EpigenomicsNet

Powered by fsg

Pregnancy and the epigenome

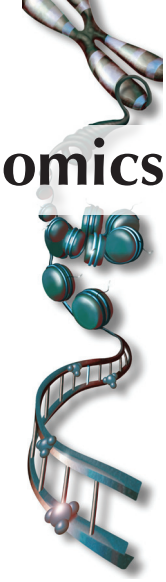
TOP ARTICLES SUPPLEMENT

CONTENTS

Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



Prenatal exposure to mixtures of xenoestrogens and genome-wide DNA methylation in human placenta

Nadia Vilahur*, Mariona Bustamante, Eva Morales, Valeria Motta, Mariana Fátima Fernandez, Lucas Andrés Salas, Georgia Escaramis, Ferran Ballester, Mario Murcia, Adonina Tardon, Isolina Riaño, Loreto Santa-Marina, Jesús Ibarluzea, Juan Pedro Arrebola, Xavier Estivill, Valentina Bollati, Jordi Sunyer & Nicolás Olea

*Author for correspondence: nvilahur@creal.cat

For full author affiliations list, please see page 54

Background: *In utero* exposure to xenostrogens may modify the epigenome. We explored the association of prenatal exposure to mixtures of xenoestrogens and genome-wide placental DNA methylation. **Materials & methods:** Sex-specific associations between methylation changes in placental DNA by doubling the concentration of TEXB-alpha exposure were evaluated by robust multiple linear regression. Two CpG sites were selected for validation and replication in additional male born placentas. **Results:** No significant associations were found, although the top significant CpGs in boys were located in the *LRPAP1*, *HAGH*, *PPARGC1B*, *KCNQ1* and *KCNQ1DN* genes, previously associated to birth weight, Type 2 diabetes, obesity or steroid hormone signaling. Neither technical validation nor biological replication of the results was found in boys for *LRPAP* and *PPARGC1B*. **Conclusion:** Some suggestive genes were differentially methylated in boys in relation to prenatal xenoestrogen exposure, but our initial findings could not be validated or replicated.

First draft submitted: 9 July 2015; **Accepted for publication:** 15 September 2015; **Published online:** 18 December 2015

Keywords: DNA methylation • endocrine disruptors • epigenome xenoestrogens • placenta • prenatal • programming • TEXB

Background

Xenoestrogens are a group of endocrine disrupting chemicals (EDCs) that specifically interfere with the endogenous estrogen hormone signaling pathways and/or metabolism [1]. Exposure to xenoestrogens during susceptible developmental stages like the prenatal period has been related to a number of adverse health outcomes in the offspring, both in humans and in animals, including alterations in birth weight, growth and body mass index, male and female reproductive abnormalities, infant neurodevelopment or increased risk for diabetes and several types of cancer among others [2–10]; with evidences for sex-specific associations [10–15].

The environmental epigenetic hypothesis suggests that the fetal epigenome may be affected by *in utero* environmental exposures, and this may play a role in later disease phenotypes [16]. The adverse effects of environmental exposures are especially relevant in the context of early exposure to EDCs, since endogenous hormones, active at extremely low concentrations, play critical developmental roles during the prenatal period [17,18]. Mice models have revealed that *in utero* exposure to xenoestrogens may disrupt DNA methylation. Bisphenol A and diethylstilbestrol, compounds with known xenoestrogenic properties, induced higher expression of the *EZH2* gene, a histone methyltransferase which resulted in increased mammary histone H3 trimethylation and triggered methylation changes in several estrogen-responsive genes [19,20]. In another study, dietary exposure to soy phytoestrogens in pregnant rats advanced sexual maturation and induced aberrant promoter methylation of skeletal α-actin, estrogen receptor-α and c-fos genes in the offspring [21,22]. In addition, the

Future Medicine part of fsg

pesticide methoxychlor induced changes in DNA methylation at a number of imprinted genes, accompanied by a substantial decrease in mice sperm count [23].

In humans, significant associations have been reported between prenatal exposure to single xenoestrogens like the persistent organic chemicals dichlorodiphenyltrichloroethane (DDT), dichlorodiphenyldichloroethylene (DDE) or polychlorinated biphenyls (PCBs) and global DNA hypomethylation, measured as DNA methylation in retrotransposon elements (LINEs and SINEs) [14,24–26], as well as at gene-specific level [27]. On the other hand, a recent study has shown that prenatal exposure to phthalates and phenols was related to methylation changes in placenta in the imprinted *H19* gene and in the *IGF2* differentially methylated region 2 (DMR2) only in boys, two regions known to play a major role in fetal and placental growth, although a further relation with birth weight could not be demonstrated [28]. Epigenetic dysregulation of the placenta, which can be caused by several environmental factors, may lead to abnormal placental development and function [29]. Even if this organ does not form part of an adult, the human placenta plays a key role in ensuring optimal fetal development and growth, with implications for newborns disease predisposition in later life.

Exposure during pregnancy to arsenic and cadmium, metals which have been suggested to interfere with estrogen signaling [30–33], has been associated to methylation changes in cord blood DNA, both globally and in specific genomic sites as revealed by epigenome-wide association studies (EWAS) [34], with stronger associations often observed in males [15,35–37].

However, exposure to single chemicals is often an unrealistic scenario, since environmental contamination, including EDCs, is rarely due to a single compound but to mixtures to which populations are exposed that may produce additive or even synergistic effects [38–40]. The use of biomarkers of cumulative exposure, such as the total effective xenoestrogen burden (TEXB) is, therefore, a more realistic approach to study the impact of co-exposure to mixtures of chemicals with estrogenic disrupting properties in a real world scenario [41,42].

We previously reported male-specific associations between placenta TEXB-alpha and children birth weight, early growth and motor development at age 1–2 [43,44], accompanied by DNA hypomethylation in *Alu* retrotransposons in placenta [45]. The aim of the present study is to perform EWAS analyzing boys and girls separately to identify differentially methylated genomic *loci* in placenta in relation to prenatal TEXB-alpha exposure.

Material & methods

Study population

The INMA – INfancia y Medio Ambiente – (Environment and Childhood) Project is a Spanish multicenter birth cohort study exploring the role of environmental pollutants on children development and health [46]. All participants involved in the study provided written consent prior to participation, and the research protocol was approved by the Ethical Committees of the Institutions and Centers from the different Spanish regions.

Two subsets of samples were analyzed in the current study, one for the discovery step and the other for the replication step. The discovery study included 181 women of Caucasian origin enrolled from November 2003 to January 2008 from four different areas of Spain: Asturias (18%), Basque Country (34%), Catalonia (37%) and Valencia (11%), who had not followed any program of assisted reproduction, gave singleton birth at the reference hospitals and had placenta collected at delivery. In the replication step, 126 women from the same cohort which had male deliveries were selected, enrolled in the study in the same period and following the same inclusion criteria from Asturias (4%), Basque Country (39%), Catalonia (37%) and Valencia (20%).

Exposure assessment

The TEXB is a biomarker of the combined estrogenic effect of environmental estrogens [47]. The detailed procedure has previously been published elsewhere [42,48]. Briefly, half of eachplacenta was homogenized, in order to obtain a sample representative of the maternal-fetal unit. Thereafter, an hexane-based solid–liquid extraction method was used to separate less lipophilic chemicals including endogenous hormones (beta fraction) from more lipophilic environmental compounds, that is, persistent organic pollutants with xenoestrogenic potential (alpha fraction). Then, the estrogenicity of the alpha fraction (i.e., TEXB-alpha) for each placenta sample was quantified using the E-Screen bioassay, a cell proliferation assay using MCF7 breast cancer cells, at the Biomedical Research Center from the University of Granada (Spain). TEXB-alpha was expressed in picomolar (pM) estradiol equivalent units (Eeq) per gram of placenta tissue (pM Eeq/g placenta).

DNA isolation & methylation genome-wide data generation

INMA placentas were stored at -80°C at the IUSC Biobank of the San Cecilio University Hospital (Granada). Later, half of each sample was homogenized for exposure assessment as described previously, and the other half was partially defrosted and biopsies of 5 cm^3

from the inner region of the placenta were conducted, approximately at a distance of 1.0–1.5 cm below the fetal membranes, which were previously removed, and at a distance of ~5 cm from the site of cord insertion, in order to obtain biopsies from the placental villous parenchyma as homogeneous as possible across samples. 25 mg of tissue was used for DNA extraction, previously rinsed twice during 5 min in 0.8 ml of 0.5X PBS in order to remove traces of maternal blood. Genomic DNA from placenta was isolated using the DNeasy® Blood and Tissue Kit (Qiagen, CA, USA) in narrow time windows and by the same person, in order to minimize technical and operator variations.

DNA quality was evaluated using a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA) and additionally 100 ng of DNA were run on 1.3% agarose gels to confirm that samples did not present visual signs of degradation (smears or bands below 10,000 bp). Isolated genomic DNA was stored at -20°C until further processing.

Genome-wide DNA methylation was measured in 202 placenta samples (including ten duplicates) using the Illumina® Infinium Human Methylation450 BeadChip, a panel which roughly spans 486,000 CpG sites in the human genome. Samples were plated on each chip, experimentally randomized with regard to sex distribution and processed blind to sample identification at the Genome Analysis Facility of the University Medical Center Groningen (UMCG) in Holland, where 500 ng of good quality DNA was used to perform bisulfite conversion followed by methylation profiling following Illumina's protocol.

BeadChips were scanned with an Illumina iScan and image data was uploaded into the Methylation Module of Illumina's analysis software GenomeStudio (lllumina, CA USA), and converted in β-values, that range from 0 (unmethylated) to 1 (fully methylated) and represent the fraction of methylation at a given CpG locus.

Methylation data quality control & normalization

Methylation data quality control (QC) was performed in several steps to exclude low quality samples and probes.

First, using the Genome Studio software, we removed samples that did not reach a call rate of 95% at a p-value below 0.05. Then, following Illumina recommendations we verified the intensities of several control probes provided by Illumina in order to: assess the quality of the experiment (sample-dependent controls); and identify problems in specific experimental steps (sample-independent controls). If a given sample failed in 3 or more Illumina controls, it was excluded from further analyses. Altogether, two samples were excluded that did not meet these two criteria.

Ten biological duplicates, distributed either in the same or in different bisulfite plates and BeadChips, were used to estimate the discrimination threshold of the Infinium450K Array using the total deviation index (TDI) [49], which was 0.059 in duplicates from different bisulfite and hybridation arrays, 0.057 in duplicates from the same bisulfite but different hybridation arrays, and 0.066 in duplicates both from the same bisulfite and hybridation array. Additionally, the 450K BeadChip features 65 control probes which assay highly-polymorphic SNPs. Consistent results were observed when we performed pairwise correlations of these genotypes in our duplicates (see Supplementary Figure 1). After these steps, one of the biological duplicate samples (randomly selected) as well as the 65 SNP probes were removed from the dataset.

Methylation patterns in chromosome X probes were used to cluster subjects according to sex by principal component analysis, and eight mismatched samples were detected in relation to the information on sex contained in our database. These samples were also excluded from further analyses.

An additional QC step was performed in R environment using the WateRmelon package [50], and one sample presenting more than 1% of sites with a detection p-value greater than 0.05 was excluded, in addition to 1859 probes, either because they occurred in more than 1% of samples with a p-value greater than 0.05 or because they presented a bead count below 3 in more than >5% of samples.

Finally, probes that ambiguously mapped to the human genome with at least 47 base pairs or more (n = 29.233) were excluded, and probes containing a SNP with a minor allele frequency (MAF) in HapMap European population (CEU) >4%, either at the extension site or in the ten nucleotides immediately before (n = 14.122), as suggested by Chen *et al.* [51]. Probes corresponding to CpG sites located in chromosomes X and Y were also excluded (n = 8.537).

A total of 433.131 CpG sites on autosomes were tested with regard to TEXB-alpha exposure in the remaining 181 samples, representing 93 boys and 88 girls. Raw methylation beta values were then normalized to reduce technical variability and four different normalization methods were compared (dasen, BMIQ, quantile normalization and Swan) using three performance metrics as proposed by Pidsley *et al.* [52]. Dasen normalization was used for further analyses as it resulted the best ranked method in our data (additional information in Supplementary Material & Methods).

Epigenome-wide association study

A robust linear regression model was employed using MASS (R package) to test the association between doubling of TEXB-alpha concentration and methylation at each CpG site in boys and in girls separately. Analyses were adjusted for area of study and two technical factors: chip and bisulfite plate. Covariates included in the adjusted model were selected by testing the difference of the correlation of p values before and after correction using a Kolmogorov–Smirnov test. Since crude and adjusted robust regression models produced very similar results, only adjusted models are presented. The false discovery rate (FDR) correction for multiple testing was calculated with the Benjamini and Hochberg (B&H) method.

CpG sites annotation

In order to obtain information on the top differentially methylated CpG sites in our study the University of California, Santa Cruz (UCSC) Genome Browser interface was used, which in addition contains ENCODE (Encyclopedia of DNA Elements) detailed information on regulatory elements, including chromatin accessibility and epigenetic marks across the genome both in DNA and in histones [53,54]. The human gene database GeneCards [55,56] was used to obtain information on the genomic location and gene (or nearest gene) function and reported disease associations, while toxicological interactions of these genes with chemical compounds were explored with the Comparative Toxicogenomics Database [57,58].

Validation & replication by pyrosequencing

Two CpGs were selected for further DNA methylation validation in the same samples as in the discovery study, and replication was conducted in 126 independent placenta samples (boys only) from the INMA cohort. For that purpose, bisulfite pyrosequencing, a highly quantitative PCR-based analysis was used. A total of 500 ng of extracted DNA was bisulfite converted using the EZ DNA Methylation-Gold™ Kit (Zymo Research, CA, USA), and 1 µl of converted DNA was further PCR amplified and sequenced. Additional information on primer design and PCR assay conditions can be found in online Supplementary Material & Methods, Supplementary Table 1. Samples were run in duplicate on a PyroMark Q96 ID pyrosequencing system (Qiagen) and the pairwise correlation between technical duplicates was 0.94 for cg05342136 and 0.80 for cg15809858 (see Supplementary Figure 2).

The association between doubling the concentration of TEXB-alpha and changes in DNA methylation was tested using adjusted linear mixed regression models including repeated measurements (pyrosequencing duplicates) as random intercept.

Results

Study population characteristics

Overall, our discovery study population did not differ from the rest of INMA cohort participants for the main pregnancy related and sociodemographic characteristics. Women in the discovery study were on average 32 years old, with a prepregnancy BMI of 23.45 kg/m². Only three children were born preterm, which represented a slightly lower (nonsignificant) percentage than what we observed in the rest of the INMA cohort participants and, concordantly, we saw less small for gestational age (SGA) children in our discovery study group when compared with the rest of INMA cohort (p-value < 0.02) (see Supplementary Material & Methods, Supplementary Table 2). Additionally, no significant differences were observed for any of the maternal and infant sociodemographic and pregnancy related characteristics when comparing the discovery and replication populations used in our study, as shown in Table 1.

TEXB-alpha exposure among the 181 participants in the discovery sample did not differ by newborn sex (Kruskal–Wallis test p-value < 0.624), and was also similar between the discovery and the replication samples (discovery: median = 0.75, iqr = 0.28 to 1.28 pM Eeq/g placenta and replication: median = 0.76, iqr: 0.40–1.41 pM Eeq/g placenta; Kruskal-Wallis test p-value < 0.438).

Association of TEXB-alpha and genome-wide DNA methylation in placenta

No epigenome-wide significant associations were found either in boys or in girls between each doubling of TEXB-alpha (pM Eeq/g placenta) and methylation at CpG sites in placenta after correcting for multiple testing. Results for the 12 most significant CpG sites in boys (n = 93) and in girls (n = 88) are presented in Table 2. Quantile–quantile (Q–Q) plots showing the observed versus expected -log10 (p-values) under the null hypothesis of no association are shown in online Supplementary Figure 3. Among the top CpG sites differentially methylated in boys we found some genes previously related to growth and steroid hormone signaling, while none of these genes were observed in girls (see Supplementary Table 3 for additional information).

Validation & replication of selected top CpGs in boys

Two CpGs differentially methylated in relation to TEXB-alpha in boys were selected for further validation by pyrosequencing in the same samples (n = 92) and replication was conducted in additional 126 placenta samples from male deliveries.

Table 1. Main characteristics and comparison between discovery and replication study mother–child pairs enrolled in the INMA Project from the four participating INMA cohorts[†].

Variables	Discovery study sample (n = 181) n (%); mean (SD) or median (IQR)	Replication study sample (n = 126) n (%); mean (SD) or median (IQR)	p-value
Maternal characteristics			
Maternal age (years)	32 (3.97)	31.70 (4.04)	0.437
Pre-pregnancy BMI (kg/m²)	23.45 (4.10)	24.01(5.30)	0.705
Type of delivery:			0.539
– Vaginal	121 (67.22)	80 (63.49)	
– Instrumental	31 (17.22)	28 (22.40)	
– Caesarean	28 (15.56)	17 (13.60)	
Parity:			0.497
– Primiparous	106 (58.56)	69 (54.76)	
– Multiparous (2+)	75 (41.44)	57 (45.24)	
Smoke during pregnancy (yes)	49 (27.37)	31 (24.60)	0.777
Maternal educational level:			0.131
– Below secondary school	34 (19)	40 (31.75)	
– Secondary school	83 (45.86)	33 (26.19)	
– University degree	64 (35.14)	53 (42.06)	
Infant characteristics			
Sex (male)	93 (51.38)	126 (100)[‡]	–
Gestational age (weeks)	39.80 (1.35)	39.85 (1.33)	0.674
Preterm (<37 weeks)	3 (1.67)	4 (3.17)	0.574
Birth weight	3299.07 (419.04)	3305.75 (447)	0.396
Small for gestational age[§]	10 (5.65)	11 (8.73)	0.385
Large for gestational age[¶]	14 (7.91)	12 (9.52)	0.259

[†]Asturias, Gipuzkoa, Sabadell and Valencia.
[‡]Only boys were included in the replication study.
[§]Small for gestational age: below the 10th percentile of birth weight, adjusted for sex and gestational age.
[¶]Large for gestational age: above the 90th percentile of birth weight, adjusted for sex and gestational age.
IQR: Interquartile range; SD: Standard deviation.

The first CpG, the top differentially methylated cg05342136 in boys, is located in the exon 1 of the low-density *LRPAP1*, a lipid-metabolism gene highly expressed in placenta that in turn interacts with TGFb1, an angiogenic factor mediating successful placentation and fetal growth via regulation of trophoblast invasion, cell differentiation, immunosuppression and apoptosis of vascular endothelial cells, which has also been associated with susceptibility to degenerative dementia [59–62]. The second selected CpG site, cg15809858, is located in the first intron of *Homo sapiens PPARGC1B*, a gene expressed in human placenta that stimulates the activity of several transcription factors and nuclear receptors, including estrogen receptor alpha, nuclear respiratory factor 1, and glucocorticoid receptor, shown to be down regulated in prediabetic and Type 2 diabetes mellitus patients and previously related to an increased risk of developing obesity [63–65].

Results from the discovery EWAS could not be neither validated nor replicated for the two CpG sites analyzed (Table 3), although a trend for technical validation (i.e., we were able to confirm the magnitude and the direction of the effect) was observed for cg05342136 (β = 0.29 in the discovery vs β = 0.25 in the validation study). Scatter plots showing the correlation between DNA methylation values measured using the Illumina 450K array platform and by bisulfite pyrosequencing in the same samples (discovery) are shown for cg05342136 and cg15809858 in online Supplementary Material & Methods, Supplementary Figure 4.

Discussion

This is the first genome-wide study analyzing site-specific DNA methylation changes in placenta tissue in relation to a biomarker of exposure to mixtures of environmental estrogens. At the interface between mother

and child, the human placenta is an organ involved in the regulation of fetal growth and development and represents a gateway for substances, including xenoestrogens, to enter fetal circulation [66,67].

In the INMA cohort we previously reported sex-specific associations between the TEXB-alpha biomarker measured in placenta and birth weight changes in boys only [48], along with lower *AluYb8* retrotransposon methylation in placental DNA in the same group, used as a surrogate to study global genomic DNA methylation [45]. In order to gain further mechanistic insight into these associations, we have performed an EWAS study in placenta, stratifying data by sex, to identify differentially methylated genes as a result of prenatal exposure to xenoestrogens. No genome-wide significant associations were found between each doubling of TEXB-alpha and DNA methylation after correcting for multiple testing in either group. However, among the top significant CpG sites differentially methylated in relation to each doubling of TEXB-alpha in boys, several were located in genes that have been related to birth weight regulation, Type 2 diabetes and obesity risk or steroid hormone metabolism, and are known to be expressed in human placenta tissue. The top significant one, cg05342136, is located in the first exon of *LRPAP1*, a gene involved in cholesterol metabolism, the primary metabolite of steroid hormone synthesis [68]. We also found an increase in DNA methylation in cg15809858, located in exon 1 of *PPARGC1B*. The protein encoded by this ubiquitously expressed gene stimulates the activity of several transcrip-

Table 2. Top CpG sites differentially methylated in placenta in relation to prenatal total effective xenoestrogen burden-alpha exposure, ranked by nominal p-value.

CpG name	Mean methylation, % (SD)	Regression β† (% methylation)	p-value	FDR‡	Chr	UCSC gene name
Boys (n = 93)						
cg05342136	89.07 (1.41)	0.29	4.72 E-07	0.20	4	*LRPAP1*
cg08983490	6.63 (0.71)	-0.12	2.25 E-06	0.49	16	*HAGH*
cg00698124	9.73 (1.07)	0.21	7.84 E-06	0.70	18	*SETBP1*
cg23261491	11.52 (1.10)	0.19	7.88 E-06	0.70	12	*OSBPL8*
cg15809858	10.55 (1.29)	0.23	8.87 E-06	0.70	5	*PPARGC1B*
cg14218861	29.81 (4.91)	-0.77	9.73 E-06	0.70	10	
cg16172549	61.33 (7.01)	-1.25	1.52 E-05	0.72	1	*PCP4L1*
cg10447095	55.83 (8.40)	-1.55	1.75 E-05	0.72	16	
cg19584136	87.49 (2.38)	-0.40	1.82 E-05	0.72	10	*MXI1*
cg00836964	91.08 (1.17)	0.22	2.22 E-05	0.72	5	
cg00957580	84.65 (2.30)	0.37	2.27 E-05	0.72	14	*NDRG2*
cg23903244	55.29 (2.82)	0.48	2.61 E-05	0.72	11	*KCNQ1*
Girls (n = 88)						
cg21877656	4.88 (0.45)	0.08	1.66 E-06	0.54	19	*ZNF329*
cg19743820	14.26 (3.82)	-0.87	3.15 E-06	0.54	5	*COX7C*
cg21690627	6.19 (0.56)	-0.13	3.73 E-06	0.54	17	*C17orf59*
cg04919579	7.60 (0.89)	-0.19	9.85 E-06	0.88	3	*RNF168*
cg23313650	12.04 (1.62)	-0.28	1.24 E-05	0.88	7	*WASL*
cg01648887	94.32 (0.84)	-0.16	2.19 E-05	0.88	16	*SPG7*
cg11804334	5.31 (0.48)	-0.10	2.96 E-05	0.88	11	*CCDC34*
cg25248213	88.40 (1.55)	0.30	3.14 E-05	0.88	11	
cg01374565	8.97 (0.95)	-0.20	3.15 E-05	0.88	11	*GDPD5*
cg16705665	8.02 (0.70)	-0.13	3.35 E-05	0.88	11	*RCOR2*
cg27489994	8.48 (1.57)	0.31	3.36 E-05	0.88	13	*TPT1*
cg10424681	83.85 (2.98)	-0.62	3.98 E-05	0.88	6	*C6orf201*

†Estimates per doubling total effective xenoestrogen burden-alpha concentration.
‡All models were adjusted for area of study, bead chip and bisulfite plate.
Chr: Chromosome; FDR: False discovery rate; SD: Standard deviation; UCSC: University of California Santa Cruz.

Table 3. Technical validation and biological replication in boys of placenta DNA methylation in selected CpG sites in relation to prenatal total effective xenoestrogen burden-alpha exposure.

CpG name	UCSC gene name	Mean methylation, % (SD)	Technical validation (discovery samples)				Biological replication (independent samples)			
			n	B†	95% CI	p-value	n	B	95% CI	p-value
cg05342136	*LRPAP1*	89.77 (2.6)	92	0.25	-0.23–0.76	0.299	125	-0.07	-0.28–0.15	0.546
cg15809858	*PPARGC1B*	1.87 (1.52)	92	0.06	-0.08–0.19	0.405	126	0.01	-0.04–0.06	0.652

†Estimates per doubling total effective xenoestrogen burden-alpha concentration. Models adjusted for area of study, gestational age, maternal age during pregnancy, smoking during pregnancy and bisulfite plate. Pyrosequencing duplicate was included as a random intercept.
SD: Standard deviation; UCSC: University of California Santa Cruz.

tion factors and nuclear receptors, including *ERα*, and may be involved in fat oxidation, nonoxidative glucose metabolism and the regulation of energy expenditure. This protein is downexpressed in prediabetic and Type 2 diabetes mellitus patients and certain allelic variations in this gene increase the risk of the development of obesity, Type 2 diabetes and breast cancer [63,69]. Moreover, it has been shown to be downregulated by the hormonally active compound benzo(a)pyrene in mice [70]. Other suggestive genes appeared among the top hits in boys, such as *HAGH*, encoding for an hydroxylase enzyme involved in the pyruvate metabolism, or *KCNQ1,* a paternally imprinted gene that although relatively low expressed in placenta, contains genetic polymorphisms related to birth weight and Type 2 diabetes [71] and is located within a cluster of imprinted genes in the chromosomal region 11p15.5, that includes *KCNQ1DN, H19*, *IGF2* and *KCNQ1OT1* among others, previously associated with fetal and placental growth [28]. Mutations and epimutations in these genes have been associated to the Beckwith–Wiedemann Syndrome (BWS), an overgrowth imprinting disorder that causes large body size and large organs in addition to other clinical manifestations present from birth [28,72–73].

Overall, although statistically nonsignificant, these findings seem to go in line with our previous results, showing that higher levels of xenoestrogens (TEXB-alpha) measured in placenta were associated with higher birth weight in boys (on average 148 grams when comparing high versus low exposed children), while no effects were found in girls [48]. However, we could not validate neither replicate our initial findings for two selected CpGs, and only for cg05342136, technical verification showed a coefficient (β value) of the same magnitude and direction as in the discovery study. In our data, we observe a poor reproducibility between DNA methylation values measured using the 450K Illumina array platform and by bisulfite pyrosequencing in the same samples, especially when CpGs are hypomethylated (as for cg15809858). Whether this lack of technical replication and biological validation of the results in our study reflects differences in the reproducibility, specificity and/or measurement sensitivity across different platforms used to measure DNA methylation (i.e., hibridation array vs bisulfite pyrosequencing), as previously demonstrated for miRNA quantitative expression data [74], a lack of statistical power to reach statistical significance (especially when the magnitude of changes might be small), or truly negative findings remains to be addressed with additional larger studies.

Lack of statistical power is a problem when analyzing -omics data, and we are likely underpowered in our study, where stratified analyses were conducted. Moreover, by using an array based approach such as

the Illumina Infinium Human Methylation450 BeadChip, covering with probes roughly a 2% of the ~28 million CpG sites described in the human genome [75], we might have missed potentially important genomic regions in our study.

Our study has two main methodological strengths: first, exposure to mixtures of xenoestrogens was measured in placenta tissue using a biomarker, and second DNA methylation changes were analyzed in the same tissue, which is relevant considering the tissue specificity of epigenetic marks, the role of this organ during prenatal development and its sensitivity to the effects of hormones [76].

The magnitude of the differences in DNA methylation that we observed was small, although similar to what has been previously shown in other EWAS in relation to prenatal exposure to other environmental chemicals, including potential xenoestrogens like cadmium [77]. To some extent, our results may have been confounded by cell type mixtures in placenta samples, or by the possible maternal cell contamination, which in both cases could have led to a possible underestimation of the effects, while we do not know whether the observed changes in DNA methylation have functional effects on gene expression, since RNA was degraded due to placenta collection conditions in our cohort. Only one biopsy for DNA extraction was conducted per sample, which could have introduced additional noise due to regional variations in DNA methylation, although some authors have suggested that this is not a major source of DNA methylation variation in human placenta [78]. Finally, some uncertainty exists on whether the TEXB-alpha biomarker, based on a lipophilic extraction of compounds (excluding endogenous steroid hormones and more polar xenoestrogens) followed by a quantification of MCF7 cell proliferation assay, is exclusively a biomarker of xenoestrogenicity, or also a biomarker of other lipophilic compounds present in the placenta that activate growth, and that might not necessarily (or uniquely) act through binding or interacting with the estrogen receptor.

