



Howe, C., Cox, B., Fore, R., Jungius, J. E., Kvist, T., Miles, H., Salas, LA., Rifas-Shiman, S. L., Starling, A., Yousefi, P. D., Ladd-Acosta, C., Baccarelli, AA., Binder, E., Chatzi, V. L., Czamara, D., Dabelea, D., DeMeo, DL., Ghantous, A., Herceg, Z., ... The Pregnancy and Childhood Epigenetics (PACE) consortium (2019). Maternal Gestational Diabetes and Newborn DNA Methylation: Findings from the Pregnancy and Childhood Epigenetics Consortium. *Diabetes Care*, Article dc190524. Advance online publication. <https://doi.org/10.2337/dc19-0524>

Peer reviewed version

Link to published version (if available):  
[10.2337/dc19-0524](https://doi.org/10.2337/dc19-0524)

[Link to publication record on the Bristol Research Portal](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via American Diabetes Association at <https://care.diabetesjournals.org/content/early/2019/10/09/dc19-0524>. Please refer to any applicable terms of use of the publisher.

## University of Bristol – Bristol Research Portal

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>

**Title:** Maternal Gestational Diabetes and Newborn DNA Methylation: Findings from the Pregnancy and Childhood Epigenetics Consortium

**Short Running Title:** Gestational Diabetes and Newborn Methylation

**Authors:** Caitlin G Howe<sup>1,\*</sup>, Bianca Cox<sup>2</sup>, Ruby Fore<sup>3</sup>, James Jungius<sup>4,5</sup>, Tuomas Kvist<sup>6</sup>, Samantha Lent<sup>3</sup>, Harriet E Miles<sup>4,5</sup>, Lucas A Salas<sup>7,8,9</sup>, Sheryl Rifas-Shiman<sup>6</sup>, Anne P Starling<sup>10,11</sup>, Paul Yousefi<sup>3,4</sup>, Christine Ladd-Acosta<sup>12</sup>, Andrea Baccarelli<sup>13</sup>, Elisabeth B Binder<sup>14,15</sup>, Vaia Lida Chatzi<sup>1,16,17</sup>, Darina Czamara<sup>14</sup>, Dana Dabelea<sup>10,11,18</sup>, Dawn L DeMeo<sup>19</sup>, Akram Ghantous<sup>20</sup>, Zdenko Herceg<sup>20</sup>, Eero Kajantie<sup>21,22,23,24</sup>, Jari MT Lahti<sup>5</sup>, Debbie A Lawlor<sup>4,4,25</sup>, Augusto Litonjua<sup>19</sup>, Tim S Nawrot<sup>2,26</sup>, Ellen A Nohr<sup>27</sup>, Emily Oken<sup>6</sup>, Costanza Pizzi<sup>28</sup>, Michelle Plusquin<sup>2</sup>, Katri Räikkönen<sup>5</sup>, Caroline L Relton<sup>4,4,25</sup>, Gemma C Sharp<sup>3</sup>, Thorkild IA Sørensen<sup>4,29,30</sup>, Jordi Sunyer<sup>8,9,31</sup>, Martine Vrijheid<sup>8,9,31</sup>, Weiming Zhang<sup>11,32</sup>, Marie-France Hivert<sup>3, 33,‡</sup>, Carrie V Breton<sup>1,‡</sup>

‡These authors contributed equally to this work

\*Corresponding author: Dr. Caitlin Howe; 2001 N Soto St, Los Angeles, CA, 90032; caitlin.howe@usc.edu

**Author Affiliations:**

<sup>1</sup>Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

<sup>2</sup>Center for Environmental Sciences, Hasselt University, Hasselt, Belgium

<sup>3</sup>Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA

<sup>4</sup>MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

<sup>5</sup>Population Health Science, Bristol Medical School, Bristol University, Bristol, UK

<sup>6</sup>Department of Psychology and Logopedics, University of Helsinki,

Finland<sup>7</sup>Department of Epidemiology, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA

<sup>8</sup>ISGlobal, Barcelona, Spain

<sup>9</sup>Universitat Pompeu Fabra (UPF), Barcelona, Spain

<sup>10</sup>Department of Epidemiology, Colorado School of Public Health, University of Colorado, Anschutz Medical Campus, Aurora, CO, USA

<sup>11</sup>Lifecourse Epidemiology of Adiposity and Diabetes (LEAD Center), University of Colorado Anschutz Medical Campus, Aurora, CO, USA

<sup>12</sup>Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

<sup>13</sup>Laboratory of Precision Environmental Biosciences, Columbia University Mailman School of Public Health, New York, NY, USA

<sup>14</sup>Department of Translational Research in Psychiatry, Max-Planck Institute of Psychiatry, Munich, Germany

<sup>15</sup>Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, GA, USA

<sup>16</sup>Department of Social Medicine, University of Crete, Heraklion, Crete, Greece

<sup>17</sup>Department of Genetics & Cell Biology, Maastricht University, Maastricht, The Netherlands.

<sup>18</sup>Department of Pediatrics, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

<sup>19</sup>Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA USA

<sup>20</sup>Epigenetics Group, International Agency for Research on Cancer (IARC), Lyon, France

<sup>21</sup>Institute for Health and Welfare, Helsinki and Oulu, Finland

<sup>22</sup>PEDEGO Research Unit, MRC Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland

<sup>23</sup>Department of Clinical and Molecular Medicine, Norwegian University for Science and Technology, Trondheim, Norway

<sup>24</sup>Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

<sup>25</sup>Bristol National Institute of Health Research Biomedical Research Centre

<sup>26</sup>Department of Public Health and Primary Care, Leuven University, Leuven, Belgium

<sup>27</sup>Research Unit for Gynaecology and Obstetrics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

<sup>28</sup>Department of Medical Sciences, University of Turin, Italy

<sup>29</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, Section on Metabolic Genetics, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>30</sup>Department of Public Health, Section of Epidemiology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>31</sup>CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

<sup>32</sup>Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora CO, USA

<sup>33</sup>Diabetes Unit, Massachusetts General Hospital, Boston, MA, USA

**Word Count:** 3,966

**Number of Tables and Figures:** 4

**Abstract:**

*Background:* Maternal gestational diabetes mellitus (GDM) has been associated with adverse outcomes in the offspring. Growing evidence suggests that the epigenome may play a role, but most previous studies have been small and adjusted for few covariates. The current study meta-analyzed the association between maternal GDM and cord blood DNA methylation in the Pregnancy and Childhood Epigenetics Consortium.

