# Epigenome Wide Association Study Reveals Methylation Loci Associated with Offspring Gestational Diabetes Mellitus Exposure and Maternal Methylome

Running title: Epigenetic changes in response to GDM exposure

Mickaël Canouil PhD [1,2*], Amna Khamis PhD † [1,2*], Elina Keikkala [3,4], Sandra Hummel PhD [5], Stephane Lobbens [1,2], Amélie Bonnefond PhD [1,2], PhD, Fabien Delahaye PhD [1,2], Evangelia Tzala [6], Sanna Mustaniemi MD [3,4], Marja Vääräsmäki MD [3,4], Marjo-Riitta Jarvelin [6,7,8,9], Sylvain Sebert PhD [4], Eero Kajantie [3,4,10,11], Philippe Froguel MD † [1,2,12], Toby Andrew PhD [12].

[1] Inserm U1283, CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, Lille, F-59000, France.

[2] University of Lille, Lille University Hospital, Lille, F-59000, France.

[3] PEDEGO Research Unit, MRC Oulu, Oulu University Hospital and University of Oulu, FI-90014 Oulu, Finland.

[4] Finnish Institute for Health and Welfare, Public Health Promotion Unit, 00271 Helsinki and 90100 Oulu, Finland.

[5] Institute of Diabetes Research, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich-Neuherberg, Germany, and Forschergruppe Diabetes, Technical University Munich, at Klinikum rechts der Isar, Munich, and Forschergruppe Diabetes e.V., Munich-Neuherberg, Germany.

[6] Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, W2 1PG, United Kingdom.

[7] Center for lifecourse Health research, Faculty of Medicine, University of Oulu, Oulu, Finland.

[8] Unit of Primary Health Care, Oulu University Hospital, OYS, Kajaanintie 50, 90220 Oulu, Finland.

[9] Department of Life Sciences, College of Health and Life Sciences, Brunel University London, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom.

[10] Norwegian University of Science and Technology, Department of Clinical and Molecular Medicine, NO-7491 Trondheim, Norway.

[11] Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland.

[12] Department of Metabolism, Digestion and Reproduction, Imperial College London, London, United Kingdom.

* Joint first authors.

† Co-corresponding authors:

Dr Amna Khamis

European Genomic Institute for Diabetes - UMR 1283/8199

Pôle Recherche - 1er étage Aile Ouest

1, Place de Verdun, 59045 LILLE CEDEX

Email: amna.khamis@cnrs.fr

Tel: +33 (0)3 74 00 8135

Prof Philippe Froguel, MD, PhD

Department of Metabolism, Imperial College Faculty of Medicine

Hammersmith Hospital Campus, Burlington-Danes building, Du Cane Road

London, W12 0NN, United Kingdom

Email: p.froguel@imperial.ac.uk

Tel: 44 (0)207 594 65 20

Word count: 3,860

Number of figures and/or tables: 3 figures and 1 table

Abstract

*Objective:* Gestational diabetes mellitus (GDM) is associated with a future offspring risk for the development of obesity and insulin resistance in Gestational diabetes mellitus (GDM) is associated with an increased risk of obesity and insulin resistance in offspring later in life, which might be explained by epigenetic changes in response to maternal hyperglycaemic exposure.

*Research Design and Methods:* We explored the association of GDM exposure on maternal blood and newborn cord-blood methylation of 536 mother-offspring pairs from the prospective FinnGeDi cohort, using Illumina's methylationEPIC BeadChip 850K arrays. We assessed two hypotheses First, we tested for shared maternal and offspring epigenetic effects due to GDM exposure. Second, we tested whether GDM exposure and maternal methylation has an epigenetic effect on the offspring.

*Results:* We did not find any epigenetic marks (differentially methylated CpG probes) with shared and consistent effects between mothers and offspring. After including maternal methylation in the model, we identified a single significant (FDR = $1.38 \times 10^{-2}$) CpG at the cg22790973 probe (*TFCP2)* associated with GDM. We identified seven additional FDR-significant interactions of maternal methylation and GDM status, with the strongest association at the same cg22790973 probe (*TFCP2)*, plus cg03456133, cg24440941 (*H3C6*), cg20002843 (*LOC127841)*, cg19107264, cg11493553 located in the *UBE3C* gene and cg17065901 in *FAM13A,* both susceptibility genes for type 2 diabetes and BMI and cg23355087, within the *DLGAP2* gene, known to be involved in insulin resistance during pregnancy.

*Conclusion:* Our study reveals the potential complexity of the epigenetic transmission between GDM mothers and their offspring, likely determined by not only GDM

exposure, but also other factors indicated by maternal epigenetic status, such as maternal metabolic history.

