




Novel epigenetic link between gestational diabetes mellitus and macrosomia

Brian T Joyce^{*,‡,1} , Huikun Liu^{‡,2}, Leishen Wang², Jun Wang¹, Yinan Zheng¹, Drew Nannini¹, Alex Drong^{3,4}, Stephanie Shiau⁵, Weiqin Li², Junhong Leng², Yun Shen^{6,7}, Ru Gao⁶, Andrea Baccarelli⁴, Gang Hu⁶ & Lifang Hou¹

¹Center for Global Oncology, Institute for Global Health, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

²Tianjin Women's & Children's Health Center, Tianjin, China

³Big Data Institute, University of Oxford, Oxford, UK

⁴Department of Environmental Health Science, Mailman School of Public Health, Columbia University, NY, 10032, USA

⁵Department of Biostatistics & Epidemiology, Rutgers School of Public Health, Piscataway, NJ, 08854, USA

⁶Pennington Biomedical Research Center, Baton Rouge, LA, 70808, USA

⁷Department of Endocrinology & Metabolism, Shanghai Jiao Tong University Affiliated Six People's Hospital, Shanghai, China

*Author for correspondence: Tel.: +1 312 503 5407; b-joyce@northwestern.edu

‡Authors contributed equally

Background & objectives: Examine maternal gestational diabetes mellitus (GDM), macrosomia and DNA methylation in candidate genes *IGF1*, *IGF2*, *H19*, *ARHGRF11*, *MEST*, *NR3C1*, *ADIPOQ* and *RETN*. **Materials & methods:** A total of 1145 children (572 GDM cases and 573 controls) from the Tianjin GDM study, including 177 with macrosomia, had blood DNA collection at median age 5.9 (range: 3.1–10.0). We used logistic regression to screen for associations with GDM and model macrosomia on 37 CpGs, and performed mediation analysis. **Results:** One CpG was associated with macrosomia at false discovery rate (FDR) <0.05 (cg14428359 in *MEST*); two (cg19466922 in *MEST* and cg26263166 in *IGF2*) were associated at $p < 0.05$ but mediated 26 and 13%, respectively. **Conclusion:** *MEST* and *IGF2* were previously identified for potential involvement in fetal growth and development (Trial Registration number: NCT01554358 [ClinicalTrials.gov]).

Lay abstract: Many women who get gestational diabetes during pregnancy go on to give birth to larger (macrosomic) babies. These babies then grow up to have greater risk of being overweight or obese, and all the health concerns this entails. We sought to examine whether epigenetic factors could help explain this link, by examining the blood of some children whose mothers were enrolled in a gestational diabetes study in China. We identified three sites on two different genes as being associated with both gestational diabetes and macrosomia. The way these genes work suggest a mechanism for how they contribute to macrosomia, providing a promising new avenue for future research, early detection and precision prevention (Trial Registration number: NCT01554358 [ClinicalTrials.gov]).

First draft submitted: 23 March 2021; Accepted for publication: 22 July 2021; Published online: 2 August 2021

Keywords: epigenetic epidemiology • epigenetics • macrosomia • metabolic diseases • pediatrics

Fetal macrosomia (defined as birth weight >4000–4500 g, or greater than 90% for gestational age) is estimated to affect 7% of all births in the US as of 2015 [1], with the crude rate of macrosomia having increased nearly 24-times from 1979 to 2010 [2]. Macrosomia is associated with increased rates of maternal [3,4] and fetal [5–7] complications as well as increased risks of obesity and Type 2 diabetes [8,9] and potentially other aging-related diseases as well [10]. Gestational diabetes mellitus (GDM) is a known risk factor for macrosomia [11], with recent estimates suggesting that between 15 and 45% of newborns of mothers with GDM are macrosomic compared with 12% in children of non-GDM mothers [12]. Epigenetic reprogramming in the in utero environment is one suspected driver of this relationship.

Inflammatory cells (including white blood cells) may participate in the etiology of pediatric obesity [13] by sustaining metabolic changes such as those due to GDM exposure in offspring [14], and contributing to obesity [15].

GDM [16–18] can affect the child's blood DNA methylation, which can in turn influence metabolic risks [17–19]. For example, West *et al.* [20] linked blood DNA methylation in children 8–12 years old to prenatal GDM exposure. Talens *et al.* found that methylation at loci sensitive to prenatal exposures was associated with myocardial infarction later in life [19]. Even if these methylation changes do not persist into later life, that can still perturb cardiometabolic phenotypes in childhood (e.g., macrosomia) and in turn influence health later in life. Thus, blood DNA derived from circulating leukocytes provides an ideal tissue for methylomic biomarkers of GDM exposure and potentially its subsequent adverse metabolic impacts.

Most studies to date of DNA methylation associated with macrosomia have focused on Beckwith–Wiedemann syndrome, a congenital disorder characterized by macrosomia and caused by genetic and epigenetic abnormalities in chromosomal region 11p15.5, which includes the imprinting genes *IGF2/H19* and *CDKN1C/KCNQ1OT1* [21]. To date, few studies have examined DNA methylation in children with macrosomia but without Beckwith–Wiedemann syndrome. Two that examined DNA methylation in cord blood found altered methylation in *IGF2/H19* [22] and *ARHGAP11* [23] associated with macrosomia in offspring exposed to in utero hyperglycemia. This study's objectives were to explore associations between both GDM and macrosomia and DNA methylation of eight genes selected for analysis based on a literature review of candidates potentially involved in GDM and obesogenic pathways [12,24].