Conclusion

We conducted a genome-wide methylation study in placental tissue in relation to prenatal exposure to mixtures of xenoestrogens using the TEXB-alpha biomarker, and although we identified some suggesting genes differentially methylated in boys, we were not able to validate neither replicate our initial findings by pyrosequencing. Future studies are warranted to confirm the observed associations and their potential to mediate the effect of prenatal exposure to mixtures of endocrine disruptors on the offspring's health.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/epi.15.91

Acknowledgements

The authors acknowledge all the study participants for their generous collaboration. A full list of INMA Study Investigators can be found at: www.proyectoinma.org/presentacion-inma/listado-investigadores/listado-investigadores.html

Financial & competing interests disclosure

This work was supported by grants from the Spanish Ministry of Health (FIS-PI042018; FIS-PI060867; FIS-PI081151; FIS-PI09/02311; FIS-PI09/02647; FIS-PI11/00610; FIS-PI13/02429); Instituto de Salud Carlos III (Red INMA G03/176 and CB06/02/0041); the EU Commission (QLK4-1999-01422, QLK4-2002-00603 and CONTAMED FP7-ENV-212502), the Generalitat de Catalunya-CIRIT (1999SGR 00241); the Fundació La Marató de TV3; the Consejería de Salud de la Junta de Andalucía (grant number 183/07 and 0675/10), the Diputación Foral de Gipuzkoa (DFG06/004), the Department of Health of the Basque Government (2005111093), the University of Oviedo, the Fundación Liberbank, and the Fundación Roger Torné.

Nadia Vilahur was supported by an FPI Grant from the Spanish Ministry of Health (BES-2009-023933) and a Formación de Personal Investigador Grant for Short Research Stays in Foreign Institutions (BES-2009-023933). The HUSC BioBank, integrated in the Andalusia Public Health System (SSPA) and the National Biobank Network, is financed by the Institute of Health Carlos III, (project RD09/0076/00148) and the Regional Government of Andalusia.

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved and/or their legal tutors.

Executive summary

Background

- The prenatal period may be a susceptible developmental window to the effects of endocrine disrupting chemicals, including xenoestrogens.

Material & methods

- Changes in human placental DNA methylation were explored using the Illumina 450K Array in relation to TEXB-alpha, a biomarker of exposure to mixtures of xenoestrogens.
- Verification and validation of initial results was conducted by bisufite pyrosequencing on two selected CpGs in the same and in additional placenta samples.

Results

- No significant epigenome-wide associations were found. However, the top differentially methylated CpGs in boys in relation to TEXB-alpha exposure were located in genes previously associated to birth weight, Type 2 diabetes, obesity or steroid hormone signaling.
- We could not validate or replicate our initial findings.

Discussion & conclusion

- Although statistically nonsignificant, our findings in boys in genes related to growth and energy expenditure seem to go in line with previous findings of male-specific positive associations between levels of TEXB-alpha and birth weight.
- Additional larger studies are needed to confirm our initial results, or provide further evidence of their truly negative nature.

References

Papers of special note have been highlighted as:
•• of considerable interest

1 Bulzomi P, Marino M. Environmental endocrine disruptors: does a sex-related susceptibility exist? *Front. Biosci. (Landmark Ed.)* 16, 2478–2498 (2011).

2 Alavanja MC, Samanic C, Dosemeci M *et al.* Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. *Am. J. Epidemiol.* 157(9), 800–814 (2003).

3 Crinnion WJ. Toxic effects of the easily avoidable phthalates and parabens. *Alt. Med. Rev.* 15(3), 190–196 (2010).

4 Govarts E, Nieuwenhuijsen M, Schoeters G *et al.* Birth weight and prenatal exposure to polychlorinated biphenyls (PCBs) and dichlorodiphenyldichloroethylene (DDE): a meta-analysis within 12 European birth cohorts. *Environ. Health Perspect.* 120(2), 162–170 (2012).

5 Toppari J, Larsen JC, Christiansen P *et al.* Male reproductive health and environmental xenoestrogens. *Environ. Health Perspect.* 104(Suppl. 4), 741–803 (1996).

6 Torres-Sanchez L, Rothenberg SJ, Schnaas L *et al. In utero* p,p'-DDE exposure and infant neurodevelopment: a perinatal cohort in Mexico. *Environ. Health Perspect.* 115(3), 435–439 (2007).

7 Valvi D, Mendez MA, Garcia-Esteban R *et al.* Prenatal exposure to persistent organic pollutants and rapid weight gain and overweight in infancy. *Obesity (Silver Spring)* 22(2), 488–496 (2013).

8 Valvi D, Mendez MA, Martinez D *et al.* Prenatal concentrations of polychlorinated biphenyls, DDE, and DDT and overweight in children: a prospective birth cohort study. *Environ. Health Perspect.* 120(3), 451–457 (2012).

9 Nadal A, Alonso-Magdalena P, Soriano S, Quesada I, Ropero AB. The pancreatic beta-cell as a target of estrogens and xenoestrogens: implications for blood glucose homeostasis and diabetes. *Mol. Cell. Endocrinol.* 304(1–2), 63–68 (2009).

10 Reed CE, Fenton SE. Exposure to diethylstilbestrol during sensitive life stages: a legacy of heritable health effects. *Birth Defects Res. C Embryo Today* 99(2), 134–146 (2013).

11 Philippat C, Mortamais M, Chevrier C *et al.* Exposure to phthalates and phenols during pregnancy and offspring size at birth. *Environ. Health Perspect.* 120(3), 464–470 (2012).

12 Vafeiadi M, Agramunt S, Papadopoulou E *et al. In utero* exposure to dioxins and dioxin-like compounds and anogenital distance in newborns and infants. *Environ. Health Perspect.* 121(1), 125–130 (2013).

13 Llop S, Lopez-Espinosa MJ, Rebagliato M, Ballester F. Gender differences in the neurotoxicity of metals in children. *Toxicology* 311(1–2), 3–12 (2013).

14 Rusiecki JA, Baccarelli A, Bollati V, Tarantini L, Moore LE, Bonefeld-Jorgensen EC. Global DNA hypomethylation is associated with high serum-persistent organic pollutants in Greenlandic Inuit. *Environ. Health Perspect.* 116(11), 1547–1552 (2008).

15 M Wagatsuma Y, Rahman A *et al.* Environmental exposure to arsenic and cadmium during pregnancy and fetal size: a longitudinal study in rural Bangladesh. *Reprod. Toxicol.* 34(4), 504–511 (2012).

16 Gluckman PD, Hanson MA, Buklijas T, Low FM, Beedle AS. Epigenetic mechanisms that underpin metabolic and cardiovascular diseases. *Nat. Rev. Endocrinol.* 5(7), 401–408 (2009).

17 Zoeller RT, Brown TR, Doan LL *et al.* Endocrine-disrupting chemicals and public health protection: a statement of principles from The Endocrine Society. *Endocrinology* 153(9), 4097–4110 (2012).

18 Schug TT, Janesick A, Blumberg B, Heindel JJ. Endocrine disrupting chemicals and disease susceptibility. *J. Steroid Biochem. Mol. Biol.* 127(3–5), 204–215 (2011).

19 Bromer JG, Wu J, Zhou Y, Taylor HS. Hypermethylation of homeobox A10 by *in utero* diethylstilbestrol exposure:

an epigenetic mechanism for altered developmental programming. *Endocrinology* 150(7), 3376–3382 (2009).

20 Doherty LF, Bromer JG, Zhou Y, Aldad TS, Taylor HS. *In utero* exposure to diethylstilbestrol (DES) or bisphenol-A (BPA) increases EZH2 expression in the mammary gland: an epigenetic mechanism linking endocrine disruptors to breast cancer. *Horm Cancer* 1(3), 146–155 (2010).

21 Guerrero-Bosagna CM, Sabat P, Valdovinos FS, Valladares LE, Clark SJ. Epigenetic and phenotypic changes result from a continuous pre and post natal dietary exposure to phytoestrogens in an experimental population of mice. *BMC Physiol.* 8, 17 (2008).

22 Ho SM, Johnson A, Tarapore P, Janakiram V, Zhang X, Leung YK. Environmental epigenetics and its implication on disease risk and health outcomes. *ILAR J.* 53(3–4), 289–305 (2012).

23 Stouder C, Paoloni-Giacobino A. Specific transgenerational imprinting effects of the endocrine disruptor methoxychlor on male gametes. *Reproduction* 141(2), 207–216 (2011).

24 Kang SC, Lee BM. DNA methylation of estrogen receptor alpha gene by phthalates. *J. Toxicol. Environ. Health A* 68(23–24), 1995–2003 (2005).

25 Kile ML, Baccarelli A, Hoffman E *et al.* Prenatal arsenic exposure and DNA methylation in maternal and umbilical cord blood leukocytes. *Environ. Health Perspect.* 120(7), 1061–1066 (2012).

26 Kim KY, Kim DS, Lee SK *et al.* Association of low-dose exposure to persistent organic pollutants with global DNA hypomethylation in healthy Koreans. *Environ. Health Perspect.* 118(3), 370–374 (2010).

27 Morales E, Bustamante M, Vilahur N *et al.* DNA hypomethylation at ALOX12 is associated with persistent wheezing in childhood. *Am. J. Respir. Crit. Care Med.* 185(9), 937–943 (2012).

28 Larocca J, Binder AM, McElrath TF, Michels KB. The impact of first trimester phthalate and phenol exposure on IGF2/H19 genomic imprinting and birth outcomes. *Environ. Res.* 133, 396–406 (2014).

29 Nelissen EC, Van Montfoort AP, Dumoulin JC, Evers JL. Epigenetics and the placenta. *Human Reprod. Update* 17(3), 397–417 (2011).

30 Davey JC, Bodwell JE, Gosse JA, Hamilton JW. Arsenic as an endocrine disruptor: effects of arsenic on estrogen receptor-mediated gene expression *in vivo* and in cell culture. *Toxicol. Sci.* 98(1), 75–86 (2007).

31 Watson WH, Yager JD. Arsenic: extension of its endocrine disruption potential to interference with estrogen receptor-mediated signaling. *Toxicol. Sci.* 98(1), 1–4 (2007).

32 Ali I, Penttinen-Damdimopoulou PE, Makela SI *et al.* Estrogen-like effects of cadmium *in vivo* do not appear to be mediated via the classical estrogen receptor transcriptional pathway. *Environ. Health Perspect.* 118(10), 1389–1394 (2010).

33 Johnson MD, Kenney N, Stoica A *et al.* Cadmium mimics the *in vivo* effects of estrogen in the uterus and mammary gland. *Nature Med.* 9(8), 1081–1084 (2003).

34 Liu X, Zheng Y, Zhang W *et al.* Blood methylomics in response to arsenic exposure in a low-exposed US population. *J. Expo. Sci. Environ. Epidemiol.* 24(2), 145–149 (2014).

35 Kippler M, Engstrom K, Mlakar SJ *et al.* Sex-specific effects of early life cadmium exposure on DNA methylation and implications for birth weight. *Epigenetics* 8(5), 494–503 (2013).

•• **This is an epigenome-wide association study evaluating the association between prenatal cadmium exposure and DNA methylation, which reports sex-specific differentially methylated genes in cord blood and their relation to birth weight.**

36 Pilsner JR, Hall MN, Liu X *et al.* Influence of prenatal arsenic exposure and newborn sex on global methylation of cord blood DNA. *PLoS ONE* 7(5), e37147 (2012).

37 Broberg K, Ahmed S, Engström K *et al.* Arsenic exposure in early pregnancy alters genome-wideDNA methylation in cord blood, particularly in boys. *J. Dev. Orig. Health Dis.* 5(4), 288–298 (2014).

38 Rajapakse N, Silva E, Kortenkamp A. Combining xenoestrogens at levels below individual no-observed-effect concentrations dramatically enhances steroid hormone action. *Environ. Health Perspect.* 110(9), 917–921 (2002).

39 Silva E, Rajapakse N, Kortenkamp A. Something from 'nothing' – eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. *Environ. Sci. Technol.* 36(8), 1751–1756 (2002).

40 Diamanti-Kandarakis E, Bourguignon JP, Giudice LC *et al.* Endocrine-disrupting chemicals: an Endocrine Society scientific statement. *Endocrine Rev.* 30(4), 293–342 (2009).

41 Soto AM, Sonnenschein C, Chung KL, Fernandez MF, Olea N, Serrano FO. The E-SCREEN assay as a tool to identify estrogens: an update on estrogenic environmental pollutants. *Environ. Health Perspect.* 103(Suppl. 7), 113–122 (1995).

42 Lopez-Espinosa MJ, Silva E, Granada A *et al.* Assessment of the total effective xenoestrogen burden in extracts of human placentas. *Biomarkers* 14(5), 271–277 (2009).

•• **This article explains in detail the standardized method used to assess the total effective xenoestrogen burden (TEXB) in human placentas.**

43 Vilahur N, Baccarelli AA, Bustamante M *et al.* Storage conditions and stability of global DNA methylation in placental tissue. *Epigenomics* 5(3), 341–348 (2013).

44 Vilahur N, Fernandez MF, Bustamante M *et al. In utero* exposure to mixtures of xenoestrogens and child neuropsychological development. *Environ. Res.* 134, 98–104 (2014).

45 Vilahur N, Bustamante M, Byun HM *et al.* Prenatal exposure to mixtures of xenoestrogens and repetitive element DNA methylation changes in human placenta. *Environ. Int.* 71, 81–87 (2014).

•• **This study evaluates global DNA methylation in repetitive elements in placenta in relation to the TEXB biomarker, and reports a male-specific hypomethylation of the AluYb8 retrotransposon.**

46 Guxens M, Ballester F, Espada M *et al.* Cohort Profile: the INMA-INfancia y Medio Ambiente–(Environment and Childhood) Project. *Int. J. Epidemiol.* 41(4), 930–940 (2012).

47 Fernandez MF, Rivas A, Olea-Serrano F *et al.* Assessment of total effective xenoestrogen burden in adipose tissue and identification of chemicals responsible for the combined estrogenic effect. *Anal. Bioanal. Chem.* 379(1), 163–170 (2004).

48 Vilahur N, Molina-Molina JM, Bustamante M *et al.* Male specific association between xenoestrogen levels in placenta and birthweight. *Environ. Int.* 51, 174–181 (2013).

•• **This study reports a significant positive association in boys between prenatal exposure to xenoestrogens and birth weight using the TEXB biomarker.**

49 Escaramis G, Ascaso C, Carrasco JL. The total deviation index estimated by tolerance intervals to evaluate the concordance of measurement devices. *BMC Med. Res. Methodol.* 10, 31 (2010).

50 Schalkwyk LC, Pidsley R, Cc W *et al.* wateRmelon: Illumina 450 methylation array normalization and metrics. *R package version* 1, 40 (2013).

51 Chen YA, Lemire M, Choufani S *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 8(2), 203–209 (2013).

52 Pidsley R, Cc YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* 14 293 (2013).

53 Kellis M, Wold B, Snyder MP *et al.* Defining functional DNA elements in the human genome. *Proc. Natl Acad. Sci. USA* 111(17), 6131–6138 (2014).

54 Kent WJ, Sugnet CW, Furey TS *et al.* The human genome browser at UCSC. *Genome Res.* 12(6), 996–1006 (2002).

55 Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. *Trends Genet.* 13(4), 163 (1997).

56 GeneCards: The Human Gene Database. www.genecards.org

57 Davis AP, King BL, Mockus S *et al.* The comparative toxicogenomics database: update 2011. *Nucleic Acids Res.* 39(Database issue), D1067–D1072 (2011).

58 Comparative Toxicogenomics Database. www.ctdbase.org

59 Aldahmesh MA, Khan AO, Alkuraya H *et al.* Mutations in LRPAP1 are associated with severe myopia in humans. *Am. J. Hum. Genet.* 93(2), 313–320 (2013).

60 Briana DD, Liosi S, Gourgiotis D *et al.* Fetal concentrations of the growth factors TGF-alpha and TGF-beta1 in relation to normal and restricted fetal growth at term. *Cytokine* 60(1), 157–161 (2012).

61 Pandey P, Pradhan S, Mittal B. LRP-associated protein gene (LRPAP1) and susceptibility to degenerative dementia. *Genes Brain Behav.* 7(8), 943–950 (2008).

62 Relf M, Lejeune S, Scott PA *et al.* Expression of the angiogenic factors vascular endothelial cell growth factor, acidic and basic fibroblast growth factor, tumor growth factor beta-1, platelet-derived endothelial cell growth factor, placenta growth factor, and pleiotrophin in human primary breast cancer and its relation to angiogenesis. *Cancer Res.* 57(5), 963–969 (1997).

63 Andersen G, Wegner L, Yanagisawa K *et al.* Evidence of an association between genetic variation of the coactivator PGC-1beta and obesity. *J. Med. Genet.* 42(5), 402–407 (2005).

•• **This article provides some evidence of association between the *PPARGC1B* gene and obesity in a Caucasian population.**

64 Esterbauer H, Oberkofler H, Krempler F, Patsch W. Human peroxisome proliferator activated receptor gamma coactivator 1 (*PPARGC1*) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression. *Genomics* 62(1), 98–102 (1999).

65 Patti ME, Butte AJ, Crunkhorn S *et al.* Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: potential role of PGC1 and NRF1. *Proc. Natl Acad. Sci. USA* 100(14), 8466–8471 (2003).

66 Barr DB, Bishop A, Needham LL. Concentrations of xenobiotic chemicals in the maternal-fetal unit. *Reprod. Toxicol.* 23(3), 260–266 (2007).

67 Leino O, Kiviranta H, Karjalainen AK *et al.* Pollutant concentrations in placenta. *Food Chem. Toxicol.* 54 59–69 (2013).

68 Dixit M, Choudhuri G, Keshri LJ, Mittal B. Association of low density lipoprotein receptor related protein-associated protein (*LRPAP1*) gene insertion/deletion polymorphism with gallstone disease. *J. Gastroenterol. Hepatol.* 21(5), 847–849 (2006).

69 Wirtenberger M, Tchatchou S, Hemminki K *et al.* Associations of genetic variants in the estrogen receptor coactivators PPARGC1A, PPARGC1B and EP300 with familial breast cancer. *Carcinogenesis* 27(11), 2201–2208 (2006).

70 Kerley-Hamilton JS, Trask HW, Ridley CJ *et al.* Inherent and benzo[a]pyrene-induced differential aryl hydrocarbon receptor signaling greatly affects life span, atherosclerosis, cardiac gene expression, and body and heart growth in mice. *Toxicol. Sci.* 126(2), 391–404 (2012).

71 Horikoshi M, Yaghootkar H, Mook-Kanamori DO *et al.* New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat. Genet.* 45(1), 76–82 (2013).

72 Weksberg R, Shuman C, Caluseriu O *et al.* Discordant KCNQ1OT1 imprinting in sets of monozygotic twins discordant for Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* 11(11), 1317–1325 (2002).

73 Pettenati MJ, Haines JL, Higgins RR, Wappner RS, Palmer CG, Weaver DD. Wiedemann-Beckwith syndrome: presentation of clinical and cytogenetic data on 22 new cases and review of the literature. *Hum. Genet.* 74(2), 143–154 (1986).

74 Mestdagh P, Hartmann N, Baeriswyl L *et al.* Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. *Nat. Methods* 11(8), 809–815 (2014).

75 Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat. Rev. Genet.* 14(8), 585–594 (2013).

76 Armstrong DA, Lesseur C, Conradt E, Lester BM, Marsit CJ. Global and gene-specific DNA methylation across multiple tissues in early infancy: implications for children's health research. *FASEB J.* 28(5), 2088–2097 (2014).

77 M Hossain MB, Lindh C *et al.* Early life low-level cadmium exposure is positively associated with increased oxidative stress. *Environ. Res.* 112, 164–170 (2012).

78 Avila L, Yuen RK, Diego-Alvarez D, Penaherrera MS, Jiang R, Robinson WP. Evaluating DNA methylation and gene expression variability in the human term placenta. *Placenta* 31(12), 1070–1077 (2010).

Affiliations

Nadia Vilahur
ISGlobal, Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; and,
Universitat Pompeu Fabra (UPF), Barcelona, Spain; and,
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Genomics & Disease Group, Bioinformatics & Genomics Program, Centre for Genomic Regulation (CRG), Barcelona, Spain

Mariona Bustamante
ISGlobal, Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; and,
Universitat Pompeu Fabra (UPF), Barcelona, Spain; and,
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Genomics & Disease Group, Bioinformatics & Genomics Program, Centre for Genomic Regulation (CRG), Barcelona, Spain

Eva Morales
ISGlobal, Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; and,
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
IMIB-Arrixaca Research Institute, Virgen de la Arrixaca University Hospital, Murcia, Spain

Valeria Motta
EPIGET – Epidemiology, Epigenetics & Toxicology Lab – Department of Clinical Sciences & Community Health, Università degli Studi di Milano, Milan, Italy

Mariana Fátima Fernandez
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Department of Radiology, University of Granada, Spain; and,
Instituto de Investigación Biosanitaria, ibs.GRANADA, Spain

Lucas Andrés Salas
ISGlobal, Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; and,
Universitat Pompeu Fabra (UPF), Barcelona, Spain; and,
CIBER Epidemiología y Salud Pública (CIBERESP), Spain

Georgia Escaramis
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Genomics & Disease Group, Bioinformatics & Genomics Program, Centre for Genomic Regulation (CRG), Barcelona, Spain

Ferran Ballester
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
FISABIO-Universitat de València – Universitat Jaume I Joint Research Unit of Epidemiology & Environmental Health, Valencia, Spain; and,
University of Valencia, Valencia, Spain

Mario Murcia
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
FISABIO-Universitat de València – Universitat Jaume I Joint Research Unit of Epidemiology & Environmental Health, Valencia, Spain

Adonina Tardon
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
University of Oviedo, Oviedo, Asturias, Spain

Isolina Riaño
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Hospital San Agustín, SESPA, Asturias, Spain

Loreto Santa-Marina
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Biodonostia, Health Research Institute, San Sebastián, Spain; and,
Public Health of Gipuzkoa, Department of Health, Government of the Basque Country, San Sebastian, Spain

Jesús Ibarluzea
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Biodonostia, Health Research Institute, San Sebastián, Spain; and,
Public Health of Gipuzkoa, Department of Health, Government of the Basque Country, San Sebastian, Spain

Juan Pedro Arrebola
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Department of Radiology, University of Granada, Spain; and,
Instituto de Investigación Biosanitaria, ibs.GRANADA, Spain

Xavier Estivill
Universitat Pompeu Fabra (UPF), Barcelona, Spain; and,
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Genomics & Disease Group, Bioinformatics & Genomics Program, Centre for Genomic Regulation (CRG), Barcelona, Spain; and,
Hospital del Mar Research Institute (IMIM), Barcelona, Spain

Valentina Bollati
EPIGET – Epidemiology, Epigenetics & Toxicology Lab – Department of Clinical Sciences & Community Health, Università degli Studi di Milano, Milan, Italy

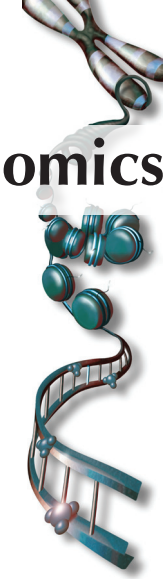
Jordi Sunyer
ISGlobal, Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain; and,
Universitat Pompeu Fabra (UPF), Barcelona, Spain; and,
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Hospital del Mar Research Institute (IMIM), Barcelona, Spain

Nicolás Olea
CIBER Epidemiología y Salud Pública (CIBERESP), Spain; and,
Department of Radiology, University of Granada, Spain; and,
Instituto de Investigación Biosanitaria, ibs.GRANADA, Spain

Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



Associative DNA methylation changes in children with prenatal alcohol exposure

Aim: Prenatal alcohol exposure (PAE) can cause fetal alcohol spectrum disorders (FASD). Previously, we assessed PAE in brain tissue from mouse models, however whether these changes are present in humans remains unknown. **Materials & methods:** In this report, we show some identical changes in DNA methylation in the buccal swabs of six children with FASD using the 450K array. **Results:** The changes occur in genes related to protocadherins, glutamatergic synapses, and hippo signaling. The results were found to be similar in another heterogeneous replication group of six FASD children. **Conclusion:** The replicated results suggest that children born with FASD have unique DNA methylation defects that can be influenced by sex and medication exposure. Ultimately, with future clinical development, assessment of DNA methylation from buccal swabs can provide a novel strategy for the diagnosis of FASD.

Keywords: cohesin • CTCF • fetal alcohol spectrum disorders • glutamatergic synapse • hippo signaling • neurodevelopment • protocadherin • synapse

Benjamin I Laufer[1], Joachim Kapalanga[2], Christina A Castellani[1], Eric J Diehl[1], Liying Yan[3] & Shiva M Singh*[,1,2,4]

[1]Molecular Genetics Unit, Department of Biology, The University of Western Ontario, London, ON, N6A 5B7, Canada

[2]Department of Pediatrics, The University of Western Ontario, London, ON, Canada

[3]EpigenDx Inc., Hopkinton, MA 01748, USA

[4]Program in Neuroscience, The University of Western Ontario, London, ON, Canada

*Author for correspondence: Tel.: +1 519 661 3135 ssingh@uwo.ca

Prenatal alcohol exposure (PAE) can cause a continuum of heterogeneous defects termed fetal alcohol spectrum disorders (FASD). Patients affected with FASD show life-long defects, particularly affecting the central nervous system. At the molecular level PAE disturbs critical signaling related to apoptosis and cell-to-cell interactions [1]. Furthermore, PAE causes changes to synaptic signaling and disrupts the neuroendocrine system, where it causes enhanced activity of the hypothalamic–pituitary–adrenal (HPA) axis [2]. Finally, the effects are confounded by genetic variation in pathways related to the immediate toxicity, which include ethanol's metabolite acetaldehyde and oxidative stress pathways [3].

The question of how fetal alcohol exposure causes long-term molecular and behavioral defects has been difficult to ascertain. It has necessitated the development of mouse models [4] and genome-wide technologies [5]. Results show that exposure to alcohol during neurodevelopment causes behavioral disabilities in the resulting offspring that mimic human FASD patients in a C57BL/6J (B6) mouse model [6]. Furthermore, the resulting mice show life long changes in brain gene expression [7] as well as 'footprints' of exposure from critical windows of epigenetic programming, particularly related to DNA methylation and noncoding RNA (ncRNA) expression [8]. This effect of PAE in rodent models has been reported from a number of independent laboratories [9–15] over the last decade. They are now viewed as a milestone in FASD research [16,17]. The genes affected participate in the critical neural processes of synaptogenesis, apoptosis, cellular identity, cell–cell adhesion, and signaling [7,8].

These results are best explained by observations (in alcoholism and FASD research) that alcohol directly affects the dynamics of the epigenome via altering one-carbon metabolic pathways from which methyl groups are derived [18–20]. Developmental alcohol exposure also results in altered expression of the genes that establish, maintain and read DNA methylation, such as *Dnmt1, Dnmt3a, Mecp2,* and *Mtrr* [7,21,22]. Furthermore, using

Future Medicine part of fsg

genome-wide assays to interrogate the cytosine methylation and ncRNA expression within the brain of adult mice prenatally exposed to moderate amounts of alcohol, we have identified a transcriptomic and epigenomic footprint from exposure during critical windows of development [5,8,23–25]. This 'footprint' is characterized by alterations to imprinted regions of the genome that encode multiple developmentally important noncoding RNAs (ncRNAs) and are regulated by CCCTC-binding factor (CTCF), a zinc finger protein that functions as a genomic insulator [26]. The canonical pathways affected include PTEN/PI3K/AKT/mTOR pathway where the alterations affect cell adhesion and proliferation [8,12,27]. These findings suggest that 'footprints' in the epigenome may represent past developmental transcriptional aberrations as well as current ones in related cells that are either active or poised for a response. In fact, these processes are developmentally integral to the outcome of larger scale cortical brain structure formation, since the events of both pre- and postnatal neurodevelopment are highly dependent on the (epi)genotype, as well as the molecular environment of the differentiating cells [28]. Taken together, current evidence suggests that ethanol can create long-term changes in the epigenetic landscape that are essential and actively needed for remaining neurodevelopment.

Of special interest to this report are the genes related to the clustered protocadherin gene family and glutamatergic synapses. Protocadherins are expressed in the nervous system where they mediate highly complex cell-to-cell interactions. The protocadherin molecules affect the adhesion of a number of synaptic events, including glutamatergic synapses. Glutamate is the most abundant excitatory neurotransmitter of the nervous system and has a role in synaptic plasticity. As such, it is a key molecular component of learning and memory. Glutamate is known for its involvement in the plasticity process of long-term potentiation and is utilized in the hippocampus and neocortex, among other regions, and the related signaling is particularly sensitive to ethanol. On the other hand, protocadherins are transmembrane proteins with repeats of the cadherin motif or the cadherin extracellular domain. This domain contains a conserved Ca^{2+} binding sequence and the number of extracellular domains may vary across different cadherin molecules. The protocadherins function as cell adhesion molecules and create individual neuronal identities that allow for large-scale network formation via controlling synaptic interactions. They have been implicated in a number of brain disorders with similar endophenotypes to FASD such as autism spectrum disorders, bipolar disorder, and schizophrenia [29].