*Methods:* Seven pregnancy cohorts (3,677 mother-newborn pairs, 317 with GDM) contributed results from epigenome wide association studies, using DNA methylation data acquired by the Infinium HumanMethylation450 array.

Associations between GDM and DNA methylation were examined using robust linear regression, adjusting for potential confounders. Fixed-effects meta-analyses were performed using METAL. Differentially methylated regions (DMRs) were identified by taking the intersection of results obtained using two regional approaches: comb-p and DMRcate.

*Results:* Two DMRs were identified by both comb-p and DMRcate. Both regions were hypomethylated in newborns exposed to GDM *in utero*, compared with controls. One DMR (chr1:248100345-248100614) was located in the *OR2L13* promoter and the other (chr10:135341870-135342620) was located in the gene body of *CYP2E1*. Individual CpG analyses did not reveal any differentially methylated loci based on a  $P_{FDR}$  threshold of 0.05.

*Conclusions:* Maternal GDM was associated with lower cord blood methylation levels within two regions, including the promoter of *OR2L13*, a gene associated

with autism spectrum disorder, and the gene body of *CYP2E1*, which is upregulated in type 1 and type 2 diabetes. Future studies are needed to understand whether these associations are causal and possible health consequences.

**Background:**

Gestational diabetes mellitus (GDM) is one of the most common pregnancy complications, with prevalence estimates ranging from 2% to 25% depending on the screening and diagnostic criteria used and the population examined (1, 2). In addition to the adverse pregnancy and delivery outcomes associated with GDM, which can include preeclampsia, macrosomia, and shoulder dystocia (3), women diagnosed with GDM are four times more likely to have children who develop metabolic syndrome later in life and twice as likely to have children who become overweight or obese (4). There is also evidence that maternal GDM during pregnancy alters fetal growth trajectories (5) and adversely affects neurodevelopment (6, 7). Thus, understanding the molecular changes related to prenatal exposure to GDM could have widespread implications for children's health.

One potential mechanism underlying such a diverse array of GDM-associated outcomes is epigenetic dysregulation. In support of this, a growing number of studies have observed associations between GDM and cord blood DNA methylation patterns (8-17). However, the majority of studies have been small (e.g., fewer than 100 participants or fewer than 30 GDM cases), adjusted for few if any covariates, and used lenient or no adjustment for multiple testing (8-10, 12-15, 17), which may have contributed to a lack of replication of results across studies.

There has therefore been a call for research on GDM and offspring DNA methylation within larger studies (18). The current study conducted a meta-



analysis of results from epigenome-wide association studies (EWAS) of GDM and cord blood DNA methylation patterns from seven cohorts participating in the Pregnancy and Childhood Epigenetics (PACE) consortium (19). Additionally, we conducted a look-up in our meta-analysis results for CpGs that were previously identified as differentially methylated in prior publications.

## **Methods:**

### *Participating Cohorts:*

All cohorts in the PACE consortium (19) were invited to participate in the current meta-analysis. Seven cohorts, representing eight countries, participated, contributing a total of 317 GDM cases and 3,360 controls (**Table 1, Table S1**). These cohorts are the Avon Longitudinal Study of Parents and Children (ALSPAC), the Genome-Wide Population-Based Association Study of Extremely Overweight Young Adults (GOYA), the Healthy Start study, the Proyecto Infancia y Medio Ambiente (INMA) study, the Prediction and Prevention of Preeclampsia and Intrauterine Growth Restriction (PREDO) study, Project Viva, and a pooled analysis of three cohorts: the Rhea study (RHEA), The ENVIRonmental influence ON early AGEing in early life (ENVIRONAGE) study, and the Piccolipiù study (RHEA/ENVIRONAGE/Piccolipiù). Cohort details are described in the “Cohort-Specific Information” section of **Supplementary Data**. Each cohort received ethics approval and informed consent from participants prior to data collection,

and the current meta-analysis was approved by the Health Sciences Institutional Review Board of the University of Southern California.

*Gestational Diabetes:*

Participants diagnosed with type 1 or type 2 diabetes prior to the index pregnancy were excluded from analyses. The criteria used to classify GDM cases are summarized by cohort in Table 1 and are also described in more detail in the Supplementary Data. All cohorts except Piccolipiù used information abstracted from medical records as their primary method for identifying GDM cases. Due to a lack of international consensus, the criteria used to classify GDM differ by country and have changed over time. In the United States and some European countries, GDM is often diagnosed using a 2-step approach, which entails universal screening with a 50-g glucose challenge test, followed by a 100-g 3-hour oral glucose tolerance test (OGTT) for those who test positive (20). In contrast, some European countries have adopted the International Associations of the Diabetes and Pregnancy Study Groups (IADPSG) guidelines (21), which recommend a 1-step approach, in which a 75-g 2-hour OGTT is performed for all women at 24-28 weeks' gestation. Furthermore, some countries use a selective approach and only administer GDM diagnostic tests to women with traditional risk factors. GDM cases from Healthy Start, Project Viva, Rhea, and ENVIRONAGE were classified based on the 2-step approach, using the Carpenter-Coustan criteria (22). GDM cases from PREDO were classified based on the IADPSG 1-step approach. Piccolipiù identified GDM cases based on self-

