation. In all models, LncRNAs were considered to be differentially transcribed based on a false discovery rate (FDR) < 0.10 [94]. Further, we fit infant-sex stratified models with similar covariate control to examine sex-specific associations. We also refit main-effect models with an additional exposure–infant-sex interaction term to test ppBMI–infant-sex or ppOB–infant-sex interactions. An FDR < 0.10 was also used to determine statistical significance of interactions terms.

We then examined the relationship between lncRNA expression and birthweight, with birthweight as the exposure and lncRNA transcripts as the outcome. This reverse model construction allowed us to take advantage of the performance properties of the *limma* package methodology, which require RNA levels to be the dependent variable[89,91]. These models included identical covariate sets to those described above. Models containing interaction of birthweight and infant-sex were also fit to examine infant-sex specific associations, followed by fitting of infant-sex stratified models. As in BMI-related models, all birthweight analyses used an FDR < 0.10 to determine statistical significance.

All analyses were performed with complete-cases, resulting in 51 subjects being excluded from CANDLE models for missing covariate information (final N = 725). Of the original 205 subjects, 46 were removed from individual GAPPS models due to a combination of missing covariate information, resulting in a final analysis sample size of N = 159. All analyses were performed using R (v 4.1.2).

3.3 Results

On average, study participant mothers were 27.3 (SD=5.5) and 30.5 (SD=5.7) years old in CANDLE and GAPPS, respectively (**Table 3-1**). Mean gestational age at birth was 39.0 (SD = 1.5 weeks) and 38.3 weeks (SD = 3.0 weeks) for CANDLE and GAPPS participants, respectively. Most CANDLE participants were Black (56%) or White (38%), while GAPPS participants were predominantly White (80%). The analytic populations of both studies were representative of their source cohorts (**Supplemental Table 1**). While average pre-pregnancy BMIs were similarly overweight (BMI ≥ 25 kg/m^2) in both cohorts (27.7 kg/m^2 [SD = 7.4] and 27.4 kg/m^2 [SD = 7.4] for CANDLE and GAPPS, respectively), a slightly – though not significantly – larger proportion of CANDLE mothers were obese pre-pregnancy (32%) compared with mothers in GAPPS (25%) (**Table 3-2**).

We found no overall associations of ppBMI and ppOB with lncRNA transcripts in CANDLE or GAPPS. In sex-stratified models, among male infants in CANDLE ppBMI was associated with three lncRNA transcripts (*ERVH48-1, AC139099.1, CEBPA-DT*) (**Table 3-3**). These transcripts were upregulated by approximately 0.9%, 2.0% and 1.1%, respectively per unit increase in ppBMI among males (all FDR values = 0.85). Among females, ppOB was associated with 55.7% downregulation of one lncRNA transcript (*ZNF225-AS1*) in CANDLE (FDR = 0.020). In GAPPS, ppBMI was associated with downregulation of two lncRNA transcripts (*AP000879.1* and *AL365203.2*) among male infants, by approximately 5.2% and 3.6%, respectively, per unit BMI (both FDR values = 0.85). We did not find associations of ppBMI or ppOB with lncRNA transcripts among female infants in GAPPS. Further, no BMI-infant-sex interaction terms were statistically significant in either cohort.

We did not find overall associations of placental lncRNA transcripts with infant birthweight in CANDLE or GAPPS. In stratified models, among male infants in CANDLE we found a positive association of one lncRNA transcript (*AC104083.1*) with infant birthweight (logFC = 4.3e-4, FDR = 0.039) (**Table 3-4**). Among female infants in GAPPS, we found associations of 17 lncRNA transcripts with infant birthweight (all FDR < 0.010). Six of these lncRNAs (*LINC02709, KANSL1-AS1, DANCR, EPB41L4A-AS1, GABPB1-AS1, CRIM1-DT*) were previously described, while the remaining 11 lncRNAs were novel. Similar to ppBMI-related analyses, none of the infant-sex–birthweight interaction terms were statistically significant in either cohort.

3.4 Discussion

In the current study of CANDLE and GAPPS, we did not find overall associations of lncRNA transcripts with either ppBMI, ppOB, or infant birthweight in either cohort. Three transcripts (*ERVH48-1*, *AC139099.1* and *CEBPA-DT*) were associated with ppBMI in CANDLE and two transcripts (*AP000879.1* and *AL365203.2*) were associated with ppBMI in GAPPS, all among male infants. One transcript (*ZNF225-AS1*) was associated with ppOB among female infants in CANDLE. Further, one lncRNA transcript (*AC104083.1*) was associated with infant birthweight among male infants in CANDLE, while 17 lncRNA transcripts (seven previously characterized and eleven novel) were associated with infant birthweight among females in GAPPS. We did not observe overlap of findings across the two cohorts and none of the infant-sex interaction terms were statistically significant.

To our knowledge, this study was the first to examine human placental lncRNA transcripts in relation to maternal ppBMI or ppOB. Prior studies of human lncRNA transcripts in adipose tissue have highlighted expression of a number of transcripts related to obesity, adipogenesis, lipogenesis and metabolic dysfunctions, including *TUG1*, *MALAT1*, *RP11-20G13.3*, *HOTAIR*, *GYG2P1*, *OLMALINC* and *MIR31HG*[142–146], though we did not see these transcripts among our findings. Of note, these prior investigations did not evaluate placental lncRNA transcripts except for one study of mouse placenta and gonadal fat that compared the impacts of high-fat and low-fat diets[147]. The study identified 52 lncRNAs that were dysregulated in placentas of high fat diet mice with roles related to inflammation, metabolism, and vascular growth. Gene-gene interaction networks with placental mRNAs and lncRNAs identified prominent gene-expression changes for extra-cellular matrix interactions and metabolic pathways[147], though no homologues of the identified lncRNA transcripts were found in our study.

ERVH48-1 and *CEBPA-DT* are lncRNAs whose functions have been extensively characterized. *ERVH48-1* (endogenous retrovirus group 48 member 1) – also known as *HERV-Fb1* – is an endogenous retroviral sequence. The otherwise non-coding *ERVH48-1* contains an exon sequence that codes for a small protein, which the Sugimoto and colleagues designated suppressyn (SUPYN) after observing that knockdown of the *HERV-Fb1* transcript triggered a large increase in cellular fusion of BeWo cells *in vitro*[148]. The same authors reported that SUPYN was primarily localized to villous and extravillous cytotrophoblasts, yet it was also detected in syncytiotrophoblasts and even in maternal decidua; in contrast, its cell-surface target ASCT2 was limited purely to cytotrophoblast lines[119]. Kudaka et al. observed that *ERVH48-1* expression was decreased in the placentas of preeclamptic or hypertensive women[118], while

Sugimoto et al. noticed differential expression under hyper- and hypoxia conditions – the former intended to mimic possible hyperoxia that occurs under preeclampsia[119]. These findings tend to indicate that the downstream consequences of obesity in pregnancy, which are known to include preeclampsia, gestational diabetes and other pregnancy complications[23], can conceivably be related to the impact of *ERVH48-1* expression (and SUPYN production) on aberrant placental fusion events. A recent investigation within the Rhode Island Child Health Study (RICHS) found higher expression of *ERVH48-1* to be significantly associated with lower birthweight among infants (β = -320.9 g, 95% CI: -487.8, -154.0)[83]. For comparison, an ad hoc analysis of *ERVH48-1* using similar linear regression models in the CANDLE and GAPPS populations (with our present covariates) estimates birthweight differences of β = -77.3 g (95% CI = -236.5, 82.1) and -184.2 g (95% CI = -555.2, 186.8), respectively, per 10x increase in *ERVH48-1* expression ($p > 0.05$ in both cases). While the direction of birthweight associations with *ERVH48-1* is similar in CANDLE and GAPPS to Hussey et al. (2020), part of the difference in scale of estimates could be contributed by the latter study's intentional over-selection of large for gestational age (LGA) and small for gestational age (SGA) infants during recruitment, resulting in relatively stronger associations of *ERVH48-1* with birthweight. However, associations were observed only in male infants in the current study, and the previous analyses did not examine sex-specific associations. One other study by Song et al. compared RNA expression profiles of the placentas of macrosomic and non-macrosomic infants, and found 2,929 lncRNAs upregulated and 2,127 lncRNAs downregulated in macrosomic infants[57]. However, no overlaps were observed between the small subset of these lncRNAs presented by the authors and current findings.

CEBPA-DT, previously designated *ADINR* for Adipogenic Differentiation Induced

Noncoding RNA, is a transcript that has been linked to adipogenesis and cancer. Investigators observed that *CEBPA-DT* is up-regulated in oral squamous cell carcinoma (OSCC) cells, whereas silencing of *CEBPA-DT* resulted in reduced proliferation and induced cells apoptosis of OSCC[149]. These effects, as well as alterations in cell migration and invasion, were theorized to be controlled by *CEBPA-DT*'s influence on *CEBPA* and subsequent *CEBPA* regulatory actions on *BCL2*, an apoptosis promoting gene[149]. This observation was replicated in another OSCC investigation by Qiao and colleagues[150]. No studies were found that directly examined *CEBPA-DT* in placenta or in relation to maternal obesity or birthweight. However, links to *MTOR* (molecular target of rapamycin) expression give some indirect evidence as to the placental relevance of *CEBPA-DT*'s proliferative properties[151–153]. Placental mTOR acts as a key regulator of cell growth, metabolism and adaptive response to cell stress, and in general, mTOR signaling is increased in placentas of higher birthweight infants and decreased in infants with lower birthweight[154]. Multiple studies have found *CEBPA* to be a regulator of *MTOR* expression: increased *CEBPA* was found to inhibit *MTOR* expression[151], while decreased *CEBPA* led to the opposite effect[152,153]. *CEBPA-DT* conceivably influences birthweight via control of placental cell proliferation under stress conditions, including maternal obesity.

While the majority of placental transcripts related to birthweight among female infants were uncharacterized, a small subset have been associated with varied processes involved in cell growth and development – proper and aberrant. Of these transcripts, *DANCR*, differentiation antagonizing non-protein coding RNA, is the best characterized. Previous investigators postulated that the transcript played a role in the transition from progenitor to differentiated cells. Further experiments in regenerated epidermal tissue found that depletion of *DANCR* resulted in expression of differentiation genes in all cell types, including epidermal basal cells that normally

have no expression of such proteins[155]. This led to the conclusion that *DANCR* is explicitly antagonistic toward differentiation processes in cell types where it is expressed[155]. Similar to *CEBPA-DT*, *DANCR* expression has also been linked to the action of the mTOR pathway[156] – though not in a placental study context. *DANCR* was found to directly bind and inhibit miR-241-5p, a microRNA that is highly expressed in placental tissues in preeclampsia[157]. Given that miR-241-5p is known to restrict trophoblast proliferation, invasion and migration – its theorized role in preeclampsia pathology[48,158] – increased expression of *DANCR* could conceivably represent part of an adaptive response to the pro-preeclamptic pressures of maternal pre-pregnancy obesity. Consequently, *DANCR*'s opposition to placental pathologies provides a plausible explanation for the observed positive association between *DANCR* expression and birthweight in female infants.

EPB41L4A-AS1, the antisense transcript to *EPB41L4A* on chromosome 5, is one of a few dozen transcripts whose expression is rapidly up-regulated during the G_0 phase of the cell cycle under conditions of cell cycle arrest[159]. CRISPRi analysis of lncRNAs in seven cell types by Liu and colleagues revealed *EPB41L4A-AS1* was differentially expressed in relation to anti-growth pathways for U87 (Uppsala 87 Malignant Glioma) and K562 (immortalized myelogenous leukemia) cells[160]; gene ontology analysis noted positive up-regulation of genes in the K572 "response to O_2 levels" pathway and slight up-regulation for U87 genes in a number of signaling pathways (p53, Integrin, VEGF), gene expression, and negative regulation of cell cycle – consistent with the later Hegre et al. 2021 study[161] – with increasing *EPB41L4A-AS1*[160]. Placental *EPB41L4A-AS1* expression was also investigated in the context of recurrent miscarriage (RM), to our knowledge the only study that has specifically examined *EPB41L4A-AS1* in placental pathology[162]. *EPB41L4A-AS1* was highly upregulated in placental tissue of

early RM pregnancies compared to normal controls and similarly exerted an inverse impact on levels of glycolysis in response to PCG-1α signaling (up-regulated in early RM). Notably, the authors also found that placentas from normal pregnancies had lower *EPB41L4A-AS1* expression than placentas from pregnancies with pathologies such as preeclampsia, intrauterine growth restriction ($p < 0.05$)[162].

The *GABPB1-AS1* transcript, positively associated with birthweight among female infants in GAPPS, has been noted as a marker of chemical-exposure induced cell stress in both cancer and stem cells[163,164]. *GABPB1-AS1* was also observed to be up-regulated by erastin and exert negative control on *GABPB1* translation via a direct binding mechanism in the context of cancer cell ferroptosis; *GABPB1* downregulation resulted in downregulation of peroxiredoxin-5 peroxidase (*PRDX5*) and subsequent accumulation of hydrogen peroxide[165]. An inverse relationship was observed between *GABPB1-AS1* expression and expression of a number of genes that promote cell-cycle progression in gliomas, contrasted by high up-regulation of *GABPB1-AS1* for certain auto-immune disorders[166]. While no studies were found linking this transcript to the placenta or birthweight, these prior findings point to a potential widespread role of *GABPB1-AS1* as a short-term regulator of cell response to stress conditions.

Finally, KANSL1 antisense RNA 1 (*KANSL1-AS1*) – associated with increased birthweight among female infants in this study – is a non-coding transcript on chromosome 17 found proximal to its namesake gene, KAT8 regulatory NSL complex subunit 1 (*KANSL1*). Only two recent investigations have explored functional roles for *KANSL1-AS1*, including a TWAS-style analysis linking it to anxiety disorders[167]. REACTOME and gene ontology analyses in that study indicated possible links of anxiety-associated transcripts to myogenesis, the Wnt signaling pathway/calcium modulating pathway and – most strongly based upon FDR

values – mRNA binding; however, the proportion of contribution *KANSL1-AS1* provided to these pathways was not clear[167]. The second investigation found that *KANSL1-AS1* expression was associated with a decrease in osteosarcoma incidence and development risk scores[168]. While these studies point to potential roles of *KANSL1-AS1* in promotion or attenuation of disease, its mechanistic relevance to placental tissue and infant birthweight remains unclear.

Despite a lack of prior studies on placental lncRNA modulation by ppBMI, LncRNAs have been shown to play roles in pathways underlying development of placental abnormalities and pregnancy complications, most clearly in conditions of hypoxia or improper uterine invasion contributing to preeclampsia[48,54,81,118,119,158], yet also in other conditions such as gestational diabetes[117,120,169]. Given lncRNAs have also been associated with some disease pathologies in a sex-specific manner[123,124], our sex-specific findings are not only plausible but conceivably relate to observed sex-specific differences in gene expression during normal placental growth[121] or response to exogenous cell stressors[30,55,170,171].

This investigation had a number of key strengths. To our knowledge, this ECHO-PATHWAYS investigation is the first TWAS to explicitly examine maternal BMI, obesity status and placental lncRNA expression in humans, as well as the largest existing TWAS cohort analyses of placental lncRNA expression and infant birthweight. Beyond its size, the other primary strength of this investigation is its robust consideration of covariates from extensive surveys and medical record abstraction. This allowed for *a priori* construction of models that include a number of possible confounders and experimental variables. We hoped the relatively large number of subjects and excellent covariate control in these cohorts would allow us to explore mediation by transcripts, yet a lack of overlapping significant associations between pre-pregnancy BMI, lncRNAs, and birthweight precluded such analysis. Also, the larger of the two

cohorts, CANDLE was mainly comprised of Black women and their infants (56.4%). This makes this analysis the largest of its kind to focus on placental lncRNA expression in the American Black population, a population that has rarely been a focus in TWAS-style analyses historically. Further, our large sample sizes provided enough power to examine sex-specific associations with lncRNAs, something not possible in many earlier analyses of this type.

Several limitations of this investigation deserve mention. While the underlying intention behind using two cohorts for this analysis was to provide an opportunity for replication of results, it is clear that the cohorts had different distributions of population characteristics and exposures – likely linked to major underlying differences due to region and demographics – which led to consistently different effect sizes for various transcripts, independent of the major difference in power between the CANDLE (N = 776) and GAPPS (N = 205) analytic samples. Additionally, since our outcome, placental lncRNA expression, is necessarily measured at birth, this analysis was cross-sectional in nature, thus we could only use our associations as a single time point inference on overall impacts of ppOB on lncRNAs. This is a key limitation, as placental transcriptome varies widely over the course of pregnancy[125]. Further, despite greater power in combined models to detect an interaction, we did not observe any statistically significant interactions for the various transcripts highlighted in stratified models. This finding – and our less stringent FDR < 0.1 – increases the possibility of false positives, yet the fact that numerous transcripts we found were clearly highlighted in growth and proliferation control roles (including under stress conditions) lends credence to our findings. Definitive confirmation of these transcripts' importance in pregnancy would require us to address the issue of placental cell heterogeneity, and the best methods to do this, including single-cell sequencing methods, were not feasible in this investigation. Finally, generalizability of our findings likely only applies to

populations with similar ppBMI, ppOB, or birthweight distributions to those in our study cohorts.

In conclusion, in this study we found several infant-sex specific associations of placental lncRNA with BMI, maternal pre-pregnancy obesity status, and infant birthweight. Aside from sex-specific findings, no transcripts were associated with maternal BMI or birthweight in main-effect or interaction models. Our work suggests that maternal pre-pregnancy BMI has potential infant-sex related impacts on placental growth and development via lncRNA expression, and that placental lncRNA may also have infant-sex specific impact on birthweight. These findings warrant future confirmatory and functional studies.