## Introduction

Gestational diabetes mellitus (GDM) is a condition in which hyperglycaemia develops during pregnancy. GDM is one of the most common metabolic complications affecting 8-25 % of pregnancies and prevalence rates show considerable variation between countries depending on the test strategies used (1). Observational studies support evidence that GDM is associated with an increased future offspring risk of obesity and insulin resistance, key risk factors of type 2 diabetes and other cardiometabolic diseases, in adulthood (2–4), suggesting that hyperglycaemia exposure *in utero* has a direct adverse impact on offspring. There is also evidence for a genetic basis to GDM with a history of diabetes in either parent increasing risk of GDM at least twofold (5) and identified shared genetic risk variants that confer risk of both GDM and type 2 diabetes (6). Epigenetic mechanisms likely contribute to GDM, since evidence suggests that risks factors arise from prenatal exposure to a diabetic intrauterine environment (5,7). Targeted and candidate gene approaches have found several differentially methylated positions (DMPs) in leukocytes associated with offspring GDM exposure (8). However, the recent increase in technological advances have allowed for a more systematic and genome-wide approach for the identification of novel methylation *loci*. Indeed, these studies have identified several DMPs associated with GDM (9–11). The most recent epigenome-wide association study (EWAS) was a meta-analysis of cord blood DNA methylation from six birth cohorts, from nine countries, including a total of 3,677 newborn offspring, of which 317 were exposed to GDM. Despite the relatively large meta-analysis sample size, the authors did not observe any individual cord-blood CpG probes associated with GDM exposure using the false discovery rate (FDR) for testing significance (12). Here, we used the Finnish Gestational Diabetes (FinnGeDi) study cohort of 298 mother-offspring pairs with GDM

and 238 without, all with maternal blood and cord blood samples (13), in order to investigate whether epigenetic changes occur in response to GDM exposure during pregnancy, both in mothers and their offspring. Specifically, we tested two hypotheses, whether exposure to GDM is associated with: 1) epigenetic changes shared by mother and offspring pairs and 2) specific epigenetic changes in the offspring.

## Methods

### Cohort

We designed a case-control study as part of the FinnGeDi (Finnish Gestational Diabetes) prospective multicentre case-control cohort, including a total of 1,072 participants (536 mother/offspring pairs), of which 298 (55.6 %) had GDM during pregnancy. As described in detail (13), GDM women were recruited from delivery units when coming to give birth and the next woman without GDM was recruited as a control. Maternal venous blood and offspring cord blood samples were collected during delivery. Exclusion criteria were multifetal births, pre-pregnancy diabetes, or smoking during pregnancy. GDM was diagnosed following the Finnish Current Care Guideline (14), that is modified from the criteria of the International Association of the Diabetes and Pregnancy Study Groups (IADPSG). The Guideline uses oral glucose tolerance test (OGTT) with an oral dose of 75 g of glucose and the following cut-off values: fasting glucose ≥ 5.3 mmol/L (95 mg/dL), 1-hour glucose ≥ 10 mmol/L (180 mg/dL), or 2-hour glucose ≥ 8.6 mmol/L (155 mg/dL). Mothers without GDM and having normal OGTT results after the 24$^{th}$ week of gestation were eligible controls. Mothers in each group were carefully group-matched by age and pre-pregnancy body mass index (BMI) to select an equal number of women in each stratum. However, there were insufficient eligible control mothers in the higher age and BMI categories

(Supplementary Table 1). Therefore, a total of 238 women were selected as controls. A summary of the cohort characteristics is shown in Table 1. The samples were collected and approved following the ethics committee of Oulu University Hospital (reference number 33/2008).

Infinium MethylationEPIC BeadChip

DNA was extracted from whole blood for the mothers and from cord blood for the offspring using the DNeasy Blood & Tissue kit (Qiagen, Germany). Bilsulfite conversion of the genomic DNA was performed using the EZ-96 DNA Methylation kit (Zymo Research) following the manufacturer's protocols and bisulfite-converted DNA was subjected to a genome-wide DNA methylation analysis performed using Illumina's Infinium '850K' MethylationEPIC BeadChip array (Illumina, Inc., San Diego, CA, USA), which covers a total of 853,307 CpG sites. All the samples were randomised across the chips and analysed on the same machine by the same technician to reduce batch effects. After single base extension and staining, the BeadChips were imaged with the Illumina iScan. Raw fluorescence intensities of the scanned images were extracted with the GenomeStudio Methylation module (Illumina, Inc., San Diego, CA, USA). The fluorescence intensity ratio was used to calculate the $\beta_{value,}$ which corresponds to the methylation score for each analysed site according to the following equation: $\beta_{value} = I_M /(I_U + I_M + 100)$, where $I_M$ is the intensity of the methylated allele and $I_U$ the intensity of the unmethylated allele. DNA methylation $\beta_{value}$ ranges from zero (completely unmethylated) to one (completely methylated). All samples had high bisulfite conversion efficiency (signal intensity > 4000) and were included for further analysis based on GenomeStudio quality control. Quality-control was performed using the R software (version 4.0.0) (15).The DNA methylation IDAT files were imported using the R package *minfi* for pre-processing and quality-control (16).