In this report we analyzed children (3–6 year olds) with a comprehensive FASD diagnosis and a maternal history of PAE (termed the discovery cohort) for genome-wide DNA methylation changes that are analogous to the results on mice. This effect is also replicated in an independent group of FASD children (termed the replication cohort) collected 6 months apart. The repeated results establish that DNA methylation is affected at specific genomic loci in children with FASD. The analysis allowed for the identification of protocadherins, glutamate, and hippo signaling factors as critical molecules in the etiology of this neurodevelopmental disorder. These results may also explain overlaps involving FASD and other neurodevelopmental disorders of unknown etiology.

Materials & methods

The results included in this report are based on two cohorts (discovery and replication) of young children clinically diagnosed with FASD by J Kapalanga in Ontario, Canada (Table 1). The discovery cohort consisted of 6 FASD patients and 5 matched controls, while the replication cohort consisted of a different set of 6 FASD children and 6 matched controls. The diagnosis shows that FASD patients from the discovery cohort are much more homogeneous and better matched to controls with respect to age (3–6 year olds), sex (males only), socioeconomic status, maternal history of drinking, life exposures, and ancestry (Table 1A). All subjects had Northern-European ancestry and the matched controls were visiting the clinic for annual check-ups or relatively minor non-neurodevelopmental or behavioral problems like asthma, allergies, musculoskeletal problems, or gastrointestinal problems. Following ethical approval and informed consent from parents or guardians, the study subjects contributed cheek swabs for analysis of buccal epithelial DNA. While four (R2, R3, R5 and R6) of the six patients in the replication cohort were 6–10 year old males and not on medication, two were older 7 and 9 year old females (R1 and R4) and on psychotropic/stimulant medication (Table 1B). Thus, while the replication cohort matches with the discovery cohort, the replication cohort is different in that some are females, on medication, and/or have a wider age range.

Human 450K Infinium Methylation BeadChip

Individual cheek swabs were used to isolate genomic DNA using the QIAamp DNA Mini Kit following manufacturer's protocol. This DNA was processed following sodium bisulfite modification at The Centre for Applied Genomics (Toronto, Canada). The genome-wide CpG methylation for each subject was assessed using the Illumina HumanMethylation450 BeadChip

Table 1. Clinical features of patients from the discovery and replication cohorts with the diagnosis of fetal alcohol spectrum disorders.

(A) Discovery cohort						
Patient ID	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**
Patient (age, years) and sex	**(6) M**	**(6) M**	**(5) M**	**(4) M**	**(3) M**	**(3) M**
Clinical features						
Developmental delay	x	x	x	x	x	x
Hyperactivity, poor impulse control	x	x	x	x	x	x
Learning disorders	x	x	x			
Short attention span and inattention	x	x		x		
Conduct disorder	x	x	x		x	
Oppositional defiant disorder	x			x	x	
Social difficulties	x			x	x	
Nervousness and anxiety		x	x	x		
Mood disorder	x			x		
Macrocephaly/macrotia		x	x			
Microcephaly/microtia				x	x	
Distinctive facial features	x			x	x	x
Stimulant/psychotrophic medication	No	No	No	No	No	No
(B) Replication cohort						
Patient ID	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**
Patient (age, years) and sex	**(7) F**	**(6) M**	**(6) M**	**(9) F**	**(6) M**	**(10) M**
Clinical features						
Developmental delay	x	x	x		x	x
Hyperactivity, poor impulse control	x	x	x	x	x	x
Learning disorders	x		x		x	x
Short attention span and inattention	x	x	x	x	x	x
Conduct disorder	x	x	x	x	x	x
Oppositional defiant disorder	x	x	x	x		
Social difficulties	x	x	x			
Nervousness and anxiety	x	x	x			x
Mood disorder			x	x		x
Macrocephaly/macrotia		x				
Microcephaly/microtia						x
Distinctive facial features	x		x			x
Stimulant/psychotrophic medication	Yes	No	No	Yes	No	No

following the manufacturer's protocol. The array quantifies the methylation status of over 480,000 CpGs in the human genome. Arrays were scanned using Illumina's GenomeStudio at The Centre for Applied Genomics (Toronto). It generated scan data (.idat files) for each human subject used in this analysis.

The .idat files were analyzed using Partek Genomics Suite® Version 6.6. A site-based DNA methylation workflow for Illumina BeadArray Methylation was performed as the primary analysis. Preprocessing using both control normalization and background subtraction was performed using Illumina's algorithm to generate β-values. β-values were then normalized, converted to a fold-change, and subjected to a one-way ANOVA analysis to identify regions of differential methylation using annotations from Hg18 for the primary analysis. Custom genome dot plots (also known as Manhattan plots) were generated from the ANOVA

results. For hierarchical clustering and gene ontologies from Enrichr [30] a stringency of $p < 0.005$ was used. For Partek Pathways a gene list with a $p < 0.05$ stringency was used. The subset-quantile within array normalization (SWAN) algorithm [31] from the minifi bioconductor package was used within Partek for a secondary analysis to compare across batches [32].

Confirmation by pyrosequencing

Using the same bisulfite converted buccal swab DNA from the discovery cohort, EpigenDx performed pyrosequencing on the PSQ96 HS System (Qiagen) following the manufacturer's instructions, using custom assays [33], and a gradient of controls with known methylation levels. This allowed for the quantification of the absolute percent methylation [34] of each CpG at select loci using QCpG software (Qiagen). The sequencing reads also allowed for the analysis of SNPs known to be within or close to the CpGs of the 450K probes.

Comparative mouse model epigenomics

All the significant ($p < 0.01$) annotated peaks from the mouse model Differential Enrichment Peak (DEP) analysis were extracted and combined for peak filtering. Peaks from the MeDIP CpG Island tilling arrays are represented as unique stretches of assembled contiguous probes showing significant alterations to DNA methylation. This near base-pair resolution data was previously generated from the mouse model using a number of bioinformatic scripts and packages [8]. The filtering criteria was set as peaks within both top 20% PeakScore and top 20% PeakDMValue. Alignment of reference genomes was done using the UCSC genome browser [35].

Results

Neurodevelopmental alcohol exposure causes differential genome-wide DNA methylation

Using a stringent but not false discovery rate (FDR) corrected CpG list (children $p < 0.005$ and mice $p < 0.01$) genome-wide differential methylation was analyzed in: mice brains subjected to neurodevelopmental alcohol; buccal swab DNA of FASD patients and matched controls; and also a replication cohort (Figure 1). In each case, the heat maps show that mice with FASD phenotypes and children diagnosed with FASD clearly group together and are distinct from matched non-FASD controls. The results have also allowed for the identification of specific genes that are concordant between the three independent sets of results. The experiments involving mice were described earlier and are based on a MeDIP-CHIP tilling array that analyzes CpG islands with greater than 2 CpGs affected per a 50–75 base pair probe and then tilled into larger contiguous sequences [8]. For the children, we used the 450K methylation array, which analyzes alterations to single CpGs via sodium bisulfite conversion and an epigenotyping approach. From theses arrays a one-way ANOVA was created to generate a list of CpGs and corresponding genes that showed significant differential methylation between FASD children and matched controls (after extensive preprocessing and data normalization). The results of the discovery and replication cohorts were analyzed separately in order to avoid any batch effects and confounding differences from sample heterogeneity. These effects appeared to reduce the statistical power more substantially than the limits of sample size in this experimental design. Therefore, in order to accommodate for some of the limitations of sample size the list for the discovery cohort was filtered down to a stringent, but not FDR corrected, 269 significant ($p < 0.005$) differentially methylated CpG sites (Supplementary Data). Interestingly, of the sites showing an increase in methylation, 21 represent sites in the regulatory features of the clustered protocadherins genes located on 5q31 (Figure 1D & Tables 2 & 3). This preferential methylation of the protocadherin cluster is repeated in the mice as well as the replication samples with some variations. Furthermore, alterations to imprinted ncRNA from the *SNRPN-UBE3A* locus (*Snord 116* [*HBII-85*] [$p < 0.005$] and *Snord 115* [*HBII-52*] [$p < 0.01$]) were present in the human buccal DNA, which was the most ubiquitous marker for prenatal alcohol exposure in our mouse models.

Finding signal(s) in the noise

We used differentially methylated genes (with a less stringent statistical cut-off, $p < 0.05$) from the discovery cohort that represented established maternal drinking and uniform manifestation of FASD (developmental delays, hyperactivity/poor impulse control, and mental deficits) in an analysis towards identification of specific biological processes, functions, and pathways affected as a result of changes in methylation. While the less stringent cut-off for this gene list doesn't account for multiple testing, the downstream ontology and pathways analysis do as they function on the principles of independent component analysis. The results (Table 4) show that the major biological processes affected include nervous ($p = 0.00000001$), anatomical ($p = 0.00000008$) and skeletal ($p = 0.006$) development, cell adhesion ($p = 0.00002$), and cell projection assembly ($p = 0.002$). These processes appear to be affected by significant alterations ($p < 0.05$) to major molecular functions including neuropeptide binding and receptors, protein interactions, and channel activity. These genes and their associated molecular functions are also involved in major canonical pathways

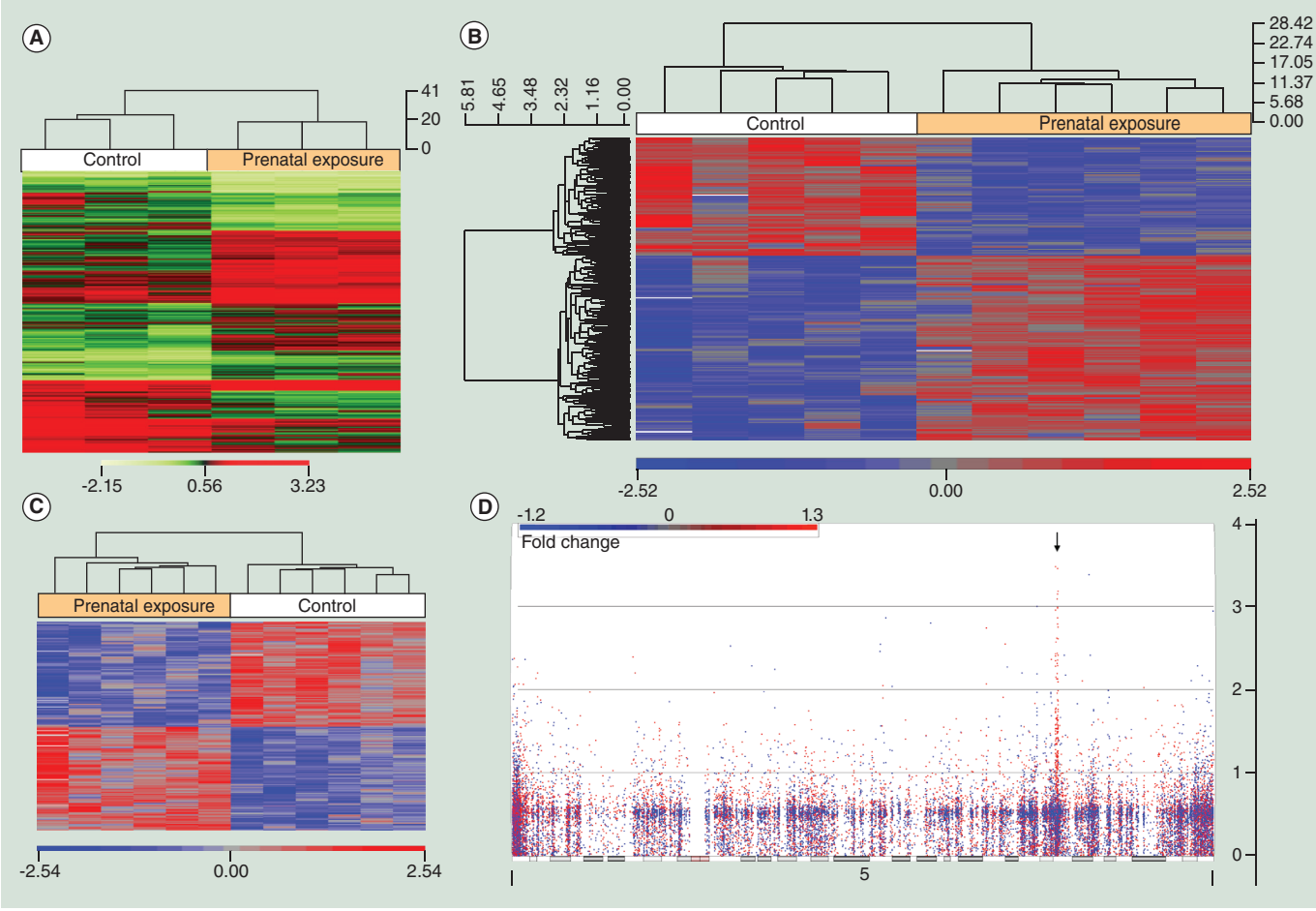


Figure 1. Methylation signals from fetal alcohol exposed mice and children diagnosed with fetal alcohol spectrum disorders. (A) Heatmap of top significant ($p < 0.01$) contigs from the adult mouse prenatally exposed to alcohol. **(B)** Heatmap of significantly ($p < 0.005$) differentially methylated targeted CpG sites generated using hierarchical clustering of buccal epithelial DNA obtained via swab from fetal alcohol spectrum disorders ($n = 6$) and matched control ($n = 5$) children from the discovery cohort. **(C)** Heatmap of significantly ($p < 0.005$) differentially methylated targeted CpG sites generated using hierarchical clustering of buccal epithelial DNA obtained via swab from fetal alcohol spectrum disorders ($n = 6$) and matched control ($n = 6$) children from the replication cohort. **(D)** Manhattan plot for human chromosome 5 from the discovery sample. Genomic location is plotted on the x-axis and $-\log_{10}$ (p-value) of differential CpG methylation from prenatal alcohol exposure is on the y-axis. Each dot represents a single CpG site. A red dot indicates an increase in methylation and a blue dot indicates a decrease in methylation. The black arrow identifies clustered protocadherin genes on 5q31.

including hippo signaling, glutamatergic synapses, serotonergic synapses, and cholinergic synapses as well as axon guidance, focal adhesion, and apoptosis. Interestingly, the top three enriched canonical pathways are hippo signaling ($p = 0.0002$), glutamatergic synapse ($p = 0.001$), and calcium signaling ($p = 0.005$).

The FASD methylation footprint includes protocadherin genes

Figure 2 shows the Genemania network generated using the stringent gene list (269 CpG sites, $p < 0.005$) that showed significant altered methylation in patients with FASD. The results reveal involvement of members of the α, β, and γ family of protocadherin genes (Figure 1D). Tables 2 & 3 include probes showing significant differential methylation from children ($p < 0.005$) and mice ($p < 0.01$) in clustered protocadherins. Tables 2 & 3 also include corresponding protocadherin genes that were differentially methylated in adult brains following neurodevelopmental alcohol exposure in mice. Figure 3A includes a plot of the raw Illumina β values from significant ($p < 0.005$) CpGs in the differentially methylated region of the protocadherin γ cluster and how they align with alternate transcripts choice. Protocadherins are involved in the generation of synaptic complexity in the developing brain and evolved primarily by tandem duplications and divergence. This complex locus contains three clusters (α, β, and γ) of protocadherin gene families that are transcriptionally complex and very similar between humans and mice (Figure 3B). The clustered protocadherin locus generates comparable transcripts in the two

Table 2. Nucleotide-specific analysis of CpGs in the protocadherin gene family clusters showing significant ($p < 0.005$) alterations to CpG methylation in human buccal epithelial DNA from children of the discovery sample (with fetal alcohol spectrum disorders compared to matched controls).

Family	p-value	CpG ID	CpG coordinate	Fold change	Average β
α	2.24E-03	cg03318614	140154796	1.1	0.55
α	4.90E-03	cg16234335	140168303	1.11	0.6
α	4.07E-03	cg25225155	140202437	1.23	0.64
α	3.79E-03	cg13619597	140237001	1.54	0.68
α	4.42E-04	cg15122993	140236606	1.17	0.49
β	1.99E-03	cg05941060	140594004	1.43	0.46
β	1.90E-03	cg27086874	140549040	1.14	0.19
γ	8.85E-04	cg25657261	140705331	1.34	0.33
γ	1.37E-03	cg01561869	140704042	1.22	0.29
γ	4.86E-03	cg26856475	140734108	1.38	0.42
γ	1.08E-03	cg22087053	140734498	1.27	0.35
γ	6.14E-05	cg24922090	140720687	1.24	0.37
γ	6.18E-04	cg18705909	140754610	1.58	0.45
γ	4.98E-03	cg24633027	140752470	1.32	0.55
γ	8.56E-04	cg03831054	140769892	1.53	0.48
γ	4.36E-04	cg21117330	140769634	1.2	0.28
γ	8.64E-04	cg21915313	140779155	1.5	0.39
γ	3.24E-03	cg02074191	140731690	1.6	0.46
γ	1.36E-04	cg06757405	140731792	1.35	0.35
γ	4.41E-03	cg08854987	140759645	1.26	0.34
γ	8.46E-04	cg18297751	140749183	1.31	0.39

species and has relatively conserved CpG sites. More importantly, a number of the CpG sites/islands are specifically sensitive to increased methylation following alcohol exposure in both species.

The top canonical pathways are related to neurodevelopment & synaptic signaling

The top canonical pathways identified in this analysis involve hippo signaling, glutamatergic synapses, and calcium signaling, all of which are fundamental pathways related to neurodevelopment. Hippo signaling (Figure 4A) had 25 out of 120 genes affected, was the most significant ($p = 0.0002$) canonical pathway, and is a highly conserved signaling pathway studied in cancer that is involved in the control of cell growth and proliferation, apoptosis, and organ size control. It is also interconnected to the TGFβ and Wnt signaling pathways. Interestingly, hippo signaling has recently been linked to chromatin modification by recruiting a histone methyltransferase complex [36] and regulates the growth of niche stem cell precursors [37]. The glutamatergic synapse pathway (Figure 4B) was the top synaptic pathway and showed a significant ($p = 0.001$) enrichment with 18 out of 77 genes belonging to it showing significant (less stringent list, $p < 0.05$) alterations to CpG methylation. A number of glutamate receptors are G-protein coupled receptors and are involved in secondary messenger signaling. Also of importance is the significant ($p = 0.005$) alteration to the calcium signaling pathway, with 23 out of 102 genes affected. Calcium signaling has numerous fundamental and diverse roles in (neural) signaling. The relevance of these pathways to neural events is an unexpected result of easily obtained buccal swabs.

The results of discovery cohort are shared by the replication cohort

The replication cohort (R) included six patients (3–10 years old) and six matched controls (Table 1B) and their cheek swab DNA was used in genome-wide DNA methylation profiles (Figure 1C). They were recruited from the same clinic and diagnosed by the same Pediatrician (J Kapalanga). However, the replication cohort was also much more heterogeneous compared with the discovery cohort with

respect to sex, medication exposure, environment, and socioeconomic status. In order to compare the cohorts across batches a different preprocessing algorithm was selected: the subset-quantile within array normalization (SWAN) algorithm [31] from the minifi [38] bioconductor package. When the two cohorts were preprocessed separately by the Illumina algorithm or together with the SWAN algorithm, they produced genome-wide results that had similarities but were not identical to the discovery cohort. There was a trend of decreased significance in the replication cohort and greater levels of variation. Despite these differences, however, the main similarities between the two cohorts were significant alterations in ontologies related to glutamate and the canonical glutamatergic synapse pathway. Additionally, there was a consistent and prominent protocadherin signal on the Manhattan plot, as well as replication of alterations to imprinted regions also seen in the mouse models and the discovery cohort.

Next we reanalyzed the replication cohort results following removal of females and children on medication (Table 1B; R1 and R4). Also, four controls of the replication cohort were removed to better match age range and medical history. The selected scans were then preprocessed using the Illumina algorithm. Upon removal, the groups cluster much more closely and the signal is almost identical to what was observed in the discovery cohort. Glutamatergic synapses, cell-adhesion, neuroactive ligand receptor interactions, MAPK signaling, axon guidance, and a number of other pathways were all significant ($p < 0.005$) even though the ranking of specific pathways varied slightly. The results argue that although DNA methylation is altered in patients with FASD, other factors, including sex, age, and medication history, may affect the final outcome.

Confirmations via pyrosequencing

Next we sought to confirm the microarray results using an independent technology that does not depend on the matched controls for modeling. For this we chose probes with different performances from the carefully matched discovery cohort. First we chose the only CpG to pass multiple testing at the FDR filtering level in this relatively small cohort. The analysis identified an enhancer of *COLEC11* (FDR $p = 1.93 \times 10^{-7}$). Pyrosequencing analysis of the affected CpG (cg15730644) revealed that the *COLEC11* CpG shows a significant increase ($p = 0.002$) with an average methylation of 94.2% in FASD compared with 79.8% in matched controls (Table 5). It also shows that all children affected with FASD consistently have higher (~94% of cells) methylation at this CpG site compared with their matched controls (~80% of cells). The sequencing run also examined SNP rs182514706 (T>G) and found that it was not present in any members of either the exposed or control groups and thus would not interfere with the confidence in the performance of this probe. *COLEC11* is involved in fundamental developmental processes and serves as a guidance cue for neural crest cell migration where aberrations are known to produce spectrum disorders with similar endophenotypes to FASD known as 3MC syndrome [39].

Table 3. Tiling-based analysis of protocadherin gene clusters showing significant ($p < 0.01$) increases in methylation in adult brain tissue from a mouse model of fetal alcohol spectrum disorders.

Family	Peak start	Peak end	Peak length
α	37090887	37091731	844
α	37099823	37099982	159
α	37100493	37101082	589
α	37120628	37121169	541
α	37121813	37122462	649
α	37128647	37129396	749
α	37148470	37149114	644
α	37153966	37154515	549
α	37159353	37160304	951
α	37166096	37167215	1119
α	37181780	37182430	650
β	37456114	37456763	649
β	37603162	37603321	159
β	37645176	37646223	1047
β	37657621	37657881	260
β	37674825	37674994	169
β	37678546	37679275	729
γ	37828934	37831185	2251
γ	37835333	37836297	964
γ	37840948	37841597	649
γ	37850740	37850974	234
γ	37854441	37854897	456
γ	37868606	37868955	349
γ	37886555	37887889	1334
γ	37892011	37893005	994
γ	37922706	37922835	129

Results are from methylated DNA immunoprecipitation (MeDIP) CpG Island tilling arrays and peaks are represented as unique stretches showing significant alterations to DNA methylation in prenatal alcohol exposed mice compared with matched controls.

Table 4. Gene ontology and KEGG-based canonical pathway analysis of significant differences in CpG methylation from the discovery cohort.

Ontologies and pathways	p-value
Biological processes	
Nervous system development (GO:0007399)	0.00000001
Anatomical structure development (GO:0048856)	0.00000008
Cell adhesion (GO:0007155)	0.00002176
Calcium-dependent cell–cell adhesion (GO:0016339)	0.00073511
Skeletal system development (GO:0001501)	0.00562206
Cell projection assembly (GO:0030031)	0.01521803
Cell–cell adhesion (GO:0016337)	0.01396605
Molecular functions	
Neuropeptide binding (GO:0042923)	0.024
Neuropeptide receptor activity (GO:0008188)	0.022
Sodium channel activity (GO:0005272)	0.016
Protein binding (GO:0005515)	0.007
Voltage-gated sodium channel activity (GO:0005248)	0.009
RNA polymerase II transcription factor activity (GO:0003702)	0.026
Cation channel activity (GO:0005261)	0.043
Canonical pathways (number of genes affected)	
Hippo signaling pathway (25 genes)	0.0002
Glutamatergic synapse (18 genes)	0.0011
Calcium signaling pathway (23 genes)	0.0051
Retrograde endocannabinoid signaling (16 genes)	0.0062
Serotonergic synapse (16 genes)	0.0092
Axon guidance (17 genes)	0.0118
Cell adhesion molecules (18 genes)	0.0132
Long-term depression (8 genes)	0.0349
Focal adhesion (21 genes)	0.0377
Dorso–ventral axis formation (5 genes)	0.0467
Cholinergic synapse (14 genes)	0.0546
Apoptosis (10 genes)	0.0608

The next two probes examined did not pass a FDR cut-off and are representative of the gene list used for higher level analysis, including heatmaps, ontologies, and pathways. HTT (Huntington) is a major hub in neurological disorders and showed a 1.5-fold increase ($p = 0.003$) in methylation in FASD patients. Pyrosequencing analysis of the affected CpG (cg26128129) revealed that the CpG in a enhancer of HTT shows a significant increase ($p = 0.001$) with an average methylation of 87.6% in FASD and an average of 50.7% in matched controls. Interestingly, there is variation among the control samples as compared with the FASD patients (Table 5). This CpG is known to contain a SNP [rs362313 (C>T)] that causes a loss of the CpG, as well as another SNP near the probe [rs147422679 A>G]; however, pyrosequencing confirmed neither SNP was present in any of the individuals examined. The final probe assessed by pyrosequencing was PCDHB18, which is a member of the protocadherin β cluster. The affected CpG (cg27086874) is in a north CpG shore that is located in a reprogramming-specific differentially methylated region (DMR). The results show an identical pattern to the arrays and the mouse models (results not shown). The available results support the suggestion that CpG specific DNA methylation is altered in the cheek swab DNA of children affected with FASD.

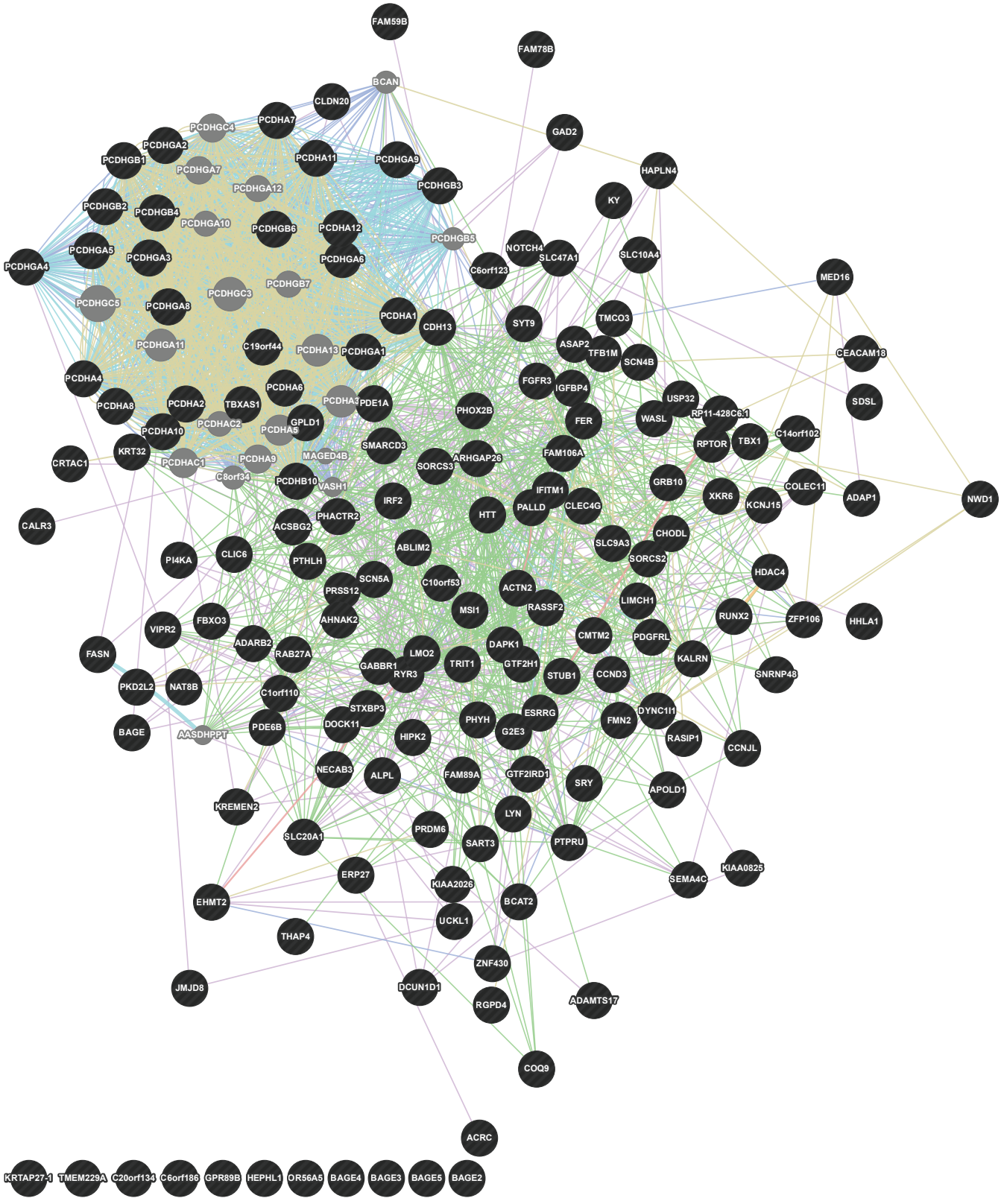


Figure 2. GeneMANIA network analysis using differentially methylated CpGs ($p < 0.005$) from cheek swabs of the discovery sample.
The network highlights preferential and synchronous alteration involving Protocadherin genes in children with fetal alcohol spectrum disorders.