3.5 Tables

Table 3-1: Select Characteristics of the Study Populations

Population Covariates	**CANDLE** (N=776)	**GAPPS** (N=205)
Continuous	**Mean (SD)**	**Mean (SD)**
Maternal Age (years)	27.3 (5.5)*	30.5 (5.7)*
Gestational Age (weeks)	39.0 (1.5)	38.3 (3.0)*
Family Income (dollars)	41400 (27600)*	58900 (27600)*
Birthweight (grams)	3310 (516)	3270 (806)
Categorical	**N (%)**	**N (%)**
Maternal Race/Ethnicity		
White	292 (37.6)	163 (79.5)†
Black/African American	438 (56.4)	4 (2.0)†
Asian	6 (0.8)	7 (3.4)†
Native Hawaiian Pacific Islander	0 (0)	0 (0)
American Indian/Alaska Native	0 (0)	2 (1.0)†
Other	37 (4.8)	12 (5.9)†
Multiple Race	3 (0.4)	3 (1.5)†
Offspring Sex		
Male	381 (49.1)	108 (52.7)
Female	395 (50.9)	97 (47.3)
Season of Birth		
Spring	166 (21.4)†	53 (25.9)
Summer	219 (28.2)†	50 (24.4)
Fall	217 (28.0)†	52 (25.4)
Winter	170 (21.9)†	50 (24.4)
Mode of Delivery		
Normal/Vaginal	470 (60.6)	134 (65.4)
C-section	306 (39.4)	71 (34.6)
Presence/Absence of Labor		
Spontaneous, spontaneous augmented, or induced labor	627 (80.8)†	180 (87.8)†
No labor	148 (19.1)†	16 (7.8)†
Maternal Smoking		
Yes/Cotinine positive (>200 ng/mL)	37 (4.8)†	0 (0)†
No/Cotinine negative (<200 ng/mL)	736 (94.8)†	2 (1.0)†

Maternal Smoking (Self-report)		
Yes	65 (8.4)	4 (2.0)†
No	711 (91.6)	199 (97.1)†
Maternal Education		
Less than high school	66 (8.5)	10 (4.9)†
High school completion	346 (44.6)	57 (27.8)†
Graduated college or technical school	260 (33.5)	85 (41.5)†
Some graduate work or professional degree	104 (13.4)	52 (25.4)†

* Mean and standard deviation values calculated after removal of missing values

† Percentages do not sum to 100% due to missing values of the covariates

Table 3-2: Maternal Pre-pregnancy Body Mass Index (BMI) Distributions

		Distribution Characteristics†						
Exposure Variable	*Obesity Status: N (%)	Mean	SD	Min	25%	Median	75%	Max
CANDLE								
Maternal Pre-Pregnancy BMI (kg/m^3)	249 (32.1)	27.7	7.4	14	22	26	32	62
GAPPS								
Maternal Pre-Pregnancy BMI (kg/m^3)	50 (25.0)	27.4	7.4	16	22	26	30	54

* Obese status is defined by a BMI ≥ 30 kg/m^3

† BMI was rounded to the nearest whole number upon initial calculation

Table 3-3: Top Sex-specific Associations of Maternal Pre-pregnancy BMI with LncRNA transcripts

ENSEMBL ID	Transcript Symbol*	Description	Infant-sex Strata	logFC	Fold Change	P-value	FDR† P-value
Continuous BMI (CANDLE)							
ENSG00000233056	ERVH48-1	Endogenous retrovirus group 48 member 1 (Chr. 21)	**Male**	**0.012**	**1.009**	**<0.001**	**0.085**
			Female	-0.001	1.000	0.813	0.961
ENSG00000262094	AC139099.1	Novel transcript (Chr. 17)	**Male**	**0.029**	**1.020**	**<0.001**	**0.085**
			Female	0.010	1.007	0.067	0.781
ENSG00000267296	CEBPA-DT	CEBPA divergent transcript (Chr. 19)	**Male**	**0.016**	**1.011**	**<0.001**	**0.085**
			Female	0.003	1.002	0.343	0.926
Continuous BMI (GAPPS)							
ENSG00000254721	AP000879.1	Novel transcript (Chr. 11)	**Male**	**-0.077**	**0.948**	**<0.001**	**0.085**
			Female	-0.005	0.997	0.738	0.986
ENSG00000273038	AL365203.2	Novel transcript (Chr. 10)	**Male**	**-0.053**	**0.964**	**<0.001**	**0.085**
			Female	-0.011	0.992	0.295	0.967
Obesity Status (CANDLE)							
ENSG00000186019	ZNF225-AS1	ZNF225 and ZNF224 antisense RNA 1 (Chr. 19)	Male	0.380	1.302	0.140	0.905
			Female	**-1.175**	**0.443**	**<0.001**	**0.020**

* Gene symbols and descriptions from HGNC database

† The FDR cutoff for significance was <0.10

Table 3-4: Top Sex-specific Associations of Birthweight with LncRNA Transcripts

ENSEMBL ID	Transcript Symbol*	Description	Infant-sex Strata	logFC	Fold Change	P-value	FDR† P-value
Continuous Birthweight (CANDLE)							
ENSG00000260244	AC104083.1	Novel transcript, overlapping GUCY1A3 (Chr. 4)	**Male**	**4.3E-04**	**1.000**	**<0.001**	**0.039**
			Female	2.2E-04	1.000	0.021	0.418
Continuous Birthweight (GAPPS)							
ENSG00000228748	AL450306.1	Novel transcript (Chr. 10)	Male	-1.2E-04	1.000	0.709	0.985
			Female	**2.0E-03**	**1.001**	**<0.001**	**0.030**
ENSG00000245522	LINC02709	Long intergenic non-protein coding RNA 2709 (Chr. 11)	Male	3.0E-04	1.000	0.034	0.933
			Female	**5.8E-04**	**1.000**	**<0.001**	**0.030**
ENSG00000261838	AC092718.6	Novel transcript, antisense to C16orf46 (Chr. 16)	Male	1.8E-04	1.000	0.390	0.979
			Female	**1.1E-03**	**1.001**	**<0.001**	**0.030**
ENSG00000262049	AC139530.1	Novel transcript, antisense to HGS (Chr. 17)	Male	-2.3E-04	1.000	0.090	0.967
			Female	**7.8E-04**	**1.001**	**<0.001**	**0.030**
ENSG00000286037	AC011479.5	Novel transcript, antisense to SPINT2 (Chr. 19)	Male	-6.2E-04	1.000	0.404	0.979
			Female	**3.9E-03**	**1.003**	**<0.001**	**0.030**
ENSG00000214401	KANSL1-AS1	KANSL1 antisense RNA 1 (Chr. 17)	Male	-2.2E-04	1.000	0.346	0.979
			Female	**1.1E-03**	**1.001**	**<0.001**	**0.036**
ENSG00000267469	AC005944.1	Novel transcript, antisense to AES (Chr. 19)	Male	9.5E-04	1.001	0.178	0.967
			Female	**3.6E-03**	**1.003**	**<0.001**	**0.036**
ENSG00000277283	AC004812.2	Novel transcript, antisense to RAB35 (Chr. 12)	Male	1.1E-04	1.000	0.447	0.979
			Female	**5.9E-04**	**1.000**	**<0.001**	**0.036**
ENSG00000225339	AL354740.1	Novel transcript, antisense to C6orf1 (Chr. 6)	Male	6.7E-04	1.000	0.011	0.843
			Female	**9.4E-04**	**1.001**	**<0.001**	**0.037**
ENSG00000226950	DANCR	Differentiation antagonizing non-protein coding RNA (Chr. 4)	Male	2.0E-05	1.000	0.861	0.988
			Female	**7.1E-04**	**1.000**	**<0.001**	**0.037**

ENSG00000258017	AC011603.3	Novel transcript, antisense to	Male	5.4E-05	1.000	0.917	0.993
		TUBA1B (Chr. 12)	**Female**	**3.2E-03**	**1.002**	**<0.001**	**0.037**
ENSG00000224032	EPB41L4A-AS1	EPB41L4A antisense RNA 1	Male	-1.5E-04	1.000	0.196	0.967
		(Chr. 5)	**Female**	**5.6E-04**	**1.000**	**<0.001**	**0.061**
ENSG00000251661	AC136475.1	Novel transcript, antisense to	Male	-1.4E-04	1.000	0.362	0.979
		IFITM3 (Chr. 11)	**Female**	**5.1E-04**	**1.000**	**<0.001**	**0.064**
ENSG00000244879	GABPB1-AS1	GABPB1 antisense RNA 1	Male	1.4E-05	1.000	0.888	0.989
		(Chr. 15)	**Female**	**4.8E-04**	**1.000**	**<0.001**	**0.070**
ENSG00000260025	CRIM1-DT	CRIM1 divergent transcript	Male	-7.3E-05	1.000	0.390	0.979
		(Chr. 2)	**Female**	**5.0E-04**	**1.000**	**0.001**	**0.070**
ENSG00000249456	AL731577.2	Novel transcript, sense	Male	-1.5E-04	1.000	0.360	0.979
		overlapping ZRANB1 (Chr. 10)	**Female**	**-9.8E-04**	**0.999**	**0.001**	**0.082**
ENSG00000259661	AC068831.5	Novel transcript, antisense to	Male	1.3E-04	1.000	0.673	0.985
		MAN2A2 (Chr. 15)	**Female**	**8.9E-04**	**1.001**	**0.001**	**0.082**

* Gene symbols and descriptions from HGNC database
† The FDR cutoff for significance was <0.10

CHAPTER 4: Associations of maternal pre-pregnancy obesity with placental DNA methylation and the placental transcriptome

Abstract

Background: Maternal pre-pregnancy overweight (ppOW) or obesity (ppOB) status impacts the course of pregnancy and increases the risk of maternal pregnancy complications and adverse fetal growth related outcomes. The placental transcriptome and related epigenetic regulatory mechanisms may mediate some of these effects, yet are understudied. We investigated overall and infant-sex specific associations of ppOB or ppOW with placental DNA methylation (DNAme), mRNA, and long non-coding RNA (lncRNA).

Methods: Mother-child participants of NICHD's Fetal Growth Studies who provided placental samples were included in genome-wide placental DNA methylation- (N = 301) and transcriptomic-specific (N = 75) analyses. Placental DNAme and RNA were assayed using Illumina Infinium Human Methylation 450 Beadchip arrays and an Illumina HiSeq2000, respectively. In linear models, CpG-site-specific methylation was included as an outcome and either categorical ppOB (BMI≥30 kg/m^2) or ppOW (BMI≥25 kg/m^2) were included as exposures. Models controlled for *a priori*-selected confounding and experimental/precision variables. We constructed similar main-effect models for transcriptome-wide expression of mRNAs or lncRNAs. Further, we considered effect modification of associations by infant-sex using infant-sex stratified models and main-effects models with an added interaction term. A false discovery rate (FDR<0.10) was applied to account for multiple testing.

Results: On average, mothers were 27.7 years old (SD=5.3) with 24.5 kg/m^2 BMI (SD=4.7) pre-pregnancy. Approximately 36% of mothers were overweight and 10% were obese. ppOB was associated with higher methylation of four CpG sites, cg11844079[*PIGM*], cg22591875[*MEDAG*], cg27278787[*DLGAP1*], cg04069951[*CD81*], and lower methylation of cg04426977[*FAM107B*] (all *BACON*-FDR<0.05). ppOB was also associated with higher expression of the ACE2 gene (logFC=0.922, FDR=0.033). In sex-stratified models, ppOB was associated with 3.1-97.6% lower methylation of cg09858237[*WDR16*; *STX8*], cg15606914[*H1FOO*] and cg00311799[*CTSF*] among male infants, and ppOB was associated with lower expression of lncRNA RP4-555D20.2 among females (logFC=-1.699, FDR=0.17). Significant infant-sex–ppOB interaction was observed for cg08956510[*NME1*] (BACON-interaction-FDR=0.017), with non-significant positive ppOB-methylation association observed among males and inverse ppOB-methylation association among females (logFCs=0.509 and -1.695, respectively). ppOW was not associated with DNA methylation, mRNA or lncRNA after FDR correction.

Conclusions: We identified associations of ppOB with placental methylation of genes related to transcription control, adipogenesis, and control of cell growth and migration. Further, we observed association of ppOB with expression of *ACE2* and the lncRNA *RP4-552D20.2*. We also found that associations may vary by infant-sex. These findings expanded on previous investigations and affirmed that epigenetic mechanisms likely play a role in the impact of ppOB on course of pregnancy and fetal development.

4.1 Introduction

Maternal pre-pregnancy obesity (ppOB) and overweight status (ppOW) have been linked to maternal pregnancy complications[23,28] and poor offspring outcomes[172]. As of 2019, approximately two-thirds of US mothers were overweight (BMI $\geq$ 25 kg/m^2) and 29% were obese (BMI $\geq$ 30 kg/m^2) at the start of pregnancy[20]. Placenta-mediated consequences of ppOB or ppOW are of key interest, as proper placental development and function is essential to both maternal health and fetal development[67,127]. Findings of prior research indicate that ppOB and ppOW are associated with the dimensions (i.e. weight, breadth and thickness) of the placenta[129,130], overall rates of cell proliferation and apoptosis in the placenta[134], as well as its capacity to properly vascularize[131–133]. The altered molecular changes in placenta related to maternal weight[28] involve broad changes in expression of genes (mRNA) in numerous pathways that influence angiogenesis, lipid transport, and stress response[34,35,37,128].

Gene expression is controlled by numerous epigenetic mechanisms that can promote or restrict the transcription of DNA into RNA or regulate the translation of mRNA into protein products. DNA methylation (DNAme) of nucleotide-level CpG sites is the quintessential example of the former, while lncRNAs (long non-coding RNAs) are an example of the latter regulation mechanisms. In several studies, including one by our group[61], placental DNAme levels have been associated with pre-pregnancy BMI (ppBMI), though most of these previous investigations were restricted to a relatively small list of candidate CpG sites[173–178], examined only global placental DNAme[179] (vs. genome-wide DNAme), or were limited in sample size[180].

LncRNAs are long RNA sequences (greater than 200 base pairs) that are not translated into proteins, instead playing regulatory roles over the course of mRNA transcription and

translation[45,135,181]. These transcripts impact placental function[49,77,79] and placental lncRNAs have been shown to be associated with pregnancy complications[54,78,81], yet their association with ppOB or ppOW is unclear. To our knowledge, no published study investigated genome-wide placental lncRNAs and mRNAs in the context of ppOB or ppOW.

We investigated overall and infant-sex specific associations of maternal ppBMI, specifically ppOB or ppOW status, with site-specific placental CpG methylation as well as transcriptome-wide lncRNA and mRNA expression. We also examined associations of differentially methylated CpGs (related to ppOB or ppOW) with expression of nearby transcriptome elements (mRNA and lncRNAs).

4.2 Methods

Study Setting and Study Participants

Participants in this study included singleton births from the National Institute of Child Health and Development's (NICHD) Fetal Growth Studies (FGS). FGS, described previously[61,66], is a longitudinal birth cohort containing 2,802 mother-child dyads recruited between 2009 and 2013 from twelve medical facilities across the states of NY, DE, NJ, SC, AL, IL, CA, RI and MA. To be eligible, women needed to plan to deliver at one of the participating hospitals, be between 8 – 13 weeks into gestation, 18 – 40 years old, and have no confirmed or suspected congenital anomalies in their pregnancies. At the time of recruitment, women were separated into two groups based upon calculated ppBMI, singleton low-risk (BMI 19.0 – 29.9 kg/m^2, N = 2,334) and singleton obese (BMI 30.0 – 45.0 kg/m^2, N = 468).

During routine prenatal study visits when fetal anthropometric information was collected, mothers were interviewed using standardized questionnaires to gather demographic information, reproductive history, and health behaviors. Maternal BMI (in kg/m^2) was computed from maternal height and pre-pregnancy weight self-reported at these study visits. Gestational age was confirmed via early pregnancy (ranging between 8-13 weeks) ultrasound. A subset of 312 women from both recruitment groups provided placental samples at time of birth, providing the basis for our investigation. All participants provided written informed consent, and all study protocols were reviewed and approved by institutional review board participating hospitals and the NICHD.

Extraction and Measurement of Placental DNAme and RNA

Placental samples were collected within one hour of delivery from the fetal side of the placenta just below the fetal surface. Biopsies 0.5 cm by 0.5 cm by 0.5 cm were placed in RNALater before freezing, and placental DNA and RNA were extracted using methods previously described[42,182]. To measure placental DNAme levels, DNA was assayed on Illumina Infinium Human Methylation 450 Beadchip arrays (Illumina Inc., San Diego, CA), and Illumina protocols were followed to normalize probes prior to computation of CpG-specific average β values for and detection *P*-values in Illumina's genome studio from methylation intensity files. CpGs or probes with detection *P*-values ≥0.05 and CpGs within 20 base pairs of known single nucleotide polymorphisms (SNPs) were removed, leaving 409,101 CpGs for potential analysis[61]. Additionally, 11 placental samples were excluded for discrepancies in phenotypic and genotypic sex, outliers in genetic clusters or mismatched sample identifiers. This left a final sample size of N = 301 for DNAme analyses that included 31 women from the

singleton obese population and 271 from the singleton low-risk population. Finally, methylation β values were logit-transformed to the M-scale for all modeling analyses, which report the change in methylation beta value between the exposed and unexposed groups or for a 1 kg/m^2 increase in BMI in ppBMI models[183].

Full-transcriptome RNA was isolated from a subset of 82 placental biopsies, of which 75 passed initial quality control (11 singleton obese, 64 singleton low-risk) using TRIZOL reagent (Invitrogen, MA, USA)[42]. RNA was enriched via Poly-A pull-down prior to preparing cDNA libraries with the Illumina TruSeq RNA kit (Illumina Inc., San Diego, CA), which were then sequenced using an Illumina HiSeq2000 machine at a setting of 100 bp paired-end reads. RTA (Illumina, San Diego, CA, USA) and bcl2fastq (version 1.8.4) were used to base-call and convert output BCL files to the FASTQ format, and reads were mapped using Tophat (version 2.0.4) to a reference genome (NCBI/build37.2). Relative expression level of genes was estimated using cufflinks (version 2.0.2) on default settings, and the output FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values were log2-transformed following quantile normalization, ultimately producing 57,773 total distinct RNA transcripts. Prior to analysis, the 20,327 available mRNAs and 13,860 available lncRNAs (across multiple lncRNA classes) were separately processed using the Bioconductor *edgeR* package's Trimmed Mean of M-values (TMM) function in R[90] and filtered by upon mean logCPM values. For mRNAs, transcripts were separated from the full 57,773 by annotation class, followed by removal of transcripts with unreliably low expression (in this case mean log counts/million (logCPM) ≤ 0), leaving 12,561 mRNAs for normalization with the TMM function. LncRNAs were first normalized using the full RNA library then similarly filtered according to an average logCPM ≤ 0, leaving 936 primarily lincRNA-type (long-intergenic non-coding RNA) transcripts available for analysis.