Materials & methods

Study population & data collection

The full study population and methods have been described previously [25]. Briefly, data were taken from the Tianjin GDM Observational Study, in the 4th-largest city in China; mothers were aged 20–49 and diagnosed with GDM between 2005 and 2009 using the 1998 WHO criteria [26]. All enrolled pregnant women at 26–30 gestational weeks first underwent a 1-h oral glucose test (OGTT) with a 50-g glucose load; women with a glucose reading ≥ 7.8 mmol/l were invited to a 2-h OGTT with a 75 g glucose load and had a result confirming either diabetes (fasting glucose >7 nmol/l or 2-h OGTT >11.1 nmol/l) or impaired glucose tolerance (2-h glucose >7.8 and <11.1 nmol/l) were defined as having GDM.

The present study included 580 women with GDM and their children; an additional 580 mother–child pairs without GDM and frequency were matched on child sex and birth date were also included [27,28]. Of these, only 15 (eight from GDM mothers, seven from the control group) mother–child pairs were excluded from epigenetic analyses during data preprocessing. This left a total of 1145 mothers for analysis, with baseline visit between 2013 and 2016 which included self-administered questionnaires and a physical examination. Blood draws for DNA methylation analysis (using the Illumina 850 K array [Illumina, CA, USA]) were collected from children at median age 5.9 years, range 3.1–10.2.

Covariate measurement

All mothers completed a questionnaire on sociodemographic variables (age, marital status, education, income and occupation), history of GDM (family history of diabetes, coronary heart disease, stroke, cancer and hypertension), medical history (hypertension, diabetes and hypercholesterolemia), prepregnancy and postpartum weight and weight gain during pregnancy, alcohol intake, smoking habits, physical activity, dietary habits (via self-administered food frequency questionnaire) and sleep status. Bodyweight, height, waist and hip circumferences, blood pressure, and heart rate were all measured according to standard protocols. Body fat was measured by a body composition analyzer (InBody J20 [InBody UK, Coalville, UK]) to the nearest 0.1% and triceps, subscapular and suprailiac skinfolds by skinfold caliper to the nearest 0.5 cm. BMI was calculated as the bodyweight in kilograms divided by the square of the height in meters. Children's information was obtained by another questionnaire completed by their mothers and included age, sex, birth date, birth length, birth weight, gestational age (copied from birth certificate), infant feeding patterns, lactation duration, routine activities (physical activities and sleep duration), dietary habits and history of diseases.

White blood cell type proportion estimation & principal components for non-negative probes

To account for cell composition variability we estimated the proportions of CD4⁺ T lymphocytes, CD8⁺ T lymphocytes, B lymphocytes, natural killer cells, monocytes and granulocytes using the method published by Houseman *et al.* [29] To account for experimental batch effects and other technical biases, we derived surrogate variables from intensity data for nonnegative internal control probes using principal components analysis [30]. The

top 16 principal components explain 95.03% of the variation across the non-negative internal control probes and can be adjusted for in multivariate analyses to account for potential batch effects.

Illumina EPIC array methylation data quality control & preprocessing

Infinium MethylationEPIC BeadChip raw data (IDAT files [Illumina]) were loaded into the R package *minfi*. Quality control and data preprocessing were conducted using the R package *ENmix* (Bioconductor, NY, USA) [30] with default parameter settings. In the quality control step, low-quality methylation measurements were identified by detection p -value $< 10^{-6}$ or number of beads < 3 [30]. We excluded 10,861 CpGs with a detection rate $< 95\%$ and 13 samples with a percentage of low-quality methylation measurements with $> 5\%$ or extremely low intensity of bisulfite conversion probes (less than $3 \times$ standard deviation of the intensity across samples below the mean intensity) [30]. We identified no sample outliers based on total intensity across CpGs. Dye bias was corrected using regression on logarithm of internal control probes (RELIC), which utilizes the intensity values of paired internal control probes that monitor the two-color channels [31]. We then separately quantile-normalized M or U intensities for Infinium I or II probes, respectively. At last, low-quality methylation values (detection $p < 10^{-6}$ or number of beads < 3) and extreme β -value outliers across samples (defined by Tukey's method) were set as missing. The final clean methylation working dataset contains 856,146 CpG probes and 1145 samples.

Statistical methods

First, we conducted a descriptive analysis comparing all variables of interest between macrosomia cases and controls; continuous variables were compared using a t -test for normally distributed variables and a Wilcoxon rank-sum test for non-normally distributed variables (i.e., skinfolds), and categorical variables using a chi-squared test. Next, based on a literature review we identified eight candidate genes with a total of 488 CpG sites available in our dataset: *ARHGEF11* (52 CpG sites), *MEST* (89 CpGs), *NR3C1* (86 CpGs), *ADIPOQ* (13 CpGs), *RETN* (11 CpGs), *IGF1* (27 CpGs), *IGF2* (151 CpGs) and *H19* (59 CpGs). For the single-CpG analysis we initially screened 488 CpG sites in a logistic regression model for associations with methylation as the outcome and GDM as the independent variable of interest at $p < 0.05$; 37 of these were significant and tested in a second logistic regression model with macrosomia as the outcome (and methylation as the exposure). Both of these models adjusted for maternal height, age, smoking status, passive smoking exposure, prepregnancy BMI, pregnancy weight gain, parity, pregnancy hypertension, child sex and gestational age at delivery, child age and weight for age z -score at methylation measurement, blood cell type proportions and 10 principal components representing processing batch effects ($> 90\%$ of variance). We used FDR adjustment to correct for multiple testing in these final macrosomia models, with FDR-adjusted $q < 0.05$ considered statistically significant. We also conducted a sensitivity analysis exploring the effect of replacing macrosomia as our outcome with birth weight as a continuous variable in our single-CpG analysis. Next, we used two methods (the Fisher method and the Empirical Brown method) [32] to combine the single-CpG raw p -values and conduct regional analyses of the promoter and body regions of each of our target genes, testing their associations with macrosomia. Finally, we conducted separate mediation analyses of our three most significant hits using the R package 'mediation' [33]. Specifically, two models are fit: one modeling the effects of GDM on DNA methylation, and a second one jointly modeling the effects of GDM (directly) and DNA methylation (indirectly) on macrosomia. Using Monte Carlo simulations, a mediation proportion is estimated, indicating how much of the effects of GDM on macrosomia could be explained by the indirect path in which GDM drives a change in DNA methylation, which then affects macrosomia. This effect is estimated in a model including all confounding variables that affect both the mediating variable (methylation) and the outcome (macrosomia).