reported questionnaire data collected at delivery, and all but one case was confirmed using medical record data (IADPSG 1-step approach (21)). GDM cases from INMA were diagnosed using a selective screening approach, where women at high risk for GDM were administered a glucose challenge test, followed by a diagnostic OGTT, using the Carpenter-Coustan criteria (22). GDM cases from ALSPAC and GOYA were diagnosed based on the practices at the time in the United Kingdom and Denmark, respectively, in which diagnostic tests were only performed for women 1) at high risk for GDM based on established risk factors or 2) with glycosuria (23, 24). Given anticipated under-reporting of GDM in the medical records, information from telephone interviews was also used to classify GDM cases in GOYA.

*Methylation Measurements:*

Cord blood DNA was bisulfite-converted using the EZ-96 DNA Methylation Kit (Zymo Research Corporation, Irvine, USA). Each cohort measured DNA methylation using the Infinium® HumanMethylation450k BeadChip array (Illumina Inc., San Diego, United States), either at Illumina or in cohort-specific laboratories, and completed its own quality control and normalization of data, as described in the Supplementary Data. Since the PACE consortium has observed that extreme outliers (> 3 times the interquartile range) can have a large impact on results, they were removed prior to analyses. For all analyses, normalized, untransformed betas were evaluated as the outcomes.

### *Cohort-Specific Statistical Analyses:*

Cohorts ran independent EWAS models according to the same analysis plan, using robust linear regression, as this method controls for possible heteroscedasticity and potential outliers. Only singleton pregnancies were included in analyses. GDM was modeled as the exposure of interest and the cord blood DNA methylation level at each CpG was modeled as the outcome. Regression models were adjusted for hypothesized confounders, which included newborn's sex, maternal age, maternal education level, maternal body mass index (pre-pregnancy or early pregnancy), maternal smoking status during pregnancy (ever versus never), and maternal genetic ancestry (if available) or maternal race/ethnicity. Cohort-specific details for covariate assessment are described in the Supplementary Data. First, we adjusted only for this baseline set of covariates (results are presented in the Supplementary Data), such that results could be compared with previous studies, which have generally not accounted for cord blood cell heterogeneity. However, our final model was additionally adjusted for cord blood cell fractions, including B cells, CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, granulocytes, natural killer cells, monocytes, and nucleated red blood cells, which were estimated using a cord blood reference panel (25). We also examined results from two of the larger participating cohorts (PREDO and Project Viva) after additional adjustment for parity. Since results were very similar (**Tables S2-S3**), parity was not included in the final model.

### *Meta-Analyses:*

METAL (26) was used to conduct inverse variance-weighted fixed effects meta-analyses, using results from the cohort-specific analyses. Control probes, probes mapping to the X and Y chromosomes, and probes that have been shown to cross-hybridize or which target polymorphic CpGs or contain SNPs at the single base pair extension (27) were excluded. A total of 380,878 CpGs were therefore included in the meta-analyses. Probes were annotated to hg19 using the IlluminaHumanMethylation450kanno.ilmn12.hg19 R package (28). After meta-analyses were complete, a second analyst ran shadow meta-analyses to rule out potential human error. CpGs were considered differentially methylated if the false discovery rate-adjusted p-value ( $P_{FDR}$ ) was  $< 0.05$ .

Potential heterogeneity between studies was assessed using Cochran's Q statistic and  $I^2$ . Additionally, leave-one-out meta-analyses (i.e., comparison of results after the sequential removal of one cohort and a meta-analysis of the remaining six cohorts) were conducted to evaluate the influence of each individual cohort on the results.

Differentially methylated regions (DMRs) were identified from meta-analysis results by taking the intersection of DMRs identified using two different software programs: comb-p (29) and DMRcate (30). Comb-p identifies regions enriched for low p-values, uses the Stouffer-Liptak method to correct for auto-correlation, and adjusts for multiple testing using the Sidak correction (29). DMRcate calculates two smoothed estimates for each chromosome (one weighted by F-statistics, one not) and uses a Satterthwaite approximation to compare these estimates; it then adjusts for multiple testing using the FDR

method (30). These approaches were selected, because they can be applied to meta-analysis results. Windows of 500 and 1,000 base pairs were compared for each approach. For comb-p, a p-value threshold of  $1 \times 10^{-3}$  was used to specify the start of each region, and a distance of 200 base pairs was selected for extending the region. For DMRcate, the default settings were used, as recommended (30), and FDR thresholds of 0.05 and 0.01 were compared.

### *Sensitivity Analyses*

Since the seven participating cohorts represent different geographic regions and differ in the timing of participant recruitment and the criteria used to classify GDM cases, we ran a series of sensitivity meta-analyses. We compared meta-analysis results after restricting to 1) cohorts with GDM cases identified by selective versus universal screening, 2) cohorts with GDM cases identified using a 1-step 75 g versus a 2-step 100 g OGTT, 3) European versus U.S. cohorts, and 4) cohorts that recruited participants prior to 2004 versus after 2004.