Placental Cell-type Estimation

Estimates of placental cell-type composition were produced for each subject by comparing their placental DNAme profiles to third-trimester reference trophoblast, stromal, Hofbauer, endothelial, syncytiotrophoblast and nRBC cell methylomes contained in the *planet* R package[184]. CpGs analyzed were restricted to those available in the *planet* third-trimester data and cell-type proportions were then estimated using the *EpiDISH* R package's robust partial correlations (epidish) function[185], one of two cell deconvolution methods suggested in *planet* documentation. These cell proportions were intended to supplement – though not replace – prior control of cell heterogeneity with surrogate variable analysis (SVA)[93].

Statistical Analyses

First, descriptive analyses were conducted to compare the DNAme and RNA populations, with t-tests and X^2 tests of independence used to measure differences in continuous and categorical covariates, respectively ($\alpha = 0.05$ for significance). Next, DNAme differences between placentas of mothers with higher and lower BMIs were compared using weighted linear regression models fit using *limma* in R. Models regressed maternal DNAme M-values at each CpG site on one exposure, either maternal overweight status (ppOW, BMI ≥ 25) compared to normal weight (ppBMI < 25) or ppOB (BMI ≥ 30) compared to non-obese weight (ppBMI < 30). Both types of models controlled for *a priori*-identified variables including maternal age (in years), maternal education (less than high school, high school, some college, bachelor's degree, master's degree or higher), maternal ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian & pacific islander), infant-sex, DNA methylation plate, ten genotype-based principal components, and estimated placental cell-type proportions. However,

syncytiotrophoblast proportions were excluded from models due to a high degree of negative correlation with other cell types in this population, particularly trophoblasts. DNAme analyses were performed with complete-cases for all exposures and covariates, so the final analytic sample size was N = 295 for all main models after 6 subjects were removed for missing maternal BMI information. Surrogate variable analysis (SVA) was added to all methylation models to further control for cell-type heterogeneity and overall variability in methylation samples[93]. All analyses used Benjamini-Hockberg correction[94] to adjust for multiple testing as a part of the *limma-voom* computational pipeline in R (FDR < 0.05), and *BACON*-adjustments were applied to FDR values to correct for genomic inflation (λ)[186]. A *BACON*-corrected FDR < 0.05 was therefore used to determine statistical significance. Similar to methods in Shrestha et al. (2020)[61], quantile-quantile (Q-Q) P-value distributions plots and λ values were compared before and after *BACON*-adjustment, and all λ post-correction values were close to 1.

Sex-specific associations of placental CpG methylation with ppOB and ppOW were then explored using two approaches. First, infant-sex stratified models were constructed using *limma* for male or female infants that regressed site-specific CpG methylation on ppOB or ppOW. These models excluded control for infant-sex but otherwise used identical covariate control to main models. Second, interaction of ppBMI exposures and infant-sex on CpG methylation was tested by adding interaction terms for ppOB/ppOW and infant-sex to main-effect models. Both model types utilized a false discovery rate and *BACON*-corrected FDR values < 0.05 to determine significance.

Next, transcriptome-wide LncRNA and mRNA expression in the RNA subset were separately regressed on ppBMI exposure variables. Each transcript that passed quality checks received its own weighted linear regression model using the *limma-voom* computational pipeline

in R[91]. While similar to the modeling methods used for CpGs, the additional *voom* step for RNA transcripts estimates observation-level weights from the mean-variance relationship of the transcripts that are subsequently added into the linear regression models. Models regressed transcript expression of RNAs or lncRNAs separately on ppOB, ppOW, or ppBMI (in kg/m^2), and likewise controlled *a priori* for maternal age, education, ethnicity, infant-sex, the ten genotype-based principal components, and estimated placental cell-type proportions (excluding syncytiotrophoblasts). SVA was also used in all models to further control for cell-type heterogeneity and overall variability in RNA samples. Given that *BACON*-correction is not typically used in RNA-seq modeling, all models simply used Benjamini-Hockberg correction for multiple comparisons (FDR < 0.05) as the measure of statistical significance. Due to missing BMI for two subjects, the final sample size for analyses was N = 73.

Additionally, infant-sex specific association of mRNA expression and lncRNA transcription was also considered in a similar format to effect-modification analyses for CpGs. This included separate models for male and female infants similarly constructed to those mentioned above for ppOB, ppOW and ppBMI, as well as models that integrated interaction terms of these variables with infant-sex into main models.

Given that the inclusion of placental cell type estimates was not a model feature available in prior FGS DNAme analyses, we also conducted a sensitivity analysis that regressed site-specific placental DNAme on continuous ppBMI, as previously seen in Shrestha et al., 2020, to compare findings of our otherwise highly similar models with those of the prior investigation[61]. BMI models likewise included SVA prior to modeling and integrated *BACON*-corrected FDR values as the measure of significance. Further, to expand upon the prior investigation, additional infant-sex stratified models and models with interaction terms for ppBMI and infant-sex were

constructed with identical covariate control to ppOW and ppOB models. As before, all models estimating CpG methylation applied a *BACON*-Corrected FDR < 0.05 to determine significance.

For any DNAme CpG sites found significant in main and secondary analyses, pre-transcriptional regulation of mRNA and lncRNA levels by DNAme was examined. In-range RNA candidates were determined by matching CpG locations in Illumina's 450K annotation files (available via the R *Bioconductor* package) to genomic start locations of mRNA and lncRNA transcripts in Ensembl's GRCH37 annotation file; transcripts within 200 kilobases up- or downstream of CpGs were considered. Weighted linear regression models were then fit for expression of mRNAs or expression of lncRNAs on DNAme levels of identified CpGs, controlling for the same covariate set as primary models and restricting results to only in-range transcripts. As in previous transcriptome analyses, an FDR < 0.05 was used to determine statistical significance, though SVA was not included at this step. As in prior RNA analyses, this analysis was restricted to subjects with both DNAme and RNA data (N = 73). All analyses were performed in R (v 4.1.2).

4.3 Results

Mothers in the DNAme analytic population had an average age at interview of 27.7 years (SD = 5.3) and gestational age of 39.5 weeks (SD = 1.1) at time of birth (**Table 4-1**). The largest proportion of the population was Hispanic mothers (34%), followed by White (26%), Black (24%) and Asian & Pacific Islander (17%). This distribution was similar in the RNA subset, though with slightly higher proportions of White and Black mothers and a decreased proportion

of Asian and Pacific Islander mothers. Male and female infants were approximately equally represented. Another recent study in this cohort compared the placental sample population (N = 301) with FGS mothers lacking placental samples (N = 2,501)[187]; except for the latter's mean gestational age (39.1 weeks, SD = 2.0, p < 0.001), no covariates differed significantly at α = 0.05. The mean ppBMI for the DNAme group and the RNA subset were 24.5 kg/m2 (SD = 4.7) and 25.1 kg/m2 (SD = 5.5), with 10.3% and 14.7% of the groups being overweight and obese, respectively (**Table 4-2**).

Compared to pregnant women with normal ppBMI, women with ppOB had higher methylation at four CpG sites (cg11844079[*PIGM*], cg22591875[*C13orf33*], cg04069951[*DLGAP1*], and cg04426977[*CD81*]; 11.2–36.9% increase) and lower methylation of one CpG site (cg04426977[*FAM107B*]; 15.3% lower methylation)(*BACON*-FDR range = 0.011 – 0.041) (**Table 4-3**). ppOW was not significantly associated with methylation at any CpG sites.

In infant-sex stratified models, ppOB was associated with lower methylation of three CpG sites (cg09858237[*WDR16*; *STX8*], cg15606914[*H1FOO*] and cg00311799[*CTSF*]; change in methylation beta estimates ranged from -3.1% for cg15606914[*H1FOO*] to -97.4% for cg09858237[*WDR16*; *STX8*] among male infants (*BACON*-FDRs = 0.002, 0.017, 0.048, respectively) (**Table 4-3**). Statistically significant interaction was observed for ppOB and infant-sex on one CpG, cg08956510[*NME1*; *NME1-NME2*] methylation levels (*BACON*-FDR = 0.017) (**Table 4-3**). In sex-stratified analyses, methylation of cg08956510[*NME1*; *NME1-NME2*] was 42.3% greater in placentas of male infants and 69.1% lower in female infants, though neither association was statistically significant after FDR correction (**Supplemental Table 3**). ppOW was not associated with CpG-site methylation.

In the RNA subset, ppOB was associated with more than two-fold higher expression of the *ACE2* mRNA (logFC = 1.052, FDR = 0.005) (**Table 4-3**). ppOB, ppOW and ppBMI were not associated with placental mRNA or lncRNA in overall or infant-sex specific models with one exception: female infants of obese mothers had 69.2% lower expression of one uncharacterized lncRNA, *RP4-55D20.2* (FDR = 0.017) (**Table 4-3**).

Sensitivity analysis examining associations of continuous BMI with CpG-site DNA methylation – with the addition of placental cell-type control to the prior investigation's models[61] – identified four CpG sites (cg09861871[*ARFGAP3*], cg21246968[*KDELR2*], cg25418406[*RANGRF*; *SLC25A35*], and cg02937776[*COX16*]) with lower methylation per 1 kg/m^2 unit higher BMI (**Table 4-4**). Secondary analysis of ppBMI found 31 CpGs were differentially methylated among male infants, including cg21246968[*KDELR2*], one of 27 CpGs whose methylation was inversely associated with maternal BMI (**Table 4-5**). Additionally, cg09861871[*ARFGAP3*] methylation was significantly inversely associated with BMI among female infants (**Table 4-5**) and the infant-sex interaction term was significant (**Table 4-4**).

4.4 Discussion

In the current study, we found overall and infant-sex specific associations of ppOB with placental DNAme, mRNA expression, and lncRNAs. We identified differential placental DNAme at five CpGs sites (cg11844079[*PIGM*], cg22591875[*C13orf33*], cg04069951[*DLGAP1*], cg04426977[*CD81*] and cg04426977[*FAM107B*]) in relation to ppOB. Additionally, we found several potential infant-sex specific associations between ppOB and

placental DNAme. Methylation at three CpG sites (cg09858237[*WDR16*; *STX8*], cg15606914[*H1FOO*] and cg00311799[*CTSF*]) was decreased in relation to ppOB among male infants, although ppOB–infant-sex interaction terms were not significant. Significant infant-sex interaction with ppOB was observed for methylation at cg08956510[*NME1*; *NME1-NME2*] with a positive relationship among male infants and an inverse relationship among female infants (both associations not statistically significant). Placental *ACE2* expression was positively associated with ppOB, while the largely uncharacterized lncRNA *RP4-555D20.2* was inversely associated with ppOB among female infants.

This study is among the first to comprehensively examine associations of ppBMI with placental EWAS and the first study, to our knowledge, to combine such analysis with transcriptome-wide examination of placental lncRNAs[61,64,173–180]. Prior work in this area included work by Michels et al. (2011), who examined global placental LINE-1 element methylation (N = 319) in relation to multiple infant and maternal factors yet did not observe maternal obesity to be significantly associated with LINE-1 methylation[173]. Two studies by Lesseur et al. subsequently examined methylation of 23 loci within the promoter region of the leptin (LEP) gene. The authors observed 2.3% increased methylation in the *LEP* gene among obese mothers in the first study (N = 79, p = 0.23)[174], and 1.17% increased methylation in obese mothers in the second (N = 473, p = 0.06)[175]. Nogues et al. expanded their DNAme comparisons of ppOB to also include adiponectin, and observed hypomethylation of the adiponectin promoter and hypermethylation of the leptin promoter, the latter supporting prior studies' findings[177]. In our investigation, we observed slight decreases in DNAme across 15 measured *LEP* CpG sites, yet none of these CpGs were observed to be *BACON*-FDR significant. More akin to our findings, another 2014 study by Haghiac and colleagues observed little to no change in

adiponectin promoter methylation between normal and obese women (N = 133)[176]. Further investigations expanded the scope of DNAme inquiry, with Nomura et al. conducting a pilot study that found placentas of obese mothers had about 2.5% increased global methylation (n = 50, p = 0.01)[179]. Similarly, in their EWAS study (N = 42), Mitsuya et al. observed a 21% increase in global placental methylation and a 31% decrease in hydroxymethylation among their obese group compared to their lean BMI group[180]. The authors identified 8 genes (*GH2*, *CSHL1*, *PSG2*, *MIA*, *OGFOD1*, *TET1*, *TET2* and *IDH2*) and 2 gene clusters (*GH/CSH* and *PSG*) associated with increased methylation, one gene with decreased methylation, *CMTM1*, and three with decreased hydroxymethylation (not clear from the text) within 100 bp of transcription start sites (TSS)[180]. None of these genes were differentially methylated in relation to ppOB in our study and, further, did not display consistent directionality of methylation with Mitsuya's findings[180]. ppOB has also been directly linked to upregulation of *ANGPTL2*, a pro-inflammatory cytokine, and downregulation of *HIFIA*, *VEGFA*, *HTA-TIP2*, and *PROK2*, genes involved in angiogenesis processes[34,35,37]. While our study did not observe any of these genes to be significantly associated with ppOB or ppOW status –our CpG and other transcriptomic associations allude to perturbations in some of these pathways. In contrast to placental gene expression, to our knowledge, there has been no published transcriptome-wide analyses of placental lncRNA transcription in relation to ppOB or ppBMI, making this the first analysis of its kind.

Several CpG sites potentially relevant to pregnancy and fetal development were associated with ppOB in our main effects models. First, cg11844079 is a CpG found in the TSS200 region of *PIGM*, a housekeeping gene whose transcription by the transcription factor *Sp1*, or lack thereof under mutated conditions[188,189], regulates levels of

glycosylphosphatidylinositol. Notably, *Sp1* also upregulates *LEP* transcription, which is upregulated in placenta in maternal obesity and altered in many pregnancy complications[28]. cg22591875 is an isolated CpG in the 3'UTR of *C13orf33*, recently classified as *MEDAG*, mesenteric estrogen-dependent adipogenesis gene. Adipocyte differentiation and nutrient uptake has been positively associated with BMI, breast cancer[190,191] and type-2 diabetes development[192]. Similarly, expression of *FAM107B*, a heat shock-inducible tumor suppressor regulated by cg04426977 – one of the CpG sites identified in our study – has been observed to oppose aberrant tumor growth in multiple types of cancers[193,194]. The other CpG we identified in our study in relation to ppOB, cg27278787, is proximal to *DLGAP1*, linked to lipoatrophy[195] or promote resistant hypertension phenotypes[196]. For example, Zhang et al. observed *Dlgap1* overexpression opposed the browning of white adipocytes *in vitro*, counterproductive in opposing the effects of obesity on adipose tissue health[197]. Finally, our findings included cg04069951, a CpG located on an island in the TSS1500 region of the *CD81* gene. CD81, a surface tetraspanin protein, serves in intercellular exosome communications in the placenta[198,199] and *CD81* expression is linked to abnormal placental function, with up-regulation inhibiting cytotrophoblast invasion of maternal decidua in early-onset severe preeclampsia[200]. Our study also observed that the *ACE2* gene, angiotensin-converting enzyme-2, to be nearly doubled in expression in obese placentas compared to normal placentas. *ACE2* codes for a transmembrane protein with a peptidase N-terminal domain that – among other roles – cleaves angiotensin II into various other combinations of angiotensin proteins, thereby triggering vasodilation and vasoprotection cascades[201]. These mechanisms have been described in relation to several pregnancy complications (including preeclampsia) that complicate pregnancies among obese and overweight women[202].

Our study also pointed to several infant-sex specific associations of ppOB with placental DNA methylation, similar to some past studies[139,203–205], as well as sex-specific association of ppOB with placental lncRNA. Male and female placentae are known to differ in placental methylation patterns without any exogenous exposure considerations[139,203,204], so our findings are not unexpected. In the current study, cg08956510, which was differently methylated with ppOB (positive among males and inverse among females, though not statistically significant), is proximal to the *NME1* gene. *NMEI* has been observed to function in cell motility and adhesion in melanoma[206,207] and, when expressed at the maternal-fetal interface, downregulates the invasiveness of human extravillous cytotrophoblasts, in part through interaction with various MAPK pathways[208]. We also found that among males, methylation in cg09858237, that is linked to *STX8* detected in blood of women who were subject to preterm birth and preeclampsia[209]. On a more sex-specific level, *WDR16*, the second gene proximal to cg09858237 is a transcript associated with flagellum development in male germ cells[210] and, at a more broad developmental level, situs inversus (the lateral flipping of all internal organs during development) when it is missing[211]. Further, *H1FOO*, a gene close to cg15606914, that was associated with ppOB among male infants, is a linker histone that establishes loose chromatin structure as a part of oocyte maturation, thereby having a more distinctly female consequence to methylation[212]. Additionally, our study also observed inverse association of the lncRNA transcript *RP4-55D20.2* with ppOB in placentas of female infants in our study. RP4-552D20.2 is an uncharacterized long intergenic transcript on the third chromosome that is noted, however, to be within the transcript family of miR-138. MiR-138 targets HIF-1α (hypoxia-inducible factor 1α) whose expression is elevated in preeclamptic pregnancies[213–215].