The following probes were excluded from further analysis: probes with a detection p-value greater than or equal to 0.01 for at least 1 sample, cross-hybridizing probes (17), probes with a bead count less than 3 in at least 5 % of the samples, non-CpG probes, and probes which lie near single nucleotide polymorphisms (18). Probes on chromosomes X and Y were used for sex estimation and then excluded from downstream analyses. Quality control identified one offspring sex-discordant sample, which was excluded from further analyses. Samples with less than 99 % probes with a detection p-value lower than 0.01 were excluded, and one offspring sample was excluded with a call-rate < 99 %. Probe-design biases and batch effects were normalised using R packages *ENmix* (19) and *sva* (20), respectively. After quality-control, 1,070 samples (534 offspring and 536 mothers) and 724,671 CpG probes were available for further downstream analyses.

Additionally, as blood and cord blood samples are expected to include a variety of cell types, which might have a potential confounding effect on DNA methylation, cell composition was estimated based on whole blood (for mothers) and cord blood (for neonates) reference panels: the Bioconductor R *EPIC* (21) and *FlowSorted.CordBloodCombined.450k* (22).

Statistical Analyses

P-values were corrected for multiple testing using the false discovery rate (FDR) method from Benjamini-Hochberg (23). For all analyses, the $\beta_{value}$ (B) has been transformed into an $M_{value}$ (M) (24), where $M = \log_2(B/1\text{-}B)$. The relationship between B and M-values follows a logistic curve with, for example, B-values of 0.2, 0.5 and 0.8 corresponding to M-values of -2, 0 and 2, respectively.

Three sets of analyses were performed to test the two study hypotheses described in Figure 1. The first hypothesis was that both offspring and mothers might share the

same epigenetic effects as a result of shared GDM exposure. The second hypothesis tested for potential epigenetic effects that are specific or unique to the offspring as a result of GDM exposure. Equations for the implemented models are described in Supplementary Table 2. The three analyses implemented were:

1) EWAS for mothers and offspring: two separate EWAS were conducted to identify epigenetic changes associated with GDM exposure in both mothers and offspring, separately, using the R package *limma* (25). A linear regression was performed using CpG methylation levels to identify differentially methylated positions in offspring and mothers with GDM exposure compared to non-exposed controls. Offspring methylation probes were adjusted using the following covariates: sex, birthweight, gestational age at delivery for the offspring (Supplementary Figure 1) and for the mothers we adjusted methylation for age, gestational weight gain and pre-pregnancy BMI. Cell composition was adjusted for cord blood cell composition for the offspring (Supplementary Figure 1) and blood cell composition for the mothers (Supplementary Figure 2).

2) Shared epigenetic associations (pooled data model): a linear mixed model was performed to identify potentially shared epigenetic changes between the maternal and offspring methylome. This was performed using the combined methylome probe data for the mothers and the offspring, accounting for the correlation between each mother and their offspring, using the R package *lme4* (26). This test was performed to identify potential shared causal paths for the mother and offspring in response to *in utero* exposure to GDM that could explain a shared effect (Figure 1A). In the first step, a linear regression was performed on the methylome of the offspring and mothers using the same covariates described above. Cell composition was performed for offspring (cord blood) and mothers (blood), separately. Residuals from the first step were then

used in a mixed linear model as the trait of interest and regressed upon the explanatory variables of GDM, offspring/maternal status and interaction between GDM and offspring/maternal status.

3) Epigenetic changes specific to offspring: a linear regression was performed to identify offspring-specific effects in response to GDM exposure and maternal environment (*i.e.,* methylation levels). This test was performed to identify whether methylation changes of the offspring is determined not only by GDM status, but also by the maternal methylation levels (Figure 1B). To do this, the offspring methylome was linearly regressed on GDM exposure status, with the methylome of the mothers (pre-adjusted for age, BMI and blood cell composition) included as a covariate. In addition, an interaction term between GDM exposure and maternal methylation status at the same probe was included, in addition to sex, birthweight, gestational week and cord blood cell composition. In order to control for outlying genomic data points, offspring-mother pairs were excluded from this model, following Tukey's method, if the $M_{value}$ for either the offspring or the mother was lower than three times the interquartile range below the first quartile or higher than three times the interquartile range above the third quartile.

Bioinformatics analyses: The Genotype-Tissue Expression (GTEx) Portal (https://gtexportal.org/home/) was used to identify the expression patterns of genes in 44 tissues. The expression level was plotted using the median TPM.