### *Look-Up Analyses:*

In an effort to replicate previous findings, a look-up of CpGs previously identified as differentially methylated by GDM status was conducted within results from the meta-analyses (both with and without adjustment for estimated cell proportions). Relevant studies were identified in PubMed using the following search terms: Gestational diabetes AND DNA methylation. We focused on studies that: 1) were not included in the current meta-analyses, 2) included > 10

GDM cases, 3) measured DNA methylation in cord blood using Illumina's 450k, EPIC, or 27k array, 4) adjusted for multiple testing using any method, and 5) provided effect estimates and p-values for individual CpGs. Two studies met these criteria (9, 12). These studies collectively reported a total of 110 differentially methylated CpGs, none of which were common. Additionally, nine CpGs within two genes (*MEST*, *NR3C1*) that were identified as differentially methylated by GDM status in both cord blood and placenta in a previous candidate gene study (11), which are represented on the 450k array, were evaluated. Of these 119 CpGs, 32 were cross-reactive or polymorphic, or the CpG probe contained a SNP at the single base pair extension (27). These 32 CpGs were therefore excluded, leaving a total of 87 CpGs for the look-up analyses.

## **Results:**

### *Study Characteristics:*

Characteristics of participating studies are shown in Table 1 and Table S1. The number (%) of GDM cases per study ranged from 12 (7.7%) for INMA to 180 (23.1%) for PREDO. The majority of participants were of European ancestry, and about half of the newborns were male (N = 1,900, 51.7%).

### *Meta-Analyses for the Individual CpG Results*

Probe numbers and the level of inflation ( $\lambda$ ) for individual cohort results are shown in **Table S4**. The  $\lambda$  for the meta-analyses was 1.15. Meta-analysis results are summarized in a Manhattan plot (**Figure 1**). No CpGs were identified as differentially methylated by GDM status based on a  $P_{FDR} < 0.05$ , but six were identified based on a  $P_{FDR} < 0.10$  (**Table 2**). While the directions of effect were generally consistent for Healthy Start; INMA; PREDO; Project Viva; and the pooled analysis of Rhea, Environage, and Piccolipiù; they often differed for ALSPAC or GOYA (Table 2). For five of the CpGs, there was not strong evidence of heterogeneity ( $I^2 < 10.0$ ,  $P_{\text{heterogeneity}} > 0.36$ ), but for one CpG (cg11723077), there was evidence of moderate heterogeneity ( $I^2=38.7$ ,  $P=0.13$ ). However, effect estimates were similar across the leave-one-out meta-analyses (results shown in **Figure S1** and **Table S5**).

#### *Look-Up Analysis Results*

The full look-up analysis results are presented in **Tables S6** and **S7** within the Supplementary Data. Of the 87 CpGs examined, four were differentially methylated (uncorrected  $P < 0.05$ ) in the same direction in the meta-analysis that accounted for cell heterogeneity (Table S7). These four CpGs (cg01203331, cg03345925, cg08471713, cg20507276) were annotated to a total of seven genes: *NOP56*, *SNORD56*, *SNORD57*, *SNORD86*; *ZC3H3*; *MEOX1*; and *OR2L13*, respectively. However, based on the 87 tests conducted, FDR-corrected p-values exceeded 0.05 for all of the CpGs evaluated.



### *DMRs Identified from the Meta-Analysis Results*

Using individual CpG results from the meta-analyses, comb-p identified five regions that were differentially methylated by GDM status (**Table S8**). Comb-p results were the same when either a 500 or 1,000 base pair window was used. DMRcate identified two DMRs when using an FDR threshold of 0.10. One DMR was identified when using the 500 base pair window (chr1:248100407-248100614) and the other when using the 1,000 base pair window (chr10:135341870-135342620) (**Table S9**). Both of these DMRs overlapped two DMRs that had also been identified by comb-p (**Table 3**). One was located in the promoter region of *OR2L13* and was also annotated to pseudogene *CLK3P2*. The second overlapped a CpG island in the gene body of *CYP2E1*. %Methylation levels in both regions were lower in the GDM case, compared with control, group, and effect estimates were generally consistent for the individual CpGs contained within each region (**Figure S2**). DMRcate did not identify any DMRs when using an FDR threshold of 0.05.

### *Sensitivity Analysis Results*

Results were generally similar for the six CpGs with a  $P_{\text{FDR}} < 0.10$ , and also for CpGs within the 2 DMRs identified by comb-p and DMRcate, when restricting to cohorts with GDM cases identified by a 1-step 75 g OGTT versus a 2-step 100 g OGTT or using selective versus universal screening. They were also generally similar for U.S. versus European cohorts and for cohorts that recruited participants prior to versus after 2004 (**Figures S3-S5**).

## **Discussion:**

While previous studies have investigated associations between maternal GDM and newborn DNA methylation (8-16), the majority have been small, used lenient or no adjustment for multiple testing, did not consider regional methylation differences, and adjusted for a limited number of covariates. In particular, few studies have adjusted for cell heterogeneity, an important source of variability in DNA methylation (31). Results have been inconsistent between these previous studies, raising questions of robustness and reproducibility. The current study therefore conducted meta-analyses of EWAS results from seven cohorts (3,677 mother-newborn pairs, 317 with GDM) participating in the PACE consortium (19), which examined associations between GDM and cord blood DNA methylation, after adjusting for a larger number of potential confounders. We evaluated methylation differences at both the regional and individual CpG level.

Using two dimension reduction approaches (comb-p (29) and DMRcate (30)), we identified two regions that are differentially methylated by GDM status. One of the DMRs identified by the meta-analysis (chr1:248100276-248100614) is located in the promoter region of *OR2L13*, a gene which codes for an olfactory receptor (9). Methylation levels in this region were lower in cord blood from GDM exposed, compared with unexposed, newborns. This finding is consistent with a previous study by Quilter et al., which observed lower cord blood methylation levels at a CpG located in this DMR (cg20507276) among GDM exposed newborns (9). This same CpG has also been identified as differentially

methylated in both blood and buccal cells from autism spectrum disorder (ASD) cases versus controls (32). While the mechanism by which *OR2L13* may contribute to ASD is currently unknown, olfactory dysfunction has been associated with more severe social impairments among individuals with ASD (32). Since children exposed to maternal GDM *in utero* have a higher risk of developing ASD (6), future investigation into the potential mediating role of *OR2L13* in GDM-associated ASD is merited. Methylation levels in the second DMR (chr10:135341933-135342560) were also lower in the GDM case, compared with control, group. This DMR is located in a CpG island within the gene body of *CYP2E1*, which codes for an enzyme that is highly expressed in the liver and metabolizes ethanol, numerous drugs, and certain protoxicants (33). Although, to our knowledge, the CpGs within this DMR have not previously been associated with *in utero* exposure to GDM, increased *CYP2E* expression has been observed in peripheral blood from individuals with type 1 and type 2 diabetes (33).