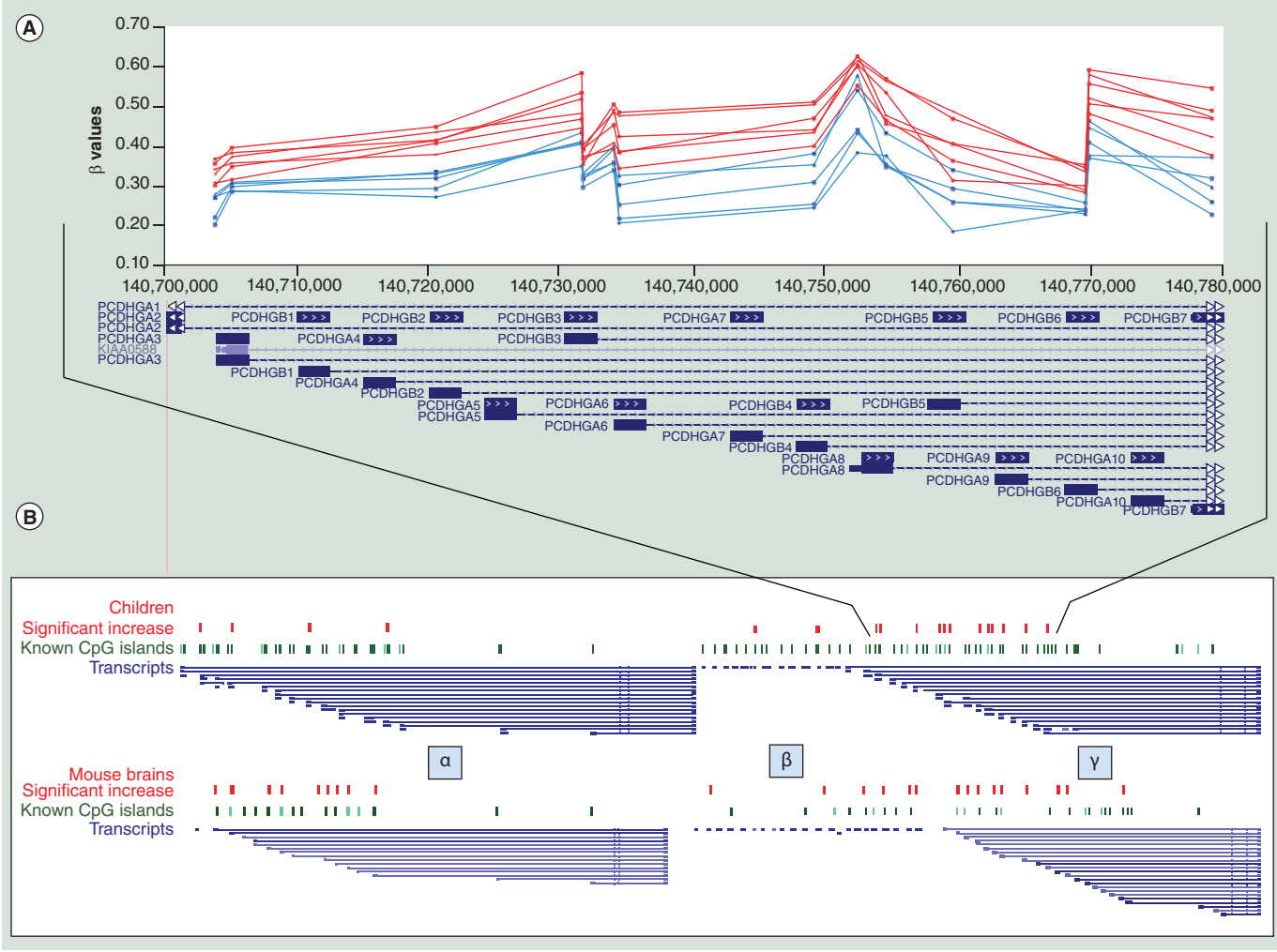


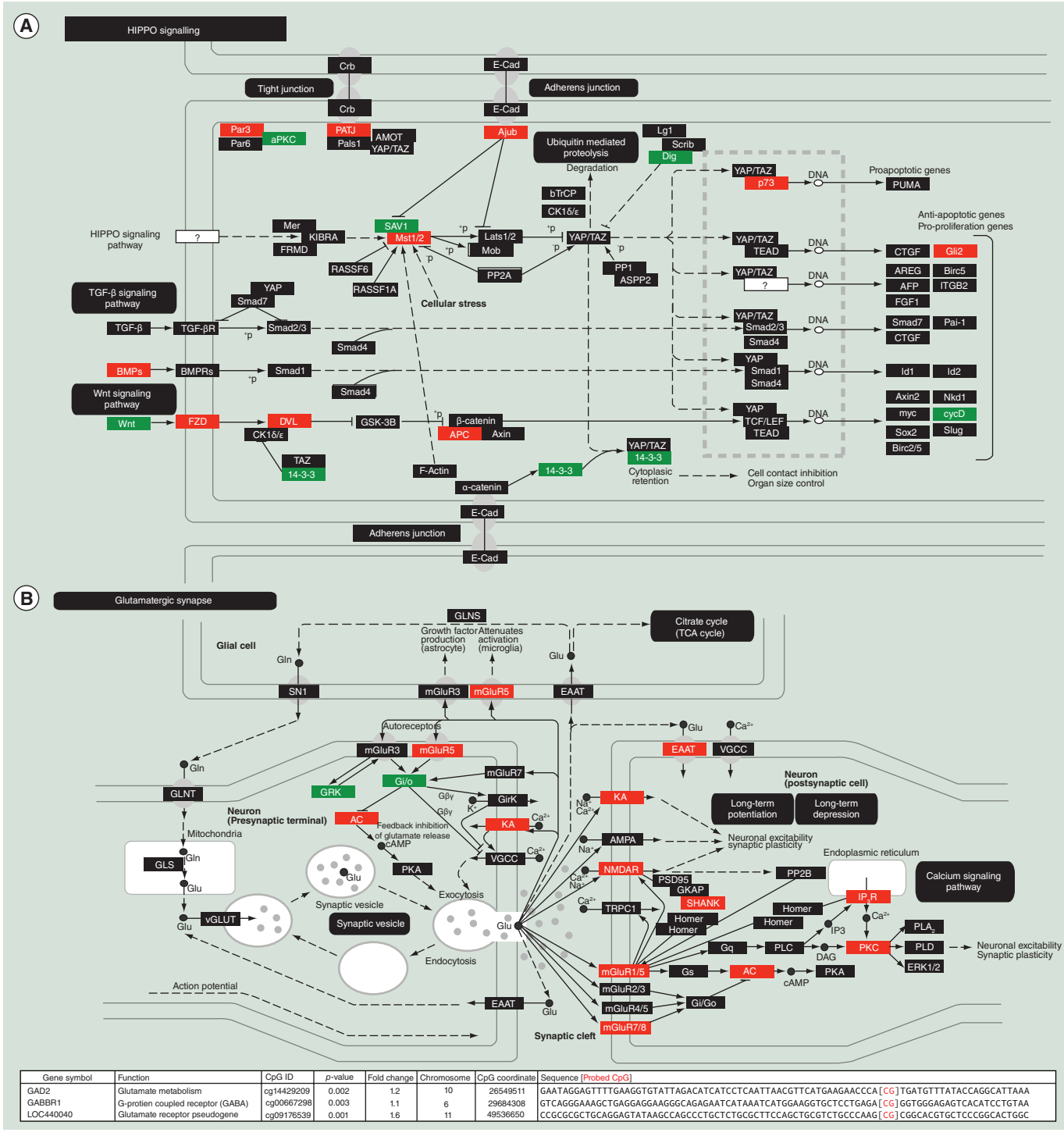
Figure 3. CpG methylation of the Protocadherins in children with fetal alcohol spectrum disorders and matched controls.
(A) Representative methylation patterns in 14 CpGs in the gamma cluster in six fetal alcohol spectrum disorders children and their five matched controls. **(B)** Significantly ($p < 0.005$) increased methylation in the clustered protocadherin region of patients (red dots), known CpG island (green dots) and potential transcripts (blue lines) along with their similarity to the mouse model ($p < 0.01$). Figure is not to scale as non-informative regions have been removed for clarity.

Discussion

Ethanol is a classic teratogen capable of inducing a wide range of developmental abnormalities. It interferes with a number of developmental processes at many levels. Of interest to this report is that PAE interferes with one-carbon metabolism [40], which is essential for DNA methylation, histone modifications, and DNA synthesis. As such, it is implicated in complex gene-environment interactions that alter gene expression patterns, especially during neurodevelopment [18,41–44]. Much of this understanding has come from animal models that have successfully generated phenotypes relevant to FASD [6,7]. The mouse results support involvement of a number of developmental pathways in the etiology of FASD [7]. Also, the underlying mechanisms may involve epigenetic remodeling that persists into adulthood [8] even though the molecular targets that could be considered as causal and 'driving' are not known. Furthermore, although the animal studies are informative, the literature lacks replication of such results in humans, which forms the primary objective of this report. The analysis of human children has also implicated a fundamental and well-known player in alcohol research: the glutamatergic synapse (Figure 4B). Additionally, the identification of hippo-signaling (Figure 4A) offers a very promising candidate for the mechanism given its role in regulating stem cells and controlling organ size. However, not much is currently known about this pathway in a context outside of cancers, which are diseases of cell differentiation, and developmental fly biology. Thus, the focus for remainder of this report is a single genomic locus, which is consistently affected in children born with FASD and in PAE mice that may also 'drive' psychiatric phenotypes, and has not yet been implicated in the literature.

Increased methylation in protocadherin genes can result in FASD phenotypes

The protocadherins are the largest subgroup of the cadherin gene superfamily of homophilic cell-adhesion proteins [45,46]. In humans, they are clustered on 5q31 and show a concerted increase in methylation in FASD (Figure 1D & Figure 3A). They are primarily expressed in the developing nervous system [47] where they are



Gene symbol	Function	CpG ID	*p*-value	Fold change	Chromosome	CpG coordinate	Sequence [Probed CpG]
GAD2	Glutamate metabolism	cg14429209	0.002	1.2	10	26549511	GAATAGGAGTTTTGAAGGTGTATTAGACATCATCCTCAATTAACGTTCATGAAGAACCCA[CG]TGATGTTTATACCAGGCATTAAA
GABBR1	G-protien coupled receptor (GABA)	cg00667298	0.003	1.1	6	29684308	GTCAGGGAAAGCTGAGGAGGAAGGGCAGAGAATCATAAATCATGGAAGGTGCTCCTGAGA[CG]GGTGGGAGAGTCACATCCTGTAA
LOC440040	Glutamate receptor pseudogene	cg09176539	0.001	1.6	11	49536650	CCGCGCGCTGCAGGAGTATAAGCCAGCCCTGCTCTGCGCTTCCAGCTGCGTCTGCCCAAG[CG]CGGCACGTGCTCCCGGCACTGGC

Figure 4. Top canonical (KEGG) pathways affected by altered CpG methylation in children with fetal alcohol spectrum disorders. Genes were identified and then assembled as protein complexes. **(A)** Hippo signaling; **(B)** glutamatergic synapse pathway. Green marks a decrease in methylation and potentially increased gene expression while red indicates a increase in methylation and potentially decreased gene expression.

Table 5. Pyrosequencing results for *COLEC11* and *HTT*.

Individual	*Colec11* panel		*HTT* panel	
ID	SNP T>G rs182514706	cg15730644	cg26128129 SNP C>T rs362313	SNP A>G rs147422679
E1	T/T	91.5	89.30	A/A
E2	T/T	96.0	86.62	A/A
E3	T/T	92.0	90.84	A/A
E4	T/T	96.9	86.63	A/A
E5	T/T	94.0	84.61	A/A
E6	T/T	94.5	87.82	A/A
C1	T/T	63.8	37.37	A/A
C2	T/T	81.9	44.32	A/A
C3	T/T	84.2	84.18	A/A
C4	T/T	82.3	44.82	A/A
C5	T/T	82.2	42.57	A/A

Sample ID represents the individuals from the discovery cohort and the metrics represent percent methylation of the CpG from the sample.

stochastically involved in a neuronal self-avoidance that allows neurons to interact with the same neuronal subtype but not themselves. Indeed, protocadherins have been observed mediating cell-to-cell interactions in dendrites, axons, synapses, growth cones, and neuronal soma [48–54]. In vertebrates, the majority of protocadherin genes are organized in three clusters: α, β, and γ (Figure 3B). Intriguingly, the cell specific variation in transcripts created by this locus parallels the complexity of the immune system. Unlike the immune system, which relies on genetic recombination, the protocadherin complexity is realized by epigenetic mechanisms that parallel the act of recombination via alternative transcript choice. The transcriptional initiation of the α and γ families is dependent on the methylation status of numerous promoters in the variable regions of the family. β, on the other hand, does not contain a shared exon and each gene appears to be under the control of its own promoter. The α and γ gene families form a DNA loop with CTCF, which has methylation sensitive binding and is proposed to bring the isoform specific promoters into the proximity of flanking enhancers and transcriptional machinery. Thus, it appears that CTCF serves as a master transcription factor for this locus and is used to determine isoform expression [55]. Furthermore, the regulation of this stochastic expression is dependent on epigenetic priming by *de novo* methylation patterns established at promoters by *Dnmt3b* during early embryonic stages, which goes on to causatively influence isoform expression in subsequent differentiated cells [56].

The isoform diversity created by these clusters creates a molecular signature on the cell surface of individual neurons that is used to interact in *trans* [54]. This diversity is accomplished by individual neurons expressing different combinations of the protocadherin isoforms on their cell surface and forms even more diversity by creating multimers (in *cis*). Ultimately, each neuron could have its own unique identity amongst other individuals of the same subtype [57,58]. This is believed to be the biological basis for the specific and yet enormous network formation of precise neuronal connections required for neurodevelopment. Given this essential role underlying neurodevelopment, it comes as no surprise that the protocadherin gene family clusters have been implicated in a number of neurodevelopmental disorders [59]. Also, an increase in methylation in some, most, or all gene promoters (Tables 2 & 3), as observed in human FASD patients and all mouse models, is expected to reduce the isoform diversity, thus restraining brain function.

We also assessed the significance of affected sequences in a transcription factor analysis [30]. It revealed a significant ($p < 0.005$) enrichment for CTCF insulator binding sites [60], which are involved in establishing chromatin domains and boundaries and result in a large-scale nuclear architecture that is specific to cell type. Another transcription factor enriched for was RAD21, which cooperates with pluripotency transcription factors, including CTCF in the maintenance of embryonic stem cell identity [61]. CTCF is known to be the master regulator of the clustered protocadherins [62] where it influences complex DNA looping (Figure 3A), and many genes under its transcriptional regulation appear deregulated in our mouse models including genomically imprinted genes [8].

The increased methylation of protocadherin genes in the cheek swab DNA of FASD patients are backed by comparable results in the brains of adult mice prenatally exposed to alcohol (Figure 3B). Among the most common is *Pcdhb18,* which was shown to have

a significant (FDR corrected $p < 0.05$) decrease (40%) in adult brain gene expression in one of our mouse models of FASD [7]. Such results have the potential to cause down-regulation of specific gene products depending on the specificity of methylated promoters in neural cells of mice as well as humans. Furthermore, such changes are expected to affect function because *Pcdhb18* has been demonstrated to control axon growth and arborization in zebrafish [63]. Finally, it has not escaped our attention that this rather synchronous alteration in DNA methylation of a specific gene family with a critical role in brain functioning in children with FASD may open the door to early diagnosis of this common and complex disorder as well as offer novel strategies for its amelioration.

Conclusion & future perspective

While the analysis was theoretically limited by small sample size, it also had the practical and unquantifiable added statistical power of extremely stringent sample matching and concordance with our multiple mouse models, which showed almost identical alterations. Furthermore, the changes could be representative of alterations to individual cellular populations, particularly neurons or glia populations, but nonetheless these observations would still be meaningful as both buccal cells and neural precursor cells are related at the ectoderm primary germ layer. The connection via neuroepithelial cells also suggests a role for the developmental events of epithelial–mesenchymal transition (EMT) and the inverse mesenchymal–epithelial transition (MET), which are used in neural crest migration and neural tube formation. Thus, despite the potential caveats, we suggest that the results included may have a number of implications. First, the methylomic profile (Figures 1 & 2) may serve as a diagnostic feature of FASD, including early diagnosis from an easily accessible peripheral tissue. Second, the identification of target loci and molecules, including protocadherins, may serve as a foundation for the development of novel strategies of treatment and application of any corrective therapy, which are most effective if applied as early as possible during neurodevelopment [64].

Supplementary Data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/epi.15.60

Author contributions

The experiments were conceived and designed by BI Laufer, J Kapalanga, and SM Singh. Children were previously diagnosed by J Kapalanga and under his care swabs were collected and matched to controls. Human DNA isolation and purification was done by BI Laufer, CA Castellani, and EJ Diehl. All bioinformatic analysis was done by BI Laufer. L Yan designed and oversaw the pyrosequencing assays. BI Laufer and SM Singh wrote the manuscript with intellectual contributions from J Kapalanga and edits from CA Castellani, EJ Diehl, J Kapalanga, and L Yan. The entire project was under the supervision of SM Singh.

Acknowledgements

The authors thank I Craig for his expertise in editing some of the figures for clarity, the TCAG for expertise in array hybridization and scanning, and A Meyer from EpigenDx for technical expertise with pyrosequencing.

Disclaimer

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Financial & competing interests disclosure

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) scholarship to BI Laufer, a scholarship from NSERC to EJ Diehl, and grants from NSERC and the Canadian Institute of Health Research (CIHR) to SM Singh. L Yan is the Founder and President of EpigenDx and offered their services at a substantial collaborative discount. The authors have no other relevant affiliations or financial involve-

Executive summary

- First analysis of alterations to genome-wide DNA methylation in human children with fetal alcohol spectrum disorders (FASD).
- Results of buccal swabs from human children with FASD have similar DNA methylation patterns at candidate loci when compared with prenatal alcohol exposure (PAE) adult mouse brains.
- Results are replicated but not identical in a much more heterogeneous human FASD cohort.
- Common candidates include the clustered protocadherin locus, glutamatergic synapse pathways, and the hippo signaling pathway.
- The regions identified are not limited to ubiquitous environmentally responsive imprinted regions or one-carbon metabolic pathways.
- The candidates will be of use for future development of molecular therapies, such as somatic epigenome editing, which involves modified genome editing systems inducing precision epigenetic alterations without altering sequence.

ment with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

Male and female C57BL/6J (B6) mice were originally obtained from Jackson Laboratories (Bar Harbor, ME) and maintained at the Health Sciences Animal Care Facility at the University of Western Ontario (London, ON, Canada). The Animal Use Subcommittee of the University of Western Ontario approved all procedures undertaken on the animals and can viewed in detail in Laufer *et al.* 2013. All human samples were collected under the care of J Kapalanga at Grey Bruce Health Services and ethical approval from Grey Bruce Health Services, the University of Western Ontario, and the parents/guardians was obtained for buccal swab collection and anonymous diagnostic information. The authors state that they have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Open access

References

Papers of special note have been highlighted as: • of interest

1 Guerri C, Bazinet A, Riley EP. Foetal alcohol spectrum disorders and alterations in brain and behaviour. *Alcohol Alcohol.* 44(2), 108–114 (2009).

2 Lee S, Choi I, Kang S, Rivier C. Role of various neurotransmitters in mediating the long-term endocrine consequences of prenatal alcohol exposure. *Ann. NY Acad. Sci.* 1144, 176–188 (2008).

3 Bosco C, Diaz E. Placental hypoxia and foetal development versus alcohol exposure in pregnancy. *Alcohol Alcohol.* 47(2), 109–117 (2012).

4 Kleiber ML, Wright E, Singh SM. Maternal voluntary drinking in C57BL/6J mice: advancing a model for fetal alcohol spectrum disorders. *Behav. Brain Res.* 223(2), 376–387 (2011).

5 Kleiber ML, Laufer BI, Wright E, Diehl EJ, Singh SM. Long-term alterations to the brain transcriptome in a maternal voluntary consumption model of fetal alcohol spectrum disorders. *Brain Res.* 1458, 18–33 (2012).

6 Mantha K, Kleiber M, Singh S. Neurodevelopmental timing of ethanol exposure may contribute to observed heterogeneity of behavioral deficits in a mouse model of Fetal Alcohol Spectrum Disorder (FASD). *J. Behav. Brain Sci.* 3, 85–99 (2013).

7 Kleiber ML, Mantha K, Stringer RL, Singh SM. Neurodevelopmental alcohol exposure elicits long-term changes to gene expression that alter distinct molecular pathways dependent on timing of exposure. *J. Neurodev. Disord.* 5(1), 6 (2013).

8 Laufer BI, Mantha K, Kleiber ML, Diehl EJ, Addison SM, Singh SM. Long-lasting alterations to DNA methylation and ncRNAs could underlie the effects of fetal alcohol exposure in mice. *Dis. Model. Mech.* 6(4), 977–992 (2013).

• **Multiple mouse model characterization of prenatal alcohol exposure (PAE) that compares the transcriptomics and methylomics of moderate alcohol consumption to transcriptional changes at binge exposure time points.**

9 Olney JW, Tenkova T, Dikranian K, Qin YQ, Labruyere J, Ikonomidou C. Ethanol-induced apoptotic neurodegeneration in the developing C57BL/6 mouse brain. *Brain Res. Dev. Brain Res.* 133(2), 115–126 (2002).

10 Choi IY, Allan AM, Cunningham LA. Moderate fetal alcohol exposure impairs the neurogenic response to an enriched environment in adult mice. *Alcohol. Clin. Exp. Res.* 29(11), 2053–2062 (2005).

11 Young C, Olney JW. Neuroapoptosis in the infant mouse brain triggered by a transient small increase in blood alcohol concentration. *Neurobiol. Dis.* 22(3), 548–554 (2006).

• **Apoptosis can be triggered in the developing mouse brain by a brief and relatively low dose alcohol exposure.**

12 Green ML, Singh AV, Zhang Y, Nemeth KA, Sulik KK, Knudsen TB. Reprogramming of genetic networks during initiation of the Fetal Alcohol Syndrome. *Dev. Dyn.* 236(2), 613–631 (2007).

13 Kajimoto K, Allan A, Cunningham LA. Fate analysis of adult hippocampal progenitors in a murine model of Fetal Alcohol Spectrum Disorder (FASD). *PLoS ONE* 8(9), e73788 (2013).

14 Subbanna S, Nagre NN, Shivakumar M, Umapathy NS, Psychoyos D, Basavarajappa BS. Ethanol induced acetylation of histone at G9a Exon1 and G9a-mediated histone H3 dimethylation leads to neurodegeneration in neonatal mice. *Neuroscience* 258, 422–432 (2013).

15 El Shawa H, Abbott CW, 3rd, Huffman KJ. Prenatal ethanol exposure disrupts intraneocortical circuitry, cortical gene expression, and behavior in a mouse model of FASD. *J. Neurosci.* 33(48), 18893–18905 (2013).

16 Kleiber ML, Diehl EJ, Laufer BI *et al.* Long-term genomic and epigenomic dysregulation as a consequence of prenatal alcohol exposure: a model for fetal alcohol spectrum disorders. *Front. Genet.* 5, 161 (2014).

17 Zhou FC. Dissecting FASD through the global transcriptome. *Alcohol. Clin. Exp. Res.* 39(3), 408–412 (2015).

• **Current state of PAE transcriptomics.**

18 Haycock PC. Fetal alcohol spectrum disorders: the epigenetic perspective. *Biol. Reprod.* 81(4), 607–617 (2009).

19 Hutson JR, Stade B, Lehotay DC, Collier CP, Kapur BM. Folic acid transport to the human fetus is decreased in

pregnancies with chronic alcohol exposure. *PLoS ONE* 7(5), e38057 (2012).

- **PAE interferes with one-carbon metabolism.**

20 Ungerer M, Knezovich J, Ramsay M. *In utero* alcohol exposure, epigenetic changes, and their consequences. *Alcohol. Res.* 35(1), 37–46 (2013).

21 Perkins A, Lehmann C, Lawrence RC, Kelly SJ. Alcohol exposure during development: impact on the epigenome. *Int. J. Dev. Neurosci.* 31(6), 391–397 (2013).

22 Veazey KJ, Carnahan MN, Muller D, Miranda RC, Golding MC. Alcohol-induced epigenetic alterations to developmentally crucial genes regulating neural stemness and differentiation. *Alcohol. Clin. Exp. Res.* 37(7), 1111–1122 (2013).

- **Characterization of alterations to post-translational histone modifications in a stem cell model of PAE.**

23 Stringer R, Laufer B, Kleiber M, Singh S. Reduced expression of brain cannabinoid receptor 1 (Cnr1) is coupled with an increased complementary micro-RNA (miR-26b) in a mouse model of fetal alcohol spectrum disorders. *Clin. Epigenetics* 5(1), 14 (2013).

24 Mantha K, Laufer BI, Singh SM. Molecular changes during Neurodevelopment following second-trimester binge ethanol exposure in a mouse model of Fetal Alcohol Spectrum Disorder: from immediate effects to long-term adaptation. *Dev. Neurosci.* 36(1), 29–43 (2014).

25 Kleiber ML, Laufer BI, Stringer RL, Singh SM. Third trimester-equivalent ethanol exposure is characterized by an acute cellular stress response and an ontogenetic disruption of genes critical for synaptic establishment and function in mice. *Dev. Neurosci.* 36(6), 499–519 (2014).

26 Laufer BI, Diehl EJ, Singh SM. Neurodevelopmental epigenetic etiologies: insights from studies on mouse models of fetal alcohol spectrum disorders. *Epigenomics* 5(5), 465–468 (2013).

27 Hard ML, Abdolell M, Robinson BH, Koren G. Gene-expression analysis after alcohol exposure in the developing mouse. *J. Lab. Clin. Med.* 145(1), 47–54 (2005).

28 Sur M, Leamey CA. Development and plasticity of cortical areas and networks. *Nat. Rev. Neurosci.* 2(4), 251–262 (2001).

29 Hirano S, Takeichi M. Cadherins in brain morphogenesis and wiring. *Physiol. Rev.* 92(2), 597–634 (2012).

30 Chen EY, Tan CM, Kou Y *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14, 128–2105–14–128 (2013).

31 Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* 13(6), R44–2012–13–6-r44 (2012).

32 Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* 14, 293–2164–14–293 (2013).

33 EpigenDx. www.EpigenDx.com

34 Lim A, Candiloro I, Wong N *et al.* Quantitative methodology is critical for assessing DNA methylation and impacts on correlation with patient outcome. *Clin. Epigenetics* 6(1), 22 (2014).

35 The UCSC Genome Browser database (2015). www.genome.ucsc.edu

36 Oh H, Slattery M, Ma L, White KP, Mann RS, Irvine KD. Yorkie promotes transcription by recruiting a histone methyltransferase complex. *Cell. Rep.* 8(2), 449–459 (2014).

37 Sarikaya DP, Extavour CG. The hippo pathway regulates homeostatic growth of stem cell niche precursors in the drosophila ovary. *PLoS Genet.* 11(2), e1004962 (2015).

38 Aryee MJ, Jaffe AE, Corrada-Bravo H *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30(10), 1363–1369 (2014).

39 Rooryck C, Diaz-Font A, Osborn DP *et al.* Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. *Nat. Genet.* 43(3), 197–203 (2011).

40 Halsted CH, Villanueva JA, Devlin AM, Chandler CJ. Metabolic interactions of alcohol and folate. *J. Nutr.* 132(8 Suppl.), S2367–S2372 (2002).

41 Sathyan P, Golden HB, Miranda RC. Competing interactions between micro-RNAs determine neural progenitor survival and proliferation after ethanol exposure: evidence from an *ex vivo* model of the fetal cerebral cortical neuroepithelium. *J. Neurosci.* 27(32), 8546–8557 (2007).

- **First evidence for the role of miRNA in PAE.**

42 Haycock PC, Ramsay M. Exposure of mouse embryos to ethanol during preimplantation development: effect on DNA methylation in the H19 imprinting control region. *Biol. Reprod.* 81(4), 618–627 (2009).