**Results**

To investigate epigenetic changes that occur in response to maternal GDM exposure, we performed a whole methylome analysis in a total of 534 mother-offspring pairs from the Finnish Gestational Diabetes (FinnGeDi) study. We addressed two questions: 1/ whether there were epigenetic changes in response to GDM exposure that were

shared between mother and offspring, using a paired mother-offspring study design, and 2/ whether any individual DMPs associated with GDM exposure were specific to the offspring. We used two approaches to address the first question (Models 1 and 2, described above), both of which yielded the same results. We did not observe any FDR-significant DMPs associated with GDM exposure in the EWAS for both the mothers (n= 536) and offspring (n = 534), separately (Supplementary Table 3 and 4), consistent with recent findings in newborn GDM (12). Using the paired mother-offspring study design, following two separate EWAS for mother and offspring, we implemented a generalised linear mixed model for the combined methylation data to formally test for DMPs due to GDM exposure that are shared by mothers and their offspring. We found no strong evidence (FDR < 0.05) to support shared epigenetic changes associated with GDM exposure (Supplementary Figure 3).

Second, we used the same paired design in a linear regression to assess whether both maternal methylation status and GDM exposure had an effect on offspring methylation. Using this model, we observed a hypomethylation at the cg22790973 probe (average methylation: 1.03 %; $M_{value}$ estimate: -1·56 ± 0·27; FDR = 1.38 x $10^{-2}$) associated with GDM exposure (Figure 2A and 2B). This probe is located at the TSS1500 (1,500 base pairs upstream of the transcription start site) of the ubiquitously expressed Transcription Factor CP2 *(TFCP2)* gene (Supplementary Figure 4). This CpG probe is also located in a CpG island (chr12:51566680-51567072), which is a region in the genome with a high CpG density and has been previously shown to regulate gene expression (27).

To test for potential offspring specific effects, we also fitted an interaction term in our model between GDM exposure and maternal methylation (*i.e.,* the impact of GDM exposure in the context of the maternal methylation status) (Supplementary Table 5;

Figure 2C-E). We observed that the same locus (*TFCP2* cg22790973), showed the highest FDR-significant association between offspring and maternal methylation (average methylation: 1.03 %; $M_{value}$ estimate: 0.52 ± 0.08; FDR = 3.59 x $10^{-4}$). We observed that as maternal GDM methylation increases at this locus, so does the GDM-exposed offspring methylation, compared to a decrease in unexposed GDM offspring (Figure 3A and 3B).

We observed seven additional DMPs that reached FDR-significance for the interaction between maternal GDM status and the methylation of the mothers (Supplementary Figure 5; Supplementary Table 5). At the cg03456133 probe, located in an intergenic region, we found an increase in methylation of the GDM-exposed offspring as the maternal methylation increased, compared to a decrease in non-exposed controls (average methylation: 96.48 %; $M_{value}$ estimate: 0.56 ± 0.09; FDR = 1.46 x $10^{-3}$). We also found a similar trend at the cg19107264 probe (average methylation: 96.64 %; $M_{value}$ estimate: 0.57 ± 0.10; FDR = 7.37 x $10^{-3}$), located within the gene body of the *DLGAP2,* encoding Disks large-associated protein 2, which is a gene most strongly expressed in the brain (Supplementary Figure 6). We also found a decrease in methylation at the cg23355087 probe in GDM offspring as GDM mother's methylation increased, compared to non-exposed controls (average methylation: 95.66 %; $M_{value}$ estimate: -0.47 ± 0.08; FDR = 8.24 x $10^{-3}$), located in an intergenic region. A similar trend was found at the cg20002843 probe, located at the *LOC127841* gene body (average methylation: 47.37 %; $M_{value}$ estimate: -0.29 ± 0.05; FDR = 9.68 x $10^{-3}$) and at the cg24440941 probe (average methylation: 1.48 %; $M_{value}$ estimate: -0.49 ± 0.09; FDR = 1.24 x $10^{-2}$), located in the TSS1500 of *H3C6* (H3 Clustered Histone 6) (Supplementary Figure 7). At the cg17065901 probe, located at the *FAM13A* gene body, we found an increase in methylation as GDM mother's methylation increased,

compared to non-GDM exposed offspring (average methylation: 98.25 %; $M_{value}$ estimate: 0.60 ± 0.11; FDR = 1.76 x $10^{-2}$). The *FAM13A* gene is highly expressed in adipose tissue (Supplementary Figure 8). Lastly, for the cg11493553 probe located in the ubiquitously expressed *UBE3C* gene, we found a decrease in methlylation in the GDM-exposed offspring as maternal methylation increased, compared to the non-exposed controls (average methylation = 97.07 %; $M_{value}$ estimate: -0.50 ± 0.09; FDR = 2.41 x $10^{-2}$; Supplementary Figure 9).

**Discussion**

We present a comprehensive and in-depth study investigating the epigenetic associations in response to GDM exposure in mother-offspring pairs using the prospective FinnGeDi cohort. Our data did not support robust epigenetic associations for women and offspring exposed to GDM during pregnancy. However, in terms of epigenetic associations that are specific to offspring, our study has identified a potentially novel perspective in maternal transmission, which includes not only GDM as the exposure, but also maternal methylation, which could have an effect on offspring methylation.