In our sensitivity analysis comparing inclusion of placental cell proportions in ppBMI

models to similar models in Shrestha et al. (2020), we observed four CpGs to be significantly associated with increasing ppBMI (cg09861871[*ARFGAP3*], cg21246968[*KDELR2*], cg25418406[*RANGRF*; *SLC25A35*], and cg02937776[*COX16*]). These CpGs were not among the three previously found to be associated with maternal BMI by Shrestha et al.[61], cg14568196[*EGFL7*], cg15339142[*VETZ*], and cg02301019[*AC092377.1*], which were all significantly associated with ppBMI by p-value in our study (p = <0.001, 0.002, and 0.002, respectively) but did not remain significant following FDR and BACON-corrections (BACON-FDRs = 0.481, 0.729, 0.737, respectively). All three CpGs were also associated with ppOB (all p-values < 0.05) and cg14568196[*EGFL7*] and cg02301019[*AC092377.1*] were associated with ppOW (p-values < 0.05), but these associations were likewise not significant after multiple testing adjustment. The lack of direct overlap between main effect findings in this study and those of Shrestha et al. suggest that the three CpGs highlighted in Shrestha et al. may represent particularly potent cell-type-specific methylation effects at extreme ppBMI levels. In contrast, our study may have identified CpGs consistently altered across the placental cell-types, though further sequencing studies to examine these distinctions are warranted. Interestingly, two main-effect associations in our sensitivity analysis, cg09861871 and cg21246968, associated with expression of genes *ARFGAP3* and *KDELR2*, respectively, were also differentially methylated in an infant-sex specific manner with increasing ppBMI (**Tables 4-4** and **4-5**). *ARFGAP3* is known to regulate golgi apparatus processes in an androgen sensitive manner[216,217], and *KDELR2* promotes cell proliferation in cancer cells[218,219]. Some of these cancer proliferation pathways integrated mTOR, a key regulator of trophoblast growth and proliferation[220] whose dysregulation has been linked to poor fetal[154] and maternal[221] outcomes in pregnancy. Beyond sensitivity analysis, our exploration of RNA transcripts within 200 kilobases of CpGs associated with

maternal obesity and maternal BMI found expression of several mRNAs to be associated at $\alpha < 0.05$ with methylation levels at their proximal CpGs. Apart from cg11844079, which was associated with expression of *TAGLN2*, *IGSF8*, and *IGSF9*, methylation of other CpGs was selectively associated with single genes (**Supplemental Table 4**). However, none of these CpG/gene-expression associations remained significant after FDR adjustment.

This study had several prominent strengths. FGS was originally assembled for the purpose of creating fetal growth and size-for-gestational age standards for the United States, and therefore intentionally selected from a national cross-section of non-Hispanic White, non-Hispanic Black, Hispanic and Asian or Pacific Islander mothers to participate. As a result, robust DNAme and transcriptomic measurements were applied to diverse populations that tend to be underrepresented in omic analyses, particularly US Hispanics and Blacks. This combination of molecular measurements also made this the first study of its kind to simultaneously conduct full EWAS and TWAS style analyses on DNAme, mRNA and lncRNAs in relation to ppOB and ppBMI. However, there are also study limitations that need to be mentioned. First, given that analyses were conducted on placental samples obtained at birth, all analyses are cross-sectional. Second, a common limitation of placental studies is placental cell heterogeneity. We attempted to account for this in our study design by including novel control for estimated placental cell proportions[184], and further added control for genotype-based principal components for population stratification[222], Bayesian methods for genomic inflation and multiple testing[186], as well SVA for control of residual model variability[93]. Despite these controls, the problem of high inter-cohort and even inter-individual variability placental signatures is well known for both DNAme[223] and lncRNA[126] analyses in particular, so it is possible some residual bias remained uncontrolled despite our efforts. The RNA subset of this

cohort was also substantially smaller than several other recent transcriptomic populations, so the lower power of the subset may account in large part for the relatively few observed mRNA and lncRNA associations. It is also important to consider that the DNAme Microarray chips utilized in this study, while covering the full breadth of the genome, only cover <2% of all CpG sites (many of which are tightly clustered in CpG islands). The relative lack of correlation between our omic measures should be interpreted in this context. Additionally, the unique design of this study population – as compared to other molecular analysis cohorts – makes it difficult to replicate previously reported findings. Our self-reported weight and height values used for BMI calculations could also introduce some inaccuracy due to reporting biases. Further, while relatively few significant covariate differences exist between the placental sample portion of FGS and the remaining FGS population it is also possible women self-selected into the placental cohort in a differential fashion, introducing selection bias. Finally, while this FGS analytic population is among the largest of its kind, even larger combined methylation and transcriptomic cohorts may be required to observe more subtle relationships between ppOB/ppOW and placental DNA methylation and transcriptomes.

In summary, we found five placental DNAme-sites that were associated with ppOB , as well as one mRNA and one lncRNA similarly associated with ppOB in TWAS models. Identified methylation sites and transcripts have roles that include histone acetylation, adipocyte differentiation, neuronal regulation, placental exosome release, the proliferation and migration of cancer cells, and vasoprotection/vasodilation. Further, we observed potential infant-sex specific associations of placental DNA methylation with ppOB for sites that have function in cancer development, neurological disorders, and early human development. Our findings support and

build upon prior literature regarding the impact of maternal BMI on placental epigenomics and transcriptomics, and offer potential targets for future replication and functional studies.

4.5 Tables

Table 4-1: Select Characteristics of the Study Population

	Analytic Samples	
Population Covariates	**DNAme** (N=301)	**RNA** (N=75)
Continuous	**Mean (SD)**	**Mean (SD)**
Maternal Age (years)	27.7 (5.3)	27.7 (6.0)
Gestational Age (weeks)	39.5 (1.1)	39.3 (1.3)
Categorical	**N (%)**	**N (%)**
Maternal Race/Ethnicity		
White	77 (25.6)	23 (30.7)
Black	72 (23.9)	20 (26.7)
Hispanic	102 (33.9)	24 (32.0)
Asian & Pacific Islander	50 (16.6)	8 (10.7)
Offspring Sex		
Male	152 (50.5)	39 (52.0)
Female	149 (49.5)	36 (48.0)
Maternal Education		
Less than High School	37 (12.3)	12 (16.0)
High School or Equivalent	56 (18.6)	21 (28.0)
Some College or Associates Degree	86 (28.6)	18 (24.0)
Bachelor's degree	74 (24.6)	16 (21.3)
Master's Degree or Higher	48 (15.9)	8 (10.7)

Table 4-2: Maternal Pre-pregnancy Body Mass Index (BMI) Distributions

			BMI Distribution						
Exposure Variable	Overweight - N(%)*	Obese - N(%)*	Mean	SD	Min	25%	Median	75%	Max
DNAme Population									
Maternal Pre-Pregnancy BMI	107 (35.5)	31 (10.3)	24.5	4.7	18.0	21.2	23.5	26.9	43.9
RNAseq Population									
Maternal Pre-Pregnancy BMI	25 (33.3)	11 (14.7)	25.1	5.5	18.4	21.1	23.4	28.2	43.9

* Overweight status is defined by a BMI ≥ 25; obese status is defined by a BMI ≥ 30

Table 4-3: Placental CpG Sites and RNA Transcripts Associated with Maternal Pre-pregnancy Obesity

CpG/ENSEMBL ID	Chr.	Position	UCSC RefGene Name/Symbol*	Infant-sex	logFC	Fold Change	P-value	FDR P-value	Bacon FDR
CpG Methylation									
cg11844079	1	160001979	PIGM		0.453	1.369	<0.001	0.014	0.011
cg22591875	13	31499033	C13orf33		0.319	1.248	<0.001	0.014	0.011
cg27278787	18	3499813	DLGAP1		0.153	1.112	<0.001	0.016	0.013
cg04069951	11	2398336	CD81		0.235	1.177	<0.001	0.053	0.041
cg04426977	10	14575365	FAM107B		-0.239	0.847	<0.001	0.058	0.041
Infant-Sex Interaction									
cg08956510	17	49231229	NME1; NME1-NME2				<0.001	0.028	0.017
Infant-Sex Specific Associations									
cg09858237	17	9479929	WDR16; STX8	**Male**	**-5.261**	**0.026**	**<0.001**	**0.014**	**0.002**
				Female	0.103	1.074	0.919	0.994	1.000
cg15606914	3	129262026	H1FOO	**Male**	**-0.203**	**0.869**	**<0.001**	**0.056**	**0.017**
				Female	-0.055	0.963	0.290	0.878	0.967
cg00311799	11	66331609	CTSF	**Male**	**-0.244**	**0.844**	**<0.001**	**0.123**	**0.048**
				Female	0.033	1.023	0.556	0.950	0.996
mRNA (Gene) Expression									
ENSG00000130234	X	15579156	ACE2		0.922	1.895	<0.001	0.033	
LncRNA Transcription (Sex-specific Associations)									
ENSG00000261786	3	44158791	RP4-555D20.2	Male	1.201	2.299	0.097	0.898	
				Female	**-1.699**	**0.308**	**<0.001**	**0.017**	

* CpG symbols and descriptions from Illumina 450k database; gene symbols and descriptions from HGNC database

Table 4-4: CpG Sites Differentially Methylated in relation to Maternal Pre-pregnancy BMI

CpG	Chr.	Position	UCSC RefGene Name	logFC	Fold Change	P-Value	FDR P-Value	Bacon FDR
Main Associations								
cg09861871	22	43253517	ARFGAP3	-0.061	0.959	<0.001	0.040	0.017
cg21246968	7	6523810	KDELR2	-0.185	0.880	<0.001	0.040	0.017
cg25418406	17	8191793	RANGRF; SLC25A35	-0.084	0.944	<0.001	0.040	0.017
cg02937776	14	70826572	COX16	-0.123	0.918	<0.001	0.091	0.044
Infant-sex Interaction								
cg11207127	4	186666487	SORBS2			<0.001	0.012	0.002
cg15805921	6	42952109	PPP2R5D			<0.001	0.012	0.002
cg05467167	17	45000793	GOSR2			<0.001	0.035	0.010
cg09861871	22	43253517	ARFGAP3			<0.001	0.035	0.010
cg13137577	11	73018312	ARHGEF17			<0.001	0.035	0.010
cg13771376	9	135546871	DDX31; GTF3C4			<0.001	0.035	0.027

* CpG symbols and descriptions from Illumina 450k database

Table 4-5: Infant-sex Specific Differentially Methylated CpG Sites Associated with Maternal Pre-pregnancy BMI

CpG	Chr.	Position	UCSC RefGene Name	Infant-sex	logFC	Fold Change	P-Value	FDR P-Value	Bacon FDR
Maternal BMI - (Male)									
cg11832210	10	91295346	SLC16A12	**Male**	**-0.174**	**0.886**	**<0.001**	**0.026**	**0.003**
				Female	-0.001	0.999	0.973	0.999	1.000
cg12803694	19	15443052		**Male**	**-0.183**	**0.881**	**<0.001**	**0.039**	**0.008**
				Female	-0.034	0.977	0.437	0.957	0.990
cg22274539	12	103696209	C12orf42	**Male**	**-0.078**	**0.947**	**<0.001**	**0.039**	**0.008**
				Female	0.003	1.002	0.811	0.988	0.996
cg11445323	19	18063178	KCNN1	**Male**	**-0.130**	**0.914**	**<0.001**	**0.040**	**0.010**
				Female	0.017	1.012	0.258	0.926	0.980
cg18090431	4	37691603		**Male**	**-0.288**	**0.819**	**<0.001**	**0.040**	**0.010**
				Female	-0.021	0.986	0.737	0.984	0.996
cg27494615	2	9614744	IAH1	**Male**	**-0.133**	**0.912**	**<0.001**	**0.040**	**0.010**
				Female	-0.024	0.984	0.369	0.948	0.988
cg16023551	5	16616834	FAM134B	**Male**	**-0.161**	**0.894**	**<0.001**	**0.040**	**0.011**
				Female	0.059	1.042	0.223	0.918	0.977
cg11869233	2	48009809	MSH6	**Male**	**0.055**	**1.039**	**<0.001**	**0.073**	**0.016**
				Female	-0.018	0.987	0.020	0.787	0.890
cg21246968	7	6523810	KDELR2	**Male**	**-0.288**	**0.819**	**<0.001**	**0.066**	**0.018**
				Female	-0.130	0.914	0.010	0.755	0.855
cg03725447	1	860008	SAMD11	**Male**	**-0.063**	**0.957**	**<0.001**	**0.073**	**0.025**
				Female	0.002	1.001	0.835	0.990	0.997
cg01100175	2	149645880	KIF5C	**Male**	**-0.048**	**0.967**	**<0.001**	**0.073**	**0.026**
				Female	0.012	1.009	0.170	0.905	0.973
cg18553354	12	96588123	ELK3	**Male**	**-0.115**	**0.924**	**<0.001**	**0.075**	**0.027**
				Female	-0.027	0.982	0.318	0.938	0.984
cg19146902	11	47516584	CUGBP1	**Male**	**0.023**	**1.016**	**<0.001**	**0.085**	**0.027**

				Female	0.001	1.001	0.850	0.991	0.997
cg12807628	8	11838843	DEFB135	**Male**	**0.030**	**1.021**	**<0.001**	**0.085**	**0.033**
				Female	0.007	1.005	0.330	0.941	0.986
cg06323897	1	98386683	DPYD	**Male**	**-0.179**	**0.883**	**<0.001**	**0.085**	**0.034**
				Female	0.061	1.043	0.034	0.819	0.921
cg12781778	6	157097800	ARID1B	**Male**	**-0.024**	**0.984**	**<0.001**	**0.085**	**0.034**
				Female	0.005	1.004	0.410	0.954	0.989
cg18143317	11	2037124		**Male**	**-0.018**	**0.988**	**<0.001**	**0.085**	**0.034**
				Female	-0.002	0.999	0.622	0.975	0.994
cg19474239	13	21348215	N6AMT2	**Male**	**-0.270**	**0.829**	**<0.001**	**0.085**	**0.034**
				Female	0.091	1.065	0.116	0.886	0.967
cg21440829	7	100808893	VGF	**Male**	**-0.136**	**0.910**	**<0.001**	**0.085**	**0.034**
				Female	0.055	1.039	0.010	0.751	0.851
cg21761988	19	39421489	MRPS12; SARS2	**Male**	**-0.107**	**0.929**	**<0.001**	**0.085**	**0.034**
				Female	-0.016	0.989	0.377	0.949	0.988
cg04842215	14	24899033	CBLN3; KHNYN	**Male**	**-0.147**	**0.903**	**<0.001**	**0.085**	**0.036**
				Female	0.020	1.014	0.309	0.937	0.984
cg05495011	2	224477964		**Male**	**-0.181**	**0.882**	**<0.001**	**0.085**	**0.038**
				Female	-0.021	0.986	0.661	0.979	0.995
cg06001237	20	62694736	TCEA2	**Male**	**-0.130**	**0.914**	**<0.001**	**0.085**	**0.038**
				Female	-0.002	0.998	0.940	0.997	0.999
cg09247949	6	24667244	ACOT13; TTRAP	**Male**	**-0.146**	**0.904**	**<0.001**	**0.085**	**0.038**
				Female	0.033	1.023	0.127	0.891	0.969
cg18326616	2	74010554	C2orf78	**Male**	**-0.038**	**0.974**	**<0.001**	**0.085**	**0.038**
				Female	0.011	1.007	0.127	0.891	0.969
cg19640589	20	57427973	GNAS	**Male**	**-0.025**	**0.983**	**<0.001**	**0.085**	**0.038**
				Female	0.002	1.001	0.717	0.983	0.995
cg14827643	4	3365442	RGS12	**Male**	**-0.036**	**0.976**	**<0.001**	**0.098**	**0.045**
				Female	-0.007	0.995	0.458	0.960	0.991

cg15469171	1	23541345		**Male**	**-0.041**	**0.972**	**<0.001**	**0.105**	**0.048**
				Female	0.026	1.018	0.039	0.827	0.928
cg17775091	2	66654137		**Male**	**-0.041**	**0.972**	**<0.001**	**0.105**	**0.048**
				Female	-0.001	0.999	0.937	0.997	0.999
cg20295671	22	22090486	YPEL1	**Male**	**-0.137**	**0.909**	**<0.001**	**0.105**	**0.048**
				Female	-0.007	0.995	0.723	0.983	0.995
cg14036479	7	155250220	EN2	**Male**	**0.060**	**1.042**	**<0.001**	**0.133**	**0.048**
				Female	-0.005	0.996	0.685	0.981	0.995
Maternal BMI - (Female)									
cg09861871	22	43253517	ARFGAP3	Male	-0.016	0.989	0.098	0.944	0.956
				Female	**-0.123**	**0.918**	**<0.001**	**0.089**	**0.029**

* CpG symbols and descriptions from Illumina 450k database

CHAPTER 5: Discussion

Overall Summary of Findings

Our investigation of $PM_{2.5}$ exposure and placental lncRNAs in the ECHO PATHWAYS cohorts yielded several overall and potentially infant-sex specific associations. In CANDLE we observed two transcripts, *LINC00702* and *AC105345.1*, to be differentially transcribed in relation to $PM_{2.5}$ in the first and second trimesters, only one of which had been previously characterized. We found significant interaction of $PM_{2.5}$ exposure during the full pregnancy period with infant-sex for one transcript, *LINC00339*. Additionally, second-trimester $PM_{2.5}$ exposure was associated with 45 lncRNAs among female CANDLE infants. This list included prominent transcripts *XIST*, *NEAT1* and *MALAT1* that have been previously associated with normal placental function and dysfunction. The only two infant-sex specific associations of $PM_{2.5}$ with lncRNA in GAPPS were observed for exposure during the first trimester among male infants. Further, placental association of lncRNA with birthweight (in the context of $PM_{2.5}$ analysis) was limited to a single uncharacterized lncRNA, *AC104083.1*, among male infants in CANDLE. The lncRNA main-effect and interaction findings associated with $PM_{2.5}$ in CANDLE were primarily linked to cancer phenotypes, yet sex-specific findings among female infants in CANDLE were found to relate to a wide variety of placental pathways involving cell growth, cell proliferation and stress response.