43 Liu Y, Balaraman Y, Wang G, Nephew KP, Zhou FC. Alcohol exposure alters DNA methylation profiles in mouse embryos at early neurulation. *Epigenetics* 4(7), 500–511 (2009).

- **First methylomic analysis of PAE.**

44 Kaminen-Ahola N, Ahola A, Maga M *et al.* Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genet.* 6(1), e1000811 (2010).

45 Hirano S, Takeichi M. Cadherins in brain morphogenesis and wiring. *Physiol. Rev.* 92(2), 597–634 (2012).

46 Chen WV, Maniatis T. Clustered protocadherins. *Development* 140(16), 3297–3302 (2013).

- **Expert review on clustered protocadherins.**

47 Sano K, Tanihara H, Heimark RL *et al.* Protocadherins: a large family of cadherin-related molecules in central nervous system. *EMBO J.* 12(6), 2249–2256 (1993).

48 Kohmura N, Senzaki K, Hamada S *et al.* Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron* 20(6), 1137–1151 (1998).

49 Wang X, Weiner JA, Levi S, Craig AM, Bradley A, Sanes JR. Gamma protocadherins are required for survival of spinal interneurons. *Neuron* 36(5), 843–854 (2002).

50 Kallenbach S, Khantane S, Carroll P *et al.* Changes in subcellular distribution of protocadherin gamma proteins accompany maturation of spinal neurons. *J. Neurosci. Res.* 72(5), 549–556 (2003).

51 Phillips GR, Tanaka H, Frank M *et al.* Gamma-protocadherins are targeted to subsets of synapses and intracellular organelles in neurons. *J. Neurosci.* 23(12), 5096–5104 (2003).

52 Junghans D, Heidenreich M, Hack I, Taylor V, Frotscher M, Kemler R. Postsynaptic and differential localization to neuronal subtypes of protocadherin beta16 in the mammalian central nervous system. *Eur. J. Neurosci.* 27(3), 559–571 (2008).

53 Lefebvre JL, Kostadinov D, Chen WV, Maniatis T, Sanes JR. Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. *Nature* 488(7412), 517–521 (2012).

54 Chen WV, Alvarez FJ, Lefebvre JL *et al.* Functional significance of isoform diversification in the protocadherin gamma gene cluster. *Neuron* 75(3), 402–409 (2012).

55 Golan-Mashiach M, Grunspan M, Emmanuel R, Gibbs-Bar L, Dikstein R, Shapiro E. Identification of CTCF as a master regulator of the clustered protocadherin genes. *Nucleic Acids Res.* 40(8), 3378–3391 (2012).

56 Toyoda S, Kawaguchi M, Kobayashi T *et al.* Developmental epigenetic modification regulates stochastic expression of clustered protocadherin genes, generating single neuron diversity. *Neuron* 82(1), 94–108 (2014).

57 Bonn S, Seeburg PH, Schwarz MK. Combinatorial expression of alpha- and gamma-protocadherins alters their presenilin-dependent processing. *Mol. Cell. Biol.* 27(11), 4121–4132 (2007).

58 Schreiner D, Weiner JA. Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *Proc. Natl Acad. Sci. USA* 107(33), 14893–14898 (2010).

59 Kalmady SV, Venkatasubramanian G. Evidence for positive selection on Protocadherin Y gene in *Homo sapiens*: implications for schizophrenia. *Schizophr. Res.* 108(1–3), 299–300 (2009).

60 Kim TH, Abdullaev ZK, Smith AD *et al.* Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128(6), 1231–1245 (2007).

61 Nitzsche A, Paszkowski-Rogacz M, Matarese F *et al.* RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. *PLoS ONE* 6(5), e19470 (2011).

62 Guo Y, Monahan K, Wu H *et al.* CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice. *Proc. Natl Acad. Sci. USA* 109(51), 21081–21086 (2012).

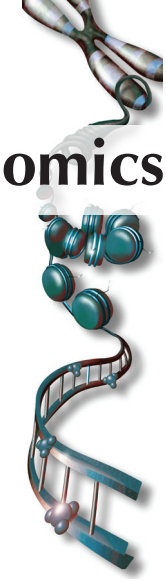
63 Biswas S, Emond MR, Duy PQ, Hao le T, Beattie CE, Jontes JD. Protocadherin-18b interacts with Nap1 to control motor axon growth and arborization in zebrafish. *Mol. Biol. Cell* 25(5), 633–642 (2014).

64 Paley B, O'Connor MJ. Intervention for individuals with fetal alcohol spectrum disorders: treatment approaches and case management. *Dev. Disabil. Res. Rev.* 15(3), 258–267 (2009).

Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



LRP1B, *BRD2* and *CACNA1D*: new candidate genes in fetal metabolic programming of newborns exposed to maternal hyperglycemia

Andrée-Anne Houde[‡,1,2], Stephanie-May Ruchat[‡,1,2], Catherine Allard[3], Jean-Patrice Baillargeon[3], Julie St-Pierre[2,4,5], Patrice Perron[2,3], Daniel Gaudet[2,6], Diane Brisson[2,6], Marie-France Hivert[3,7,8] & Luigi Bouchard*[,‡1,2]

[1]Department of Biochemistry, Université de Sherbrooke, Sherbrooke, QC, Canada
[2]ECOGENE-21 & Clinical Research Center & Lipid Clinic, Chicoutimi Hospital, Saguenay, QC, Canada
[3]Department of Medicine, Division of Endocrinology, Université de Sherbrooke, Sherbrooke, QC, Canada
[4]Department of Pediatrics, Chicoutimi Hospital, Saguenay, QC, Canada
[5]Department of Health Sciences, Université du Québec à Chicoutimi, Saguenay, QC, Canada
[6]Department of Medicine, Université de Montréal, Montréal, QC, Canada
[7]Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, MA, USA
[8]General Medicine Division, Massachusetts General Hospital, Boston, MA, USA
*Author for correspondence:
Tel.: +1 418 541 1234 ext. 2475
Fax: +1 418 541 1116
luigi.bouchard@usherbrooke.ca
[‡]Authors contributed equally

Aim: To assess the associations between gestational diabetes mellitus (GDM) and DNA methylation levels at genes related to energy metabolism. **Patients & methods:** Ten loci were selected from our recent epigenome-wide association study on GDM. DNA methylation levels were quantified by bisulfite pyrosequencing in 80 placenta and cord blood samples (20 exposed to GDM) from an independent birth cohort (Gen3G). **Results:** We did not replicate association between DNA methylation and GDM. However, in normoglycemic women, glucose levels were associated with DNA methylation changes at *LRP1B* and *BRD2* and at *CACNA1D* and *LRP1B* gene loci in placenta and cord blood, respectively. **Conclusion:** These results suggest that maternal glucose levels, within the normal range, are associated with DNA methylation changes at genes related to energy metabolism and previously associated with GDM. Maternal glycemia might thus be involved in fetal metabolic programming.

Keywords: cord blood • DNA methylation • epigenetics • gestational diabetes • placenta

Gestational diabetes mellitus (GDM) is the most common pregnancy complication, with a prevalence generally ranging from 2 to 6%, but reaching up to 20% in specific high-risk populations [1]. GDM is defined as impaired glucose tolerance first recognized during pregnancy and arises when pancreatic β-cells cannot produce sufficient insulin to compensate for pregnancy-induced insulin resistance. Consequently, blood glucose levels rise, resulting in hyperglycemia and GDM [2]. A number of epidemiological studies have demonstrated that *in utero* exposure to maternal hyperglycemia is associated with fetal overgrowth and newborn excessive birth weight, as well as with obesity and associated metabolic complications during childhood, adolescence or adulthood [3–7], supporting the Developmental Origin of Health and Disease (DOHaD) hypothesis. However, parental confounding factors exist [8] and should be carefully considered in DOHaD studies, whenever it is possible. The assessment of the molecular mechanisms underlying the metabolic health programming of newborn exposed to variations in maternal glycemia may help better understanding its roles in DOHaD. This is particularly important as the number of children exposed to increased maternal glycemia will likely increase in the near future as obesity, a major risk factor for GDM [9,10], is now affecting a rising number of women of childbearing age.

In the last few years, we and others have reported evidence supporting the role of epigenetics in fetal metabolic programming of the newborn exposed to maternal hyperglycemia [11–18]. DNA methylation is the most stable and studied epigenetic marks in the context of fetal programming. During embryonic development, DNA methylation marks are reprogrammed and have thus been shown to be particularly sensitive to *in utero* environment insults, including maternal metabolic challenges [19]. Interestingly, the DNA methylation status of several genes appears to be relatively stable over mitosis, suggesting that the epigenetic signature acquired in the womb might have

Future Medicine part of fsg

long-term adverse consequences on gene expression, thus on cellular and tissue phenotypes [20–22].

Using an epigenome-wide approach in the E-21 birth cohort [18], we have recently identified a number of gene loci that were potentially differentially methylated in placenta and/or cord blood samples exposed or not to GDM. Interestingly, 115 of these loci, common to both placenta and cord blood, harbored genes related to metabolic diseases pathway (identified by Ingenuity Pathway Analyses [IPA] software) including diabetes mellitus and glucose metabolism disorders. These results suggested that the epigenetic adaptations to GDM are not randomly distributed across the epigenome, but target more specifically genes involved in the regulation of metabolic diseases. More recently, Finer *et al.* reported similar results [23]. Our epigenome-wide association study (EWAS) was the first to support that epigenomic research might be helpful in identifying genes involved in fetal metabolic programming in response to GDM. However, these results need to be confirmed as none of the lead signals reached statistical significance after accounting for the number of probes tested (Infinium HumanMethylation450K beadchips >485,000 probes). The aim of the current study was thus to assess in Gen3G, an independent birth cohort, some of the associations we recently reported in our EWAS. We have selected the ten most promising genes out of the 115 involved in the metabolic diseases pathway.

Patients & methods

Gen3G birth cohort

Sample & clinical data

The Gen3G birth cohort comprises singleton pregnant women of the Sherbrooke City area (QC, Canada) who were recruited during their first trimester of pregnancy. Women were excluded from this study if they were diagnosed with pregestational Type 1 or 2 diabetes or other major disorders known to affect glucose metabolism, or presented a history of alcohol and/or drug abuse during the current pregnancy. Although women with pregnancy complicated with Type 1 and 2 diabetes are of potential interest in fetal epigenetic programming, they were excluded because they are hyperglycemic from the beginning of pregnancy and DNA methylation variability was previously shown to be dependent on gestational timing of the exposure [24]. In addition, women with Type 1 and 2 diabetes have their own clinical follow-up which could potentially bias our results. All participants gave a written informed consent before their inclusion in the Gen3G birth cohort, and all clinical data were de-identified. This project was approved by the Centre hospitalier universitaire de Sherbrooke (CHUS) ethics committee and was conducted in accordance with the Declaration of Helsinki.

Women were met during the first trimester (5–13 weeks, study entry) and the second trimester (24–29 weeks) of pregnancy at the CHUS. At study entry, demographic characteristics and personal/familial medical history were collected. Anthropometric variables were measured in the first and the second trimester using the standardized procedures [25]. In the first trimester, nonfasting blood samples were obtained to measure glycated hemoglobin (Hb_{A1c}). Venous blood samples were also drawn 1 h after a nonfasting 50 g glucose challenge test (50 g GCT) to assess plasma glucose levels. In the second trimester, each participant performed a 75 g oral glucose tolerance test (OGTT) under fasting state (≥8 h). GDM was diagnosed according to the International Association of Diabetes and Pregnancy Study Groups (IADPSG) guidelines (fasting glucose ≥5.1 mmol/l; or 1 h glucose ≥10.0 mmol/l; or 2 h glucose ≥8.5 mmol/l) [26]. Fasting insulin and glucose were used to calculate the homeostatic model assessment for insulin resistance (HOMA-IR = fasting glucose [mmol/l] × fasting insulin [mU/l]/22.5) [27]. At delivery, clinical information was collected from medical files and included maternal body weight at last prenatal visit, newborn gestational age, sex and anthropometric measures. Birth weight z-scores were assessed with the Fenton preterm growth charts [28].

Eighty women and their respective newborns were selected for this study. Twenty women with GDM (12 treated with diet only and 8 with diet and insulin) and for whom both placenta and cord blood samples were available were first selected (cases). Then, three normal glucose-tolerant women (NGT, controls) per case were matched for age (± 3 years) and BMI (± 0.7 kg/m^2) at first trimester of pregnancy. The set of data collected is complete for the 80 women, except for fasting insulin levels measured at the second trimester that were missing for one NGT woman.

Biochemical analysis

Plasma glucose concentrations were measured by the glucose hexokinase method (Roche Diagnostics). Fasting total cholesterol, high-density lipoprotein-cholesterol (HDL-C) and triglycerides were measured by colorimetric methods (Johnson & Johnson Clinical Diagnostics). Then, low-density lipoprotein-cholesterol (LDL-C) levels were estimated using the Friedewald formula [29]. Hb_{A1c} was measured by high-performance liquid chromatography (HPLC; VARIANT; Biorad). Insulin concentrations were measured using multiplexed particle-based flow cytometric assay; Milliplex map kits were purchased from Millipore.

Placenta & cord blood sampling

Placenta and cord blood were sampled according to the same standardized procedure as previously described for the E-21 birth cohort used in our EWAS [18]. Briefly, samples were collected by well-trained clinicians (MD) within minutes of delivery and placenta expulsion. Placenta biopsies were taken on the fetal side near the insertion point of the umbilical cord and comprised the intervillous tissues and chorionic villi. The biopsies were washed in PBS 1× to remove cord/maternal blood, dissected to remove conjunctive tissues and kept in RNAlater (Qiagen) at -80°C. Aliquots of cord blood samples were stored at -80°C in cryovials and PAXGene blood RNA Tubes (Qiagen) until nucleic acid extraction

Nucleic acid extraction

DNA and RNA were purified from placenta biopsies using the all-prep DNA/RNA/Protein Mini Kit (Qiagen). DNA and RNA were purified from cord blood samples with the Gentra Puregene Blood Kit (Qiagen) and PAXgene Blood RNA Kit (preAnalytiX), respectively. RNA quality was assessed with Agilent 2100 Bioanalyzer RNA Nano Chips (Agilent Technologies, Santa Clara, CA, USA). Sixty-five placenta RNA samples (GDM = 17; NGT = 48) and 58 cord blood RNA samples (GDM = 11; NGT = 47) showed good quality (RNA integrity number (RIN) ≥7.0) and were used for subsequent mRNA analyses. RIN ranged from 7.4 to 10.0 (mean: 8.6 ± 0.6) for placenta and from 7.1 to 9.9 (mean: 9.0 ± 0.6) for cord blood.

DNA methylation analysis

As previously mentioned, the EWAS we have recently conducted in the E-21 birth cohort identified 115 genes that showed DNA methylation differences (p-values ≤ 0.05) in both placenta and cord blood samples in response to GDM exposure and were related to metabolic disease pathways (Supplementary Table 1) [18]. Among these 115 genes, we have selected the ten most promising genes. For the genes for which we identified more than one CpG, we selected the CpG with the lowest p-value (Supplementary Table 2). Although from the same ten genes, different CpG sites were selected in the placenta and cord blood (Supplementary Table 1). Consequently, DNA methylation levels were assessed at ten distinct CpG sites in placenta and cord blood samples in the Gen3G birth cohort.

Pyrosequencing technology was used to quantify DNA methylation levels at the targeted CpGs. pyrosequencing assays combine sodium bisulfite DNA conversion chemistry (EpiTech Bisulfite Kits; Qiagen), PCR amplification (Pyromark PCR Kit; Qiagen) and sequencing by synthesis assay (Pyromark Gold Q24 Reagents; Qiagen). Sodium bisulfite preferentially deaminates unmethylated cytosine residues to thymines (after PCR amplification), whereas methylcytosines remain unmodified. Each of the pyrosequencing runs performed included a negative pCR control and a sequencing primer control as well as sodium bisulfite conversion controls. Additionally, a pyrosequencing quality control (peak height, deviation from reference sequence pattern and unexpected peak height) was assessed for each sample using the Pyromark Q24 Analysis Software (v1.0.10.134). pCR primers were designed to cover the 20 CpG sites selected for replication. PyroMark Assay Design software v2.0.1.15 was used to design the primers. Five CpGs were excluded because the design of specific pyrosequencing assays was restricted by: highly polymorphic regions (*HLA-DQB1* in both placenta and cord blood), SNPs (C/T and G/A) at the targeted CpG sites (cg26729380 (*TNF*) and cg19285525 (*RBMS1*) in placenta) and high complexity regions (stretch of thymines after bisulfite conversion, cg10673915 (*RBMS1*) in the cord blood). The PCR and pyrosequencing primers used for DNA methylation analysis of the other 15 CpG sites are described in Supplementary Table 3.

Gene expression analysis

mRNA levels were quantified in the Gen3G birth cohort for genes for which the results were replicated (*BRD2, LRP1B* and *CACNA1D*). Complementary DNA (cDNA) was generated from total RNA using a random primer hexamer provided with the High Capacity cDNA Archive kit (Life Technologies, #4368814). Equal amounts of cDNA were run in duplicate and amplified in a 20 µl reaction containing 10 µl of 2× Universal PCR Master Mix (Life Technologies, #4366072). Primers and Taqman probes were designed to cover exon boundaries (Life Technologies, *BRD2*: Hs01121984_m1, *LRP1B*: Hs01069153_m1, *CACNA1D*: Hs01073321_m1). Each placenta sample was calibrated to the *YWHAZ* housekeeping gene (endogenous control; *YWHAZ*: Hs00237047_m1), and each cord blood sample was calibrated to the *GAPDH* housekeeping gene (endogenous control; *GAPDH*: Hs99999905_m1). Relative quantification estimations were performed using an Applied Biosystems 7500 Real Time PCR System as recommended by the manufacturer. *BRD2*/*YWHAZ* and *LRP1B*/*YWHAZ* C_t ratio (1/x) values were used for expression analyses in placenta and *BRD2*/*GAPDH*, *LRP1B*/*GAPDH* and *CACNA1D*/*GAPDH* C_t ratio (1/×) values were used in cord blood.

USCS genome browser and a PubMed literature search were used to identify consensus sequences of transcription factor binding sites at gene loci for which

DNA methylation levels were found to be significantly correlated with mRNA levels [30].

Statistical analyses

Statistical analyses were performed using nonparametric tests as numerous variables were not normally distributed (assessed using the Kolmogorov–Smirnov test). Residual-based linear regression analysis was used to adjust DNA methylation for the following confounding variables: maternal BMI at first trimester, gestational weight gain [31], previous history of GDM and newborn's sex and gestational age. Spearman rank correlation coefficients were used to investigate the association between DNA methylation residual score and variables of interest (fasting glucose, HOMA-IR, glycemia 2 h post-OGTT at second trimester). The association between DNA methylation and mRNA expression levels a *BRD2* and *CACNA1D* gene loci were also assessed with Spearman rank correlation coefficients. DNA methylation levels residual for the 15 CpG sites, maternal and newborn characteristics were compared between women with and without GDM using Mann–Whitney *U* test or the Pearson χ^2 test for categorical variable. Results were considered statistically significant when p-values were <0.05 (two-sided). All statistical analyses were performed with the IBM SPSS Statistics 20 software (release 20.0.0, SPSS, IL, USA).

Results

The characteristics of the 80 mother and newborn pairs according to maternal glucose status (NGT vs GDM) are shown in Table 1. At first trimester of pregnancy, NGT and GDM women presented similar characteristics: median age was 30.3 years old and median BMI was in the overweight range (>25 kg/m^2). As expected, GDM women had significantly higher levels of fasting glucose as well as higher glucose levels 2 h post-OGTT at the second trimester of pregnancy.

Association between GDM & DNA methylation levels

To confirm our previous findings, we first compared DNA methylation levels between placenta and cord blood samples exposed or not exposed to GDM. The association reported in placenta at DLGAP2 cg05951364 (GDM: 94.0 ± 10.2% vs NGT: 86.6 ± 16.5%; p = 0.009) was the only one statistically significant (Table 2).

Correlation between DNA methylation levels & maternal glucose metabolism at second trimester

Given that maternal hyperglycemia, below the diagnostic threshold for GDM, was found to be linearly correlated with newborn health outcomes (birth weight, adiposity and fetal insulin production [7,32]) and that GDM diagnostic criteria were different between Gen3G (IADPSG [26]) and E-21 (WHO guidelines [33]) birth cohorts, which might have biased the results along with the treatment for GDM, we decided to conduct secondary analyses to assess whether DNA methylation levels at these gene loci were correlated with maternal glucose levels 2 h post-OGTT in NGT women only.

We report significant correlations between maternal glycemia 2 h post-OGTT and DNA methylation levels for 4 out of the 15 CpGs tested (27%) (Figures 1 & 2). Higher maternal glucose levels were associated with lower DNA methylation levels at *BRD2* (cg08491668: r = -0.293; p = 0.02) and *LRP1B* (cg19355806; r = -0.290; p = 0.03) gene loci in the placenta (Figure 1A & B) and at *CACNA1D* (cg13413750: r = -0.276; p = 0.04) and *LRP1B* (cg26413307: r = -0.334; p = 0.01) gene loci in cord blood (Figure 2B & C). In cord blood, a trend for significant association between maternal glycemia 2 h post-OGTT and DNA methylation levels was also observed at *BRD2* gene locus (cg11439393: r = -0.247; p = 0.06) (Figure 2A).

In support of these findings, we also observed that maternal glycemia 1 h post-OGTT and Hb_{A1c} at first trimester of pregnancy are associated with placental DNA methylation levels at *BRD2* gene locus (cg08491668: r = -0.472; p = 0.0002 and r = -0.352; p = 0.006; Supplementary Table 4). In cord blood, additional correlations (trends included) were observed between maternal glycemia 1 h post-OGTT and DNA methylation at *BRD2* (cg11439393: r = -0.235; p = 0.07), *LRP1B* (cg26413307: r = -0.351; p = 0.007) and *CACNA1D* (cg13413750: r = -0.337; p = 0.01) gene loci; Supplementary Table 4). Importantly, 1 h and 2 h post OGTT glucose levels are only partially correlated (r = 0.640, means 41% of covariance; p <0.0001) with none of the two measures being correlated with Hb_{A1c} (1 h post-OGTT: r = 0.161, p = 0.220; 2 h post-OGTT: r = 0.114; p = 0.384) thus providing independent supportive evidence for association between maternal glycemic control in pregnancy and *BRD2* (placenta), *CACNA1D* (cord blood) and *LRPIB* (cord blood) DNA methylation levels.

Associations between DNA methylation & mRNA levels

Finally, we assessed the potential functional impact of DNA methylation marks identified in the current study. First, *LRP1B* transcripts were undetectable in both the placenta and cord blood. In cord blood, a trend of a positive correlation was observed between *BRD2*

Table 1. Comparison of the characteristics of mothers and newborns according to glucose tolerance status.

Characteristics	NGT (n = 60)		GDM (n = 20)		p-value
	Median	IQR	Median	IQR	
Personal prior history of GDM	n = 11 (18.3%)		n = 7 (35.0%)		0.12
First trimester of pregnancy (5–13 weeks)					
Age (years)	30.5	8.0	30.5	6.0	0.91
BMI (kg/m^2)	25.7	11.2	26.4	12.0	0.42
HbA1c (%)	5.30	0.30	5.50	0.50	0.07
Second trimester of pregnancy (24–29 weeks)					
Fasting glucose (mmol/l)	4.20	0.50	4.60	0.90	0.004
HOMA-IR	1.61	0.97	1.90	2.39	0.29
2 h post-OGTT glucose (mmol/l)	6.00	2.30	8.50	1.50	<0.001
Third trimester of pregnancy/childbirth					
Gestational weight gain (kg)[†]	12.3	7.8	11.4	8.5	0.38
Newborn					
Sex:					
– Male	n = 32 (53.3%)		n = 6 (30.0%)		0.07
– Female	n = 28 (46.7%)		n = 14 (70.0%)		
Birth weight z-score	0.00	1.14	0.22	1.17	0.69

[†]Maternal weight gain between first trimester of pregnancy and within 11 weeks before childbirth.
GDM: Gestational diabetes mellitus; HBA1C: Glycosylated hemoglobin; HOMA-IR: Homeostasis Model Assessment-Estimated Insulin Resistance; IQR: Interquartile range; NGT: Normal glucose tolerance; OGTT: Oral glucose tolerance test.

DNA methylation and mRNA levels (n = 58; r = 0.235; p = 0.08). This correlation might be specific to newborn girls (n = 29; r = 0.477; p = 0.009) since no correlation was observed in boys (n = 29; r = -0.025; p = 0.90).

BRD2 cg11439393 probe set is located within intron 2 and 1108 bp downstream a CpG island. Four other CpGs that were associated with maternal glycemia in cord blood in our previous EWAS [18] (Figure 3 & Supplementary Table 2) are closed to cg11439393. DNA methylation levels at these five CpGs are highly correlated (r ≥ 0.725; p <0.001). Interestingly, these CpGs are found in a region that is highly conserved across placental mammals as well as near potential binding sites for transcriptional factors such as *CEBP, OLF1* and *IRF7* (Figure 3).

Discussion

The current study is a first attempt to confirm in an independent cohort the associations between GDM and DNA methylation levels reported in a previous EWAS. Replication of (epi)genome-wide association study results faces a number of challenges such as differences in study design, population assessed and trait definition between the cohorts/studies [34]. In addition to these factors, the treatment prescribed after the diagnosis of GDM (diet alone or diet and insulin) is another challenge we encountered in the current study. Nevertheless, we report that variations in maternal glucose levels, within the normal range, and Hb_{A1c} are associated with DNA methylation changes in placenta and cord blood at three gene loci (*BRD2*, *LRP1B* and *CACNA1D*) we identified in a previous EWAS [18]. Among the 15 loci analyzed, four were negatively correlated with maternal glucose levels 2 h post-OGTT, which is more than what would be expected by chance alone. Moreover, these results are biologically concordant with our previous EWAS reporting lower DNA methylation levels at these loci in placenta and cord blood samples exposed to GDM in comparison to those exposed to normoglycemia [18]. More importantly, our findings suggest that the newborn's epigenome is sensitive to maternal glycemia even when glucose levels are

below the diagnostic threshold for GDM. Likewise, it has been shown that a modest increase in maternal glucose levels within the normal or GDM range of glycemia is associated with higher birth weight and adiposity, as well as with an increased incidence of obesity and cardiometabolic complications in children, adolescents and adults [7,32,35–37]. The results of the current study therefore support that a small increase in maternal glycemia remaining within the normal range might have deleterious effects on the offspring's short- and long-term metabolic health via epigenetic modifications. The absence of correlation between glucose and DNA methylation levels in GDM women might be explained by the possible large variability in glycemic control following GDM treatment (diet alone or diet and insulin). We might also speculate that an effective control of maternal glycemia, from GDM diagnosis to delivery, alleviates, at least partially the impact of GDM on the newborn's epigenome.

The three genes we have identified are important candidates for obesity and cardiometabolic diseases. *BRD2* encodes the bromodomain-containing protein 2 (BRD2), which regulates the transcription and splicing of more than 1500 genes [38]. In mice and *in vitro* models, disruption of *Brd2* promotes adipogenic differentiation and adipose tissue expansion [39]. DNA methylation was recently shown to be involved in *BRD2* gene expression regulation and differentiation of 3T3-L1 cells into adipocytes [40]. In mice, the disruption of *Brd2* has been associated with a metabolically healthy obesity phenotype: obesity with hyperadiponectinemia, improved glucose tolerance and reduced inflammatory response [41]. The pancreatic expression of *Cacna1d*, a voltage-dependent calcium channel, was found to be decreased in intrauterine growth restricted sheep and correlated with glucose-stimulated disposition in the adult sheep [42]. Genetic variants in *CACNA1D* were also associated with hypertension and primary aldosteronism [43,44]. Finally, *LRP1B* encodes the low-density lipoprotein receptor-related protein 1B that mediates cellular cholesterol uptake [45]. *LRP1B* was recently identified as a determinant of cholesterol concentrations in low-density lipoproteins in rats [46]. Furthermore, polymorphisms in the *LRP1B* gene were found to be associated with insulin resistance, uncontrolled emotional eating and child BMI [47–49]. The

Table 2. Mann–Whitney *U* comparisons of DNA methylation levels (%) in placenta and cord blood samples not exposed to gestational diabetes mellitus (normal glucose tolerance, n = 60) or exposed to gestational diabetes mellitus (n = 20) in the Gen3G birth cohort.