Results

Participant characteristics by maternal GDM status have been previously reported [34]. Table 1 shows our participant characteristics and descriptive analysis by child macrosomia status. Briefly, we confirmed the association between GDM and macrosomia. Of 573 children born to mothers without GDM, 63 met the criteria for macrosomia and 510 did not; of 572 children born to mothers with GDM, 114 met the criteria for macrosomia and 458 did not ($p \leq 0.001$). Mothers with greater pregnancy weight gain and prepregnancy BMI also tended to give birth to macrosomic children ($p < 0.001$ for both). Additionally, female children tended to be macrosomic ($p = 0.03$) as did those with an older gestational age at delivery ($p < 0.001$); median birth weight was 765 g heavier for macrosomic children ($p < 0.001$). Finally, macrosomic children tended to have greater BMI z -score ($p < 0.001$), weight for age

Table 1. Participant characteristics by child macrosomia status.

	Non-macrosomia (n = 968)	Macrosomia (n = 177)	p-value
Maternal characteristics			
GDM status			<0.001
– No	510 (51.26%)	63 (35.59%)	
– Yes	458 (48.74%)	114 (64.41%)	
Total parity			0.17
– 1	952 (98.35%)	177 (100%)	
– 2	16 (1.65%)	0 (0%)	
Smoking status			0.09
– Never	925 (95.56%)	166 (93.79%)	
– Former	15 (1.55%)	7 (3.95%)	
– Current	28 (2.89%)	4 (2.26%)	
Pregnancy-induced hypertension			0.56
– No	931 (96.18%)	168 (94.92%)	
– Yes	37 (3.82%)	9 (5.08%)	
Maternal height (cm)	117.5 (93.5–155.4)	118.5 (97.0–144.8)	0.17
Age at delivery	29.58 (22.6–49.4)	29.43 (23.5–41.2)	0.35
Pregnancy weight gain (kg)	15 (4–45)	20 (10–43)	<0.001
Prepregnancy BMI (kg/m ²)	21.37 (15.4–45.0)	23.28 (17.7–33.2)	<0.001
Child characteristics at birth			
Child sex			0.03
– Male	492 (50.83%)	63 (35.59%)	
– Female	476 (49.17%)	114 (64.41%)	
Child's birth weight (g)	3385 (1500–3980)	4150 (4000–5300)	<0.001
Gestational age at delivery (weeks)	39 (28–43)	40 (35–43)	<0.001
Child characteristics at methylation measurement			
Age (years, mean [SD])	5.89 (1.26)	5.81 (1.20)	0.437
Z-score for BMI (mean [SD])	0.11 (1.30)	0.60 (1.43)	<0.001
Weight for age Z-score (mean [SD])	0.50 (1.23)	1.01 (1.19)	<0.001
Bodyfat (% , mean [SD])	19.77 (7.63)	21.33 (8.61)	0.015
Skinfolds at triceps (mm, median [IQR])	11.00 [9.00, 15.00]	11.00 [9.50, 17.00]	0.222
Skinfolds at subscapular (mm, median [IQR])	6.00 [5.00, 8.00]	6.50 [5.50, 9.00]	0.07
Skinfolds at suprailiac (mm, median [IQR])	8.50 [6.00, 12.50]	9.00 [7.00, 14.00]	0.034

GDM: Gestational diabetes mellitus; IQR: Interquartile range; SD: Standard deviation.

z-score ($p < 0.001$), body fat percentage ($p = 0.02$) and suprailiac skinfold ($p = 0.03$) at the time of methylation measurement. We identified no other significant differences in participant characteristics by macrosomia status.

Table 2 shows the most significant CpG sites in our single-CpG analysis. One CpG site on *MEST* was significantly associated with macrosomia at $q < 0.05$; an additional two CpGs (one on each of *MEST* and *IGF2*) were associated with macrosomia at $p < 0.05$ but $q > 0.05$. Replacing macrosomia with birth weight in our single-CpG analysis did not substantially change our results (data available upon request). Table 3 shows the results of our regional analyses. Using the Fisher method, we identified differential methylation in the promoter region of *IGF2* ($p = 0.04$), and marginally differential methylation in the body region of *MEST* ($p = 0.06$). Of note, one of the hits in our single-CpG analysis was located in the body of *MEST* (cg19466922). Figure 1 shows the results of the mediation analyses of our three significant single-CpG hits. Interestingly, both cg19466922 and cg26263166 were significant mediators of the GDM-Macrosomia association ($p = 0.02$ for both), mediating 26% and 13% of the relationship, respectively. Our third hit, cg14428359, was not a significant mediator ($p = 0.68$).