In CANDLE and GAPPS investigations of associations of pre-pregnancy ppOB and ppBMI with lncRNAs, we observed no overall associations. In infant-sex stratified models, three lncRNAs (*ERVH48-1*, *AC139099.1*, and *CEBPA-DT*) implicated in placental growth, trophoblast differentiation and gene expression control were associated with ppBMI among male

infants in CANDLE, while one lncRNA, *ZNF225-AS1*, was associated with ppOB among female infants. Additionally, two transcripts (*AP000879.1* and *AL365203.2*) were associated with ppBMI among male infants in GAPPS. While no infant-sex and BMI or birthweight interactions on lncRNAs were significant, we found infant-sex specific associations of 17 lncRNAs with birthweight among female infants in GAPPS. These female-related transcripts, which included previously described lncRNAs such as *KANSL1-AS1, DANCR, GABPB1-AS1 and EPB41L4A-AS1*, have been linked to gene expression control, cell-cycle control, cell differentiation, and cancer development.

Finally, epigenome- and transcriptome-wide investigation of the FGS cohort revealed associations of maternal ppOB with placental CpG methylation. For instance, differential methylation of 5 CpG sites, cg11844079[*PIGM*], cg22591875[C13orf33], cg04069951[*DLGAP1*], cg04426977[*CD81*] and cg04426977[*FAM107B*], was observed in ppOB placentas compared to non-ppOB placentas. Further, we found sex-specific associations of ppOB with methylation at three CpGs, cg09858237[*WDR16*; *STX8*], cg15606914[*H1FOO*] and cg00311799[*CTSF*], with ppOB among male infants, and infant-sex significantly interacted with ppOB on methylation of cg08956510[*NME1*; *NME1-NME2*]. Transcriptomic associations were more limited: one gene, *ACE2*, was observed to be differentially expressed in ppOB placentas in main-effect models, while one uncharacterized lncRNA, *RP4-555D20.2*, was inversely associated with ppOB among male infants.

Strengths and Limitations

This dissertation project had several strengths. First, each of these cohorts were large, well-characterized populations with a robust array of covariates prospectively collected during

pregnancy and at birth. This allowed for control for a wide range of potential confounding variables based on *a priori* specifications in models. Second, the CANDLE and GAPPS cohorts utilized state-of-the-art spatiotemporal models to produce accurate estimates of residential $PM_{2.5}$ exposure across multiple exposure windows throughout pregnancy, a feature that has not been consistently present in prior investigations. The availability of short exposure windows (2 weeks) has also allowed examination of potential critical perinatal windows of exposure. Each of these cohorts featured a broad array of placental biomarkers, whether full-transcriptomic RNA measurements in CANDLE and GAPPS or full-transcriptome RNA and full-epigenome DNAme in FGS, allowing for broad TWAS and EWAS analyses of the placenta. To our knowledge, this investigation included the first TWAS analyses of placental lncRNA levels in relation to both $PM_{2.5}$ and ppOB, as well as the first genome-wide investigation of ppOB's combined impact on lncRNAs and DNAme in the placenta. Further, the sample size of these investigations was similar to or greatly surpassed that of previous investigations of this type, providing greater power to the analyses. Finally, the study cohorts, particularly CANDLE and FGS, included participants from diverse populations, including community members that are typically underrepresented in transcriptomic and epigenetic studies of the placenta. Additionally, the diversity of study participants across geographic regions bodes well for potential generalizability of findings.

This dissertation project also has several limitations that deserve mention. First, placental sampling occurred at birth. Prior research has confirmed that the placental transcriptome changes as fetal and maternal requirements over the course of pregnancy[125], so it is not clear whether findings that relate to late-pregnancy (near delivery) can be extrapolated to impacts on placental transcriptome or epigenome in early pregnancy, particularly for early pregnancy

exposures. In addition, while we controlled for mode of delivery, we cannot rule out whether the associations are related to placental processes modulated specifically by factors related to delivery (e.g. medications). Second, the placenta is a heterogenous mixture of multiple cell types whose prominence varies across the duration of pregnancy, and given that methods in all three studies involved homogenization of placental samples, it is difficult to determine the potential cell-type specific impacts of our exposures of interest. This was something that all three aims attempted to control for to varying degrees with statistical methods, e.g. through use of surrogate variable analysis[93] or, in the case of DNAme, direct estimation of placental cell type proportions[184]. Third, inter-cohort variation in several population parameters may have contributed to the differences in observed associations in this investigation. This has limited potential scenarios of replication. While we attempted to control for confounding and precision factors in models, DNAme of specific CpGs[223] and lncRNA transcription of select transcripts[126] are known to vary on an inter-individual basis, which can contribute to variable findings in even apparently similar populations. Differences in study power may also have contributed to Type II errors and some of the observed differences in statistically significance. While CANDLE and FGS are among the larger studies of lncRNAs and DNAme, respectively, FGS and GAPPS both had lesser power by comparison in transcriptome analyses, which could have contributed to Type I error and inability to identify subtle – yet true – associations of lncRNAs with our exposures of interest. Further, given that each of the study populations with placental samples was a subset of larger pregnancy cohorts, it is possible that some degree of selection bias occurred despite the apparent similarity of these subsets to their parent cohorts. Possible inaccurate self-reporting of exposures (e.g. BMI) as well as other covariates could also have led to misclassification and biased findings. Finally, these cohorts were also not tailored toward downstream functional

analyses, which requires future investigations in other cohorts. Findings of these individual studies will only likely be generalizable to populations with pre-pregnancy obesity, $PM_{2.5}$ exposure or birthweight distributions and similar population characteristics.

Future Research Recommendations

The investigations conducted as part of this dissertation project and its findings put forth a number of opportunities and suggestions for future research. First, replication studies among similarly sampled populations are required to both confirm and clarify the associations observed in our investigation. Second, comprehensive lncRNA-mRNA co-expression and gene-network analyses (akin to that of de Goede et al.[45]) should be conducted in CANDLE, GAPPS and FGS to build upon prior findings and quantify broad trends in RNA transcription and its regulation that are linked to $PM_{2.5}$, ppOB, and/or birthweight. Third, future investigations should examine placental cellular localization for all uncharacterized lncRNA transcripts highlighted in both PATHWAYS and FGS cohorts using fluorescence tags, which will allow for generation of hypotheses as to the molecular targets – if any – of these transcripts. Fourth, in the case that clear localization patterns are identified for various lncRNA transcripts, antibody elution investigations can be conducted *in vitro* based upon candidate binding targets. In the case of clear protein matches, transcripts can then be examined in either knockdown or over-expression follow-up experiments to determine the functional role of transcripts in the regulation of their binding partners. Finally, future cohort studies should include examination of associations of exposure-related placental transcriptome and epigenome markers with other placental phenotypes (e.g. size, pathology, etc.), and offspring outcomes. This will provide critical mechanistic information to better understand disease pathogenesis. While it was not feasible

during the development of our cohorts, future cohorts should also aim for use of single-cell sequencing methods to provide clearer mechanisms for morphological changes in the placenta. It is also possible that circulating blood markers might also be examined as a reflection of the placental level transcriptome and epigenome markers, providing information on the development and functions of the placenta before delivery (including in early pregnancy)[224]. Further, any associations of transcriptome or epigenome markers with placental dimensions should also be examined in an infant-sex specific manner.

In conclusion, the findings of this dissertation project provide evidence regarding placental transcriptomic and regulatory epigenetic processes related to maternal $PM_{2.5}$ exposure and ppOB. These mechanistic insights on relationships between these exposures and fetal growth and development can inform preventive and therapeutic strategies. Promoting an optimal course and outcomes of pregnancy can lead to favorable outcomes that have implications for lifelong health.

REFERENCES

1. Hanson MA, Gluckman PD. Early Developmental Conditioning of Later Health and Disease: Physiology or Pathophysiology? *Physiological Reviews*. 2014;94(4):1027-1076. doi:10.1152/physrev.00029.2013

2. Burton GJ, Jauniaux E. What is the placenta? *American Journal of Obstetrics and Gynecology*. 2015;213(4, Supplement):S6.e1-S6.e4. doi:10.1016/j.ajog.2015.07.050

3. Barker DJP, Thornburg KL, Osmond C, Kajantie E, Eriksson JG. Beyond birthweight: the maternal and placental origins of chronic disease. *Journal of Developmental Origins of Health and Disease*. 2010;1(6):360-364. doi:10.1017/S2040174410000280

4. Jansson T, Powell TL. Role of the placenta in fetal programming: underlying mechanisms and potential interventional approaches. *Clinical Science*. 2007;113(1-2):1-13. doi:10.1042/CS20060339

5. Burton GJ, Fowden AL, Thornburg KL. Placental Origins of Chronic Disease. *Physiological Reviews*. 2016;96(4):1509-1565. doi:10.1152/physrev.00029.2015

6. Goldstein JA, Gallagher K, Beck C, Kumar R, Gernand AD. Maternal-Fetal Inflammation in the Placenta and the Developmental Origins of Health and Disease. *Front Immunol*. 2020;11:531543. doi:10.3389/fimmu.2020.531543

7. Bongaerts E, Nawrot TS, Van Pee T, Ameloot M, Bove H. Translocation of (ultra)fine particles and nanoparticles across the placenta; a systematic review on the evidence of in vitro, ex vivo, and in vivo studies. *Part Fibre Toxicol*. 2020;17(1):56. doi:10.1186/s12989-020-00386-8

8. Tekin D, Kayaaltı Z, Aliyev V, Söylemezoğlu T. The effects of metallothionein 2A polymorphism on placental cadmium accumulation: is metallothionein a modifiying factor in transfer of micronutrients to the fetus? *Journal of Applied Toxicology*. 2012;32(4):270-275. doi:10.1002/jat.1661

9. Kalisch-Smith JI, Simmons DG, Dickinson H, Moritz KM. Review: Sexual dimorphism in the formation, function and adaptation of the placenta. *Placenta*. 2017;54:10-16. doi:10.1016/j.placenta.2016.12.008

10. Misra DP, Salafia CM, Miller RK, Charles AK. Non-Linear and Gender-Specific Relationships Among Placental Growth Measures and The Fetoplacental Weight Ratio☆. *Placenta*. 2009;30(12):1052-1057. doi:10.1016/j.placenta.2009.09.008

11. Cassee FR, Héroux ME, Gerlofs-Nijland ME, Kelly FJ. Particulate matter beyond mass: recent health evidence on the role of fractions, chemical constituents and sources of emission. *Inhalation Toxicology*. 2013;25(14):802-812. doi:10.3109/08958378.2013.850127

12. Feng S, Gao D, Liao F, Zhou F, Wang X. The health effects of ambient PM2.5 and potential mechanisms. *Ecotoxicology and Environmental Safety*. 2016;128:67-74. doi:10.1016/j.ecoenv.2016.01.030

13. Stieb DM, Chen L, Eshoul M, Judek S. Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. *Environmental Research*. 2012;117:100-111. doi:10.1016/j.envres.2012.05.007

14. Bekkar B, Pacheco S, Basu R, DeNicola N. Association of Air Pollution and Heat Exposure With Preterm Birth, Low Birth Weight, and Stillbirth in the US A Systematic Review. *JAMA Netw Open*. 2020;3(6):e208243. doi:10.1001/jamanetworkopen.2020.8243

15. Li X, Huang S, Jiao A, et al. Association between ambient fine particulate matter and preterm birth or term low birth weight: An updated systematic review and meta-analysis. *Environ Pollut*. 2017;227:596-605. doi:10.1016/j.envpol.2017.03.055

16. Shah PS, Balkhair T. Air pollution and birth outcomes: A systematic review. *Environ Int*. 2011;37(2):498-516. doi:10.1016/j.envint.2010.10.009

17. Luyten LJ, Saenen ND, Janssen BG, et al. Air pollution and the fetal origin of disease: A systematic review of the molecular signatures of air pollution exposure in human placenta. *Environmental Research*. 2018;166:310-323. doi:10.1016/j.envres.2018.03.025

18. Dadvand P, Parker J, Bell ML, et al. Maternal exposure to particulate air pollution and term birth weight: a multi-country evaluation of effect and heterogeneity. *Environ Health Perspect*. 2013;121(3):267-373. doi:10.1289/ehp.1205575

19. Pedersen M, Giorgis-Allemand L, Bernard C, et al. Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med*. 2013;1(9):695-704. doi:10.1016/S2213-2600(13)70192-9

20. Driscoll AK, Gregory ECW. *Increases in Prepregnancy Obesity: United States, 2016–2019*. National Center for Health Statistics, U.S. Department of Health and Human Services; 2020:8. https://www.cdc.gov/nchs/products/index.htm

21. Chen C, Xu X, Yan Y. Estimated global overweight and obesity burden in pregnant women based on panel data model. *PLOS ONE*. 2018;13(8):e0202183. doi:10.1371/journal.pone.0202183

22. Thangaratinam S, Rogozinska E, Jolly K, et al. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. *BMJ-British Medical Journal*. 2012;344:e2088. doi:10.1136/bmj.e2088

23. Huda SS, Brodie LE, Sattar N. Obesity in pregnancy: prevalence and metabolic consequences. *Seminars in Fetal and Neonatal Medicine*. 2010;15(2):70-76. doi:10.1016/j.siny.2009.09.006

24. Yu Z, Han S, Zhu J, Sun X, Ji C, Guo X. Pre-Pregnancy Body Mass Index in Relation to Infant Birth Weight and Offspring Overweight/Obesity: A Systematic Review and Meta-Analysis. Baradaran HR, ed. *PLoS ONE*. 2013;8(4):e61627. doi:10.1371/journal.pone.0061627

25. Boney CM, Verma A, Tucker R, Vohr BR. Metabolic syndrome in childhood: Association with birth weight, maternal obesity, and gestational diabetes mellitus. *Pediatrics*. 2005;115(3):E290-E296. doi:10.1542/peds.2004-1808

26. Ziemke F, Mantzoros CS. Adiponectin in insulin resistance: lessons from translational research1234. *Am J Clin Nutr*. 2010;91(1):258S-261S. doi:10.3945/ajcn.2009.28449C

27. Higgins L, Greenwood SL, Wareing M, Sibley CP, Mills TA. Obesity and the placenta: A consideration of nutrient exchange mechanisms in relation to aberrant fetal growth. *Placenta*. 2011;32(1):1-7. doi:10.1016/j.placenta.2010.09.019

28. Kelly AC, Powell TL, Jansson T. Placental function in maternal obesity. *Clinical Science*. 2020;134(8):961-984. doi:10.1042/CS20190266

29. Saenen ND, Plusquin M, Bijnens E, et al. In Utero Fine Particle Air Pollution and Placental Expression of Genes in the Brain-Derived Neurotrophic Factor Signaling Pathway: An ENVIRONAGE Birth Cohort Study. *Environmental Health Perspectives*. 2015;123(8):834-840. doi:10.1289/ehp.1408549

30. Kingsley SL, Deyssenroth MA, Kelsey KT, et al. Maternal residential air pollution and placental imprinted gene expression. *Environ Int*. 2017;108:204-211. doi:10.1016/j.envint.2017.08.022

31. Kim W, Cho Y, Song MK, et al. Effect of particulate matter 2.5 on gene expression profile and cell signaling in JEG-3 human placenta cells. *Environmental Toxicology*. 2018;33(11):1123-1134. doi:10.1002/tox.22591

32. Deyssenroth MA, Rosa MJ, Eliot MN, et al. Placental gene networks at the interface between maternal PM2.5 exposure early in gestation and reduced infant birthweight. *Environmental Research*. 2021;199:111342. doi:10.1016/j.envres.2021.111342

33. Enquobahrie DA, MacDonald J, Hussey M, et al. Prenatal exposure to particulate matter and placental gene expression. *Environment International*. 2022;165:107310. doi:10.1016/j.envint.2022.107310

34. Saben J, Lindsey F, Zhong Y, et al. Maternal obesity is associated with a lipotoxic placental environment. *Placenta*. 2014;35(3):171-177. doi:10.1016/j.placenta.2014.01.003

35. Lassance L, Haghiac M, Leahy P, et al. Identification of early transcriptome signatures in placenta exposed to insulin and obesity. *American Journal of Obstetrics and Gynecology*. 2015;212(5):647.e1-647.e11. doi:10.1016/j.ajog.2015.02.026

36. Bari MF, Ngo S, Bastie CC, Sheppard AM, Vatish M. Gestational diabetic transcriptomic profiling of microdissected human trophoblast. *Journal of Endocrinology*. 2016;229(1):47-59. doi:10.1530/JOE-15-0424

37. Altmäe S, Segura MT, Esteban FJ, et al. Maternal Pre-Pregnancy Obesity Is Associated with Altered Placental Transcriptome. *PLOS ONE*. 2017;12(1):e0169223. doi:10.1371/journal.pone.0169223

38. Sureshchandra S, Marshall NE, Wilson RM, et al. Inflammatory Determinants of Pregravid Obesity in Placenta and Peripheral Blood. *Frontiers in Physiology*. 2018;9. doi:10.3389/fphys.2018.01089

39. Cox B, Tsamou M, Vrijens K, et al. A Co-expression Analysis of the Placental Transcriptome in Association With Maternal Pre-pregnancy BMI and Newborn Birth Weight. *Frontiers in Genetics*. 2019;10. doi:10.3389/fgene.2019.00354

40. Roberts TC, Coenen-Stass AML, Wood MJA. Assessment of RT-qPCR Normalization Strategies for Accurate Quantification of Extracellular microRNAs in Murine Serum. *PLOS ONE*. 2014;9(2):e89237. doi:10.1371/journal.pone.0089237

41. Badon SE, Littman AJ, Chan KCG, et al. Physical activity and epigenetic biomarkers in maternal blood during pregnancy. *Epigenomics*. 2018;10(11):1383-1395. doi:10.2217/epi-2017-0169

42. Delahaye F, Do C, Kong Y, et al. Genetic variants influence on the placenta regulatory landscape. *PLOS Genetics*. 2018;14(11):e1007785. doi:10.1371/journal.pgen.1007785

43. Shrestha D, Rahman ML, Hinkle SN, Workalemahu T, Tekola-Ayele F. Maternal BMI-Increasing Genetic Risk Score and Fetal Weights among Diverse US Ethnic Groups. *Obesity*. 2019;27(7):1150-1160. doi:10.1002/oby.22499

44. Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. *Nat Rev Mol Cell Biol*. 2018;19(3):143-157. doi:10.1038/nrm.2017.104

45. de Goede OM, Nachun DC, Ferraro NM, et al. Population-scale tissue transcriptomics maps long non-coding RNAs to complex disease. *Cell*. 2021;184(10):2633-2648.e19. doi:10.1016/j.cell.2021.03.050

46. El Gayed EMA, Zaid IF, El Gayed AMA, Zaki AMM, Fouda EA allah M. Biochemical study on long non coding RNA gene expression in women having preeclampsia. *Gene Reports*. 2021;24:101257. doi:10.1016/j.genrep.2021.101257

47. Chen H, Meng T, Liu X, et al. Long non-coding RNA MALAT-1 is downregulated in preeclampsia and regulates proliferation, apoptosis, migration and invasion of JEG-3 trophoblast cells. *Int J Clin Exp Pathol*. 2015;8(10):12718-12727.