Genes		NGT		GDM		p-value†
	Probe ID	Median	IQR	Median	IQR	
Placenta						
BRD2	cg08491668	57.2	11.1	58.3	10.9	0.23
CACNA1D	cg10409919	36.0	14.4	38.5	14.6	0.67
DLGAP2	cg05951364	86.6	16.5	94.0	10.2	0.009
KCNQ1	cg20828897	64.6	10.0	64.2	9.2	0.44
LRP1B	cg19355808	53.9	18.2	56.4	16.8	0.87
PDE4D	cg03154690	23.4	12.7	24.3	11.4	0.71
TNF	cg26729380	–	–	–	–	–
TNFRSF1B	cg09081517	55.4	9.0	56.1	7.7	0.55
Cord blood						
BRD2	cg11439393	15.1	3.7	15.0	5.2	0.52
CACNA1D	cg13414750	72.8	6.7	72.8	4.9	0.98
DLGAP2	cg11192059	92.0	3.7	90.2	4.5	0.19
KCNQ1	cg27115973	54.8	5.4	54.0	5.7	0.91
LRP1B	cg26413307	62.7	8.3	62.9	9.4	0.16
PDE4D	cg06895913	7.7	2.0	8.3	3.6	0.23
TNF	cg20477259	48.2	7.7	53.3	6.8	0.08
TNFRSF1B	cg22677556	49.3	4.8	48.8	6.5	0.39

†p-value adjusted for history of gestational diabetes mellitus, maternal BMI at first trimester, gestational weight gain, newborn's sex and gestational age at birth.
GDM: Gestational diabetes mellitus; IQR: Interquartile range; NGT: Normal glucose tolerance.

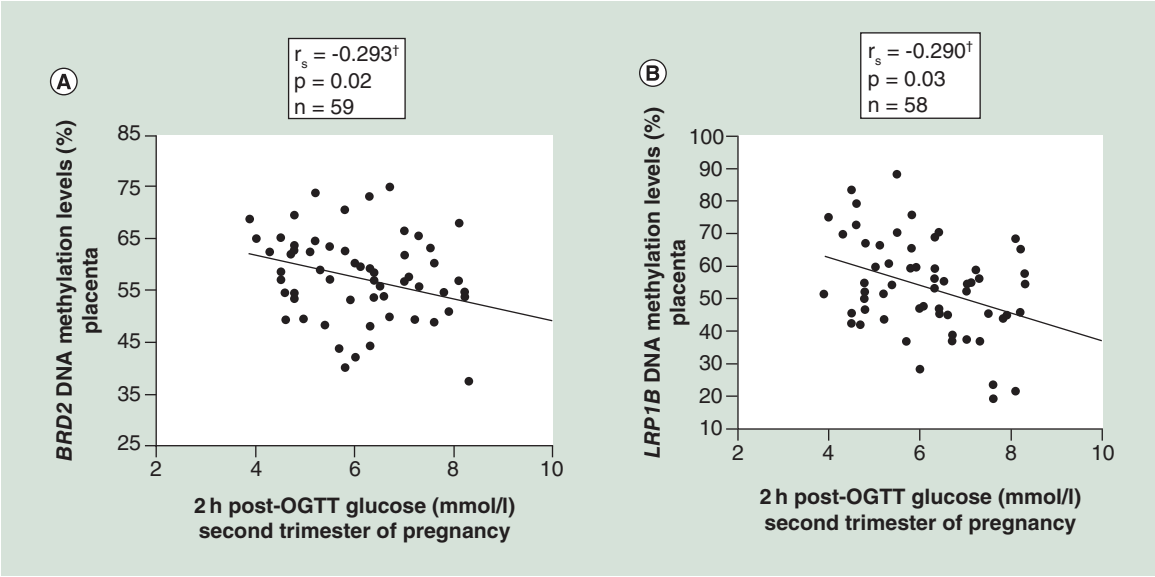


Figure 1. Association plots between maternal glucose levels 2 h post-oral glucose tolerance test and DNA methylation levels in placenta of normal glucose tolerance women. Higher maternal glucose levels 2 h post-OGTT were associated with lower DNA methylation levels at **(A)** *BRD2* (cg08491668) and **(B)** *LRP1B* (cg19355806) gene loci in placenta.
†Spearman's correlation coefficients adjusted for history of gestational diabetes, maternal BMI at first trimester, gestational weight gain, newborn sex and gestational age.
OGTT: Oral glucose tolerance test.

biological functions of these three genes are thus highly supportive of their implication in metabolic health programming of the newborn exposed to variations in maternal glycemia.

We have also reported a correlation between *BRD2* DNA methylation and mRNA levels in cord blood, which could be specific to girls only keeping in mind that this is a subgroup analysis. Sex-specific associations between DNA methylation, mRNA levels and the phenotypic variables have been reported in several studies, suggesting that DNA methylation contributes to phenotypic sexual dimorphism [50,51]. Although DNA methylation levels at other gene loci were not correlated with mRNA levels in the placenta or cord blood, these epigenetic modifications might have a functional impact in other tissues such as the adipose tissue, the skeletal muscle or the brain, where they could trigger newborns' long-term susceptibility to obesity or cardiometabolic complications. In agreement with this, we have previously reported that placental DNA methylation changes at *LPL* and *ABCA1* gene loci have tissue-specific impacts on gene expression regulation [13–14,52–53]. Further studies are needed to determine whether DNA methylation variations observed at *BRD2, CACNA1D* and *LRP1B* gene loci in the placenta and cord blood mirror those of other tissues and correlate with their mRNA levels.

Our study has strengths and limitations. First, the longitudinal follow-up of the women from first trimester of pregnancy to delivery allowed us to accurately identify women who developed GDM during pregnancy and therefore avoid their misclassification with women who might have been diabetic prepregnancy. Second, our study design permitted the inclusion of confounding variables in our statistical models. The sample size is fairly large compared with other studies but correlations/differences of smaller effect size might have been missed. One limitation is that we were not able to quantify DNA methylation levels using pyrosequencing at CpG sites within *HLA-DQB1* and *RBMS1* gene loci in the placenta and cord blood, as well as within *TNF* in the placenta. These genes have been associated with insulin resistance and the onset of Type 2 diabetes and are thus of great interest in the context of fetal metabolic programming. Other technical approaches such as Sequenom Epityper system could potentially be used to assess DNA methylation levels at these five CpG sites. In addition, we cannot exclude that neither cellular heterogeneity nor other potential confounding factors such as paternal BMI might have biased our findings. Since our sample size was relatively small and the associated statistical power limited, the association between *BRD2* DNA methylation levels in placenta and glucose levels 1 h post-OGTT is the only one remaining statistically significant ($p = 0.0002$) after multiple testing correction ($p < 0.003$; 0.05/16 and 17 tests for placenta and cord blood, respectively). Accordingly, we cannot exclude that some of the other associations reported are false-

positive. Further functional and longitudinal studies in larger cohorts are thus required to confirm our findings and assess whether these epigenetic marks are stable over time and associated with the development of obesity and obesity-related cardiometabolic complications in the long term.

Conclusion

In conclusion, DNA methylation levels at four gene loci we have identified as among our strongest candidates based on a previous EWAS of GDM cases–controls are associated with variations in maternal glucose levels. Although our findings need to be confirmed in a larger cohort, they support that epigenetic variations at *BRD2, CACNA1D* and *LRP1B,* genes associated with the development of obesity or cardiometabolic complications, are involved in fetal metabolic programming.

Future perspective

The current study, along with few others, is supportive of the role of epigenetic modifications in fetal metabolic health programming of the newborn exposed to suboptimal *in utero* environment including maternal

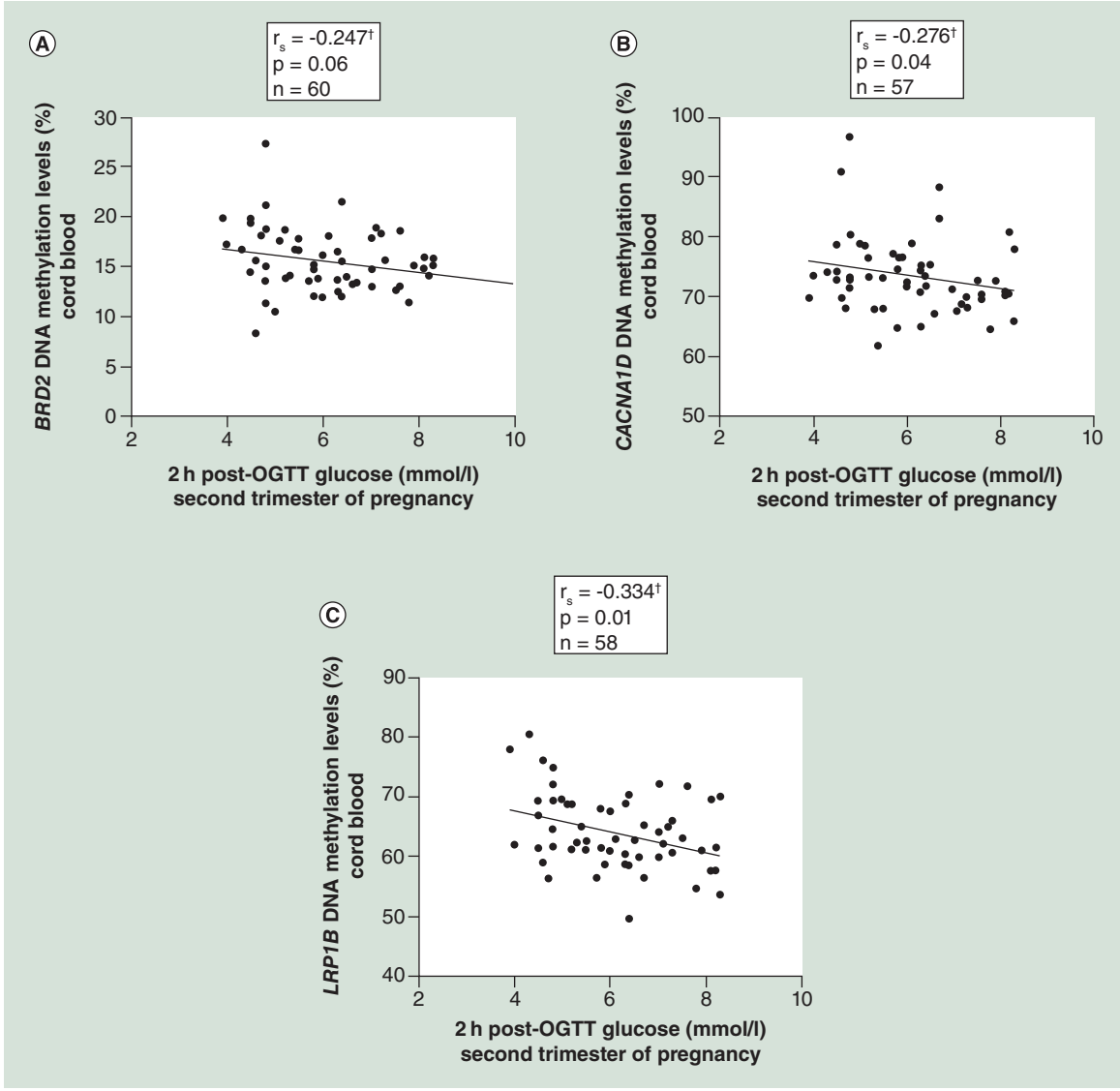


Figure 2. Association plots between maternal glucose levels 2 h post-oral glucose tolerance test and DNA methylation levels in cord blood of normal glucose tolerance women. Higher maternal glucose levels 2 h post-OGTT were associated with lower DNA methylation levels at **(A)** *BRD2* (cg11439393), **(B)** *CACNA1D* (cg10409919) and **(C)** *LRP1B* (cg26413307) gene loci in cord blood.
†Spearman's correlation coefficients adjusted for history of gestational diabetes, maternal BMI at first trimester, gestational weight gain, newborn sex and gestational age.
OGTT: Oral glucose tolerance test.

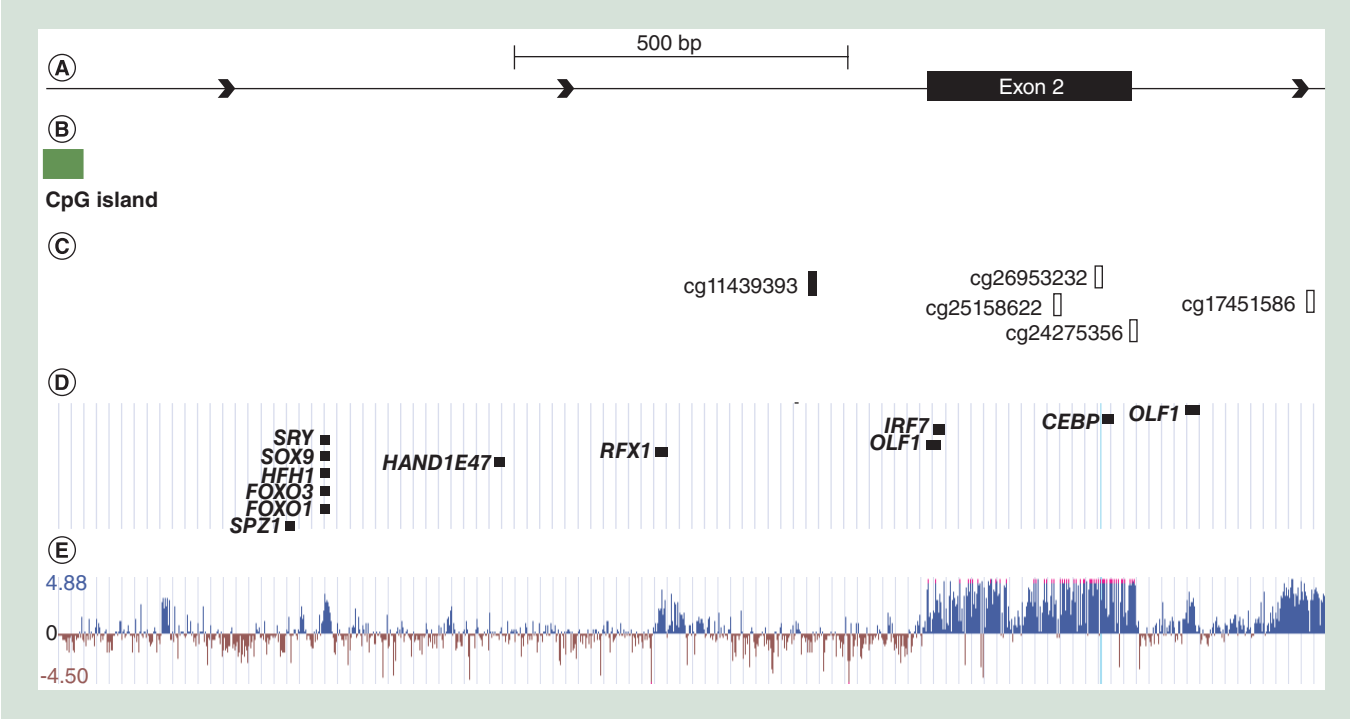


Figure 3. Location of CpGs, transcription factor binding sites and highly conserved regions within *BRD2* gene locus, adapted from UCSC genome browser tracks. Track **(A)** *BRD2* intron 2 to intron 3 region. Track **(B)** Localization of the second CpG island within *BRD2* gene locus. Track **(C)** identification of CpG sites that were found to be associated with maternal hyperglycemia (black: CpG epigenotyped in Gen3G and E-21 birth cohorts; white: CpG epigenotyped in E-21 birth cohort only). Track **(D)** localization of transcription factor binding sites. Track **(E)** conservation track for mammals.
UCSC genome browser was accessed on 27 August 2014.

hyperglycemia. Nevertheless, whether the epigenetic modifications observed at birth are stable over time and associated with the future development of obesity and cardiometabolic complications is unknown. Longitudinal studies are thus needed to assess the stability of these epigenetic marks and their association with metabolic health indicators (such as body fat percentage, blood glucose and lipid profile) in child, adolescent and adult populations. If such associations were observed, epigenetic marks might, in the long run, be used as clinical markers of obesity susceptibility at birth.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/epi.15.72

Acknowledgements

The authors acknowledge the Blood Sampling Clinic in pregnancy at Centre Hospitalier Universitaire de Sherbrooke (CHUS), which supports research activities integrated to the Blood Sampling in pregnancy Clinic; the CHUS laboratory for performing blood glucose and lipid profile analyses; and the contribution of the clinical research nurses (M Gérard, CHUS; G Proulx, CHUS; S Hayes, CHUS; and M-J Gosselin, CHUS) and research assistants (C Rousseau, CHUS; P Brassard, CHUS; J Ménard, CHUS; MC Battista, CHUS and S Claveau, ECOGENE-21 laboratory) and graduate students (L Guillemette, M Lacroix, J Patenaude) for their dedicated work in this study. The authors also express their gratitude to C Bélanger, Chicoutimi Hospital, for her thoughtful language revision of the manuscript.

Financial & competing interests disclosure

This project was supported by ECOGENE-21 Clinical Research Center, the Canadian Institutes of Health Research (CIHR), the Fonds de Recherche du Québec en Santé (FRQS) and Diabète Québec. L Bouchard is a Junior Research Scholar from the Fonds de la Recherche en Santé du Québec (FRQS). M-F Hivert was an FRSQ Research Scholar and was awarded Clinical Scientist award by the Canadian Diabetes Association (CDA); she is now supported by American Heart Association (AHA) and American Diabetes Association (ADA) awards. L Bouchard and M-F Hivert are members of the FRQS-funded Centre de recherche Clinique Étienne-Le Bel (affiliated with Centre Hospitalier Universitaire de Sherbrooke [CHUS]). A-A Houde is the recipient of doctoral training award from the FRQS and the Faculté de Médecine et des Sciences de la Santé de l'Université de Sherbrooke. S-M Ruchat has received a postdoctoral fellowship from

the CDA. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

Background

- *In utero* exposure to maternal hyperglycemia is associated with an increased long-term susceptibility to obesity and metabolic diseases.
- Epidemiologic evidence and animal models suggest that DNA methylation modifications are involved in this fetal metabolic programming.
- DNA methylation marks are sensitive to environmental stimuli during embryonic development.
- DNA methylation status of several genes appears to be stable and might have long-term adverse consequences on gene expression and cellular phenotype.

Associations between maternal glycemia & DNA methylation levels in placenta & cord blood

- In a recent epigenome-wide association study, we have identified 115 genes involved in the metabolic disease pathway differentially methylated in placenta and/or cord blood samples exposed and not exposed to gestational diabetes mellitus.
- In the current study, we confirmed, in an independent birth cohort (Gen3G), that maternal glycemia is negatively correlated with DNA methylation levels at *LRP1B* and *BRD2* gene loci in placenta and at *LRP1B, BRD2* and *CACNA1D* gene loci in cord blood.

Associations between DNA methylation & mRNA levels

- In cord blood of newborn girls, *BRD2* DNA methylation levels are positively correlated with *BRD2* mRNA levels.

Conclusion

- Our results suggest that *BRD2, CACNA1D* and *LRP1B* are epigenetically programmed by maternal glycemia.
- DNA methylation perturbations at these three candidate genes, known to be involved in metabolic diseases, might influence the long-term metabolic health of the newborn.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Galtier F. Definition, epidemiology, risk factors. *Diabetes Metab.* 36(6 Pt 2), 628–651 (2010).

2 Catalano PM. Trying to understand gestational diabetes. *Diabet. Med.* 31(3), 273–281 (2014).

3 Clausen TD, Mathiesen ER, Hansen T *et al.* High prevalence of Type 2 diabetes and pre-diabetes in adult offspring of women with gestational diabetes mellitus or Type 1 diabetes: the role of intrauterine hyperglycemia. *Diabetes Care* 31(2), 340–346 (2008).

4 Clausen TD, Mathiesen ER, Hansen T *et al.* Overweight and the metabolic syndrome in adult offspring of women with diet-treated gestational diabetes mellitus or Type 1 diabetes. *J. Clin. Endocrinol. Metab.* 94(7), 2464–2470 (2009).

5 Hillier TA, Pedula KL, Schmidt MM, Mullen JA, Charles MA, Pettitt DJ. Childhood obesity and metabolic imprinting: the ongoing effects of maternal hyperglycemia. *Diabetes Care* 30(9), 2287–2292 (2007).

6 Hill JC, Krishnaveni GV, Annamma I, Leary SD, Fall CH. Glucose tolerance in pregnancy in south india: relationships to neonatal anthropometry. *Acta Obstet. Gynecol. Scand.* 84(2), 159–165 (2005).

7 Catalano PM, Mcintyre HD, Cruickshank JK *et al.* The hyperglycemia and adverse pregnancy outcome study: associations of GDM and obesity with pregnancy outcomes. *Diabetes Care* 35(4), 780–786 (2012).

• **This study reports continuous associations between maternal glucose levels, and adverse newborn health outcomes, independently of maternal obesity.**

8 Donovan LE, Cundy T. Does exposure to hyperglycaemia *in utero* increase the risk of obesity and diabetes in the offspring? A critical reappraisal. *Diabet. Med.* 32(3), 295–304 (2015).

9 Kim SY, England L, Wilson HG, Bish C, Satten GA, Dietz P. Percentage of gestational diabetes mellitus attributable to overweight and obesity. *Am. J. Public Health* 100(6), 1047–1052 (2010).

10 Kim SY, Saraiva C, Curtis M, Wilson HG, Troyan J, Sharma AJ. Fraction of gestational diabetes mellitus attributable to overweight and obesity by race/ethnicity, California, 2007–2009. *Am. J. Public Health* 103(10), e65–e72 (2013).

11 Bouchard L, Hivert MF, Guay SP, St-Pierre J, Perron P, Brisson D. Placental adiponectin gene DNA methylation

levels are associated with mothers' blood glucose concentration. *Diabetes* 61(5), 1272–1280 (2012).

12 Bouchard L, Thibault S, Guay SP *et al.* Leptin gene epigenetic adaptation to impaired glucose metabolism during pregnancy. *Diabetes Care* 33(11), 2436–2441 (2010).

13 Houde AA, Guay SP, Desgagne V *et al.* Adaptations of placental and cord blood *ABCA1* DNA methylation profile to maternal metabolic status. *Epigenetics* 8(12), 1289–1302 (2013).

14 Houde AA, St-Pierre J, Hivert MF *et al.* Placental lipoprotein lipase DNA methylation levels are associated with gestational diabetes mellitus and maternal and cord blood lipid profiles. *J. Dev. Orig. Health Dis.* 5(2), 132–141 (2014).

15 El Hajj N, Pliushch G, Schneider E *et al.* Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes* 62(4), 1320–1328 (2013).

16 Desgagne V, Hivert MF, St-Pierre J *et al.* Epigenetic dysregulation of the IGF system in placenta of newborns exposed to maternal impaired glucose tolerance. *Epigenomics* 6(2), 193–207 (2014).

17 Petropoulos S, Guillemin C, Ergaz Z *et al.* Gestational diabetes alters offspring DNA methylation profiles in human and rat: Identification of key pathways involved in endocrine system disorders, insulin signaling, diabetes signaling and IL-K signaling. *Endocrinology* 156(6), 2222–2238 (2015).

18 Ruchat SM, Houde AA, Voisin G *et al.* Gestational diabetes mellitus epigenetically affects genes predominantly involved in metabolic diseases. *Epigenetics* 8(9), 935–943 (2013).

•• **This study is the first epigenome-wide association study assessing the impact of gestational diabetes mellitus (GDM) on DNA methylation in placenta and cord blood. The results reported suggest that GDM epigenetically affects genes mainly involved in metabolic disesase pathway. The ten genes analyzed in the current study were selected from this epigenome-wide association study.**

19 Hanson M, Godfrey KM, Lillycrop KA, Burdge GC, Gluckman PD. Developmental plasticity and developmental origins of non-communicable disease: theoretical considerations and epigenetic mechanisms. *Prog. Biophys. Mol. Biol.* 106(1), 272–280 (2011).

20 Feinberg AP, Irizarry RA, Fradin D *et al.* Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci. Transl. Med.* 2(49), 49ra67 (2010).

21 Flanagan JM, Brook MN, Orr N *et al.* Temporal stability and determinants of white blood cell DNA methylation in the breakthrough generations study. *Cancer Epidemiol. Biomarkers Prev.* 24(1), 221–229 (2014).

22 Clarke-Harris R, Wilkin TJ, Hosking J *et al.* PGC1alpha promoter methylation in blood at 5–7 years predicts adiposity from 9 to 14 years (earlybird 50). *Diabetes* 63(7), 2528–2537 (2014).

23 Finer S, Mathews C, Lowe R *et al.* Maternal gestational diabetes is associated with genome-wide DNA methylation variation in placenta and cord blood of exposed offspring. *Hum. Mol. Genet.* 24(11), 3021–3029. (2015).

24 Tobi EW, Lumey LH, Talens RP *et al.* DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.* 18(21), 4046–4053 (2009).

25 Lohman T, Roache A, Martorell R. Anthropometric standardization reference manual. *Medicine Sci. Sports Exercise* 24(8), 952 (1992).

26 Metzger BE, Gabbe SG, Persson B *et al.* International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 33(3), 676–682 (2010).

27 Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: Insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28(7), 412–419 (1985).

28 Fenton TR, Kim JH. A systematic review and meta-analysis to revise the fenton growth chart for preterm infants. *BMC Pediatr.* 13, 59 (2013).

29 Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin. Chem.* 18(6), 499–502 (1972).

30 Kent WJ, Sugnet CW, Furey TS *et al.* The human genome browser at UCSC. *Genome Res.* 12(6), 996–1006 (2002).

31 Institute of Medicine (US) and National Research Council (US) Committee to Reexamine IOM Pregnancy Weight Guidelines. *Weight Gain During Pregnancy: Reexamining The Guidelines*. Rasmussen KM, Yaktine AL (Eds). National Academy of Sciences, Washington, DC, USA (2009).

32 Hyperglycemia and adverse pregnancy outcome (HAPO) study: associations with neonatal anthropometrics. *Diabetes* 58(2), 453–459 (2009).

• **This study highlights that maternal hyperglycemia less severe than gestational diabetes mellitus is associated with fetal overgrowth and more specifically with newborn adiposity.**

33 WHO World Health Organization. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. *Part 1: Diagnosis and Classification of Diabetes Mellitus. WHO/NCD/NCS/99 (2nd Edition)*. World Health Organization, Geneva, Switzerland (1999). www.staff.ncl.ac.uk/philip.home/who_dmg.pdf

34 Bush WS, Moore JH. Chapter 11: genome-wide association studies. *PLoS Comput. Biol.* 8(12), e1002822 (2012).

35 Metzger BE, Lowe LP, Dyer AR *et al.* Hyperglycemia and adverse pregnancy outcomes. *N. Engl. J. Med.* 358(19), 1991–2002 (2008).

• **This study reports continuous associations between maternal glucose levels, below those used for GDM diagnosis, and newborn health outcomes (birth weight, adiposity, C-peptide levels).**

36 Deierlein AL, Siega-Riz AM, Chantala K, Herring AH. The association between maternal glucose concentration and child BMI at age 3 years. *Diabetes Care* 34(2), 480–484 (2011).

37 Walsh JM, Mahony R, Byrne J, Foley M, Mcauliffe FM. The association of maternal and fetal glucose homeostasis with fetal adiposity and birthweight. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 159(2), 338–341 (2011).

38 Hnilicova J, Hozeifi S, Stejskalova E *et al.* The c-terminal domain of brd2 is important for chromatin interaction and regulation of transcription and alternative splicing. *Mol. Biol. Cell.* 24(22), 3557–3568 (2013).

39 Wang F, Liu H, Blanton WP, Belkina A, Lebrasseur NK, Denis GV. Brd2 disruption in mice causes severe obesity without Type 2 diabetes. *Biochem. J.* 425(1), 71–83 (2010).

40 Sun R, Wu Y, Wang Y *et al.* DNA methylation regulates bromodomain-containing protein 2 expression during adipocyte differentiation. *Mol. Cell. Biochem.* 402(1–2), 23–31 (2015).

41 Wang F, Deeney JT, Denis GV. Brd2 gene disruption causes "metabolically healthy" obesity: epigenetic and chromatin-based mechanisms that uncouple obesity from Type 2 diabetes. *Vitam. Horm.* 91, 49–75 (2013).

42 Gatford KL, Mohammad SN, Harland ML *et al.* Impaired beta-cell function and inadequate compensatory increases in beta-cell mass after intrauterine growth restriction in sheep. *Endocrinology* 149(10), 5118–5127 (2008).

43 Fischer E, Beuschlein F. Novel genes in primary aldosteronism. *Curr. Opin. Endocrinol. Diabetes Obes.* 21(3), 154–158 (2014).

44 Lu X, Wang L, Lin X *et al.* Genome-wide association study in chinese identifies novel loci for blood pressure and hypertension. *Hum. Mol. Genet.* 24(3), 865–874 (2015).

45 Li Y, Knisely JM, Lu W *et al.* Low density lipoprotein (LDL) receptor-related protein 1b impairs urokinase receptor regeneration on the cell surface and inhibits cell migration. *J. Biol. Chem.* 277(44), 42366–42371 (2002).

46 Hodulova M, Sedova L, Krenova D *et al.* Genomic determinants of triglyceride and cholesterol distribution into lipoprotein fractions in the rat. *PLoS ONE* 9(10), e109983 (2014).