Discussion

In this study we found limited evidence of an epigenetic basis for part of the well-documented association between GDM and macrosomia. While little research has been done examining epigenetics in association with

Table 2. Top CpGs associated with gestational diabetes mellitus at $p < 0.05$ and their associations with macrosomia.

CpG information			Macrosomia n = 37 CpGs				GDM n = 488 CpGs	
CpG	Gene	Region	β	SE	p-value	q	β	SE
cg14428359	<i>MEST</i>	3'UTR	-0.012	0.003	<0.001 [†]	<0.001 [‡]	-0.006	0.003
cg19466922	<i>MEST</i>	Body	-0.010	0.004	0.006 [†]	0.089	-0.017	0.004
cg26263166	<i>IGF2</i>	3'UTR	0.003	0.001	0.017 [†]	0.169	0.003	0.001
cg05345888	<i>ARHGEF11</i>	Body	0.001	0.001	0.06	0.324	0.003	0.001
cg15886040	<i>H19</i>	TSS1500	0.006	0.003	0.064	0.324	0.008	0.004
cg07492680	<i>ARHGEF11</i>	Body	-0.004	0.002	0.064	0.324	-0.006	0.002
cg21585493	<i>ARHGEF11</i>	TSS200	0.001	0.001	0.09	0.341	0.002	0.001
cg11322849	<i>INS-IGF2</i>	TSS1500	-0.004	0.002	0.09	0.341	-0.006	0.003
cg07589972	<i>NR3C1</i>	TSS1500	0.002	0.001	0.147	0.492	-0.004	0.001
cg01785096	<i>ARHGEF11</i>	Body	0.003	0.003	0.187	0.563	0.007	0.003
cg18146873	<i>NR3C1</i>	1st Exon	0.000	0.000	0.21	0.576	0.001	0.000
cg15922305	<i>H19</i>	TSS200	-0.004	0.004	0.298	0.685	0.01	0.004
cg04811924	<i>ARHGEF11</i>	TSS200	0.001	0.001	0.33	0.685	0.001	0.001
cg20148994	<i>MEST</i>	TSS1500	0.001	0.001	0.375	0.685	0.004	0.002
cg25587594	<i>IGF1</i>	Body	0.001	0.001	0.377	0.685	0.002	0.001
cg12381613	<i>ARHGEF11</i>	Body	-0.002	0.002	0.39	0.685	-0.007	0.003
cg12888360	<i>NR3C1</i>	Body	0.001	0.001	0.444	0.685	0.004	0.002
cg15495409	<i>ADIPOQ</i>	3'UTR	0.000	0.000	0.468	0.685	0.001	0.001
cg22648984	<i>ARHGEF11</i>	Body	-0.001	0.002	0.474	0.685	-0.005	0.002
cg21167159	<i>H19</i>	Body	0.001	0.002	0.496	0.685	0.005	0.002
cg22424892	<i>H19</i>	Body	0.000	0.000	0.539	0.685	-0.001	0.000
cg25852472	<i>H19</i>	Body	0.002	0.003	0.543	0.685	-0.007	0.003
cg00629244	<i>NR3C1</i>	TSS200	0.000	0.000	0.554	0.685	0.000	0.000
cg19457823	<i>NR3C1</i>	Body	-0.002	0.003	0.555	0.685	-0.008	0.003
cg16224829	<i>NR3C1</i>	5'UTR	-0.001	0.002	0.569	0.685	-0.006	0.002
cg20101237	<i>ARHGEF11</i>	Body	-0.001	0.002	0.661	0.712	-0.005	0.003
cg02166532	<i>INS-IGF2</i>	Body	-0.001	0.001	0.672	0.712	-0.003	0.001
cg13791131	<i>IGF2AS</i>	Body	-0.001	0.001	0.675	0.712	-0.004	0.001
cg07342901	<i>H19</i>	Body	0.000	0.001	0.726	0.712	-0.003	0.001
cg16641322	<i>INS-IGF2</i>	Body	0.000	0.001	0.75	0.712	-0.003	0.001
cg15508379	<i>IGF2AS</i>	Body	0.000	0.001	0.751	0.712	0.002	0.001
cg00286878	<i>MEST</i>	TSS1500	0.001	0.002	0.756	0.712	0.007	0.002
cg12347392	<i>MEST</i>	TSS1500	-0.001	0.006	0.805	0.735	0.012	0.006
cg11607484	<i>ARHGEF11</i>	3'UTR	0.000	0.001	0.829	0.735	0.002	0.001
cg08162473	<i>INS-IGF2</i>	Body	0.000	0.000	0.937	0.807	0.000	0.000
cg13993218	<i>INS-IGF2</i>	Body	0.00E + 00	0.003	0.979	0.81	-0.008	0.004
cg07409197	<i>MEST</i>	TSS1500	0.00E + 00	0.003	0.995	0.81	0.006	0.003

[†] Significant at $p < 0.05$.

[‡] Significant at $q < 0.05$.

All models adjusted for maternal height, age, smoking status, passive smoking exposure, prepregnancy BMI, pregnancy weight gain, parity, pregnancy hypertension; child sex and gestational age at delivery; child age and weight for age z-score at methylation measurement; blood cell type proportions; and ten principal components representing processing batch effects (>90% of variance).

GDM: Gestational diabetes mellitus; SE: Standard error.