48. Gong F, Cheng H, Shi Y, Cui L, Jia G. LncRNA TDRG1/miR-214-5p axis affects preeclampsia by modulating trophoblast cells. *Cell Biochemistry and Function*. 2020;38(4):352-361. doi:https://doi.org/10.1002/cbf.3480

49. Chu T, Mouillet JF, Cao Z, Barak O, Ouyang Y, Sadovsky Y. RNA Network Interactions During Differentiation of Human Trophoblasts. *Front Cell Dev Biol*. 2021;09:677981. doi:10.3389/fcell.2021.677981

50. Fan X, Lou J, Zheng X, et al. Interference with lncRNA NEAT1 promotes the proliferation, migration, and invasion of trophoblasts by upregulating miR-411-5p and inhibiting PTEN expression. *Immunopharmacology and Immunotoxicology*. 2021;43(3):334-342. doi:10.1080/08923973.2021.1910834

51. Wu HY, Wang XH, Liu K, Zhang JL. LncRNA MALAT1 regulates trophoblast cells migration and invasion via miR-206/IGF-1 axis. *Cell Cycle*. 2019;19(1):39-52. doi:10.1080/15384101.2019.1691787

52. Lapehn S, Paquette AG. The Placental Epigenome as a Molecular Link Between Prenatal Exposures and Fetal Health Outcomes Through the DOHaD Hypothesis. *Curr Envir Health Rpt*. 2022;9(3):490-501. doi:10.1007/s40572-022-00354-8

53. Machtinger R, Zhong J, Mansur A, et al. Placental lncRNA Expression Is Associated With Prenatal Phthalate Exposure. *Toxicological Sciences*. 2018;163(1):116-122. doi:10.1093/toxsci/kfy013

54. Long W, Rui C, Song X, et al. Distinct expression profiles of lncRNAs between early-onset preeclampsia and preterm controls. *Clinica Chimica Acta*. 2016;463(Supplement C):193-199. doi:10.1016/j.cca.2016.10.036

55. Paquette AG, MacDonald J, Lapehn S, et al. A Comprehensive Assessment of Associations between Prenatal Phthalate Exposure and the Placental Transcriptomic Landscape. *Environ Health Perspect*. 2021;129(9):097003. doi:10.1289/EHP8973

56. Hussey MR, Burt A, Deyssenroth MA, et al. Placental lncRNA expression associated with placental cadmium concentrations and birth weight. *Environ Epigenet*. 2020;6(1). doi:10.1093/eep/dvaa003

57. Song GY, Na Q, Wang D, Qiao C. Microarray expression profile of lncRNAs and mRNAs in the placenta of non-diabetic macrosomia. *Journal of Developmental Origins of Health and Disease*. 2018;9(2):191-197. doi:10.1017/S2040174417000927

58. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacol*. 2013;38(1):23-38. doi:10.1038/npp.2012.112

59. Janssen BG, Godderis L, Pieters N, et al. Placental DNA hypomethylation in association with particulate air pollution in early life. *Part Fibre Toxicol*. 2013;10:11. doi:10.1186/1743-8977-10-22

60. Saenen ND, Vrijens K, Janssen BG, et al. Lower Placental Leptin Promoter Methylation in Association with Fine Particulate Matter Air Pollution during Pregnancy and Placental Nitrosative Stress at Birth in the ENVIRONAGE Cohort. *Environ Health Perspect*. 2017;125(2):262-268. doi:10.1289/EHP38

61. Shrestha D, Ouidir M, Workalemahu T, Zeng X, Tekola-Ayele F. Placental DNA methylation changes associated with maternal prepregnancy BMI and gestational weight gain. *Int J Obes*. 2020;44(6):1406-1416. doi:10.1038/s41366-020-0546-2

62. Monteagudo-Sanchez A, Sanchez-Delgado M, Hernandez Mora JR, et al. Differences in expression rather than methylation at placenta-specific imprinted loci is associated with intrauterine growth restriction. *Clin Epigenetics*. 2019;11:35. doi:10.1186/s13148-019-0630-4

63. Maghbooli Z, Hossein-nezhad A, Adabi E, et al. Air pollution during pregnancy and placental adaptation in the levels of global DNA methylation. *PLoS One*. 2018;13(7):e0199772. doi:10.1371/journal.pone.0199772

64. Opsahl JO, Moen GH, Qvigstad E, Böttcher Y, Birkeland KI, Sommer C. Epigenetic signatures associated with maternal body mass index or gestational weight gain: a systematic review. *J Dev Orig Health Dis*. 2020;12(3):373-383. doi:10.1017/S2040174420000811

65. LeWinn KZ, Karr CJ, Hazlehurst M, et al. Cohort profile: the ECHO prenatal and early childhood pathways to health consortium (ECHO-PATHWAYS). *BMJ Open*. 2022;12(10):e064288. doi:10.1136/bmjopen-2022-064288

66. Grewal J, Grantz KL, Zhang C, et al. Cohort Profile: NICHD Fetal Growth Studies–Singletons and Twins. *International Journal of Epidemiology*. 2018;47(1):13.

67. Gude NM, Roberts CT, Kalionis B, King RG. Growth and function of the normal human placenta. *Thrombosis Research*. 2004;114(5-6):397-407. doi:10.1016/j.thromres.2004.06.038

68. Klepac P, Locatelli I, Korošec S, Künzli N, Kukec A. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environmental Research*. 2018;167:144-159. doi:10.1016/j.envres.2018.07.008

69. Bearblock E, Aiken CE, Burton GJ. Air pollution and pre-eclampsia; associations and potential mechanisms. *Placenta*. 2021;104:188-194. doi:10.1016/j.placenta.2020.12.009

70. Saenen ND, Vrijens K, Janssen BG, et al. Placental Nitrosative Stress and Exposure to Ambient Air Pollution During Gestation: A Population Study. *Am J Epidemiol*. 2016;184(6):442-449. doi:10.1093/aje/kww007

71. Veras MM, Damaceno-Rodrigues NR, Caldini EG, et al. Particulate Urban Air Pollution Affects the Functional Morphology of Mouse Placenta1. *Biology of Reproduction*. 2008;79(3):578-584. doi:10.1095/biolreprod.108.069591

72. Yue H, Ji X, Zhang Y, Li G, Sang N. Gestational exposure to PM2.5 impairs vascularization of the placenta. *Science of The Total Environment*. 2019;665:153-161. doi:10.1016/j.scitotenv.2019.02.101

73. Soto S de F, de Melo JO, Marchesi GD, et al. Exposure to fine particulate matter in the air alters placental structure and the renin-angiotensin system. *PLoS One*. 2017;12(8):e0183314. doi:10.1371/journal.pone.0183314

74. Naav A, Erlandsson L, Isaxon C, et al. Urban PM2.5 Induces Cellular Toxicity, Hormone Dysregulation, Oxidative Damage, Inflammation, and Mitochondrial Interference in the HRT8 Trophoblast Cell Line. *Front Endocrinol*. 2020;11:75. doi:10.3389/fendo.2020.00075

75. Qin Z, Hou H, Fu F, et al. Fine particulate matter exposure induces cell cycle arrest and inhibits migration and invasion of human extravillous trophoblast, as determined by an iTRAQ-based quantitative proteomics strategy. *Reproductive Toxicology*. 2017;74:10-22. doi:10.1016/j.reprotox.2017.08.014

76. Barouki R, Melen E, Herceg Z, et al. Epigenetics as a mechanism linking developmental exposures to long-term toxicity. *Environ Int*. 2018;114:77-86. doi:10.1016/j.envint.2018.02.014

77. McAninch D, Roberts CT, Bianco-Miotto T. Mechanistic Insight into Long Noncoding RNAs and the Placenta. *Int J Mol Sci*. 2017;18(7):1371. doi:10.3390/ijms18071371

78. Xufei F, Xiujuan Z, Jianyi L, Liyan Y, Ting Y, Min H. Up-regulation of LncRNA NEAT1 induces apoptosis of human placental trophoblasts. *Free Radical Research*. 2020;54(8-9):678-686. doi:10.1080/10715762.2020.1826468

79. Zhang Y, Zou Y, Wang W, et al. Down-Regulated Long Non-Coding RNA MEG3 and its Effect on Promoting Apoptosis and Suppressing Migration of Trophoblast Cells. *Journal of Cellular Biochemistry*. 2015;116(4):542-550. doi:10.1002/jcb.25004

80. Basak T, Ain R. Long non-coding RNAs in placental development and disease. *Non-coding RNA Investig*. 2019;3:14-14. doi:10.21037/ncri.2019.03.01

81. Guo Y, Gao Y, Liu S. lncRNA XIST is associated with preeclampsia and mediates trophoblast cell invasion via miR-340-5p/KCNJ16 signaling pathway. *Transplant Immunology*. 2022;74:101666. doi:10.1016/j.trim.2022.101666

82. He X, He Y, Xi B, et al. LncRNAs Expression in Preeclampsia Placenta Reveals the Potential Role of LncRNAs Contributing to Preeclampsia Pathogenesis. *PLoS One*. 2013;8(11):e81437. doi:10.1371/journal.pone.0081437

83. Hussey MR, Burt A, Deyssenroth MA, et al. Placental lncRNA expression associated with placental cadmium concentrations and birth weight. Hoyo C, ed. *Environmental Epigenetics*. 2020;6(1):dvaa003. doi:10.1093/eep/dvaa003

84. Sontag-Padilla L, Burns R, Shih R, et al. *The Urban Child Institute CANDLE Study: Methodological Overview and Baseline Sample Description*. RAND Corporation; 2015. doi:10.7249/RR1336

85. Keller JP, Olives C, Kim SY, et al. A unified spatiotemporal modeling approach for predicting concentrations of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. *Environ Health Perspect*. 2015;123(4):301-309. doi:10.1289/ehp.1408145

86. Kirwa K, Szpiro AA, Sheppard L, et al. Fine-Scale Air Pollution Models for Epidemiologic Research: Insights From Approaches Developed in the Multi-ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Curr Environ Health Rep*. 2021;8(2):113-126. doi:10.1007/s40572-021-00310-y

87. Brown J, Pirrung M, McCue LA. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*. 2017;33(19):3137-3139. doi:10.1093/bioinformatics/btx373

88. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*. 2016;34(5):525-527. doi:10.1038/nbt.3519

89. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*. 2015;4:1521. doi:10.12688/f1000research.7563.2

90. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25. doi:10.1186/gb-2010-11-3-r25

91. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29. doi:10.1186/gb-2014-15-2-r29

92. Schick SF, Blount BC, Jacob P, et al. Biomarkers of exposure to new and emerging tobacco delivery products. *Am J Physiol Lung Cell Mol Physiol*. 2017;313(3):L425-L452. doi:10.1152/ajplung.00343.2016

93. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28(6):882-883. doi:10.1093/bioinformatics/bts034

94. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

95. Huang Q, Chi Y, Deng J, et al. Fine particulate matter 2.5 exerted its toxicological effect by regulating a new layer, long non-coding RNA. *Sci Rep*. 2017;7(1):9392. doi:10.1038/s41598-017-09818-6

96. Ling XX, Zhang HQ, Liu JX, et al. LncRNA LUCAT1 Activation Mediated by the Down-regulation of DNMT1 Is Involved in Cell Apoptosis Induced by PM2.5. *Biomedical and Environmental Sciences*. 2018;31(8):608-612. doi:10.3967/bes2018.082

97. Yang J, Huo T, Zhang X, et al. Oxidative stress and cell cycle arrest induced by short-term exposure to dustfall PM2.5 in A549 cells. *Environ Sci Pollut Res*. 2018;25(23):22408-22419. doi:10.1007/s11356-017-0430-3

98. Wang Y, Tang M. Integrative analysis of mRNAs, miRNAs and lncRNAs in urban particulate matter SRM 1648a-treated EA.hy926 human endothelial cells. *Chemosphere*. 2019;233:711-723. doi:10.1016/j.chemosphere.2019.05.294

99. Jiang P, Hao S, Xie L, et al. LncRNA NEAT1 contributes to the acquisition of a tumor like-phenotype induced by PM 2.5 in lung bronchial epithelial cells via HIF-1α activation. *Environ Sci Pollut Res*. 2021;28(32):43382-43393. doi:10.1007/s11356-021-13735-7

100. Costanzo V, Bardelli A, Siena S, Abrignani S. Exploring the links between cancer and placenta development. *Open Biol*. 2018;8(6):180081. doi:10.1098/rsob.180081

101. Chen BJ, Byrne FL, Takenaka K, et al. Transcriptome landscape of long intergenic non-coding RNAs in endometrial cancer. *Gynecol Oncol*. 2017;147(3):654-662. doi:10.1016/j.ygyno.2017.10.006

102. Pan W, Han J, Wei N, Wu H, Wang Y, Sun J. LINC00702-mediated DUSP1 transcription in the prevention of bladder cancer progression: Implications in cancer cell proliferation and tumor inflammatory microenvironment. *Genomics*. 2022;114(4):110428. doi:10.1016/j.ygeno.2022.110428

103. Yu S, Li N, Wang J, et al. Correlation of Long Noncoding RNA SEMA6A-AS1 Expression with Clinical Outcome in HBV-Related Hepatocellular Carcinoma. *Clinical Therapeutics*. 2020;42(3):439-447. doi:10.1016/j.clinthera.2020.01.012

104. Wang L, Ye TY, Wu H, Chen SY, Weng JR, Xi XW. LINC00702 accelerates the progression of ovarian cancer through interacting with EZH2 to inhibit the transcription of KLF2. *Eur Rev Med Pharmacol Sci*. 2019;23(3):201-208. doi:10.26355/eurrev_201908_18648

105. Li T, Ren J, Ma J, et al. LINC00702/miR-4652-3p/ZEB1 axis promotes the progression of malignant meningioma through activating Wnt/β-catenin pathway. *Biomedicine & Pharmacotherapy*. 2019;113:108718. doi:10.1016/j.biopha.2019.108718

106. Yu W, Li D, Ding X, et al. LINC00702 suppresses proliferation and invasion in non-small cell lung cancer through regulating miR-510/PTEN axis. *Aging (Albany NY)*. 2019;11(5):1471-1485. doi:10.18632/aging.101846

107. Liu Z, Kang Z, Dai Y, Zheng H, Wang Y. Long noncoding RNA LINC00342 promotes growth of infantile hemangioma by sponging miR-3619-5p from HDGF. *American Journal*

of Physiology-Heart and Circulatory Physiology. 2019;317(4):H830-H839. doi:10.1152/ajpheart.00188.2019

108. Miao Z, Liu S, Xiao X, Li D. LINC00342 regulates cell proliferation, apoptosis, migration and invasion in colon adenocarcinoma via miR-545-5p/MDM2 axis. *Gene*. 2020;743:144604. doi:10.1016/j.gene.2020.144604

109. Wu Y, Du J. Downregulated Reprimo by LINC00467 participates in the growth and metastasis of gastric cancer. *Bioengineered*. 2022;13(5):11893-11906. doi:10.1080/21655979.2022.2063662

110. Zhang Y, Zhang Y, Wang S, et al. LINC00467 facilitates the proliferation, migration and invasion of glioma via promoting the expression of inositol hexakisphosphate kinase 2 by binding to miR-339-3p. *Bioengineered*. 2022;13(2):3370-3382. doi:10.1080/21655979.2021.2018098

111. Wu Z, Zhang S, Guo W, He Y. LINC00339: An emerging major player in cancer and metabolic diseases. *Biomedicine & Pharmacotherapy*. 2022;149:112788. doi:10.1016/j.biopha.2022.112788

112. Albertsen HM, Chettier R, Farrington P, Ward K. Genome-Wide Association Study Link Novel Loci to Endometriosis. *PLoS One*. 2013;8(3):e58257. doi:10.1371/journal.pone.0058257

113. Powell JE, Fung JN, Shakhbazov K, et al. Endometriosis risk alleles at 1p36.12 act through inverse regulation of CDC42 and LINC00339. *Human Molecular Genetics*. 2016;25(22):5046-5058. doi:10.1093/hmg/ddw320

114. Fung JN, Montgomery GW. Genetics of endometriosis: State of the art on genetic risk factors for endometriosis. *Best Practice & Research Clinical Obstetrics & Gynaecology*. 2018;50:61-71. doi:10.1016/j.bpobgyn.2018.01.012

115. Holdsworth-Carson SJ, Churchill M, Donoghue JF, et al. Elucidating the role of long intergenic non-coding RNA 339 in human endometrium and endometriosis. *Molecular Human Reproduction*. 2021;27(3):gaab010. doi:10.1093/molehr/gaab010

116. Boeren J, Gribnau J. Xist-mediated chromatin changes that establish silencing of an entire X chromosome in mammals. *Current Opinion in Cell Biology*. 2021;70:44-50. doi:10.1016/j.ceb.2020.11.004

117. Guiyu S, Quan N, Ruochen W, et al. LncRNA-SNX17 Promotes HTR-8/SVneo Proliferation and Invasion Through miR-517a/IGF-1 in the Placenta of Diabetic Macrosomia. *Reprod Sci*. 2022;29(2):596-605. doi:10.1007/s43032-021-00687-z