47 Burgdorf KS, Gjesing AP, Grarup N *et al.* Association studies of novel obesity-related gene variants with quantitative metabolic phenotypes in a population-based sample of 6,039 danish individuals. *Diabetologia* 55(1), 105–113 (2012).

48 Cornelis MC, Rimm EB, Curhan GC *et al.* Obesity susceptibility loci and uncontrolled eating, emotional eating and cognitive restraint behaviors in men and women. *Obesity (Silver Spring)* 22(5), E135–E141 (2014).

49 Namjou B, Keddache M, Marsolo K *et al.* EMR-linked GWAS study: Investigation of variation landscape of loci for body mass index in children. *Front. Genet.* 4, 268 (2013).

50 Laguna-Barraza R, Bermejo-Alvarez P, Ramos-Ibeas P *et al.* Sex-specific embryonic origin of postnatal phenotypic variability. *Reprod. Fertil. Dev.* 25(1), 38–47 (2012).

51 Gabory A, Attig L, Junien C. Sexual dimorphism in environmental epigenetic programming. *Mol. Cell. Endocrinol.* 304(1–2), 8–18 (2009).

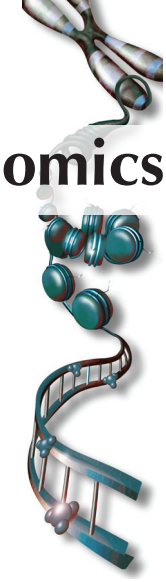
52 Guay SP, Brisson D, Lamarche B *et al.* DNA methylation variations at *CETP* and *LPL* gene promoter loci: new molecular biomarkers associated with blood lipid profile variability. *Atherosclerosis* 228(2), 413–420 (2013).

53 Guay SP, Brisson D, Munger J, Lamarche B, Gaudet D, Bouchard L. ABCA1 gene promoter DNA methylation is associated with hdl particle profile and coronary artery disease in familial hypercholesterolemia. *Epigenetics* 7(5), 464–472 (2012).

Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics



Altered miRNA expression in the cervix during pregnancy associated with lead and mercury exposure

Aim: Toxic metals including lead and mercury are associated with adverse pregnancy outcomes. This study aimed to assess the association between miRNA expression in the cervix during pregnancy with lead and mercury levels. **Materials & methods:** We obtained cervical swabs from pregnant women (n = 60) and quantified cervical miRNA expression. Women's blood lead, bone lead and toenail mercury levels were analyzed. We performed linear regression to examine the association between metal levels and expression of 74 miRNAs adjusting for covariates. **Results:** Seventeen miRNAs were negatively associated with toenail mercury levels, and tibial bone lead levels were associated with decreased expression of miR-575 and miR-4286. **Conclusion:** The findings highlight miRNAs in the human cervix as novel responders to maternal chemical exposure during pregnancy.

Keywords: blood • bone • cervix • delivery • labor • lead • mercury • miRNA • patella • tibia

Alison P Sanders[‡,1], Heather H Burris[*,‡,2], Allan C Just[3], Valeria Motta[4,5], Chitra Amarasiriwardena[1,4], Katherine Svensson[1], Emily Oken[6], Maritsa Solano-Gonzalez[7], Adriana Mercado-Garcia[7], Ivan Pantic[4,8], Joel Schwartz[3], Martha M Tellez-Rojo[7], Andrea A Baccarelli[3,4] & Robert O Wright[1]

[1]Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[2]Department of Neonatology, Beth Israel Deaconess Medical Center & Division of Newborn Medicine, Boston Children's Hospital & Harvard Medical School, 330 Brookline Ave, RO 318, Boston, MA 02215, USA

[3]Department of Environmental Health, Harvard TH Chan School of Public Health, Boston, MA 02115, USA

[4]Laboratory of Environmental Epigenetics, Exposure Epidemiology & Risk Program, Harvard TH Chan School of Public Health, Boston, MA 02115, USA

[5]Department of Clinical Sciences & Community Health University of Milan – Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy

[6]Department of Population Medicine, Harvard Medical School & Harvard Pilgrim Health Care Institute, Boston, MA 02215, USA

[7]Center for Nutrition & Health Research, National Institute of Public Health, Cuernavaca, Morelos, Mexico

[8]Division of Research in Public Health, National Institute of Perinatology, Mexico City, Mexico

*Author for correspondence:
Tel.: +1 617 667 3276
Fax: +1 617 667 7040
heburris@bidmc.harvard.edu

‡Authors contributed equally

Future Medicine part of fsg

Background

Environmental contaminants are associated with numerous adverse pregnancy outcomes including infertility, pregnancy loss and preterm delivery [1]. Specifically, maternal exposure to lead and mercury may contribute to increased risk of spontaneous abortion [2–4] and preterm delivery [5,6]. Exposure to lead and mercury is widespread in US and Mexican populations [7,8]. In Mexico, lead exposure occurs mainly through use of glazed pottery, lead paint and air pollution [8,9]. Maternal mercury exposure commonly occurs through maternal diet (fish consumption), coal combustion and personal care product use in Mexico [10–12].

Environmental exposures likely contribute to tissue-specific pathophysiology in the placenta and uterus, including the cervix. Studies have demonstrated that reproductive tissues are susceptible to environmental contaminants [13,14]. Epigenetic changes could contribute to the onset of labor via regulation of uterine contraction and quiescence [15,16]. Cervical dilation is a necessary step for labor and parturition to occur, yet few studies have examined the impact of environmental exposures on epigenetic changes in the cervix during pregnancy.

Many environmental contaminants including toxic metals have been associated with epigenetic modifications including DNA methylation, histone modification and miRNA expression [17,18]. MiRNAs are small noncoding RNAs that post-transcriptionally control gene expression. In previous reports, miR-16, miR-21 and miR-146a were downregulated in placentas exposed to cigarette smoke compared with unexposed controls [19,20]. The same investigators have linked these changes to reduced fetal growth [21]. Moreover, a number of miRNAs have been identified as associated with ovarian and cervical cancers (as reviewed in [22]) demonstrating that these biomarkers are biologically active in the reproductive system. Previously, we have shown that metal exposure affects the expression of specific miRNAs in blood from occupationally exposed individuals [23] and that miRNAs in the cervix during pregnancy are associated with subsequent gestational age at delivery [24], but metal-altered miRNA expression has not

yet been investigated with respect to tissues involved in parturition and labor.

In this study, we assessed the association between maternal lead and mercury exposures and miRNA expression in the cervix during the second trimester of pregnancy in a subset of 60 women enrolled in a prospective cohort. We used swabs to collect cervical cells during a Papanicolaou (Pap) smear and analyzed the expression profiles of 800 miRNA using the NanoString nCounter assay. This study aimed to identify miRNAs that were significantly associated with blood lead, patellar or tibial bone lead, or toenail mercury. Lead was selected as a prioritized metal of interest based on previously observed associations between prenatal lead exposure and preterm birth [6,25–26]. Mercury was also prioritized since limited literature suggests a relationship may exist with increased preterm birth risk [5]. The findings herein have the potential to enhance the current understanding of the complex molecular systems that govern environmentally associated alterations in the cervix during pregnancy.

Materials & methods

Study design

This study was conducted on a subcohort of 60 Mexican women aged 18–40 years participating in the PROGRESS birth cohort in Mexico City, and who consented to a cervical swab during pregnancy thereby participating in the PROGRESS Cervix Study. Full details of enrollment for the parent cohort and subcohort are published elsewhere [24,27]. Briefly, women in their second trimester were recruited between 2007 and 2011 through the Mexican social security system (Instituto Mexicano del Seguro Social). The parent cohort consists of 1054 mothers and 948 mother–infant pairs. Because the Cervix Study was conceived after the majority of recruitment had occurred only the last 100 women enrolled in the parent cohort were approached. Eighty enrollees provided written informed consent for an obstetrician to obtain a cervical swab mid-pregnancy (16–19-week gestation) for miRNA analysis. Women were offered a free Pap smear in return for their participation, which is the standard of care during pregnancy in the USA, but is not routinely offered in Mexico to pregnant women. The IRBs of the participating institutions approved this study. For funding reasons, 60 (80%) of the 80 samples collected were randomly selected for miRNA profiling analysis. All 60 women had blood lead exposure data and were included in this analysis. Forty (67%), 44 (73%) and 45 (75%) women had toenail mercury, patella and tibia lead data and were included in each subanalysis.

Participant data collection

Demographics and birth outcomes were collected as part of the parent study including maternal age, parity, self-reported prepregnancy BMI and years of education. Staff conducted in-person interviews, which included a question about household smoke exposure. Household smoke exposure was dichotomized as yes/no based on the mother's report that at least one household member smoked. All participants reported that they did not smoke during pregnancy. Parity was dichotomized as multiparous, if a woman had a prior live-born infant versus nulliparous if she had not. A histopathologic assessment for evidence of inflammation on the Pap smear served as a proxy for shifts in the cell-type mixture. The cervical swab sample collection contains mostly cervical epithelial cells (ectocervical and endocervical cells), and also leukocytes, which can be categorically quantified by a trained histopathologist. Although there were too few cells to perform detailed cell sorting, the clinical pathologist provided a qualitative assessment of the inflammatory cells noted on the Pap smear thereby allowing a subanalysis to adjust for cell type. We dichotomized Pap smear inflammation as yes/no based on the histopathologist's blinded interpretation of the smear. Twenty-four women (44%) had evidence of inflammation on the Pap smear. We performed a sensitivity analysis and confirmed that regression results were unchanged when adjusting for Pap inflammation (data not shown). This is an important consideration in future studies given that cell type and/or inflammation could impact miRNA expression profiles.

Lead & mercury exposure assessment

Venous whole blood samples and toenail samples were prepared and analyzed at the Trace Metals Laboratory of Harvard TH Chan School of Public Health in Boston as we have previously reported [28,29]. Lead concentrations were measured with a dynamic reaction cell-inductively coupled plasma mass spectrometer (Elan 6100; PerkinElmer, CT, USA). Toenail mercury analysis was performed by using direct mercury analysis methods for atomic absorption (DMA-80, Milestone Inc., CT, USA). Maternal tibia (cortical) and patella (trabecular) bone lead levels were measured within 1 month of delivery using a spot-source ^{109}Cd K-shell x-ray fluorescence (K-XRF) instrument (ABIOMED, MA, USA) [30]. Bone lead measurements were calculated as the average of two measures (one from each leg), weighted by the inverse proportion of measurement error. This process can generate negative values, which were included in the analyses, as we have previously reported in other studies [31].

Cervical miRNA collection & extraction

Cervical cells were collected in a method similar to a standard Pap smear protocol, where a cotton swab was used to collect cells from the endocervix of the external os. The sample was immediately placed in RNALater (Qiagen, CA, USA) and the specimen was frozen at -80°C until subsequent analysis. Total RNA was extracted using the Exiqon miRCURY kit (Exiqon, MA, USA) according to the manufacturer's protocol. A cleanup step was then performed by using an Amicon Ultra 0.5 ml cleanup kit (EMD Millipore, MA, USA). MiRNAs were quantified by using a NanoPhotometer P-300 (Implent GmbH, CA, USA).

NanoString nCounter assay for miRNA expression

MiRNA expression was assessed using the NanoString nCounter system (NanoString Technologies, WA, USA). This method enables multiplexed direct digital counting of miRNA molecules [32]. This method measured a total of 800 probes that were available for analysis at the time of this study, and included both endogenous human-associated miRNAs as well as viral miRNAs that are expressed in human cells [33–35]. We performed a feasibility pilot with ten initial samples, including two technical replicates, which we previously reported showed strong correlation (r = 0.98) [24].

The raw count data from the 60 samples were normalized by using the NanoStringNorm R package [36]. Data were background-corrected by subtracting the mean of the six negative controls included on the platform, and normalized using the geometric mean of the ten probes with the lowest coefficients of variation – which were used to calculate a scaling factor as suggested by the package guidelines. *A priori* we required that probes be detectable in at least 60% of the samples. This resulted in 74 probes that were included in the analyses. Individual probes with expression levels below the LOD were assigned a nominal value of 1. Note that the proportion of samples below detect for each miRNA is reported in Supplementary Table 2. The distributions of miRNA expression and model residuals showed that our selection of a linear model was appropriate for these data.

Statistical analysis

To examine the association between miRNA expression levels and biomonitored lead or mercury, we used linear regression models. We used separate models to estimate the mean doubling of expression ($\log_2$ unit increase) of each miRNA associated with a unit (bone measures) or tenfold unit (toenail mercury and blood lead) increase metal exposure. To satisfy linear model assumptions, blood lead and toenail mercury levels were log transformed whereas untransformed patellar and tibial bone lead measures were normally distributed and thus we kept them untransformed. We chose covariates *a priori* and included maternal age, education, parity and smoke exposure inside the home. Parity and smoke exposure were dichotomized as described above. Maternal education was categorized as having less than high school, high school or greater than high school education. Both adjusted and unadjusted regression models were performed to examine the association between each metal exposure and the 74 probes' $\log_2$-transformed miRNA levels. The reported beta coefficients represent the fold change (doubling or halvings) in expression per unit change in exposure. To transform from fold-change into raw expression change, the beta coefficients can be back transformed by using the antilog of 2. For example, the equation $(2^{beta} - 1) \times 100$ yields the percent change in raw miRNA expression. Because expression data are conventionally interpreted in $\log_2$, we present the beta coefficients that are easily interpreted in doublings or halvings. p-values and Storey's false discovery rate (FDR) q-values were calculated to estimate significance [37]; $p < 0.05$ was considered statistically significant. MiRNAs with an FDR q-value <0.1 that met these requirements in the adjusted model were retained for downstream target prediction and pathways enrichment analyses. A sensitivity analysis was performed that also adjusted for evidence of inflammation on the Pap smear. An additional sensitivity analysis with only the two bone lead measures also adjusted for prepregnancy BMI categorized as normal (BMI: 18.5–25 kg/m^2), overweight (BMI: 25–30 kg/m^2) and obese (≥30 kg/m^2).

We performed two sensitivity analyses to assess the influence of probes with expression levels below the LOD. First, we treated probes with expression levels below the LOD as missing and excluded these probes from the analysis. Second, we assigned probes below the LOD to the minimum expression measured for a specific miRNA.

Prediction of miRNA targets

To predict downstream mRNA targets, the set of differentially expressed miRNAs which passed $p < 0.05$ and FDR q-value < 0.1 in the adjusted linear regression model were uploaded into the Ingenuity Pathway Analysis (IPA) tool (Ingenuity® Systems, CA, USA). Putative miRNA–mRNA relationships were identified using the IPA miRNA Target Filter, based on a knowledge-base of predicted and experimentally observed relationships. Where possible, we stringently selected for only the experimentally observed miRNA–mRNA relationships, and the resulting target gene list was analyzed for functional network and pathway analy-

sis. If no experimentally observed relationships were identified, we analyzed the highly predicted mRNA targets.

Functional network & pathway enrichment analysis of miRNA transcripts

Functional analysis was carried out to identify molecular networks and biological functions significantly associated with the mRNA target gene set. Analysis of molecular network mapping, physiological system function enrichment and reproductive system disease and function enrichment was performed by using the IPA tool. Within the reproductive system disease and function enrichment analysis, the results were further filtered by excluding explicitly male reproductive categories (e.g., prostate and sperm) or breast-related results. The IPA proprietary database curates gene–phenotype associations, molecular interactions, regulatory events and chemical knowledge to provide a global molecular network. Related networks were algorithmically constructed based on connectivity. Statistical significance of each biological function was calculated by using Fisher's exact text with an alpha set at 0.05.

Results

Characteristics of study participants

Demographics of the 60 Mexican women participating in the PROGRESS Cervix Study are presented in Table 1. Generally, this was a lean cohort of women, the majority of whom reported no smoke exposure in the home, had at least 12 years of education and were multiparous. Second trimester whole blood lead levels were detectable in every sample (n = 60). Six women (10%) had blood lead levels greater than 5 μg/dl, the US Centers for Disease Control and Prevention reference level for pregnant women [38]. Postpartum bone lead levels were measured in 44 and 45 women, respectively, and toenail mercury levels were measured in a subset of 40 women. Spearman correlation of the exposures showed that patella and blood lead levels were positively correlated ($r = 0.49$; $p = 0.0008$); none of the other pair-wise correlations of exposure measures were significantly related (Supplementary Table 1). Descriptive statistics for each of the 74 miRNAs measured in cervical cells are provided in Supplementary Table 2.

Mid-pregnancy cervical miRNA expression was associated with mercury exposure

Using adjusted linear regression models, we identified 17 miRNAs that were associated with toenail mercury levels, which included let-7a and let-7b, and miRs-205, 125b, 200c, 342, 203, 24, 22, 23b, 375, 23a, 210, 200b, 99a, 21 and 193b ($p < 0.05$, and FDR q-value < 0.07) (Table 2). All of the miRNAs had lower expression among the higher mercury-exposed women. The beta coefficients that indicate the respective fold-change in miRNA expression per tenfold increment of mercury exposure and their interpreted values are presented in Table 2. For example, a beta coefficient of -4.0 for miR-205 represents a fourfold decrease (i.e., four halvings) in expression per tenfold increase in toenail mercury level. For completeness, the beta coefficients can be back transformed by using the antilog of 2. For example, $2^{-4.03}$ yields a value of 0.062, which is interpreted as a 94% decrease in raw miRNA expression. Because expression data are conventionally interpreted in $\log_2$, we present the beta coefficients that are interpreted as doublings or halvings. When we compare the 25–75th percentiles (interquartile range) of toenail mercury for the significant miRNAs, the differences in miRNA expression correspond to a one- to twofold change (data not shown). Representative plots of the four most significant mercury-associated miRNAs, miR-205, miR-125b, let-7b and miR-200c are shown in Figure 1.

We performed two sensitivity analyses that treated probes with expression levels below the LOD as missing and excluded these probes from the analysis or assigned probes below the LOD to the minimum expression measured for a specific miRNA. The results were similar for the second sensitivity analysis; however, some miRNAs lost statistical significance when probes below detect were excluded from analysis probes (e.g., miR-125b and let-7a) (Supplementary Table 3).

Mid-pregnancy cervical miRNA expression was associated with lead biomarkers

We identified distinct sets of miRNAs that were associated with maternal blood or bone lead (Table 2). Using adjusted linear regression models, we identified two miRNAs, miR-297 and miR-188, that had increased expression associated with blood lead levels ($p < 0.05$). We identified seven miRNAs that were associated with patellar bone lead levels including miR-320e, miR-22, miR-93, miR-769, miR-297, miR-425 and miR-361. All but miR-297 had decreased expression with increasing patellar lead levels. Notably, miR-297 was common to both the blood lead and patellar bone subsets and had increased expression with increasing blood lead and patellar bone lead. To show that the difference in identified miRNAs associated with blood lead from bone lead measures was not due to shifts in the sample population, we performed a subanalysis of blood lead associations restricted to the 44 individuals with both bone lead measures and the results were similar (Supplementary Table 4). None of the blood or patella lead-associated miRNAs reached statistical significance after FDR correction. In addition, we identified six miRNAs associ-

ated with tibial bone lead that included miR-575, miR-4286, miR-15a, miR-142, miR-193b and miR-494 ($p < 0.05$). MiR-575, miR-4286, miR-193b and miR-630 showed decreased expression with increasing tibial lead levels, and miR-15a and miR-142 had increased expression. Both miR-575 and miR-4286 had significantly decreased expression with increasing tibial lead levels ($p < 0.05$ and FDR q-value < 0.1) and thus we retained them for downstream functional analysis. The beta coefficients in Table 2 indicate the respective fold-change in miRNA expression per one unit increment of bone lead exposure. For example, a beta coefficient of -0.06 for miR-575 represents approximately a 0.06-fold decrease or 4.1% decrease in raw expression per unit increase in tibia lead level (Table 2). A beta coefficient of 0.10 for miR-15a represents a 0.1-fold increase or a 7.2% increase in raw expression. When we compare the 25th to 75th percentiles (interquartile range) of bone lead for the significant miRNAs, the differences in miRNA expression correspond to modest approximately 0.5-fold changes (data not shown).

Because the relationship between miRNA expression and bone lead measures may additionally be confounded by BMI, we performed a sensitivity analysis adjusted for prepregnancy BMI in addition to the primary set of covariates (Supplementary Table 5); the top identified miRNAs were similar.

Metal-associated miRNAs have known & predicted mRNA targets

For each of the metal-associated miRNAs, we report the number of experimentally observed or highly predicted mRNA targets regulated by the metal-associated miRNAs (Table 2). The blood lead, patella lead, tibia lead and toenail mercury-associated miRNAs had a total of 78, 2793, 3050 and 8446 downstream gene targets. Note that among the 8446 downstream gene targets of the 17 mercury-associated miRNAs, three pairs of miRNAs share sequence homology and have identical target mRNAs including let-7b and let-7a, miR-23a and miR-23b and miR-200b and miR-200c. Together these six miRNAs potentially coregulate over 3000 genes.

Functional network & pathway analysis of miRNA target genes

The 17 mercury-associated and two tibia lead-associated miRNAs that passed FDR correction (q-value < 0.1) in the linear model were selected for subsequent functional pathway analysis. We identified over 8000 mRNAs that were experimentally observed or highly predicted targets of the 17 miRNAs (Table 2). When the data were stringently filtered for experimentally observed targets only, we found that 15 of the miRNAs are known to regulate a total of 362 genes consisting of 330 unique mRNA targets. The 330 mercury-associated mRNA targets mapped to several molecular networks (Supplementary Table 6). The top two networks were enriched for organismal, cellular and cardiovascular system development ($p = 1 \times 10^{-55}$), as well as cell cycle, cancer and gene expression ($p = 1 \times 10^{-38}$). Physiological system development and function analysis showed enrichment for organismal survival ($p < 5.0 \times 10^{-18}$), cardiovascular system development and function ($p < 4.3 \times 10^{-18}$), organismal development (4.9×10^{-13}), tissue morphology ($p < 6.2 \times 10^{-14}$) and embryonic development ($p < 4.9 \times 10^{-13}$). When the

Table 1. Maternal demographics for 60 pregnant women participating in the PROGRESS cervix study.

Participant characteristics	n (%)
Education:	
– <12 years	19 (32)
– 12 years	24 (40)
– >12 years	17 (28)
Smoke in home:	
– No	42 (70)
– Yes	18 (30)
Parity:	
– Multiparous	34 (57)
– Nulliparous	26 (43)
BMI:	
– Normal	30 (50)
– Overweight	19 (32)
– Obese	11 (18)
Maternal age (years), mean ± SD (range)	27.9 ± 5.7 (18–40)
Blood lead† (μg/dl), mean ± SD (range)	2.85 ± 1.63 (0.87–9.38)
Patella lead† (n = 44), mean ± SD (range)	4.16‡ ± 6.99 (-6.85–20.90)
Tibia lead† (n = 45), mean ± SD (range)	1.45‡ ± 8.39 (-32.60–19.45)
Toenail† mercury (n = 40); μg/g, mean ± SD (range)	0.17 ± 0.09 (0.03–0.47)

†The number of samples measured for each exposure marker varied as follows: blood lead (n = 60), patella lead (n = 44), tibia lead (n = 45) and toenail mercury (n = 40).
‡The reported mean and standard deviation for patellar and tibial lead include measurements that had negative-weighted average values (n = 15). The mean and standard deviation excluding negative values were 7.28 ± 6.20 and 5.66 ± 4.18 for patellar and tibial lead.
SD: Standard deviation.

Table 2. Differentially expressed miRNA associated with blood lead, bone lead or toenail mercury ($p < 0.05$).[†]

Biomarker	miRNA	Adjusted beta (95% CI)	Interpreted beta as % expression change (95% CI)	p-value	q-value	Predicted targets (n)
Blood lead (n = 60)	hsa-miR-297	0.88 (0.21–1.55)	84.0 (15.7–192.8)	0.01	0.45	38
	hsa-miR-188	0.57 (0.11–1.03)	48.5 (7.9–104.2)	0.01	0.45	40
Patella lead (n = 44)	hsa-miR-320e	-0.07 (-0.11 to -0.02)	-4.7 (-7.3 to -1.4)	0.003	0.17	64
	hsa-miR-22–3p	-0.07 (-0.13 to -0.01)	-4.7 (-8.6 to -0.7)	0.020	0.23	597
	hsa-miR-93–5p	-0.10 (-0.19 to -0.01)	-6.7 (-12.3 to -0.7)	0.025	0.23	1235
	hsa-miR-769–5p	-0.08 (-0.15 to -0.01)	-5.4 (-9.9 to -0.7)	0.030	0.23	82
	hsa-miR-297	0.03 (0.00–0.06)	2.1 (0.0–4.2)	0.031	0.23	38
	hsa-miR-425–5p	-0.10 (-0.19–0.00)	-6.7 (-12.3–0.0)	0.046	0.23	238
	hsa-miR-361–3p	0.04 (0.00–0.08)	2.8 (0.0–5.7)	0.047	0.23	539
Tibia lead (n = 45)	hsa-miR-575	-0.06 (-0.1 to -0.02)	-4.1 (-6.7 to -1.4)	0.003	0.08	82
	hsa-miR-4286	-0.13 (-0.21 to -0.05)	-8.6 (-13.5 to -3.4)	0.003	0.08	227
	hsa-miR-15a-5p	0.10 (0.02–0.19)	7.2 (1.4–14.1)	0.018	0.36	1452
	hsa-miR-142–3p	0.08 (0.01–0.16)	5.7 (0.7–11.7)	0.029	0.40	368
	hsa-miR-193b-3p	-0.11 (-0.20 to -0.01)	-7.3 (-12.9 to -0.7)	0.033	0.40	340
	hsa-miR-494	-0.06 (-0.12–0.00)	-4.1 (-8.0–0.0)	0.044	0.46	581
Toenail mercury (n = 40)	hsa-miR-205–5p	-4.03 (-6.48 to -1.58)	-93.9 (-98.9 to -66.6)	0.001	0.02	426
	hsa-miR-125b-5p	-5.86 (-9.59 to -2.14)	-98.3 (-99.9 to -77.3)	0.002	0.02	985
	hsa-let-7b-5p	-2.90 (-4.85 to -0.94)	-86.6 (-96.5 to -47.9)	0.004	0.02	1180[‡]
	hsa-miR-200c-3p	-3.52 (-5.97 to -1.08)	-91.3 (-98.4 to -52.7)	0.005	0.02	1074[‡]
	hsa-miR-342–3p	-3.10 (-5.39 to -0.81)	-88.3 (-97.6 to -43.0)	0.008	0.03	337
	hsa-miR-203	-3.39 (-5.95 to -0.83)	-90.5 (-98.4 to -43.7)	0.009	0.03	870
	hsa-let-7a-5p	-4.21 (-7.42 to -0.99)	-94.6 (-99.4 to -49.7)	0.010	0.03	1180[‡]
	hsa-miR-24–3p	-3.79 (-6.99 to -0.60)	-92.8 (-99.2 to -34.0)	0.020	0.05	741
	hsa-miR-22–3p	-2.81 (-5.21 to -0.41)	-85.7 (-97.3 to -24.7)	0.022	0.05	597
	hsa-miR-23b-3p	-4.14 (-7.76 to -0.52)	-94.3 (-99.5 to -30.3)	0.025	0.05	1172[‡]
	hsa-miR-375	-2.29 (-4.32 to -0.25)	-79.6 (-95.0 to -15.9)	0.027	0.05	233
	hsa-miR-23a-3p	-3.50 (-6.64 to -0.37)	-91.2 (-99.0 to -22.6)	0.028	0.05	1172[‡]
	hsa-miR-210	-3.77 (-7.20 to -0.34)	-92.7 (-99.3 to -21.0)	0.031	0.05	78
	hsa-miR-200b-3p	-3.33 (-6.40 to -0.26)	-90.1 (-98.8 to -16.5)	0.034	0.05	1074[‡]
	hsa-miR-99a-5p	-3.55 (-6.89 to -0.20)	-91.5 (-99.2 to -12.9)	0.038	0.05	73
	hsa-miR-21–5p	-5.17 (-10.20 to -0.14)	-97.2 (-99.9 to -9.2)	0.044	0.05	340
	hsa-miR-193b-3p	-3.77 (-7.53 to -0.01)	-92.7 (-99.5 to -0.7)	0.050	0.06	340

Each multivariable regression model was adjusted for maternal age, education, smoke exposure in the home and parity.
[†]The adjusted beta values represent the fold-change in miRNA per tenfold increase in toenail mercury and blood lead, or per one unit increase for bone lead measures.
[‡]Let-7b and let-7a, miR-23a and miR-23b, and miR-200b and miR-200c share sequence homology and have identical target mRNAs.

list of associated diseases and functions was filtered for reproductive system disease or function, a group of 60 genes was enriched for reproductive system development ($p < 5.76 \times 10^{-23}$) and another group of 40 genes was enriched for abnormal morphology of the reproductive system ($p < 5.05 \times 10^{-15}$). Specifically, a majority (56%, n = 26) of the gene subset enriched for abnormal morphology is regulated by three miRNAs: let-7, miR-125b and miR-24. Also of note, 18 miRNA-regulated molecules are involved in aryl hydrocarbon receptor (AHR) signaling ($p = 6 \times 10^{-12}$) (Supplementary Table 7). The AHR signaling pathway has known links to environmentally mediated epigenetic modification [39] and plays a critical role in the

female reproductive system [40]. Previously, we have shown that cervical miRNAs associated with subsequent gestational age were also enriched for the AHR pathway signaling [24].