GDM and macrosomia, a recent meta-analysis found limited evidence for associations between cord blood DNA methylation and maternal GDM [35]. Our single-CpG analysis found three CpGs in the genes *MEST* and *IGF2* to be associated with both GDM and macrosomia, limited evidence for regional associations with GDM and macrosomia in the promoters of *IGF2* and *ARHGEF11* as well as the body of *MEST*, and moderate mediation effects for two of the CpGs identified in our single-CpG analysis. These findings indicate that *MEST* and *IGF2* may be particularly important for maternal GDM phenotypes to influence macrosomia and potentially similar

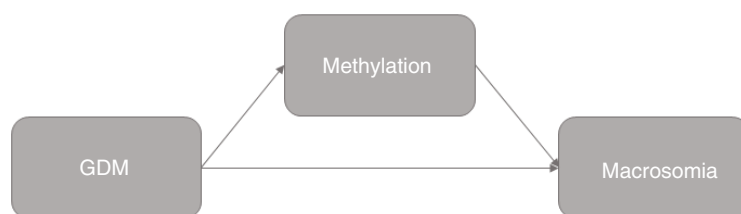
Table 3. Gene regional methylation associated with gestational diabetes mellitus and macrosomia.

Gene/region	CpGs	P _{EBM}	P _{Fisher}
IGF2_promoter	133	0.16	0.04 [†]
MEST.Body	8	0.08	0.06
ARHGEF11_promoter	20	0.13	0.09
IGF2.Body	18	0.21	0.12
ARHGEF11.Body	32	0.38	0.34
IGF1.Body	12	0.38	0.38
ADIPOQ.Body	2	0.45	0.45
MEST_promoter	81	0.48	0.53
ADIPOQ_promoter	11	0.71	0.75
NR3C1.Body	30	0.69	0.83
NR3C1_promoter	56	0.73	0.85
RETN_promoter	11	0.64	0.86
H19.Body	12	0.82	0.88
H19_promoter	47	0.78	0.96
IGF1_promoter	15	0.96	0.97

[†]Significant at p < 0.05.

All models adjusted for maternal height, age, smoking status, passive smoking exposure, prepregnancy BMI, pregnancy weight gain, parity, pregnancy hypertension; child sex and gestational age at delivery; child age and weight for age z-score at methylation measurement; blood cell type proportions; and ten principal components representing processing batch effects (>90% of variance).

EBM: Empirical Brown method; GDM: Gestational diabetes mellitus.



CpG	mediated (%)	p-value
cg14428359	0.9	0.676
cg19466922	26.4	0.018
cg26263166	12.7	0.016

Figure 1. Results of mediation analysis of three CpGs associated with macrosomia and GDM at p < 0.05. All models adjusted for maternal height, age, smoking status, passive smoking exposure, prepregnancy BMI, pregnancy weight gain, parity, pregnancy hypertension, child sex and gestational age at delivery, child age and weight for age z-score at methylation measurement, blood cell type proportions and ten principal components representing processing batch effects (>90% of variance). GDM: Gestational diabetes mellitus.

cardiometabolic outcomes (e.g., childhood obesity) later in life. Additional research is needed to further characterize this association (e.g., in terms of maternal hyperglycemia or glycemic control, when in pregnancy potential epigenetic reprogramming may occur), and assess potential interventions to ‘interrupt’ the epigenetic mechanism and prevent macrosomia.

IGF2 is an imprinted gene that manufactures the protein IGF-2 and is heavily involved in fetal growth and development. We found hypermethylation of one locus in the 3’UTR region of *IGF2* may be associated with both GDM and macrosomia, and an association with *IGF2* promoter methylation in our regional analysis. However, the link between methylation in the 3’UTR gene region and gene expression is unclear, thus without gene expression data we cannot posit a direct mechanistic link. Other studies have found changes in *IGF2* expression and/or methylation in association with GDM and macrosomia. A previous study by Su *et al.* in another Chinese population found increased expression of *IGF2* in placental tissue and cord blood of fetuses whose mothers had been diagnosed with GDM, and in children born with macrosomia [22]. Their results also suggested a joint action between *IGF2* and the gene *H19* also located in the IC1 complex; however, our analysis suggests that DNA methylation of *H19* is not involved. Another study by Zhang *et al.* found reduced *IGF2* expression in children born small for gestational age but no association with *IGF2* methylation [36]. Hoyo *et al.* likewise found an association between lower *IGF2* methylation in cord blood and elevated IGF2 plasma protein concentrations, which were

in turn associated with higher birth weight [37]. Our study found one CpG in the 3'UTR region of *IGF2* to be associated with macrosomia; until a relationship to gene expression can be established for this region (or CpG), the extent to which our finding reflects those of prior studies will be unclear. Nonetheless our findings add to the evidence that exposure to GDM *in utero* may alter *IGF2* methylation, and potentially increase *IGF2* expression and therefore fetal growth, ultimately facilitating increased birth weight. Data are needed to compare our DNA methylation results with *IGF2* expression, thus confirming the functional impact of our associations; however, if consistent with the prior literature our findings in childhood methylation suggest that the previously identified effects of GDM (and/or related variables such as maternal hyperglycemia) on *IGF2* methylation may persist after birth.