118. Kudaka W, Oda T, Jinno Y, Yoshimi N, Aoki Y. Cellular Localization of Placenta-Specific Human Endogenous Retrovirus (HERV) Transcripts and their Possible Implication in Pregnancy-Induced Hypertension. *Placenta*. 2008;29(3):282-289. doi:10.1016/j.placenta.2007.11.009

119. Sugimoto J, Schust DJ, Kinjo T, Aoki Y, Jinno Y, Kudo Y. Suppressyn localization and dynamic expression patterns in primary human tissues support a physiologic role in human placentation. *Sci Rep*. 2019;9:19502. doi:10.1038/s41598-019-55933-x

120. Tang L, Li P, Li L. Whole transcriptome expression profiles in placenta samples from women with gestational diabetes mellitus. *Journal of Diabetes Investigation*. 2020;11(5):1307-1317. doi:10.1111/jdi.13250

121. Eriksson JG, Kajantie E, Osmond C, Thornburg K, Barker DJP. Boys live dangerously in the womb. *Am J Hum Biol*. 2010;22(3):330-335. doi:10.1002/ajhb.20995

122. Gonzalez TL, Sun T, Koeppel AF, et al. Sex differences in the late first trimester human placenta transcriptome. *Biol Sex Differ*. 2018;9(1):4. doi:10.1186/s13293-018-0165-y

123. Fallah H, Azari I, Neishabouri SM, Oskooei VK, Taheri M, Ghafouri-Fard S. Sex-specific up-regulation of lncRNAs in peripheral blood of patients with schizophrenia. *Sci Rep*. 2019;9(1):12737. doi:10.1038/s41598-019-49265-z

124. Issler O, van der Zee YY, Ramakrishnan A, et al. Sex-specific role for the long non-coding RNA LINC00473 in depression. *Neuron*. 2020;106(6):912-926.e5. doi:10.1016/j.neuron.2020.03.023

125. Sitras V, Fenton C, Paulssen R, Vårtun Å, Acharya G. Differences in Gene Expression between First and Third Trimester Human Placenta: A Microarray Study. Oudejans C, ed. *PLoS ONE*. 2012;7(3):e33294. doi:10.1371/journal.pone.0033294

126. Ponting CP, Haerty W. Genome-Wide Analysis of Human Long Noncoding RNAs: A Provocative Review. *Annu Rev Genom Hum Genet*. 2022;23(1):annurev-genom-112921-123710. doi:10.1146/annurev-genom-112921-123710

127. Barouki R, Gluckman PD, Grandjean P, Hanson M, Heindel JJ. Developmental origins of non-communicable disease: Implications for research and public health. *Environ Health*. 2012;11(1):42. doi:10.1186/1476-069X-11-42

128. Fowden AL, Camm EJ, Sferruzzi-Perri AN. Effects of Maternal Obesity On Placental Phenotype. *CVP*. 2020;19(2):113-131. doi:10.2174/1570161118666200513115316

129. Winder NR, Krishnaveni GV, Veena SR, et al. Mother's lifetime nutrition and the size, shape and efficiency of the placenta. *Placenta*. 2011;32(11):806-810. doi:10.1016/j.placenta.2011.09.001

130. Swanson LD, Bewtra C, Swanson LD, Bewtra C. Increase in normal placental weights related to increase in maternal body mass index. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2008;21(2):111-113. doi:10.1080/14767050701866963

131. Bar J, Kovo M, Schraiber L, Shargorodsky M. Placental maternal and fetal vascular circulation in healthy non-obese and metabolically healthy obese pregnant women. *Atherosclerosis*. 2017;260:63-66. doi:10.1016/j.atherosclerosis.2017.03.006

132. Bar J, Schreiber L, Saruhanov E, Ben-Haroush A, Golan A, Kovo M. Placental histopathological findings in obese and nonobese women with complicated and uncomplicated pregnancies. *Arch Gynecol Obstet*. 2012;286(6):1343-1347. doi:10.1007/s00404-012-2450-z

133. Kovo M, Zion-Saukhanov E, Schreiber L, et al. The Effect of Maternal Obesity on Pregnancy Outcome in Correlation With Placental Pathology. *Reprod Sci*. 2015;22(12):1643-1648. doi:10.1177/1933719115592712

134. Higgins L, Mills TA, Greenwood SL, Cowley EJ, Sibley CP, Jones RL. Maternal obesity and its effect on placental cell turnover. *J Matern Fetal Neonatal Med*. 2013;26(8):783-788. doi:10.3109/14767058.2012.760539

135. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. 2021;22(2):96-118. doi:10.1038/s41580-020-00315-9

136. Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011;25(18):1915-1927. doi:10.1101/gad.17446611

137. Majewska M, Lipka A, Paukszto L, et al. Preliminary RNA-Seq Analysis of Long Non-Coding RNAs Expressed in Human Term Placenta. *Int J Mol Sci*. 2018;19(7):1894. doi:10.3390/ijms19071894

138. Lipka A, Jastrzebski JP, Paukszto L, et al. Sex-Biased lncRNA Signature in Fetal Growth Restriction (FGR). *Cells*. 2021;10(4):921. doi:10.3390/cells10040921

139. Martin E, Smeester L, Bommarito PA, et al. Sexual epigenetic dimorphism in the human placenta: implications for susceptibility during the prenatal period. *Epigenomics*. 2017;9(3):267-278. doi:10.2217/epi-2016-0132

140. Buckberry S, Bianco-Miotto T, Bent SJ, Dekker GA, Roberts CT. Integrative transcriptome meta-analysis reveals widespread sex-biased gene expression at the human fetal–maternal interface. *Mol Hum Reprod*. 2014;20(8):810-819. doi:10.1093/molehr/gau035

141. Clark J, Eaves LA, Gaona AR, et al. Pre-pregnancy BMI-associated miRNA and mRNA expression signatures in the placenta highlight a sexually-dimorphic response to maternal underweight status. *Sci Rep*. 2021;11(1):15743. doi:10.1038/s41598-021-95051-1

142. Ebrahimi R, Toolabi K, Jannat Ali Pour N, et al. Adipose tissue gene expression of long non-coding RNAs; MALAT1, TUG1 in obesity: is it associated with metabolic profile and lipid homeostasis-related genes expression? *Diabetol Metab Syndr*. 2020;12(1):36. doi:10.1186/s13098-020-00544-0

143. Liu Y, Ji Y, Li M, et al. Integrated analysis of long noncoding RNA and mRNA expression profile in children with obesity by microarray analysis. *Sci Rep*. 2018;8(1):8750-8750. doi:10.1038/s41598-018-27113-w

144. Divoux A, Karastergiou K, Xie H, et al. Identification of a novel lncRNA in gluteal adipose tissue and evidence for its positive effect on preadipocyte differentiation. *Obesity*. 2014;22(8):1781-1785. doi:10.1002/oby.20793

145. Huang Y, Jin C, Chanyuan Jin, et al. Knockdown of lncRNA MIR31HG inhibits adipocyte differentiation of human adipose-derived stem cells via histone modification of FABP4. *Scientific Reports*. 2017;7(1):8080-8080. doi:10.1038/s41598-017-08131-6

146. Puthanveetil P, Chen S, Feng B, Gautam A, Chakrabarti S. Long non-coding RNA MALAT1 regulates hyperglycaemia induced inflammatory process in the endothelial cells. *Journal of Cellular and Molecular Medicine*. 2015;19(6):1418-1425. doi:10.1111/jcmm.12576

147. Huang C, Huang B bin, Niu J min, et al. Global mRNA and Long Non-Coding RNA Expression in the Placenta and White Adipose Tissue of Mice Fed a High-Fat Diet During Pregnancy. *CPB*. 2018;50(6):2260-2271. doi:10.1159/000495086

148. Sugimoto J, Sugimoto M, Bernstein H, Jinno Y, Schust D. A novel human endogenous retroviral protein inhibits cell-cell fusion. *Sci Rep*. 2013;3:1462. doi:10.1038/srep01462

149. Guo Y, Ma Y, Hu X, Song R, Zhu L, Zhong M. Long non-coding RNA CEBPA-AS1 correlates with poor prognosis and promotes tumorigenesis via CEBPA/Bcl2 in oral squamous cell carcinoma. *Cancer Biol Ther*. 2018;19(3):205-213. doi:10.1080/15384047.2017.1416276

150. Qiao X, Liu J, Zhu L, Song R, Zhong M, Guo Y. Long noncoding RNA CEBPA-DT promotes cisplatin chemo-resistance through CEBPA/BCL2 mediated apoptosis in oral squamous cellular cancer. *Int J Med Sci*. 2021;18(16):3728-3737. doi:10.7150/ijms.64253

151. Wang L, Xiao H, Zhang X, Liao W, Fu S, Huang H. Restoration of CCAAT enhancer binding protein α P42 induces myeloid differentiation and overcomes all-trans retinoic acid resistance in human acute promyelocytic leukemia NB4-R1 cells. *International Journal of Oncology*. 2015;47(5):1685-1695. doi:10.3892/ijo.2015.3163

152. Yoon MS, Zhang C, Sun Y, Schoenherr CJ, Chen J. Mechanistic target of rapamycin controls homeostasis of adipogenesis. *Journal of Lipid Research*. 2013;54(8):2166-2173. doi:10.1194/jlr.M037705

153. Okawa ER, Gupta MK, Kahraman S, et al. Essential roles of insulin and IGF-1 receptors during embryonic lineage development. *Molecular Metabolism*. 2021;47:101164. doi:10.1016/j.molmet.2021.101164

154. Dong J, Shin N, Chen S, Lei J, Burd I, Wang X. Is there a definite relationship between placental mTOR signaling and fetal growth? *Biology of Reproduction*. 2020;103(3):471-486. doi:10.1093/biolre/ioaa070

155. Kretz M, Webster DE, Flockhart RJ, et al. Suppression of progenitor differentiation requires the long noncoding RNA ANCR. *Genes Dev*. 2012;26(4):338-343. doi:10.1101/gad.182121.111

156. Akbari M, Eshghyar F, Gholipour M, et al. Expression analysis of mTOR-associated lncRNAs in multiple sclerosis. *Metab Brain Dis*. 2022;37(6):2061-2066. doi:10.1007/s11011-022-01010-8

157. Zhang Q, Wang Z, Cheng X, Wu H. lncRNA DANCR promotes the migration an invasion and of trophoblast cells through microRNA-214-5p in preeclampsia. *Bioengineered*. 2021;12(2):9424-9434. doi:10.1080/21655979.2021.1988373

158. Gong F, Chai W, Wang J, et al. miR-214-5p suppresses the proliferation, migration and invasion of trophoblast cells in pre-eclampsia by targeting jagged 1 to inhibit notch signaling pathway. *Acta Histochemica*. 2020;122(3):151527. doi:10.1016/j.acthis.2020.151527

159. Yabuta N, Onda H, Watanabe M, et al. Isolation and characterization of the TIGA genes, whose transcripts are induced by growth arrest. *Nucleic Acids Res*. 2006;34(17):4878-4892. doi:10.1093/nar/gkl651

160. Liu SJ, Horlbeck MA, Cho SW, et al. CRISPRi-based genome-scale identification of functional long non-coding RNA loci in human cells. *Science*. 2017;355(6320):aah7111. doi:10.1126/science.aah7111

161. Hegre SA, Stovner EB. Joint changes in RNA, RNA polymerase II, and promoter activity through the cell cycle identify non-coding RNAs involved in proliferation. *Sci Rep*. 2021;11(1):18952. doi:10.1038/s41598-021-97909-w

162. Zhu Y, Liu Q, Liao M, et al. Overexpression of lncRNA EPB41L4A-AS1 Induces Metabolic Reprogramming in Trophoblast Cells and Placenta Tissue of Miscarriage. *Molecular Therapy - Nucleic Acids*. 2019;18:518-532. doi:10.1016/j.omtn.2019.09.017

163. Tani H, Torimura M. Identification of short-lived long non-coding RNAs as surrogate indicators for chemical stress response. *Biochemical and Biophysical Research Communications*. 2013;439(4):547-551. doi:10.1016/j.bbrc.2013.09.006

164. Tani H, Onuma Y, Ito Y, Torimura M. Long Non-Coding RNAs as Surrogate Indicators for Chemical Stress Responses in Human-Induced Pluripotent Stem Cells. *PLoS One*. 2014;9(8):e106282. doi:10.1371/journal.pone.0106282

165. Qi W, Li Z, Xia L, et al. LncRNA GABPB1-AS1 and GABPB1 regulate oxidative stress during erastin-induced ferroptosis in HepG2 hepatocellular carcinoma cells. *Sci Rep*. 2019;9:16185. doi:10.1038/s41598-019-52837-8

166. Chen X, Cheng Q, Du Y, Liu L, Wu H. Differential long non-coding RNA expression profile and function analysis in primary Sjogren's syndrome. *BMC Immunol*. 2021;22:47. doi:10.1186/s12865-021-00439-3

167. Su X, Li W, Lv L, et al. Transcriptome-Wide Association Study Provides Insights Into the Genetic Component of Gene Expression in Anxiety. *Frontiers in Genetics*. 2021;12. doi:10.3389/fgene.2021.740134

168. Liu H, Zong C, Sun J, et al. Bioinformatics analysis of lncRNAs in the occurrence and development of osteosarcoma. *Translational Pediatrics*. 2022;11(7):1182198-1181198. doi:10.21037/tp-22-253

169. Wang Q, Lu X, Li C, et al. Down-regulated long non-coding RNA PVT1 contributes to gestational diabetes mellitus and preeclampsia via regulation of human trophoblast cells. *Biomed Pharmacother*. 2019;120:109501. doi:10.1016/j.biopha.2019.109501

170. Everson TM, Marsit CJ. Integrating -Omics Approaches into Human Population-Based Studies of Prenatal and Early-Life Exposures. *Curr Envir Health Rpt*. 2018;5(3):328-337. doi:10.1007/s40572-018-0204-1

171. Winterbottom EF, Ban Y, Sun X, et al. Transcriptome-wide analysis of changes in the fetal placenta associated with prenatal arsenic exposure in the New Hampshire Birth Cohort Study. *Environ Health*. 2019;18(1):100. doi:10.1186/s12940-019-0535-x

172. Voerman E, Santos S, Patro Golab B, et al. Maternal body mass index, gestational weight gain, and the risk of overweight and obesity across childhood: An individual participant data meta-analysis. *PLoS Med*. 2019;16(2):e1002744. doi:10.1371/journal.pmed.1002744

173. Michels KB, Harris HR, Harris HR, Barault L. Birthweight, maternal weight trajectories and global DNA methylation of LINE-1 repetitive elements. *PLOS ONE*. 2011;6(9). doi:10.1371/journal.pone.0025254

174. Lesseur C, Armstrong DA, Paquette AG, Koestler DC, Padbury JF, Marsit CJ. Tissue-specific Leptin promoter DNA methylation is associated with maternal and infant perinatal factors. *Molecular and Cellular Endocrinology*. 2013;381(1):160-167. doi:10.1016/j.mce.2013.07.024

175. Lesseur C, Armstrong DA, Paquette AG, Li Z, Padbury JF, Marsit CJ. Maternal obesity and gestational diabetes are associated with placental leptin DNA methylation. *American Journal of Obstetrics and Gynecology*. 2014;211(6):654.e1-654.e9. doi:10.1016/j.ajog.2014.06.037

176. Haghiac M, Basu S, Presley L, Serre D, Catalano PM, Hauguel-de Mouzon S. Patterns of Adiponectin Expression in Term Pregnancy: Impact of Obesity. *The Journal of Clinical Endocrinology & Metabolism*. 2014;99(9):3427-3434. doi:2019041113392120900

177. Nogues P, Esther Dos Santos, Esther Dos Santos, et al. Maternal obesity influences expression and DNA methylation of the adiponectin and leptin systems in human third-trimester placenta. *Clinical Epigenetics*. 2019;11(1):20-20. doi:10.1186/s13148-019-0612-6

178. Workalemahu T, Shrestha D, Tajuddin SM, Tekola-Ayele F. Maternal cardiometabolic factors and genetic ancestry influence epigenetic aging of the placenta. *Journal of*

Developmental Origins of Health and Disease. 2021;12(1):34-41. doi:10.1017/S2040174419000801

179. Nomura Y, Lambertini L, Rialdi A, et al. Global Methylation in the Placenta and Umbilical Cord Blood From Pregnancies With Maternal Gestational Diabetes, Preeclampsia, and Obesity. *Reproductive Sciences*. 2014;21(1):131-137. doi:10.1177/1933719113492206

180. Mitsuya K, Parker AN, Liu L, Ruan J, Vissers MCM, Myatt L. Alterations in the placental methylome with maternal obesity and evidence for metabolic regulation. *PLOS ONE*. 2017;12(10). doi:10.1371/journal.pone.0186115

181. Ferrè F, Colantoni A, Helmer-Citterich M. Revealing protein-lncRNA interaction. *Brief Bioinform*. 2016;17(1):106-116. doi:10.1093/bib/bbv031

182. Tekola-Ayele F, Workalemahu T, Gorfu G, et al. Sex differences in the associations of placental epigenetic aging with fetal growth. *Aging-US*. 2019;11(15):5412-5432. doi:10.18632/aging.102124

183. Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010;11:587. doi:10.1186/1471-2105-11-587

184. Yuan V, Hui D, Yin Y, Peñaherrera MS, Beristain AG, Robinson WP. Cell-specific characterization of the placental methylome. *BMC Genomics*. 2021;22(1):6. doi:10.1186/s12864-020-07186-6

185. Teschendorff AE, Breeze CE, Zheng SC, Beck S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinformatics*. 2017;18(1):105. doi:10.1186/s12859-017-1511-5

186. van Iterson M, van Zwet EW, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol*. 2017;18:19. doi:10.1186/s13059-016-1131-9

187. Tesfaye M, Wu J, Biedrzycki RJ, Grantz KL, Joseph P, Tekola-Ayele F. Prenatal social support in low-risk pregnancy shapes placental epigenome. *BMC Med*. 2023;21:12. doi:10.1186/s12916-022-02701-w