Our study suggests that although there is no strong direct association between GDM exposure and the methylome of offspring, the context of maternal environment (*i.e.,* maternal methylome) may contribute to a multiplicative causal effect. As we included maternal methylation in our model, we were able to account for the possible modifying effect of maternal methylation. This is illustrated with the observed significant association of the methylation at *TFCP2* gene with offspring exposure to GDM. *TFCP2* is ubiquitously expressed and elucidating the function of this gene in various tissues has been a challenge. However, some studies have identified roles of *TFCP2* in reproduction and embryonic development (28), which suggests pleiotropic effects of

this transcription factor later in adulthood. We found that this CpG also had an interaction association between maternal methylation and GDM status. Although the methylation was lower in offspring exposed to maternal GDM, we found that the methylation at *TFCP2* in offspring increased as maternal methylation increased. Interestingly, the reverse was observed for the non-exposed controls. This suggests that the higher the maternal GDM methylation at this *locus*, the smaller the effect is on the GDM-exposed offspring. It is important to note that for each locus, extreme methylation values within the offspring and mothers could lead to bias, mainly in the interaction term, and were thus excluded. To avoid this bias, more than 20 % of the sample pairs were excluded for *TFCP2* cg22790973 probe, based on Tukey's method. Altogether, our data suggests that the maternal gestational environment, *i.e.*, GDM and the methylation status, has an impact on the offspring. This is demonstrated in seven further DMPs, where the interaction between the mother's methylation and GDM status were FDR-significant, within *DLGAP2, H3C6, FAM13A, LOC127841, UBE3C* and two *loci* within intergenic regions. These CpGs were only FDR-significant for the interaction association and not GDM alone.

Interestingly, a few of these genes have previously been identified to have a role in diabetes, namely *DLGAP2* and *FAM13A*. For both CpGs, we found a similar methylation trend as in *TFCP2*, *i.e.*, the methylation of the offspring was increased as a function of the methylation status of the GDM mothers, compared to the non-exposed offspring. A recent study found that methylation changes in the *DLGAP2* gene was associated with maternal insulin sensitivity in pregnancy (29). The authors found that a hypermethylation in the CpG was associated with a decrease in Matsuda index, which is a measure of insulin sensitivity. This is consistent with our data, where although not the same CpG, we found that a hypermethylation was associated with

GDM, suggesting that maternal methylation may mediate an effect on insulin sensitivity in future offspring risk.

Another study found that *DLGAP2* is involved in determining beta cell fate through AMP-activated protein kinase (AMPK) signalling (30), which has a role in maintaining insulin sensitivity and glucose homeostasis (31). Moreover, a recent study showed that *FAM13A* represses hepatic AMPK activity, thereby inducing insulin resistance in mice.(32) Single nucleotide polymorphisms in *FAM13A,* in addition to *UBE3C,* another gene identified in our analysis, have previously been found to be associated with type 2 diabetes (33,34), BMI and lipid traits (29,30), highlighting a possible role of their methylation in the future risk of obesity and type 2 diabetes in offspring. Interestingly, two separate meta-analyses in neonates identified methylation changes in both *FAM13A* and *DLGAP2* as associated with environmental changes during pregnancy, *i.e.,* prenatal air pollution and smoking, respectively (37,38).

Our results did not identify any individual CpGs directly associated with GDM exposure in our EWAS stratified for both offspring and mothers. This was consistent with our shared (*i.e.*, pooled) approach of mother-offspring pairs, which increased the sample size (as both mothers and offspring were included in our model) and thus increased the statistical power. Therefore, our study strongly suggests that for this relatively homogenous population, there is no strong association with GDM exposure alone, irrespective of maternal environmental factors.

This study exhibits several strengths. To the best of our knowledge, this is the largest, single cohort study investigating GDM epigenetic associations. Moreover, with the carefully selected mother-offspring pairs, we were able to investigate the effect of the maternal methylation on offspring in response to maternal GDM exposure. In addition, the individuals in our cohort are from a relatively homogeneous population and all were

non-smokers, thereby by-passing any confounding associations due to ethnicity or any obvious confounding health risk factors. As most published studies and cohorts have solely investigated offspring methylation, this study, indeed, warrants further investigations and provides a basis for follow up studies. In addition to strengths, this study has some limitations, which are important to consider. Given the absence of a comparable study design (*i.e*., sample size and paired mother-offspring cohort), our study does not include a replication cohort and, therefore, the results in this study should be interpreted with caution. Although our study is the largest, single population GDM epigenetic study, it is possible that we still did not have the power to detect an association with GDM exposure alone, which might be too small to detect at birth. It is important to note that our two-step approach, in addition to the different adjustments for the tissues in our study (*i.e.*, mother's blood and offspring cord blood), may have resulted in a loss of power in favour of reduced biases. It is also important to note that Finland has very high standards for GDM care and maternal glucose levels are likely to be well controlled. As a result, any shared effects of GDM exposure on mother and offspring methylome might be modest, despite the relatively large sample size for our study. Although our study design matched individuals by age and pre-pregnancy BMI, our cohort had insufficient numbers of controls with higher BMI. However, these individuals were included as they were representative of patients in Finnish clinics.