MEST is another imprinted gene that encodes a member of the alpha/beta hydrolase superfamily, which may play a role in fetal development. We found that hypomethylation of two loci, one in the 3'UTR region and one in the body of *MEST*, was associated with both GDM and macrosomia. A prior study by Kappil *et al.* found increased *MEST* expression in placental tissue of children born small for gestational age [38], while Deysenroth *et al.* found a differentially methylated region in *MEST* that was common to children in both small and large for gestational age groups [39]. Gonzalez-Nahm *et al.* found a similar, hypermethylated differentially methylated region in *MEST* associated with lower weight [40]. El Hajj *et al.* also found hypomethylation of *MEST* in newborns of mothers with GDM [41]. As with *IGF2*, gene expression data are needed to confirm the mechanistic nature of our findings, but as one of our identified loci is located in the body of *MEST*, the inverse association identified is generally consistent with the literature. Functional validation of our findings with gene expression data is necessary to confirm a mechanistic hypothesis, but if confirmed the timing of our methylation data collection suggests that there are persistent effects of GDM (and/or related variables such as maternal hyperglycemia/glycemic control) on the epigenome into childhood. As this population matures we can also explore the possibility that these mechanisms translate into elevated BMI in adolescence and/or adulthood, and potentially related phenotypes such as obesity and diabetes.

Also of note is our mediation analysis, which found that the loci cg26263166 (in the 3'UTR of *IGF2*) and cg19466922 (in the body of *MEST*), mediated 13 and 26%, respectively, of the association between GDM and macrosomia, which has been well-documented elsewhere in the literature. However in light of our weak results after FDR adjustment, and the timing of blood sample collection in our study, this finding should be interpreted with caution until it can be confirmed in additional larger populations. One of the assumptions upon which this analysis rests is that DNA methylation remains correlated between birth and the age of DNA collection (median 5.9 years). To our knowledge, few studies have repeatedly examined DNA methylation in children; however, genetic influences on DNA methylation have been shown to be stable throughout the life course [42]. If environmental influences remain relatively stable as well, then it is possible that maternal epigenetic influences can still be detected years after birth. Further epigenetic studies in children are warranted to test this possibility. Our mediation findings may also reflect methylation at other loci correlated with those we identified. Thus, we cannot rule out the possibility that these results represent other epigenetic effects, or the results of other childhood exposures such as diet. Additional, better powered research using cord blood DNA and/or DNA collected at or soon after birth is warranted to confirm our mediation findings, but if successful this would point to a potential mechanism for the effects of GDM and related variables on growth and a potential point of intervention to prevent macrosomia and/or other downstream effects (e.g., elevated BMI).

This study has many strengths: a unique, well-characterized population, a prospective design and rigorous confounder control. However, this study is also subject to limitations. As described above, the lack of gene expression data and timing of the DNA methylation measurement relative to our phenotypes of interest limit our ability to make causal inferences on the basis of our findings. The lack of DNA methylation data taken from cord blood (or at birth) is a limitation of this study, as unmeasured confounders occurring during childhood may influence our results. Additionally, this was a Chinese population, which may be subject to genetic and/or environmental effects that prevent our findings from being generalizable to other populations. This study was also under powered to detect epigenome-wide effects due to the relatively small number of macrosomia cases ($n = 177$ across GDM and non-GDM groups). This prevented an epigenome-wide study of DNA methylation markers of macrosomia, and reduced our ability to detect effects even in our candidate gene study after multiple testing and robust covariate adjustment. Future research with additional macrosomia cases will help address these limitations, and may yield more findings due to improved statistical power. Another limitation is the lack of data on maternal hyperglycemia or glycemic control during pregnancy. This prevented us from doing a more in-depth study of

important GDM-related variables which may be the primary drivers of the epigenetic relationships observed in this study. Future research should examine these possibilities. Additionally, researchers should explore other potential unmeasured confounders from this study (e.g., diet), and seek to confirm these findings using DNA methylation measured in cord blood and/or closer to birth. Finally, particularly given our findings in two untranslated regions, future research using gene expression data is necessary to place our results in their proper mechanistic context.

Conclusion

Nonetheless, this unique cohort enabled us to identify DNA methylation loci that may play a role in the relationship between GDM (or related variables) and macrosomia. These loci were associated with GDM despite being measured during childhood, and if confirmed in future studies (in particular, using blood DNA from closer to birth, and using gene expression data to confirm functional effects) this may point to mechanisms for the development of macrosomia and potentially other obesity-related phenotypes. Further investigation with more macrosomia cases, gene expression data, DNA methylation collected closer to birth, more detailed maternal data, and phenotypes collected later in life will help place our findings in an appropriate context and could potentially facilitate the development of early detection, prevention, and treatment initiatives.

Future perspective

Traditional interventions remain the best choice for controlling gestational diabetes, and may help mitigate risk of macrosomia in children. However the persistently strong associations between GDM and macrosomia suggest a form of biological transmission. This study provides modest evidence of that transmission being mediated in part by DNA methylation changes. These changes occur on genes known to be associated with fetal growth and development, further strengthening the case for a mechanistic association. Nonetheless these associations are difficult to measure and subject to a great many potentially confounding factors. Further studies to completely explore this topic are necessary to strengthen the case for an epigenetic basis for GDM-macrosomia transmission. However, if validated, these findings could lead to useful early detection tools and/or therapeutic and preventive targets in the epigenome for macrosomia in particular but potentially other related metabolic conditions as well such as obesity and type two diabetes.