None of the mRNA targets of the tibia lead-associated miRNAs, miR-575 and miR-4286 had experimentally observed relationships with target genes. Therefore, we analyzed a total of 309 highly predicted targets for the two miRNAs (Table 2). The tibia lead-associated predicted gene targets showed enrichment for several networks including carbohydrate metabolism, small-molecule biochemistry and DNA replication ($p = 1 \times 10^{-43}$), as well as behavior, lipid metabolism and posttranslational modification ($p = 1 \times 10^{-27}$) (Supplementary Table 8). Physiological system development and function analysis showed enrichment for organ and tissue development ($p < 0.02$), hematological system development and function ($p < 0.02$), hematopoiesis ($p < 0.02$) and cardiovascular system development and function ($p < 0.01$). When we filtered the list of associated disease and functions for reproductive system disease or function, a group of 14 mRNA targets was enriched for preeclampsia ($p < 8.80 \times 10^{-3}$) [41] and two genes were associated with formation of placenta in mouse models [42,43]. Ten of the preeclampsia-associated genes are the predicted targets of miR-4286, and four are the predicted targets of miR-575. Interestingly, miR-4286 is predicted to target three genes involved in AHR signaling including AHR repressor, prostaglandin E synthase 3, and tumor protein P73.

Discussion

Maternal exposure to lead or mercury has been associated with numerous adverse birth outcomes [1]; however, the molecular mechanisms by which lead or mercury influence the female reproductive system during pregnancy are unknown. Changes in cervical miRNA expression are a potential mechanism that could alter gene expression leading to aberrant changes in cervix tissue function and subsequently impact parturition. In this cohort study of pregnant women, we identified distinct miRNAs measured in cervical samples during pregnancy that are associated with maternal metal exposure. Our study identified 17 miRNAs that were significantly associated with maternal mercury exposure in the second trimester. Separate and largely distinct subsets of two, seven and six miRNAs were also identified in association with maternal second trimester blood lead, and postpartum lead measures in patellar and tibial bone, respectively. We computationally predicted the downstream mRNA targets for the most significantly metal-associated miRNAs, and subsequent functional enrichment analyses revealed that mercury-associated miRNA gene targets are involved

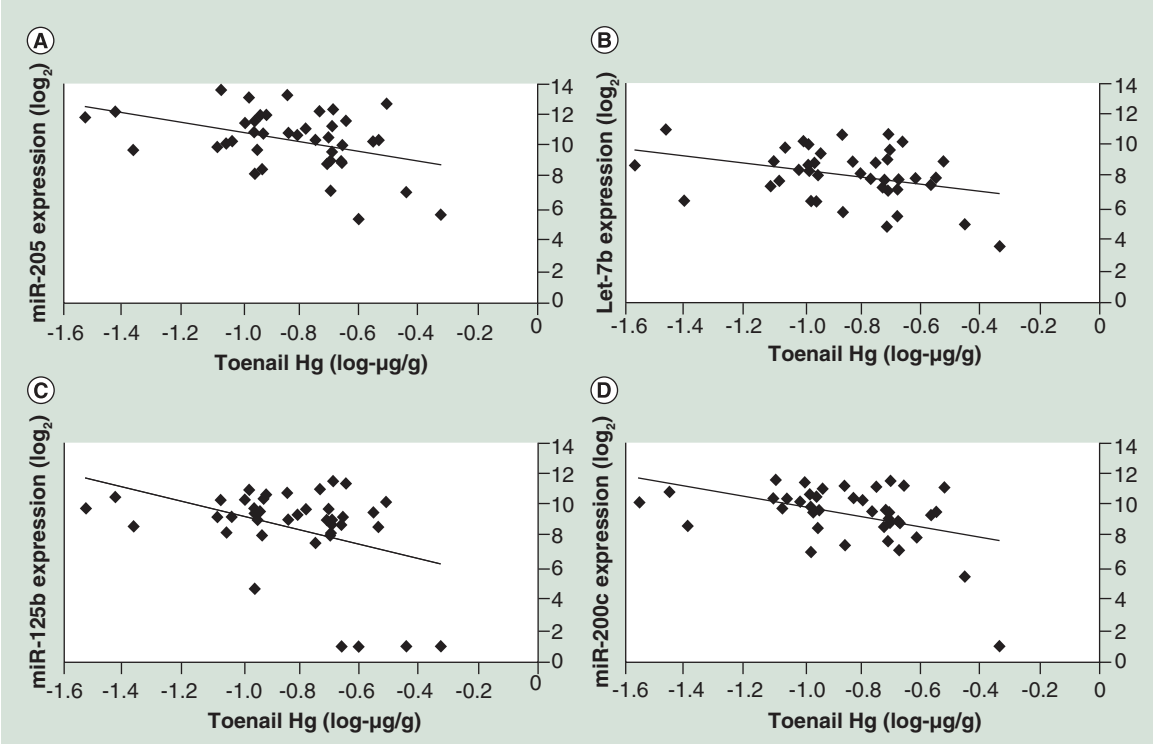


Figure 1. Representative plots of the top four mercury-associated miRNAs. Plots shown include: **(A)** miR-205; **(B)** let-7b; **(C)** miR-125b and **(D)** miR-200c. Unadjusted regression lines are shown and Pearson correlation estimates are: **(A)** -0.41; **(B)** -0.35; **(C)** -0.40 and **(D)** -0.43.

in reproductive system development and morphology, lead-associated gene targets are enriched for preeclampsia, and molecular targets of both the mercury- and lead-associated miRNAs are involved in the AHR signaling pathway.

No previous studies have evaluated the effects of metals on cervical miRNA expression; however, some similar patterns have been identified between other tissue-specific or circulating miRNAs and metals. For example, we report that miR-21 had decreased expression in association with maternal mercury levels. While miR-21 has not been previously studied with respect to mercury, miR-21 was downregulated in placentas exposed to cigarette smoke, which contains lead and cadmium, compared with controls [19]. Additionally, urinary miR-21 had a negative relationship with blood lead levels in adolescents [44]. In the present study, cervical miR-21 had marginally significantly decreased expression with respect to patella lead and was not associated with tibia lead or blood lead levels. Previously, we have shown that miR-146a expression in blood was significantly decreased in association with lead- and cadmium-rich particulate matter in adult males [45]; however, miR-146a was not detected in cervical samples in the present study. A study of fibroblasts exposed to a mixture of arsenic, cadmium and lead computationally predicted, then confirmed upregulation of six candidate miRNAs: miRs-154, 222, 379, 10, 375, 204 and 133 [46]. We report that miR-375 was significantly negatively associated with toenail mercury levels, and found no association with blood or bone lead. The only previous study of mercury-altered miRNAs exposed neuronal/glial cells derived from NT2 carcinoma pluripotent stem cells to methyl mercury chloride and showed upregulated levels of miR-302b, 367, 372, 196b and 141 [47], none of which overlap with miRNAs reported here.

Our study identified 17 miRNAs that have significantly decreased expression with maternal mercury exposure. Notably, miR-205, miR-125b and let-7 are miRNAs known as oncomirs, which are involved in a number of human cancers and directly regulate oncogenes including phosphatase and tensin homolog, tumor protein P53 and *RAS*, respectively [48]. Supported by our functional enrichment analysis, the concerted action of these oncomirs and the other identified miRNAs have known impacts on a number of cell cycle and proliferation pathways that could affect parturition.

It is not surprising then that a number of the mercury-associated miRNAs and their mRNA targets have also been identified with respect to reproductive system diseases. Our functional analysis showed distinct subsets of mRNA targets that are involved in reproductive system development and morphology. Specifically, the gene subset enriched for abnormal morphology had mRNA targets predominantly regulated by let-7, miR-125b and miR-24. For example, miR-205, miR-24 and miR-21 are highly expressed in cervical cancers compared with normal controls [49]; and miR-205 is thought to play a role in cervical cancer or serve as a diagnostic marker in plasma [22,50]. Downregulation of miR-203 and upregulation of let-7 and miR-21 are associated with early stage and invasive cervical carcinoma [51]. Furthermore, we noted a negative association between miRs-200c and 200b with increasing maternal mercury levels. The miR-200 family has been shown to increase in myometrium expression levels as pregnancies approach term, and in turn negatively regulate several contraction-associated genes including zinc finger E-box binding homeobox-*1* and *-2*, and connexin 43 [15]. We note that although miRNAs typically act to negatively regulate gene expression, evidence also demonstrates that some miRNAs act by increasing target gene expression [15]. Given the paucity of data on normal cervix expression outside of cancer-based studies with environmental contaminants, additional studies on the biological role of the identified miRNAs and their role in normal pregnancies and parturition are needed.

Analysis of the blood and bone lead exposures identified largely distinct subsets of lead-associated miRNAs with the exception of miR-297, which had increased expression with increasing blood lead and patellar bone lead. Additional studies are needed to ascertain whether miR-297 is a robust biomarker for acute and/or chronic lead exposure. Tibia lead was more strongly associated with changes in cervical miRNA expression than patellar lead. Among the tibia lead-associated targets, miR-575 and miR-4286 had significantly decreased expression. It is notable that 14 of the downstream targets we identified were differentially expressed in a study of genome-wide expression using decidua basilis (placenta) tissue from preeclamptic versus normal pregnancies [41]. Moreover, miR-4286 is predicted to regulate three genes involved in AHR signaling, which is a vital pathway in the female reproductive system that is influenced by environmental contaminants such as cigarette smoke and dioxins [40]. Taken together, these findings are suggestive of a possible environmentally mediated mechanism of miRNA gene regulation during pregnancy. Additional mechanistic studies using tandem mRNA and miRNA transcriptional profiling are needed.

The tibia consists mainly of cortical bone, which has a lower turnover rate and longer half-life with respect to lead compared with the patella, which consists mostly of trabecular bone [30]. Therefore, patella lead is likely to track closely with bone remodeling, while tibia lead represents cumulative exposure. Our data support these conclusions given the positive correlation between blood

lead and patella lead as bone lead stores are a major source of maternal blood lead. Mobilization of lead stores via bone turnover is considered the major source of circulating lead in absence of ongoing external exposure sources [30]. Thus, during pregnancy, patellar lead might exert the potential greatest impact on miRNA expression, if the mechanism required only mobilization as a result of bone demineralization. In contrast, our data support a stronger association between tibia lead (chronic lead exposure) and miRNA expression, suggesting that cumulative lead exposure among women may explain our findings more than concurrent pregnancy exposures.

To our knowledge, this is the first study to examine the association between cervical miRNA expression and environmental exposures and has many strengths. We sampled the cervix during the second trimester, a relevant time point for miRNAs potentially involved in pregnancy outcomes or reproductive system pathophysiology. miRNA expression is tissue-specific, and investigating a key tissue for delivery initiation is critical to understand mechanisms that may regulate signaling cascades, as we have previously reported in this cohort and the association with subsequent gestational age [24]. Our primary outcome of interest in the previous study was gestational age, because pathophysiologically we hypothesized that metal-induced alteration in the cervix would not be associated with fetal growth, unlike those in the placenta, which biologically might mediate a metal-fetal growth relationship. We prospectively enrolled women during pregnancy who were participating in an ongoing population-based cohort study with carefully collected covariate data. Given the cohort design of our study, we were able to adjust for potential confounding variables. We assessed total mercury in toenails, which is a valid biomarker for chronic mercury exposure that correlates with methylmercury levels from hair [52,53]. Additionally, in this study, we examined human cervical miRNA responses to acute (i.e., blood levels) and chronic (i.e., tibia bone levels) environmentally relevant levels of lead. Studies that compare the acute versus chronic miRNA response to exposure, as we observed here between the lead biomarkers are useful in understanding the impact and contributions of each type of population-level exposure. The levels of metals assessed in this study are comparable to those reported in previous cohorts. For example, the levels observed were slightly higher than the geometric mean of 0.68 μg/dl among 253 pregnant women in the USA participating in the Fourth Report on Human Exposure to Environmental Chemicals (NHANES IV) [54]. The levels of toenail mercury reported here were comparable to levels observed in two previous studies among adult populations, which reported average mercury toenail levels of 0.212 μg/g (geometric mean) [55] and 0.25 μg/g [56]. Only six women in this study had blood lead levels greater than 5 μg/dl; the Centers for Disease Control and Prevention reference level for pregnant women [38], there is no current guideline for toenail mercury levels during pregnancy.

Our study also has several limitations. The PROGRESS cohort is a racially/ethnically homogeneous population, which may limit external generalizability, although results may likely be better generalized to Mexican–Americans. We sampled the cervix just once mid-pregnancy and thus cannot identify from this study the impact of toxic metals on miRNA expression throughout pregnancy. Differences in the miRNAs identified in this study compared with previous studies may originate from differences in the tissue type, exposure levels and source populations, as well as the type of quantification platform used, and time of sample collection (during pregnancy vs not). Our source population was largely a healthy group of pregnant women in Mexico City. Our selection of miRNAs was informed by the NanoString platform. The NanoString platform analyzed 800 human and viral miRNA at the time of this study. Many of these targets (n = 726) were detectable in fewer than 60% of cervical samples, and thus were not included in our analyses. Our study, therefore, was unable to detect subtle differences in less highly expressed miRNAs. This is a known characteristic of the NanoString platform that has less sensitive probe detection than other platforms [57]. Therefore, we chose to retain miRNAs that were below the limit of detection in less than 40% of our participants and we performed sensitivity analyses to examine how our approach affected the magnitude and direction of observed associations. However, a strength of the NanoString platform is that it provides quantitative assessment of miRNAs from samples with low overall RNA yields such as the samples obtained with a single cervical swab. We did not have enough biologic material to analyze our samples in duplicate using another platform. While rigorous computational methods were employed to identify potential miRNA-mediated mRNA expression, this study did not directly measure the downstream changes on RNA expression. The future characterization of miRNA control of gene expression and protein translation in cervical tissue and *in vivo* models will provide information on how cervical cells respond to lead or mercury that could contribute to adverse pathophysiology. Additional studies from larger populations and mechanistic studies are needed to replicate the findings here.

Conclusion

Our findings support a miRNA-mediated response to environmental contaminants in the cervix, in an obtainable tissue type during human pregnancy. We observed mercury- and lead-associated miRNAs in the cervix during pregnancy. To our knowledge, this is one of the first studies to link molecular changes in cervical tissue during pregnancy to maternal toxic metal exposure.

Future perspective

Understanding the pathophysiological role of mercury- and lead-altered miRNAs in the cervix and other relevant tissues is a priority for future studies. Replication in human cohorts followed by *in vitro/vivo* mechanistic studies should be prioritized to further validate these findings. The contributions of environmentally altered molecular programming in the cervix to adverse birth outcomes such as spontaneous abortion and preterm delivery, as well as the role in transgenerational reproductive system effects remain to be studied.

Acknowledgements

The authors thank Erroll Reuckert at NanoString Technologies for his assistance with the miRNA expression profiling and analysis. The authors also thank the ABC Medical Center in Mexico City for providing facilities during data collection.

Financial & competing interests disclosure

This work was supported in part by Pilot Project funding from the HSPH-NIEHS Center for Environmental Health (ES000002) and NIH/NIEHS: K23ES022242, K99ES023450, P42ES016454, P30ES23515, R01ES013744, R01ES020268, R01ES021357, the Klarman Scholars Program at Beth Israel Deaconess Medical Center, the Harvard Catalyst D-MaPS Program and the National Institute of Public Health/Ministry of Health of Mexico. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The IRBs of the participating institutions approved this study: Brigham and Women's Hospital # 2006-P-001416 and P001792, Icahn School of Medicine at Mount Sinai human subjects management #12-00751 and Instituto Nacional de Salud Publica project #560. Written informed consent was obtained from women participating in the PROGRESS study.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/epi.15.54

Executive summary

- Seventeen mercury-associated miRNAs were identified from the cervix during pregnancy ($p < 0.05$; false discovery rate q-value < 0.1).
- Two tibia lead-associated miRNAs were identified ($p < 0.05$; false discovery rate q-value < 0.1).
- The mercury-associated miRNA gene targets were involved in reproductive system development and morphology.
- The tibia lead-associated gene targets were enriched for preeclampsia.
- Molecular targets of both the mercury- and lead-associated miRNAs are involved in the aryl hydrocarbon receptor signaling pathway.
- Our findings indicate a miRNA-mediated response to environmental contaminants in the cervix, in an obtainable tissue type during human pregnancy.
- The contributions of environmentally altered molecular programming in the cervix to adverse birth outcomes such as spontaneous abortion and preterm delivery are a priority for future studies.

References

Papers of special note have been highlighted as:
• of interest; •• of considerable interest

1 Silbergeld EK, Patrick TE. Environmental exposures, toxicologic mechanisms, and adverse pregnancy outcomes. *Am. J. Obstet. Gynecol.* 192(5 Suppl.), S11–S21 (2005).

2 Savitz DA, Sonnenfeld NL, Olshan AF. Review of epidemiologic studies of paternal occupational exposure and spontaneous abortion. *Am. J. Ind. Med.* 25(3), 361–383 (1994).

3 Aschengrau A, Zierler S, Cohen A. Quality of community drinking water and the occurrence of late adverse pregnancy outcomes. *Arch. Environ. Health* 48(2), 105–113 (1993).

4 Anttila A, Sallmen M. Effects of parental occupational exposure to lead and other metals on spontaneous abortion. *J. Occup. Environ. Med.* 37(8), 915–921 (1995).

5 Xue F, Holzman C, Rahbar MH, Trosko K, Fischer L. Maternal fish consumption, mercury levels, and risk of preterm delivery. *Environ. Health. Perspect.* 115(1), 42–47 (2007).

6 Perkins M, Wright RO, Amarasiriwardena CJ, Jayawardene I, Rifas-Shiman SL, Oken E. Very low maternal lead level in pregnancy and birth outcomes in an Eastern Massachusetts population. *Ann. Epidemiol.* 24(12), 915–919 (2014).

7 Centers for Disease Control and Prevention. Fourth National Report on human exposure to environmental chemicals, updated tables (2013). www.cdc.gov

8 Romieu I, Palazuelos E, Hernandez Avila M *et al.* Sources of lead exposure in Mexico City. *Environ. Health. Perspect.* 102(4), 384–389 (1994).

9 Caravanos J, Dowling R, Tellez-Rojo MM *et al.* Blood lead levels in Mexico and pediatric burden of disease implications. *Ann. Glob. Health.* 80(4), 269–277 (2014).

10 Weldon MM, Smolinski MS, Maroufi A *et al.* Mercury poisoning associated with a Mexican beauty cream. *West. J. Med.* 173(1), 15–18; discussion 19 (2000).

11 Agency for Toxic Substances and Disease Registry (ATSDR). Toxicological profile for Mercury (1999). http://www.atsdr.cdc.gov

12 Oken E, Guthrie LB, Bloomingdale A *et al.* Assessment of dietary fish consumption in pregnancy: comparing one-, four- and thirty-six-item questionnaires. *Public. Health. Nutr.* 17(9), 1949–1959 (2014).

13 Leino O, Kiviranta H, Karjalainen AK *et al.* Pollutant concentrations in placenta. *Food. Chem. Toxicol.* 54, 59–69 (2013).

14 Boverhof DR, Kwekel JC, Humes DG, Burgoon LD, Zacharewski TR. Dioxin induces an estrogen-like, estrogen receptor-dependent gene expression response in the murine uterus. *Mol. Pharmacol.* 69(5), 1599–1606 (2006).

15 Renthal NE, Chen CC, Williams KC, Gerard RD, Prange-Kiel J, Mendelson CR. miR-200 family and targets, ZEB1 and ZEB2, modulate uterine quiescence and contractility during pregnancy and labor. *Proc. Natl Acad. Sci. USA* 107(48), 20828–20833 (2010).

16 Williams KC, Renthal NE, Condon JC, Gerard RD, Mendelson CR. MicroRNA-200a serves a key role in the decline of progesterone receptor function leading to term and preterm labor. *Proc. Natl Acad. Sci. USA* 109(19), 7529–7534 (2012).

17 Ray PD, Yosim A, Fry RC. Incorporating epigenetic data into the risk assessment process for the toxic metals arsenic, cadmium, chromium, lead, and mercury: strategies and challenges. *Front. Genet.* 5, 201 (2014).

18 Marsit CJ. Influence of environmental exposure on human epigenetic regulation. *J. Exp. Biol.* 218(Pt 1), 71–79 (2015).

19 Maccani MA, Avissar-Whiting M, Banister CE, Mcgonnigal B, Padbury JF, Marsit CJ. Maternal cigarette smoking during pregnancy is associated with downregulation of miR-16, miR-21, and miR-146a in the placenta. *Epigenetics* 5(7), 583–589 (2010).

• **This study highlights the downregulation of placental miRNA expression in response to maternal cigarette exposure.**

20 Maccani MA, Knopik VS. Cigarette smoke exposure-associated alterations to non-coding RNA. *Front. Genet.* 3, 53 (2012).

21 Maccani MA, Marsit CJ. Exposure and fetal growth-associated miRNA alterations in the human placenta. *Clin. Epigenetics* 2(2), 401–404 (2011).

•• **The authors of this review propose a novel hypothesis that placental miRNAs may serve as mediators between environmental exposures and fetal growth.**

22 Logan M, Hawkins SM. Role of microRNAs in cancers of the female reproductive tract: insights from recent clinical and experimental discovery studies. *Clin. Sci. (Lond.)* 128(3), 153–180 (2015).

23 Motta V, Angelici L, Nordio F *et al.* Integrative analysis of miRNA and inflammatory gene expression after acute particulate matter exposure. *Toxicol. Sci.* 132(2), 307–316 (2013).

24 Sanders AP, Burris HH, Just AC *et al.* microRNA expression in the cervix during pregnancy is associated with length of gestation. *Epigenetics* 10(3), 221–228 (2015).

• **We observed associations between miRs-21, 30e, 142, 148b, 29b and 223 obtained from cervical samples in pregnancy and subsequent length of gestation in the PROGRESS Cervix Study.**

25 Andrews KW, Savitz DA, Hertz-Picciotto I. Prenatal lead exposure in relation to gestational age and birth weight: a review of epidemiologic studies. *Am. J. Ind. Med.* 26(1), 13–32 (1994).

26 Jelliffe-Pawlowski LL, Miles SQ, Courtney JG, Materna B, Charlton V. Effect of magnitude and timing of maternal pregnancy blood lead (Pb) levels on birth outcomes. *J. Perinatol.* 26(3), 154–162 (2006).

27 Burris HH, Braun JM, Byun HM *et al.* Association between birth weight and DNA methylation of IGF2, glucocorticoid receptor and repetitive elements Line-1 and Alu. *Epigenomics* 5(3), 271–281 (2013).

28 Mordukhovich I, Wright RO, Hu H *et al.* Associations of toenail arsenic, cadmium, mercury, manganese, and lead with blood pressure in the normative aging study. *Environ. Health. Perspect.* 120(1), 98–104 (2012).

29 Crinnion WJ. The CDC Fourth National Report on human exposure to environmental chemicals: what it tells us about our toxic burden and how it assist environmental medicine physicians. *Altern. Med. Rev.* 15(2), 101–109 (2010).

30 Hu H, Rabinowitz M, Smith D. Bone lead as a biological marker in epidemiologic studies of chronic toxicity: conceptual paradigms. *Environ. Health. Perspect.* 106(1), 1–8 (1998).

31 Tellez-Rojo MM, Hernandez-Avila M, Lamadrid-Figueroa H *et al.* Impact of bone lead and bone resorption on plasma and whole blood lead levels during pregnancy. *Am. J. Epidemiol.* 160(7), 668–678 (2004).

32 Geiss GK, Bumgarner RE, Birditt B *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* 26(3), 317–325 (2008).

33 Stern-Ginossar N, Elefant N, Zimmermann A *et al.* Host immune system gene targeting by a viral miRNA. *Science* 317(5836), 376–381 (2007).

34 Cullen BR. Five questions about viruses and microRNAs. *PLoS Pathog.* 6(2), e1000787 (2010).

35 Grundhoff A, Sullivan CS. Virus-encoded microRNAs. *Virology* 411(2), 325–343 (2011).

36 Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. Nanostringnorm: an extensible R package for the pre-processing of nanostring mRNA and miRNA data. *Bioinformatics* 28(11), 1546–1548 (2012).

37 Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* 100(16), 9440–9445 (2003).

38 Centers for Disease Control and Disease. Low level lead exposure harms children: a renewed call for primary prevention (2012). www.cdc.gov

39 Joubert BR, Haberg SE, Nilsen RM *et al.* 450k epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health. Perspect.* 120(10), 1425–1431 (2012).

40 Hernandez-Ochoa I, Karman BN, Flaws JA. The role of the aryl hydrocarbon receptor in the female reproductive system. *Biochem. Pharmacol.* 77(4), 547–559 (2009).

41 Loset M, Mundal SB, Johnson MP *et al.* A transcriptional profile of the decidua in preeclampsia. *Am. J. Obstet. Gynecol.* 204(1), 84.e1–e27 (2011).

42 Morasso MI, Grinberg A, Robinson G, Sargent TD, Mahon KA. Placental failure in mice lacking the homeobox gene Dlx3. *Proc. Natl Acad. Sci. USA* 96(1), 162–167 (1999).

43 Van Nes J, De Graaff W, Lebrin F, Gerhard M, Beck F, Deschamps J. The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development* 133(3), 419–428 (2006).

44 Kong AP, Xiao K, Choi KC *et al.* Associations between microRNA (miR-21, 126, 155 and 221), albuminuria and heavy metals in Hong Kong Chinese adolescents. *Clin. Chim. Acta.* 413(13–14), 1053–1057 (2012).

45 Bollati V, Marinelli B, Apostoli P *et al.* Exposure to metal-rich particulate matter modifies the expression of candidate microRNAs in peripheral blood leukocytes. *Environ. Health. Perspect.* 118(6), 763–768 (2010.

- **A prior study of metal-altered miRNAs in an adult population found that miR-222 and miR-146a were differentially expressed with respect to lead exposure.**

46 Martinez-Pacheco M, Hidalgo-Miranda A, Romero-Cordoba S, Valverde M, Rojas E. mRNA and miRNA expression patterns associated to pathways linked to metal mixture health effects. *Gene* 533(2), 508–514 (2014).

47 Pallocca G, Fabbri M, Sacco MG *et al.* miRNA expression profiling in a human stem cell-based model as a tool for developmental neurotoxicity testing. *Cell. Biol. Toxicol.* 29(4), 239–257 (2013).

48 Johnson CD, Esquela-Kerscher A, Stefani G *et al.* The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer. Res.* 67(16), 7713–7722 (2007).

49 Wang X, Tang S, Le SY *et al.* Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth. *PLoS ONE* 3(7), e2557 (2008).

50 Zheng H, Zhang L, Zhao Y *et al.* Plasma miRNAs as diagnostic and prognostic biomarkers for ovarian cancer. *PLoS ONE* 8(11), e77853 (2013).

51 Lee JW, Choi CH, Choi JJ *et al.* Altered microRNAs expression in cervical carcinomas. *Clin. Cancer. Res.* 14(9), 2535–2542 (2008).

52 Bjorkman L, Lundekvam BF, Laegreid T *et al.* Mercury in human brain, blood, muscle and toenails in relation to exposure: an autopsy study. *Environ. Health.* 6, 30 (2007).

53 Alfthan GV. Toenail mercury concentration as a biomarker of methylmercury exposure. *Biomarkers* 2(4), 233–238 (1997).

54 Woodruff TJ, Zota AR, Schwartz JM. Environmental chemicals in pregnant women in the United States: NHANES 2003–2004. *Environ. Health. Perspect.* 119(6), 878–885 (2011).

55 Xun P, Liu K, Morris JS, Jordan JM, He K. Distributions and determinants of mercury concentrations in toenails among American young adults: the CARDIA Trace Element Study. *Environ. Sci. Pollut. Res. Int.* 20(3), 1423–1430 (2013).

56 Gibb H, Haver C, Kozlov K *et al.* Biomarkers of mercury exposure in two Eastern Ukraine cities. *J. Occup. Environ. Hyg.* 8(4), 187–193 (2011).

57 Knutsen E, Fiskaa T, Ursvik A *et al.* Performance comparison of digital microRNA profiling technologies applied on human breast cancer cell lines. *PLoS ONE* 8(10), e75813 (2013).