188. Georgiou E, Layton M, Karadimitris A. Inherited GPI deficiency: A disorder of histone hypoacetylation. *Birth Defects Research Part C: Embryo Today: Reviews*. 2009;87(4):327-334. doi:10.1002/bdrc.20166

189. Costa JR, Caputo VS, Makarona K, et al. Cell-type–specific transcriptional regulation of PIGM underpins the divergent hematologic phenotype in inherited GPl deficiency. *Blood*. 2014;124(20):3151-3154. doi:10.1182/blood-2014-09-598813

190. Li Z, Li C, Wu Q, et al. MEDAG enhances breast cancer progression and reduces epirubicin sensitivity through the AKT/AMPK/mTOR pathway. *Cell Death Dis*. 2021;12(1):1-15. doi:10.1038/s41419-020-03340-w

191. Coradini D, Gambazza S, Oriana S, Ambrogi F. Gene expression profile of normal breast tissue and body mass index. *Breast Cancer*. 2021;28(2):488-496. doi:10.1007/s12282-020-01183-z

192. Yang J, Yu P. Identification of MEDAG as a Hub Candidate Gene in the Onset and Progression of Type 2 Diabetes Mellitus by Comprehensive Bioinformatics Analysis. *BioMed Research International*. 2021;2021:e3947350. doi:10.1155/2021/3947350

193. Nakajima H, Koizumi K, Tanaka T, et al. Loss of HITS (FAM107B) expression in cancers of multiple organs: tissue microarray analysis. *International Journal of Oncology*. 2012;41(4):1347-1357. doi:10.3892/ijo.2012.1550

194. Guo J, Bian Y, Wang Y, Chen L, Yu A, Sun X. FAM107B is regulated by S100A4 and mediates the effect of S100A4 on the proliferation and migration of MGC803 gastric cancer cells. *Cell Biology International*. 2017;41(10):1103-1109. doi:10.1002/cbin.10816

195. Uttayamakul S, Oudot-Mellakh T, Nakayama EE, et al. Genome-Wide Association Study of HIV-Related Lipoatrophy in Thai Patients: Association of a DLGAP1 Polymorphism with Fat Loss. *AIDS Research and Human Retroviruses*. 2015;31(8):792-796. doi:10.1089/aid.2014.0266

196. Takahashi Y, Yamazaki K, Kamatani Y, Kubo M, Matsuda K, Asai S. A genome-wide association study identifies a novel candidate locus at the DLGAP1 gene with susceptibility to resistant hypertension in the Japanese population. *Sci Rep*. 2021;11(1):19497. doi:10.1038/s41598-021-98144-z

197. Zhang J, Yang J, Yang N, et al. Dlgap1 negatively regulates browning of white fat cells through effects on cell proliferation and apoptosis. *Lipids Health Dis*. 2020;19(1):1-10. doi:10.1186/s12944-020-01230-w

198. Burkova EE, Dmitrenok PS, Bulgakov DV, Vlassov VV, Ryabchikova EI, Nevinsky GA. Exosomes from human placenta purified by affinity chromatography on sepharose bearing immobilized antibodies against CD81 tetraspanin contain many peptides and small proteins. *IUBMB Life*. 2018;70(11):1144-1155. doi:10.1002/iub.1928

199. Burkova EE, Grigor'eva AE, Bulgakov DV, et al. Extra Purified Exosomes from Human Placenta Contain an Unpredictable Small Number of Different Major Proteins. *International Journal of Molecular Sciences*. 2019;20(10):2434. doi:10.3390/ijms20102434

200. Shen L, Diao Z, Sun HX, et al. Supplemental.Up-regulation of CD81 inhibits cytotrophoblast invasion and mediates maternal endothelial cell dysfunction in preeclampsia. *Proc Natl Acad Sci U S A*. 2017;114(8):1940-1945. doi:10.1073/pnas.1617601114

201. Lima RS, Rocha LPC, Moreira PR. Genetic and epigenetic control of ACE2 expression and its possible role in COVID-19. *Cell Biochemistry and Function*. 2021;39(6):713-726. doi:10.1002/cbf.3648

202. Tamanna S, Clifton VL, Rae K, van Helden DF, Lumbers ER, Pringle KG. Angiotensin Converting Enzyme 2 (ACE2) in Pregnancy: Preeclampsia and Small for Gestational Age. *Front Physiol*. 2020;11. doi:10.3389/fphys.2020.590787

203. Mayne BT, Leemaqz SY, Smith AK, Breen J, Roberts CT, Bianco-Miotto T. Accelerated placental aging in early onset preeclampsia pregnancies identified by DNA methylation. *Epigenomics*. 2017;9(3):279-289. doi:10.2217/epi-2016-0103

204. Inkster AM, Yuan V, Konwar C, Matthews AM, Brown CJ, Robinson WP. A cross-cohort analysis of autosomal DNA methylation sex differences in the term placenta. *Biol Sex Differ*. 2021;12:38. doi:10.1186/s13293-021-00381-4

205. Gong S, Johnson MD, Dopierala J, et al. Genome-wide oxidative bisulfite sequencing identifies sex-specific methylation differences in the human placenta. *Epigenetics*. 2018;13(3):228-239. doi:10.1080/15592294.2018.1429857

206. Kaetzel DM, Leonard MK, Cook GS, et al. Dual functions of NME1 in suppression of cell motility and enhancement of genomic stability in melanoma. *Naunyn-Schmiedeberg's Arch Pharmacol*. 2015;388(2):199-206. doi:10.1007/s00210-014-1010-4

207. Novak M, Leonard MK, Yang XH, Kowluru A, Belkin AM, Kaetzel DM. Metastasis suppressor NME1 regulates melanoma cell morphology, self-adhesion and motility via induction of fibronectin expression. *Experimental Dermatology*. 2015;24(6):455-461. doi:10.1111/exd.12697

208. Xie KM, Hou XF, Li MQ, Li DJ. NME1 at the human maternal–fetal interface downregulates titin expression and invasiveness of trophoblast cells via MAPK pathway in early pregnancy. *Reproduction*. 2010;139(4):799-808. doi:10.1530/REP-09-0490

209. Fang Y, Wan C, Wen Y, et al. Autism-associated synaptic vesicle transcripts are differentially expressed in maternal plasma exosomes of physiopathologic pregnancies. *J Transl Med*. 2021;19(1):1-11. doi:10.1186/s12967-021-02821-6

210. Tapia Contreras C, Hoyer-Fender S. The WD40-protein CFAP52/WDR16 is a centrosome/basal body protein and localizes to the manchette and the flagellum in male germ cells. *Sci Rep*. 2020;10(1):14240. doi:10.1038/s41598-020-71120-9

211. Postema MC, Carrion-Castillo A, Fisher SE, Vingerhoets G, Francks C. The genetics of situs inversus without primary ciliary dyskinesia. *Sci Rep*. 2020;10(1):3677. doi:10.1038/s41598-020-60589-z

212. Funaya S, Kawabata Y, Sugie K, et al. Involvement of the linker histone H1Foo in the regulation of oogenesis. *Reproduction*. 2022;164(2):19-29. doi:10.1530/REP-21-0233

213. Song T, Zhang X, Wang C, et al. MiR-138 Suppresses Expression of Hypoxia-inducible factor 1α (HIF-1α) in Clear Cell Renal Cell Carcinoma 786-O Cells. *Asian Pacific journal of cancer prevention : APJCP*. 2011;12(5). https://pubmed.ncbi.nlm.nih.gov/21875287/

214. Liu W, Wang SJ, Lin QD. Study on the Expressions of PHD and HIF in Placentas from Normal Pregnant Women and Patients with Preeclampsia. *International Journal of Biological Sciences*. 2014;10(3):278-284. doi:10.7150/ijbs.6375

215. Zhao L, Ma R, Zhang L, et al. Inhibition of HIF-1a-mediated TLR4 activation decreases apoptosis and promotes angiogenesis of placental microvascular endothelial cells during severe pre-eclampsia pathogenesis. *Placenta*. 2019;83:8-16. doi:10.1016/j.placenta.2019.06.375

216. Liu XQ, Zhang CG, Xing GC, Chen QT, He FC. Cloning, expression and biochemical activity of ARFGAP3, a regulator of intracellular transport. *Prog Biochem Biophys*. 2001;28(4):507-513.

217. Obinata D, Takayama K ichi, Urano T, et al. ARFGAP3, an androgen target gene, promotes prostate cancer cell proliferation and migration. *International Journal of Cancer*. 2012;130(10):2240-2248. doi:10.1002/ijc.26224

218. Bajaj R, Kundu ST, Grzeskowiak CL, et al. IMPAD1 and KDELR2 drive invasion and metastasis by enhancing Golgi-mediated secretion. *Oncogene*. 2020;39(37):5979-5994. doi:10.1038/s41388-020-01410-z

219. Liao Z, She C, Ma L, et al. KDELR2 Promotes Glioblastoma Tumorigenesis Targeted by HIF1a via mTOR Signaling Pathway. *Cell Mol Neurobiol*. 2019;39(8):1207-1215. doi:10.1007/s10571-019-00715-2

220. Wen HY, Abbasi S, Kellems RE, Xia Y. mTOR: A placental growth signaling sensor. *Placenta*. 2005;26:S63-S69. doi:10.1016/j.placenta.2005.02.004

221. Liu J, Zhang Q, Ma N. LncRNA GASAL1 Interacts with SRSF1 to Regulate Trophoblast Cell Proliferation, Invasion, and Apoptosis Via the mTOR Signaling Pathway. *Cell Transplant*. 2020;29:0963689720965182. doi:10.1177/0963689720965182

222. Barfield RT, Almli LM, Kilaru V, et al. Accounting for Population Stratification in DNA Methylation Studies. *Genet Epidemiol*. 2014;38(3):231-241. doi:10.1002/gepi.21789

223. Wilson SL, Robinson WP. Utility of DNA methylation to assess placental health. *Placenta*. 2018;64:S23-S28. doi:10.1016/j.placenta.2017.12.013

224. Abdelazim SA, Shaker OG, Aly YAH, Senousy MA. Uncovering serum placental-related non-coding RNAs as possible biomarkers of preeclampsia risk, onset and severity revealed MALAT-1, miR-363 and miR-17. *Sci Rep*. 2022;12:1249. doi:10.1038/s41598-022-05119-9

SUPPLEMENT

Supplemental Table 1: Comparison of Analytic Sample Characteristics with Respective Cohort Populations

			Analytic Samples	
Population Covariates	**CANDLE** (N=1503)	**GAPPS** (N=467)	**CANDLE** (N=776)	**GAPPS** (N=205)
Continuous	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Maternal Age (years)	26.3 (5.5)*	30.9 (5.6)*	27.3 (5.5)*	30.5 (5.7)*
Gestational Age (weeks)	38.9 (1.9)*	38.3 (3.0)*	39.0 (1.5)	38.3 (3.0)*
Pre-Pregnancy BMI (kg/m^3)	27.5 (7.5)*	27.3 (7.6)*	27.7 (7.4)*	26.3 (9.1)*
Family Income (dollars)	36800 (27500)*	61400 (26600)*	41400 (27600)*	58900 (27600)*
Birthweight (grams)	3230 (566)	3250 (782)	3310 (516)	3270 (806)
Categorical	**N (%)**	**N (%)**	**N (%)**	**N (%)**
Maternal Race/Ethnicity				
White	467 (31.1)†	354 (75.8)†	292 (37.6)	163 (79.5)†
Black/African American	936 (62.3)†	10 (2.1)†	438 (56.4)	4 (2.0)†
Asian	13 (0.9)†	13 (2.8)†	6 (0.8)	7 (3.4)†
Native Hawaiian Pacific Islander	1 (0.1)†	0 (0)	0 (0)	0 (0)
American Indian/Alaska Native	1 (0.1)†	4 (0.9)†·	0 (0)	2 (1.0)†
Other	77 (5.1)†	37 (7.9)†	37 (4.8)	12 (5.9)†
Multiple Race	6 (0.4)†	12 (2.6)†	3 (0.4)	3 (1.5)†
Offspring Sex				
Male	736 (49.0)†	240 (51.4)	381 (49.1)	108 (52.7)
Female	727 (48.4)†	227 (48.6)	395 (50.9)	97 (47.3)
Season of Birth				
Spring	319 (21.2)†	125 (26.8)	166 (21.4)†	53 (25.9)
Summer	415 (27.6)†	121 (25.9)	219 (28.2)†	50 (24.4)

Fall	414 (27.5)†	107 (22.9)	217 (28.0)†	52 (25.4)
Winter	295 (19.6)†	114 (24.4)	170 (21.9)†	50 (24.4)
Mode of Delivery				
Normal/Vaginal	916 (60.9)†	280 (60.0)†	470 (60.6)	134 (65.4)
C-section	540 (35.9)†	159 (34.0)†	306 (39.4)	71 (34.6)
Presence/Absence of Labor				
Spontaneous, spontaneous augmented, or induced labor	1206 (80.2)†	369 (79.0)†	627 (80.8)†	180 (87.8)†
No labor	249 (16.6)†	43 (9.2)†	148 (19.1)†	16 (7.8)†
Maternal Smoking				
Yes/Cotinine positive (>200 ng/mL)	77 (5.1)†	0 (0.0)†	37 (4.8)†	0 (0)†
No/Cotinine negative (<200 ng/mL)	1085 (72.2)†	4 (0.9)†	736 (94.8)†	2 (1.0)†
Maternal Smoking (Self-report)				
Yes	151 (10.0)†	12 (2.6)†	65 (8.4)	4 (2.0)†
No	1351 (89.9)†	414 (88.7)†	711 (91.6)	199 (97.1)†
Maternal Education				
Less than high school	184 (12.2)†	16 (3.4)†	66 (8.5)	10 (4.9)†
High school completion	709 (47.2)†	112 (24.0)†	346 (44.6)	57 (27.8)†
Graduated college or technical school	437 (29.1)†	187 (40.0)†	260 (33.5)	85 (41.5)†
Some graduate work or professional degree	171 (11.4)†	111 (23.8)†	104 (13.4)	52 (25.4)†

* Mean and standard deviation values calculated after removal of missing values

† Percentages do not sum to 100% due to missing values of the covariates

Supplemental Table 2: Select Infant-sex Stratified Associations of $PM_{2.5}$ Exposure with LncRNAs (Infant-sex Interaction Transcripts)

ENSEMBL ID	Transcript Symbol*	Infant-sex Strata	logFC	Fold Change	P-value	FDR† P-value
Full Pregnancy $PM_{2.5}$ (CANDLE)						
ENSG00000218510	LINC00339	Male	0.003	1.002	0.948	0.998
		Female	0.041	1.029	0.368	0.997

* Transcript symbols and descriptions from HGNC database

† The FDR cutoff for significance was <0.10

Supplemental Table 3: Select Infant-sex Stratified Associations of Pre-pregnancy BMI with CpG methylation (Infant-sex Interaction CpGs)

CpG	Chr.	Position	UCSC RefGene Name	Infant-sex	logFC	Fold Change	P-Value	FDR P-Value	Bacon FDR
Maternal Obesity									
cg08956510	17	49231229	NME1; NME1-NME2	Male	0.509	1.423	0.106	0.881	0.959
				Female	-1.695	0.309	<0.001	0.231	0.356
Maternal BMI									
cg11207127	4	186666487	SORBS2	Male	0.005	1.003	0.343	0.981	0.987
				Female	-0.029	0.980	<0.001	0.348	0.274
cg15805921	6	42952109	PPP2R5D	Male	0.025	1.018	0.007	0.772	0.678
				Female	-0.111	0.926	<0.001	0.204	0.133
cg05467167	17	45000793	GOSR2	Male	0.004	1.003	0.559	0.990	0.995
				Female	-0.059	0.960	<0.001	0.156	0.092
cg09861871	22	43253517	ARFGAP3	Male	-0.016	0.989	0.098	0.944	0.956
				Female	-0.123	0.918	<0.001	0.089	0.029
cg13137577	11	73018312	ARHGEF17	Male	0.015	1.010	0.313	0.979	0.984
				Female	-0.102	0.932	<0.001	0.361	0.292
cg13771376	9	135546871	DDX31; GTF3C4	Male	-0.010	0.993	0.023	0.864	0.853
				Female	0.009	1.006	0.011	0.760	0.862

* CpG symbols and descriptions from Illumina 450k database

Supplemental Table 4: Top mRNA transcripts proximal (within 200 kilobases) to BMI-associated CpG sites

CpG	ENSEMBL ID	Gene	Description	Infant-sex	logFC	Fold Change	P-Value	FDR P-Value
Maternal Obesity								
cg11844079	ENSG00000158710	TAGLN2	Transgelin 2		-0.434	0.740	0.007	0.363
	ENSG00000162729	IGSF8	Immunoglobulin superfamily, member 8		-0.573	0.672	0.018	0.363
	ENSG00000085552	IGSF9	Immunoglobulin superfamily, member 9		-0.472	0.721	0.034	0.385
Maternal Obesity – Infant-Sex Specific Associations								
cg09858237	ENSG00000154914	USP43	Ubiquitin specific peptidase 43	Male	-0.042	0.971	0.044	0.877
Maternal BMI – Infant-Sex Interaction								
cg09861871	ENSG00000242247	ARFGAP3	ADP-ribosylation factor GTPase activating protein 3	Male	-0.285	0.821	0.044	1.000
				Female	0.119	1.086	0.544	0.911
Maternal BMI – Infant-Sex Specific Associations (Male)								
cg03725447	ENSG00000237330	RNF223	Ring finger protein 223	Male	-1.471	0.361	0.010	1.000
cg04842215	ENSG00000139899	CBLN3	Cerebellin 3 precursor	Male	0.190	1.141	0.008	1.000
cg09247949	ENSG00000111802	TDP2	Tyrosyl-DNA phosphodiesterase 2	Male	-0.081	0.946	0.013	0.401
cg15469171	ENSG00000179546	HTR1D	5-hydroxytryptamine (serotonin) receptor 1D, G protein-coupled	Male	-1.467	0.362	0.037	1.000

* CpG symbols and descriptions from Illumina 450k database; gene symbols and descriptions from HGNC database