Summary points

- Fetal macrosomia is an important birth complication associated with later-life cardiometabolic diseases, as well as in utero exposure to gestational diabetes mellitus (GDM).
- DNA methylation of white blood cells is a plausible mechanism through which GDM exposure may influence macrosomia development.
- For the first time, this study examined DNA methylation changes in childhood associated with both GDM and macrosomia, to assess whether epigenetic mechanisms potentially involved in this process can still be detected years after birth.
- We found that one CpG (cg14428359) in *MEST* was hypomethylated among macrosomic children; two others (cg19466922 in *MEST* and cg26263166 in *IGF2*) were also associated before adjustment for multiple testing.
- Of these three CpGs, one (cg19466922) is located in the gene body of *MEST* and thus may be associated with gene expression. The other two are located in untranslated regions; gene expression studies are needed to assess potential functional effects.
- Our regional analysis identified associations between macrosomia and the promoter of *IGF2*, suggesting an overall functional effect for this gene despite the lack of significant single-CpG findings. The body of *MEST* was also marginally associated.
- Our mediation analysis of these three CpGs identified two significant mediators—cg19466922 and cg26263166 in *MEST* and *IGF2*, respectively. They were found to mediate 26.4 and 12.7% of the association between GDM and macrosomia, respectively, in separate models.
- If validated, these findings suggest potential utility for methylomic biomarkers in the early detection and prevention of macrosomia, and potentially related metabolic diseases, particularly in cases of *in utero* GDM exposure.

Financial & competing interests disclosure

This study was supported by grants from the European Foundation for the Study of Diabetes (EFSD)/Chinese Diabetes Society (CDS)/Lilly programme for Collaborative Research between China and Europe, and the National Institute of Diabetes and Digestive

and Kidney Diseases (R01DK100790). This project also used core facilities supported by the NORC Center Grant P30 DK072476, and the COBRE Center Grant P30 GM118430. G Hu was partly supported by the grant from the National Institute of General Medical Sciences (U54GM104940). A Baccarelli was partially supported by a grant from the National Institute of Environmental Health Studies (P30ES009089). L Hou was partially supported by the American Heart Association Children's Strategically Focused Research Network. BT Joyce was partially supported by an American Heart Association Career Development Award (Grant number 19CDA34630050). H Liu was supported by the Natural Science Foundation of Tianjin, China (19JCYBJC28000). 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

This study was approved by the Human Subjects Committee of the Tianjin Women's and Children's Health Center, and all participants provided written informed consent. This original trial is registered with ClinicalTrials.gov, number NCT01554358.

Data sharing statement

The authors certify that this manuscript reports the secondary analysis of clinical trial data that have been shared with them, and that the use of this shared data is in accordance with the terms (if any) agreed upon their receipt. The source of this data is: Tianjin Gestational Diabetes Prevention Program (NCT01554358).

References

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

- Martin JA, Hamilton BE, Osterman MJ, Driscoll AK, Mathews TJ. Births: final data for 2015. *Natl Vital. Stat. Rep.* 66(1), 1 (2017).
- Lavery JA, Friedman AM, Keyes KM, Wright JD, Ananth CV. Gestational diabetes in the United States: temporal changes in prevalence rates between 1979 and 2010. *BJOG* 124(5), 804–813 (2017).
- Türkmen S, Johansson S, Dahmoun M. Foetal macrosomia and foetal–maternal outcomes at birth. *J. Pregnancy* 2018, 4790136 (2018).
- Crosby DA, Ahmed S, Razley A, Morrison JJ. Obstetric and neonatal characteristics of pregnancy and delivery for infant birthweight ≥ 5.0 kg. *J. Matern. Fetal Neonatal. Med.* 30(24), 2961–2965 (2017).
- Mendez-Figueroa H, Truong VTT, Pedroza C, Chauhan SP. Large for gestational age infants and adverse outcomes among uncomplicated pregnancies at term. *Am. J. Perinatol.* 34(7), 655–662 (2017).
- Yamamoto JM, Kallas-Koeman MM, Butalia S, Lodha AK, Donovan LE. Large-for-gestational-age (LGA) neonate predicts a 2.5-fold increased odds of neonatal hypoglycaemia in women with type 1 diabetes. *Diab. Metab. Res. Rev.* 33(1), doi: 10.1002/dmrr.2824 (2017) (Epub ahead of print).
- Cordero L, Paetow P, Landon MB, Nankervis CA. Neonatal outcomes of macrosomic infants of diabetic and non-diabetic mothers. *J. Neonatal Perinat. Med.* 8(2), 105–112 (2015).
- Johnsson IW, Haglund B, Ahlsson F, Gustafsson J. A high birth weight is associated with increased risk of type 2 diabetes and obesity. *Pediatric Obesity* 10(2), 77–83 (2015).
- Yu Z, Han S, Zhu G *et al.* Birth weight and subsequent risk of obesity: a systematic review and meta-analysis. *Obesity Rev.* 12(7), 525–542 (2011).
- Vaiserman AM. Birth weight predicts aging trajectory: a hypothesis. *Mech. Ageing Dev.* 173, 61–70 (2018).
- Collier A, Abraham EC, Armstrong J, Godwin J, Monteath K, Lindsay R. Reported prevalence of gestational diabetes in Scotland: the relationship with obesity, age, socioeconomic status, smoking and macrosomia, and how many are we missing? *J. Diabetes Investig.* 8(2), 161–167 (2017).
- Kc K, Shakya S, Zhang H. Gestational diabetes mellitus and macrosomia: a literature review. *Ann. Nutr. Metab.* 66(Suppl. 2), 14–20 (2015).
- Inzaugarat ME, Billordo LA, Vodanovich F *et al.* Alterations in innate and adaptive immune leukocytes are involved in paediatric obesity. *Pediatr. Obes.* 9(5), 381–390 (2013).
- Frias AE, Grove KL. Obesity: a transgenerational problem linked to nutrition during pregnancy. *Semin. Reprod. Med.* 30(6), 472–478 (2012).
- **Review articles discussing the link between GDM (gestational diabetes mellitus) and macrosomia, its consequences, and potential mechanisms.**
- Hotamisligil GS. Inflammation and metabolic disorders. *Nature* 444(7121), 860–867 (2006).
- 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).