The clinical and biological effects of these loci need to be further investigated in target tissues to explore the biological and clinical relevance of this finding. Previous evidence demonstrates that changes in DNA methylation in the blood, although being well associated with tissue specific changes, can be less intense. Indeed, in 2016, a study reported that methylation changes in the pancreatic islet were associated with age, implicating genes that had important roles in the pancreatic islet and insulin

secretion (39). The majority of these associations were also observed, to a lesser extent, in blood. So although we observed small methylation differences in cord blood, it could be speculated that this change could also occur with more substantial effects in target tissues. Therefore, in our study, it is indeed possible that the effects in target genes are more pronounced and their role needs to be further elucidated, particularly for the very small effect size observed for *TFCP2*, which is a ubiquitous gene and may have multiple roles in different tissues.

The biological significance of the *loci* we have identified to interact, suggesting a moderation by maternal methylation, are open to interpretation. There are precedents to our study that support plausible biological significance. For instance, one study observed a similar two-hit hypothesis where the level of an individual's MTHFR enzyme activity was determined by an interaction between DNA methylation status at the MTHFR gene and folate levels. The authors observed that individuals with compromised enzyme activity had the same methylation levels if the subjects were folate sufficient (and vice versa). Therefore, only individuals with both alterations resulted in depletion of genome-wide methylation levels.

**Conclusions**

To our knowledge, to date this is the largest GDM methylation study for a single cohort with access to phenotypic and methylation data from both mothers and offspring. The current study complements previous EWAS studies by incorporating the exposure and/or possible modulating effects of maternal DNA methylation status and found that there is no strong direct association between GDM exposure and the methylome of offspring, however, the context of maternal environment may contribute to a moderating effect. Our study reveals the potential complexity of the epigenetic transmission between GDM mothers and their offspring, likely determined by not only

GDM exposure, but also other factors indicated by maternal epigenetic status, which are involved in establishing epigenetic signatures in offspring.

Acknowledgements:

The study was designed by T.A., P.F. and M.C. S.L. performed the wet lab experiments to generate the methylationEPIC array data. M.V., E.Ka., S.M. and E.Ke provided clinical data. A.B., S.H., M.V., S.S., F.B. and E.Ka., E.T., M.R.J., P.F. and T.A. contributed data. M.C. has verified underlying data. M.C. has performed the statistical analyses and generated the associated figures and tables. A.K. has performed the literature search, results interpretation and written the manuscript with M.C. All authors critically reviewed and edited the manuscript. T.A. is the guarantor of this study. The authors declare that they have no competing interests.

This study was funded by the French National Research Agency (Agence Nationale de la Recherche; ANR). This study has been granted from the FFRD, sponsored by Fédération Française des Diabétiques (FFD), Abbott, AstraZeneca, Eli Lilly, Merck Sharp & Dohme (MSD) and Novo Nordisk. The co-authors were supported by PRECISE, LONGITOOLS and The Medical Research Council, UK (grants no. MR/M013138/1, MRC/BBSRC MR/S03658X/1 (JPI HDHL)). FinnGeDi study was supported by: Academy of Finland, Signe and Ane Gyllenberg Foundation, Sigrid Jusélius Foundation, Foundation for Pediatric Research, Finnish Diabetes Research Foundation, Novo Nordisk Foundation.

Data sharing statement:

The 850K methylationEPIC array data will be made available upon request to Eero Kajantie, Philippe Froguel or Toby Andrew.

Code availability:

Code to perform the analyses in this manuscript are available at https://github.com/umr1283/EpxGDM (https://doi.org/10.5281/zenodo.4709136).

References:


1. Guariguata L, Linnenkamp U, Beagley J, Whiting DR, Cho NH. Global estimates of the prevalence of hyperglycaemia in pregnancy. Diabetes Res Clin Pract. 2014 Feb;103(2):176–85.

2. Boerschmann H, Pflüger M, Henneberger L, Ziegler A-G, Hummel S. Prevalence and predictors of overweight and insulin resistance in offspring of mothers with gestational diabetes mellitus. Diabetes Care. 2010 Aug;33(8):1845–9.

3. Nolan CJ, Damm P, Prentki M. Type 2 diabetes across generations: from pathophysiology to prevention and management. Lancet Lond Engl. 2011 Jul 9;378(9786):169–81.