In contrast with the DMR results, we did not identify any individual differentially methylated CpGs when using a conservative  $P_{FDR}$  threshold of 0.05. When using a more lenient  $P_{FDR}$  threshold of 0.10, six individual CpGs (cg00812770, cg11723077, cg22791932, cg17588003, cg11187204, cg10139436) were identified as differentially methylated by GDM status, none of which had been identified in the previous studies that we reviewed. Three of these CpGs (cg11723077, cg22791932, and cg17588003) were annotated to genes: *SYNJ2*, *ZFPM1*, and *C17orf87*, respectively, and a fourth CpG

(cg00812770) was located in a long intergenic non-coding RNA (*LINC01342*). The remaining two CpGs were not annotated to any genes, and the potential consequences of altered methylation at these loci are currently unclear.

The 13 CpGs comprising the two DMRs identified by both comb-p and DMRcate were not identified as differentially methylated in individual CpG meta-analyses, likely due to the greater statistical power of the DMR approaches. Additionally, the six CpGs identified as differentially methylated based on a  $P_{FDR} < 0.10$  in the individual CpG analyses were not identified by either comb-p or DMRcate. It is possible that these six CpGs are false positives, since they did not reach statistical significance after applying a more conservative threshold of  $P_{FDR} < 0.05$ . However, two of these CpGs (cg11187204, cg10139436) also resided in intergenic regions that are either CpG-poor or sparsely represented on the 450k array, which would have precluded their identification using regional approaches.

In our examination of 87 CpGs that have previously been associated with GDM status (9, 11, 12), only four were found to be differentially methylated in the same direction in the current meta-analysis, based on an uncorrected  $P < 0.05$ . Since these previous studies were similarly conducted in predominately European populations, differences in race or ethnicity are likely not driving these discrepancies. However, some of the prior findings may be false positives due to small samples sizes; insufficient control for multiple testing; or a lack of adjustment for important confounding factors, such as maternal BMI. Other potential explanations for the lack of replication include differences in exclusion

criteria and the fact that these previous studies stratified by fetal sex (9) or GDM treatment type (11, 12), which was not feasible for the current meta-analysis.

Importantly, the seven cohorts participating in the current meta-analysis represent eight countries and multiple time periods. Since the criteria used to classify GDM differ by country and have changed across time, the severity of disease among GDM cases, and the proportion of controls with undiagnosed GDM or hyperglycemia, may have varied between cohorts. Nevertheless, we did not observe evidence of heterogeneity for the majority of meta-analysis results. Furthermore, results were generally similar across a series of sensitivity analyses, which stratified cohorts based on geographic location, time, and the criteria used for GDM classification. It is therefore possible that there may be a linear relationship between maternal glucose levels and cord blood DNA methylation. However, while there is some evidence for this (34), additional studies are needed to determine if maternal glucose is the main mechanism through which GDM alters DNA methylation and, if so, whether or not there is a clear threshold below which maternal glucose does not alter cord blood methylation.

The current study had many notable strengths. By meta-analyzing results from multiple cohorts, we were able to increase the statistical power of the study and adjust for a large number of potential confounders, including estimated cell fractions. We also used stringent adjustments for multiple testing to reduce the chance of identifying false positives. Another strength of the study was the

evaluation of DMRs (using two different approaches) in addition to individual CpGs.

However, our meta-analyses also had limitations. First, there may have been an overall underestimation of GDM cases, since GDM cases from several cohorts were diagnosed based on a selective approach. This may have resulted in some participants being misclassified as controls, which would have biased results toward the null. Another important consideration is that regression models were adjusted for maternal BMI, because it is a risk factor for GDM (35) and may impact cord blood DNA methylation (36). However, this may have also biased results toward the null, since GDM cases from several cohorts were diagnosed selectively based on traditional risk factors, including obesity. Importantly, women with GDM may have utilized different strategies to manage their disease. However, this information was not available for all cohorts, and the number of GDM cases adhering to particular management strategies or treatments was very small for most cohorts, so these differing subsets of GDM cases could not be evaluated separately. We also could not evaluate potential differences by fetal sex due to the small number of GDM cases per cohort. Additionally, since the gestational age at OGTT was not available for all participants, we were unable to adjust for this covariate. Another potential limitation was our focus on cord blood DNA methylation, which may not reflect methylation patterns in other tissues. However, cord blood DNA methylation has been associated with several outcomes that have been associated with *in utero* exposure to GDM, such as early childhood weight and adiposity (37) and ASD (38). While we excluded

CpGs that overlapped SNPs and also CpG probes with SNPs at the single base pair extension (27), we cannot rule out the possibility that some of the differentially methylated CpGs and regions identified in this meta-analysis may be driven by genetic, rather than epigenetic, differences between GDM cases and controls, which merits future investigation. Finally, since the majority of individuals in the seven participating cohorts were of European ancestry, results from the current meta-analysis may not be generalizable to other populations.