17. Heijmans BT, Tobi EW, Stein AD *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA* 105(44), 17046–17049 (2008).
18. 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).
19. Talens RP, Jukema JW, Trompet S *et al.* Hypermethylation at loci sensitive to the prenatal environment is associated with increased incidence of myocardial infarction. *Int. J. Epidemiol.* 41(1), 106–115 (2012).
20. West NA, Kechris K, Dabelea D. Exposure to maternal diabetes in utero and DNA methylation patterns in the offspring. *Immunometabol.* 1, 1–9 (2013).
21. Soejima H, Higashimoto K. Epigenetic and genetic alterations of the imprinting disorder Beckwith-Wiedemann syndrome and related disorders. *J. Hum. Genet.* 58(7), 402–409 (2013).
22. Su R, Wang C, Feng H *et al.* Alteration in expression and methylation of IGF2/H19 in placenta and umbilical cord blood are associated with macrosomia exposed to intrauterine hyperglycemia. *PLoS ONE* 11(2), e0148399 (2016).
23. Yan J, Su R, Zhang W *et al.* Epigenetic alteration of Rho guanine nucleotide exchange factor 11 (ARHGEF11) in cord blood samples in macrosomia exposed to intrauterine hyperglycemia. *J. Maternal Fetal Neonat. Med.* 34(3), 422–431 (2021).
24. Chiavaroli V, Derraik JG, Hofman PL, Cutfield WS. Born large for gestational age: bigger is not always better. *J. Pediatr.* 170, 307–311 (2016).
25. Hu G, Tian H, Zhang F *et al.* Tianjin Gestational Diabetes Mellitus Prevention Program: study design, methods, and 1-year interim report on the feasibility of lifestyle intervention program. *Diab. Res. Clin. Pract.* 98(3), 508–517 (2012).
26. Alberti KGMM, Zimmer PF. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. *Diab. Med.* 15(7), 539–553 (1998).
27. Wang J, Wang L, Liu H *et al.* Maternal gestational diabetes and different indicators of childhood obesity – a large study. *Endocr. Connect.* 7(12), 1464–1471 (2018).
28. Liang Z, Liu H, Wang L *et al.* Maternal MTNR1B genotype, maternal gestational weight gain, and childhood obesity. *Am. J. Clin. Nutr.* 111(2), 360–368 (2020).
29. Houseman EA, Accomando WP, Koestler DC *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform.* 13, 86 (2012).
30. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucleic Acids Res.* 44(3), e20 (2016).
31. Xu Z, Langie SA, De Boever P, Taylor JA, Niu L. RELIC: a novel dye-bias correction method for Illumina Methylation BeadChip. *BMC Genomics* 18(1), 4 (2017).
32. Poole W, Gibbs DL, Shmulevich I, Bernard B, Knijnenburg TA. Combining dependent P-values with an empirical adaptation of Brown's method. *Bioinform.* 32(17), i430–i436 (2016).
33. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: r package for causal mediation analysis. 59(5), doi:10.18637/jss.v059.i05 (2014) (Epub ahead of print).
34. Shiau S, Wang L, Liu H *et al.* Prenatal gestational diabetes mellitus exposure and accelerated offspring DNA methylation age in early childhood. *Epigenetics* 16(2), 186–195 (2021).
35. Howe CG, Cox B, Fore R *et al.* Maternal gestational diabetes mellitus and newborn DNA methylation: findings from the pregnancy and childhood epigenetics consortium. *Diabetes Care* 43(1), 98–105 (2020).
- **Recent meta-analysis finding associations between GDM and DNA methylation in cord blood.**
36. Zhang S, Zhai G, Wang J, Shi W, Zhang R, Chen C. IGF-II expression and methylation in small for gestational age infants. *J. Pediatr. Endocrinol. Metabol.* 28(5–6), 613–618 (2015).
37. Hoyo C, Fortner K, Murtha AP *et al.* Association of cord blood methylation fractions at imprinted insulin-like growth factor 2 (IGF2), plasma IGF2, and birth weight. *Cancer Causes Control* 23(4), 635–645 (2012).
38. Kappil MA, Green BB, Armstrong DA *et al.* Placental expression profile of imprinted genes impacts birth weight. *Epigenetics* 10(9), 842–849 (2015).
39. Deyssenroth MA, Marsit CJ, Chen J, Lambertini L. In-depth characterization of the placental imprintome reveals novel differentially methylated regions across birth weight categories. *Epigenetics* 15(1–2), 47–60 (2019).
40. Gonzalez-Nahm S, Mendez MA, Benjamin-Neelon SE *et al.* DNA methylation of imprinted genes at birth is associated with child weight status at birth, 1 year, and 3 years. *Clin. Epigenetics* 10, 90 (2018).
41. El Hajj N, Plushch G, Schneider E *et al.* Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes* 62(4), 1320–1328 (2013).
42. Gaunt TR, Shihab HA, Hemani G *et al.* Systematic identification of genetic influences on methylation across the human life course. *Genome Biol.* 17, 61 (2016).