4. Vääräsmäki M, Pouta A, Elliot P, Tapanainen P, Sovio U, Ruokonen A, et al. Adolescent manifestations of metabolic syndrome among children born to women with gestational diabetes in a general-population birth cohort. Am J Epidemiol. 2009 May 15;169(10):1209–15.

5. Williams MA, Qiu C, Dempsey JC, Luthy DA. Familial aggregation of type 2 diabetes and chronic hypertension in women with gestational diabetes mellitus. J Reprod Med. 2003 Dec;48(12):955–62.

6. Kwak SH, Kim S-H, Cho YM, Go MJ, Cho YS, Choi SH, et al. A genome-wide association study of gestational diabetes mellitus in Korean women. Diabetes. 2012 Feb;61(2):531–41.

7. 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 Sep;8(9):935–43.

8. El Hajj N, Pliushch G, Schneider E, Dittrich M, Müller T, Korenkov M, et al. Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. Diabetes. 2013 Apr;62(4):1320–8.

9. Hjort L, Martino D, Grunnet LG, Naeem H, Maksimovic J, Olsson AH, et al. Gestational diabetes and maternal obesity are associated with epigenome-wide methylation changes in children. JCI Insight [Internet]. 2018 Sep 6 [cited 2020 Aug 3];3(17). Available from: https://insight.jci.org/articles/view/122572

10. Haertle L, El Hajj N, Dittrich M, Müller T, Nanda I, Lehnen H, et al. Epigenetic signatures of gestational diabetes mellitus on cord blood methylation. Clin

Epigenetics [Internet]. 2017 Mar 27 [cited 2020 Aug 3];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5368916/

11. Dias S, Adam S, Rheeder P, Louw J, Pheiffer C. Altered Genome-Wide DNA Methylation in Peripheral Blood of South African Women with Gestational Diabetes Mellitus. Int J Mol Sci [Internet]. 2019 Nov 20 [cited 2020 Aug 3];20(23). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6928622/

12. Howe CG, Cox B, Fore R, Jungius J, Kvist T, Lent S, et al. Maternal Gestational Diabetes Mellitus and Newborn DNA Methylation: Findings From the Pregnancy and Childhood Epigenetics Consortium. Diabetes Care. 2020;43(1):98–105.

13. Keikkala E, Mustaniemi S, Koivunen S, Kinnunen J, Viljakainen M, Männisto T, et al. Cohort Profile: The Finnish Gestational Diabetes (FinnGeDi) Study. Int J Epidemiol. 2020 Jun 1;49(3):762–763g.

14. Working Group Established by the Finnish Medical Society Duodecim, the Medical Advisory Board of the Finnish Diabetes Association and the Finnish Gynecological Association. Current Care Guideline. Gestational Diabetes. Helsinki: Finnish Medical Society Duodecim. 2008.

15. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.R-project.org/

16. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinforma Oxf Engl. 2014 May 15;30(10):1363–9.

17. Nordlund J, Bäcklin CL, Wahlberg P, Busche S, Berglund EC, Eloranta M-L, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 2013;14(9):r105.

18. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2016 Oct;gkw967.

19. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic Acids Res. 2016;

20. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012 Mar 15;28(6):882–3.

21. Salas LA, Koestler DC. FlowSorted.Blood.EPIC: Illumina EPIC data on immunomagnetic sorted peripheral adult blood cells [Internet]. 2020. Available from: https://github.com/immunomethylomics/FlowSorted.Blood.EPIC

22. Salas LA, Gervin K, Jones MC. FlowSorted.CordBloodCombined.450k: Illumina 450k/EPIC data on FACS and MACS umbilical blood cells [Internet]. 2020.

Available from:
https://github.com/immunomethylomics/FlowSorted.CordBloodCombined.450k

23. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995 Jan;57(1):289–300.

24. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010 Dec;11(1):587.

25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47–e47.

26. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw. 2015 Oct 7;67(1):1–48.

27. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011 May 15;25(10):1010–22.

28. Taracha A, Kotarba G, Wilanowski T. Neglected Functions of TFCP2/TFCP2L1/UBP1 Transcription Factors May Offer Valuable Insights into Their Mechanisms of Action. Int J Mol Sci [Internet]. 2018 Sep 20 [cited 2020 Aug 3];19(10). Available from:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6213935/

29. Hivert M-F, Cardenas A, Allard C, Doyon M, Powe CE, Catalano PM, et al. Interplay of Placental DNA Methylation and Maternal Insulin Sensitivity in Pregnancy. Diabetes. 2020;69(3):484–92.

30. Kone M, Pullen TJ, Sun G, Ibberson M, Martinez-Sanchez A, Sayers S, et al. LKB1 and AMPK differentially regulate pancreatic β-cell identity. FASEB J Off Publ Fed Am Soc Exp Biol. 2014 Nov;28(11):4972–85.