### **Conclusions:**

In a meta-analysis of integrated EWAS results from seven pregnancy cohorts, comprising data from 3,677 mother-newborn pairs, GDM was associated with lower cord blood methylation levels within the promoter region of *OR2L13* and the gene body of *CYP2E1*. Given that reduced methylation in the *OR2L13* promoter has previously been associated with both GDM status and ASD, its potential role in mediating this relationship should be evaluated in future studies. Additionally, since *CYP2E1* is upregulated in peripheral blood from individuals with type 1 and type 2 diabetes, the impact of reduced methylation within this gene among GDM-exposed newborns on subsequent health merits future investigation. Finally, the inability to replicate many results from previous studies of GDM exposure and cord blood DNA methylation highlights the importance of conducting EWAS meta-analyses, using data from multiple cohorts.

## **Acknowledgements:**

Acknowledgements for each study can be found in the “Supplemental Acknowledgements” section of the Supplementary Data File.

*Statement of Author Contributions:* CGH and CVB conceptualized the study; CGH, CVB, M-FH, and CL-A drafted the meta-analysis plan; CGH conducted the meta-analysis and wrote the first draft of the manuscript; CVB and M-FH supervised the meta-analysis; BC, RF, JJ, TK, SL, HEM, LAS, SR-S, APS, and PY contributed to cohort-specific statistical analyses; AB, EBB, VLC, DC, DD, DLD, AG, ZH, EK, JM TL, DAL, AL, TSN, EAN, EO, CP, MP, KR, CLR, GCS, TIAS, JS, MV, WZ, and MFH oversaw cohort-specific analyses; RF conducted the shadow meta-analyses; all co-authors provided feedback on the manuscript.

*Funding:* Funding for individual cohorts can be found in the “Funding Support” section of the Supplementary Data file.

*Conflict of Interest Statements:* DAL has received support from numerous national and international government and charity funders and from Roche Diagnostics and Medtronic for work unrelated to that presented here.

DLD has received grants from the NIH and personal fees from Novartis.

All other authors declare no conflicts of interest.

*Guarantor Statement:* CGH is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Disclaimer:* Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are



responsible for the views expressed in this article, and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

## References:

1. DeSisto CL, Kim SY, Sharma AJ. Prevalence estimates of gestational diabetes mellitus in the United States, Pregnancy Risk Assessment Monitoring System (PRAMS), 2007-2010. *Prev Chronic Dis.* 2014;11:E104.
2. Sacks DA, Hadden DR, Maresh M, Deerochanawong C, Dyer AR, Metzger BE, et al. Frequency of gestational diabetes mellitus at collaborating centers based on IADPSG consensus panel-recommended criteria: the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study. *Diabetes Care.* 2012;35(3):526-8.
3. Mack LR, Tomich PG. Gestational Diabetes: Diagnosis, Classification, and Clinical Care. *Obstet Gynecol Clin North Am.* 2017;44(2):207-17.
4. Clausen TD, Mathiesen ER, Hansen T, Pedersen O, Jensen DM, Lauenborg J, 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.* 2009;94(7):2464-70.
5. Brand JS, West J, Tuffnell D, Bird PK, Wright J, Tilling K, et al. Gestational diabetes and ultrasound-assessed fetal growth in South Asian and White European women: findings from a prospective pregnancy cohort. *BMC Med.* 2018;16(1):203.
6. Wan H, Zhang C, Li H, Luan S, Liu C. Association of maternal diabetes with autism spectrum disorders in offspring: A systemic review and meta-analysis. *Medicine (Baltimore).* 2018;97(2):e9438.
7. Fraser A, Almqvist C, Larsson H, Langstrom N, Lawlor DA. Maternal diabetes in pregnancy and offspring cognitive ability: sibling study with 723,775 men from 579,857 families. *Diabetologia.* 2014;57(1):102-9.
8. Finer S, Mathews C, Lowe R, Smart M, Hillman S, Foo L, et al. Maternal gestational diabetes is associated with genome-wide DNA methylation variation in placenta and cord blood of exposed offspring. *Human molecular genetics.* 2015;24(11):3021-9.
9. Quilter CR, Cooper WN, Cliffe KM, Skinner BM, Prentice PM, Nelson L, et al. Impact on offspring methylation patterns of maternal gestational diabetes mellitus and intrauterine growth restraint suggest common genes and pathways linked to subsequent type 2 diabetes risk. *The FASEB Journal.* 2014;28(11):4868-79.
10. Ruchat S-M, Houde A-A, Voisin G, St-Pierre J, Perron P, Baillargeon J-P, et al. Gestational diabetes mellitus epigenetically affects genes predominantly involved in metabolic diseases. *Epigenetics.* 2013;8(9):935-43.
11. El Hajj N, Pliushch G, Schneider E, Dittrich M, Muller T, Korenkov M, et al. Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes.* 2013;62(4):1320-8.
12. Haertle L, El Hajj N, Dittrich M, Muller T, Nanda I, Lehnen H, et al. Epigenetic signatures of gestational diabetes mellitus on cord blood methylation. *Clin Epigenetics.* 2017;9:28.
13. Kang J, Lee CN, Li HY, Hsu KH, Lin SY. Genome-wide DNA methylation variation in maternal and cord blood of gestational diabetes population. *Diabetes Res Clin Pract.* 2017;132:127-36.
14. Yang IV, Zhang W, Davidson EJ, Fingerlin TE, Kechris K, Dabelea D. Epigenetic marks of in utero exposure to gestational diabetes and childhood adiposity outcomes: the EPOCH study. *Diabet Med.* 2018;35(5):612-20.

15. Weng X, Liu F, Zhang H, Kan M, Wang T, Dong M, et al. Genome-wide DNA methylation profiling in infants born to gestational diabetes mellitus. *Diabetes Res Clin Pract.* 2018;142:10-8.
16. Chen D, Zhang A, Fang M, Fang R, Ge J, Jiang Y, et al. Increased methylation at differentially methylated region of GNAS in infants born to gestational diabetes. *BMC Med Genet.* 2014;15:108.
17. Kang J, Lee CN, Li HY, Hsu KH, Wang SH, Lin SY. Association of Interleukin-10 Methylation Levels With Gestational Diabetes in a Taiwanese Population. *Front Genet.* 2018;9:222.
18. Bouchard L. Epigenetics and fetal metabolic programming: a call for integrated research on larger cohorts. *Diabetes.* 2013;62(4):1026-8.
19. Felix JF, Joubert BR, Baccarelli AA, Sharp GC, Almqvist C, Annesi-Maesano I, et al. Cohort Profile: Pregnancy And Childhood Epigenetics (PACE) Consortium. *International journal of epidemiology.* 2018;47(1):22-3u.
20. Moyer VA. Screening for gestational diabetes mellitus: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2014;160(6):414-20.
21. Metzger BE, Gabbe SG, Persson B, Buchanan TA, Catalano PA, Damm P, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care.* 2010;33(3):676-82.
22. Carpenter MW, Coustan DR. Criteria for screening tests for gestational diabetes. *American Journal of Obstetrics & Gynecology.* 1982;144(7):768-73.
23. Fraser A, Nelson SM, Macdonald-Wallis C, Lawlor DA. Associations of existing diabetes, gestational diabetes, and glycosuria with offspring IQ and educational attainment: the Avon Longitudinal Study of Parents and Children. *Exp Diabetes Res.* 2012;2012:963735.
24. Olsen SF, Houshmand-Oeregaard A, Granstrom C, Langhoff-Roos J, Damm P, Bech BH, et al. Diagnosing gestational diabetes mellitus in the Danish National Birth Cohort. *Acta Obstet Gynecol Scand.* 2017;96(5):563-9.
25. Andrews SV, Bakulski KM. FlowSorted. *CordBlood. 450k User's Guide Methylation Dataset on Sorted Cord Blood Cells.*
26. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26(17):2190-1.
27. Chen Y-a, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8(2):203-9.
28. Hansen K. *IlluminaHumanMethylation450kanno.ilmn12.hg19: Annotation for Illumina's 450k methylation arrays. R package version 060.* 2016.
29. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics.* 2012;28(22):2986-8.
30. Peters T, Peters MT, biocViews DifferentialMethylation G, GenomeAnnotation D, OneChannel T, MultipleComparison Q. Package 'DMRcate'. 2016.
31. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome biology.* 2014;15(2):R31.
32. Dall'Aglio L, Muka T, Cecil CAM, Bramer WM, Verbiest M, Nano J, et al. The role of epigenetic modifications in neurodevelopmental disorders: A systematic review. *Neurosci Biobehav Rev.* 2018;94:17-30.
33. Wang Z, Hall SD, Maya JF, Li L, Asghar A, Gorski JC. Diabetes mellitus increases the in vivo activity of cytochrome P450 2E1 in humans. *Br J Clin Pharmacol.* 2003;55(1):77-85.

34. Houde AA, Ruchat SM, Allard C, Baillargeon JP, St-Pierre J, Perron P, et al. LRP1B, BRD2 and CACNA1D: new candidate genes in fetal metabolic programming of newborns exposed to maternal hyperglycemia. *Epigenomics*. 2015;7(7):1111-22.
35. Ramos GA, Caughey AB. The interrelationship between ethnicity and obesity on obstetric outcomes. *Am J Obstet Gynecol*. 2005;193(3 Pt 2):1089-93.
36. Sharp GC, Salas LA, Monnereau C, Allard C, Yousefi P, Everson TM, et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation: findings from the pregnancy and childhood epigenetics (PACE) consortium. *Hum Mol Genet*. 2017;26(20):4067-85.
37. Kresovich JK, Zheng Y, Cardenas A, Joyce BT, Rifas-Shiman SL, Oken E, et al. Cord blood DNA methylation and adiposity measures in early and mid-childhood. *Clin Epigenetics*. 2017;9:86.
38. Andrews SV, Ellis SE, Bakulski KM, Sheppard B, Croen LA, Hertz-Picciotto I, et al. Cross-tissue integration of genetic and epigenetic data offers insight into autism spectrum disorder. *Nat Commun*. 2017;8(1):1011.

**Table 1.** Characteristics of Participating Cohorts\*

<b>Cohort</b>	<b>Location</b>	<b>Participant Enrollment Years</b>	<b>GDM Screening Approach</b>	<b>GDM Classification Criteria and Source of Information</b>	<b>Number of GDM Cases</b>	<b>Number of Controls</b>
<b>ALSPAC</b>	United Kingdom	1991-1992	Selective	Physician-diagnosed GDM (Medical Records)	22	867
<b>GOYA</b>	Denmark	1996-2002	Selective	Physician-diagnosed GDM based on a 1-Step 75 g OGTT (Medical Records) + Self-Report	28	404
<b>Healthy Start</b>	United States	2009-2014	Universal	Physician-diagnosed GDM based on a 2-Step 100 g OGTT, Carpenter and Coustan Criteria (Medical Records)	32	534
<b>INMA</b>	Spain	2004-2007	Selective	2-Step 100 g OGTT, Carpenter and Coustan Criteria (Medical Records)	12	144

<b>PREDO</b>	Finland	2006-2010	Universal	1-Step 75 g OGTT, IADPSG Criteria (Medical Records)	180	600
<b>RHEA/ENVIRONAGE/Piccolipiù (Pooled)</b>	Greece/Belgium/Italy	2007-2008/2010-2011-2015	Universal/Universal/Universal	2-Step 100 g OGTT, Carpenter and Coustan Criteria (Medical Records)/2-Step 100 g OGTT, Carpenter and Coustan Criteria (Medical Records)/Self-Report	20	352
<b>Project Viva</b>	United States	1999-2002	Universal	2-Step 100 g OGTT, Carpenter and Coustan Criteria (Medical Records)	23	459

Abbreviations used: GDM, gestational diabetes; IADPSG, *International Association of the Diabetes and Pregnancy Study Groups*; OGTT, oral glucose tolerance test

\*Additional details on GDM classification and other characteristics of each cohort are included in the “Cohort-Specific Methods” section of the Supplementary Data

**Table 2.** CpGs with a  $P_{FDR} < 0.10$  in the Meta-Analysis of Maternal GDM Exposure\*

CpG	Genomic Position	%Methylation Difference (95%CI) <sup>†</sup>	Direction by Cohort <sup>‡</sup>	Raw P-Value	FDR-Corrected P-Value	Heterogeneity P-Value <sup>§</sup>	I <sup>2</sup>	Relation to CpG Island <sup>  </sup>	Target Gene <sup>¶</sup>
cg00812770	chr1:1073510	0.8% (0.5%, 1.1%)	+++++++	1.7 x 10 <sup>-7</sup>	0.06	0.55	0.0	South Shore	<i>LINC01342</i>
cg11723077	chr6:158508188	-1.0% (-1.4%, -0.6%)	-+-----	1.3 x 10 <sup>-6</sup>	0.09	0.13	38.7	South Shore	<i>SYNJ2</i>
cg22791932	chr16:88537374	0.8% (0.4%, 1.1%)	+++++++	1.2 x 10 <sup>-6</sup>	0.09	0.74	0.0	Island	<i>ZFPM1</i>
cg17588003	chr17:5138696	-1.4% (-2.0%, -0.8%)	+-----	1.4 x 10 <sup>-6</sup>	0.09	0.67	0.0	Open Sea	<i>C17orf87</i>
cg11187204	chr17:36480526	-1.6% (-2.1%, -1.0%)	++-----	5.2 x 10 <sup>-8</sup>	0.09	0.57	0.0	Open Sea	N/A
cg10139436	chr19:30219558	-0.4% (-0.5%, -0.2%)	++-----	1.3 x 10 <sup>-6</sup>	0.09	0.36	9.7	South Shelf	N/A

\*Results are from inverse variance-weighted fixed effects meta-analyses, conducted using METAL. Each cohort independently ran robust linear regression models, adjusting for newborn's sex, maternal age (in years), maternal body mass index (early pregnancy or pre-pregnancy), maternal smoking status during pregnancy, maternal education, maternal genetic ancestry (if available) or maternal race/ethnicity, and estimated proportions of B cells, CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, granulocytes, natural killer cells, monocytes, and nucleated red blood cells in cord blood.

<sup>†</sup>%Difference in newborn DNA methylation and 95% confidence interval, comparing the gestational diabetes case group to the control group

<sup>‡</sup>Direction of association between gestational diabetes and methylation at the locus of interest by cohort, ordered as follows: ALSPAC, GOYA, Healthy Start, INMA, PREDO, RHEA/ENVIRONAGE/Piccolipiù, Project Viva

<sup>§</sup>The heterogeneity p-value and I<sup>2</sup> were calculated by METAL using Cochran's Q-test for heterogeneity.

<sup>||</sup>Relationship to CpG islands from the UCSC database, annotated using the Illumina 450k manifest (28)

<sup>¶</sup>Target gene name(s) from the UCSC database, annotated using the Illumina 450k manifest (28)

**Table 3.** Differentially Methylated Regions Identified By Both Comb-p and DMRcate\*

DMR	450k CpGs	Direction of association	Nearby Genes	Regulatory Feature Group/Gene Group/Relation to Island
<b>Comb-p (500 and 1,000 bp window)</b>				
chr1:248100345-248100614	cg00785941, cg03748376, cg04028570, cg08260406, cg08944170, cg20434529, cg20507276	-	<i>OR2L13,CLK3P2</i>	Promoter associated/1 <sup>st</sup> Exon:5'UTR or TSS200/Island or North Shore
chr10:135342218-135342413	cg10862468, cg25330361	-	<i>CYP2E1</i>	NA/Body/Island
<b>DMRcate (500 bp window)</b>				
chr1:248100407-248100614	cg00785941, cg03748376, cg04028570, cg08260406, cg08944170, cg20507276	-	<i>OR2L13</i>	Promoter Associated/1 <sup>st</sup> Exon/5'UTR/Island
<b>DMRcate (1,000 bp window)</b>				
chr10:135341870-135342620	cg00321709, cg10862468, cg19469447, cg23400446, cg24530264, cg25330361	-	<i>CYP2E1</i>	Unclassified/Body/Island

Abbreviations Used: bp, base pairs

\*Differentially methylated regions were identified from meta-analysis results for individual CpGs, which used results from robust linear regression models that were adjusted for newborn's sex, maternal age (in years), maternal body mass index (pre-pregnancy or early pregnancy), maternal



education, maternal smoking status during pregnancy (ever vs. never), maternal genetic ancestry (if available) or maternal race/ethnicity, and estimated cord blood cell fractions.

**Figure 1.** Manhattan plot summarizing results for meta-analyses of the associations between maternal gestational diabetes mellitus and cord blood DNA methylation. Meta-analyses were run using METAL on results from robust linear regression models, which adjusted for newborn's sex, maternal age, maternal education, maternal body mass index (pre-pregnancy or in early pregnancy), maternal smoking status during pregnancy (ever versus never), maternal genetic ancestry (if available) or maternal race/ethnicity, and estimated cord blood cell fractions. Blue and red lines indicate  $\log_{10}(\text{p-values})$  that are equivalent to a  $P_{\text{FDR}}$  of 0.10 and a  $P_{\text{FDR}}$  of 0.05, respectively.