31. Coughlan KA, Valentine RJ, Ruderman NB, Saha AK. AMPK activation: a therapeutic target for type 2 diabetes? Diabetes Metab Syndr Obes Targets Ther. 2014 Jun 24;7:241–53.

32. Lin X, Liou Y-H, Li Y, Gong L, Li Y, Hao Y, et al. FAM13A Represses AMPK Activity and Regulates Hepatic Glucose and Lipid Metabolism. iScience [Internet]. 2020 Feb 22 [cited 2020 Aug 3];23(3). Available from:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7063182/

33. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes. 2017 Nov;66(11):2888–902.

34. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet. 2018 Nov;50(11):1505.

35. Ji Y, Yiorkas AM, Frau F, Mook-Kanamori D, Staiger H, Thomas EL, et al. Genome-Wide and Abdominal MRI Data Provide Evidence That a Genetically Determined Favorable Adiposity Phenotype Is Characterized by Lower Ectopic Liver Fat and Lower Risk of Type 2 Diabetes, Heart Disease, and Hypertension. Diabetes. 2019;68(1):207–19.

36. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Mägi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. 2015 Feb 12;518(7538):187–96.

37. Gruzieva O, Xu C-J, Yousefi P, Relton C, Merid SK, Breton CV, et al. Prenatal Particulate Air Pollution and DNA Methylation in Newborns: An Epigenome-Wide Meta-Analysis. Environ Health Perspect. 2019;127(5):57012.

38. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. Am J Hum Genet. 2016 Apr 7;98(4):680–96.

39. Bacos K, Gillberg L, Volkov P, Olsson AH, Hansen T, Pedersen O, et al. Blood-based biomarkers of age-associated epigenetic changes in human islets associate with insulin secretion and diabetes. Nat Commun. 2016 Mar 31;7:11089.

40. Sankilampi U, Hannila M-L, Saari A, Gissler M, Dunkel L. New population-based references for birth weight, length, and head circumference in singletons and twins from 23 to 43 gestation weeks. Ann Med. 2013 Sep;45(5–6):446–54.

**Table 1: Cohort characteristics of mother/offspring pairs used in this study**

| Characteristic | Control (N = 238)[1] | GDM (N = 298)[1] | p-value[2] |
|---|---|---|---|
| **Age of the mother (years)** | 31.5 (5.2) | 32.5 (5.3) | 0.042 |
| **BMI of the mother (kg/m²)** | 25.6 (4.8) | 27.9 (6.1) | <0.001 |
| **Gestation (weeks)** | 40.32 (1.19) | 39.60 (1.33) | <0.001 |
| **Sex of the offspring** | | | 0.3 |
| *Male* | 109 (46) | 152 (51) | |
| *Female* | 129 (54) | 146 (49) | |
| **Birthweight (g)** | 3,703 (473) | 3,705 (474) | >0.9 |
| **Birthweight SD score** | 0.15 (1.00) | 0.38 (1.08) | 0.011 |
| **Primipara (yes)** | 119 (50) | 127 (43) | 0.10 |
| **Previous deliveries** | 1.28 (2.29) | 1.46 (2.27) | 0.4 |
| **Mode of delivery (vaginal)** | 198 (83) | 245 (82) | 0.8 |
| **Large gestational age 90[3]** | 24 (10) | 50 (17) | 0.032 |
| **Small gestational age 90[4]** | 15 (6.3) | 20 (6.7) | >0.9 |
| **Maternal weight gain (kg)** | 15.1 (5.6) | 11.9 (5.6) | <0.001 |
| *Unknown* | 4 | 20 | |
| **Participant's mother GDM same pregnancy (yes)** | 5 (2.5) | 9 (3.5) | 0.6 |
| *Unknown* | 37 | 44 | |
| **Participant's mother GDM any pregnancy (yes)** | 9 (100) | 23 (100) | >0.9 |
| *Unknown* | 229 | 275 | |
| **Mother's father diabetes (yes)** | 30 (100) | 48 (100) | >0.9 |
| *Unknown* | 208 | 250 | |
| **Mother's mother diabetes (yes)** | 13 (100) | 50 (100) | >0.9 |
| *Unknown* | 225 | 248 | |
| **Maternal socioeconomic status** | | | 0.4 |
| *1 = Highest* | 53 (28) | 65 (26) | |
| *2* | 73 (39) | 112 (45) | |
| *3* | 13 (7.0) | 22 (8.8) | |
| *4* | 47 (25) | 50 (20) | |
| *Unknown* | 52 | 49 | |
| **Maternal education** | | | 0.027 |
| *1 = Basic* | 2 (0.9) | 11 (3.9) | |
| *2 = Secondary* | 81 (37) | 127 (45) | |
| *3 = Lower-level tertiary* | 71 (33) | 77 (28) | |
| *4 = Upper-level tertiary* | 64 (29) | 65 (23) | |
| *Unknown* | 20 | 18 | |

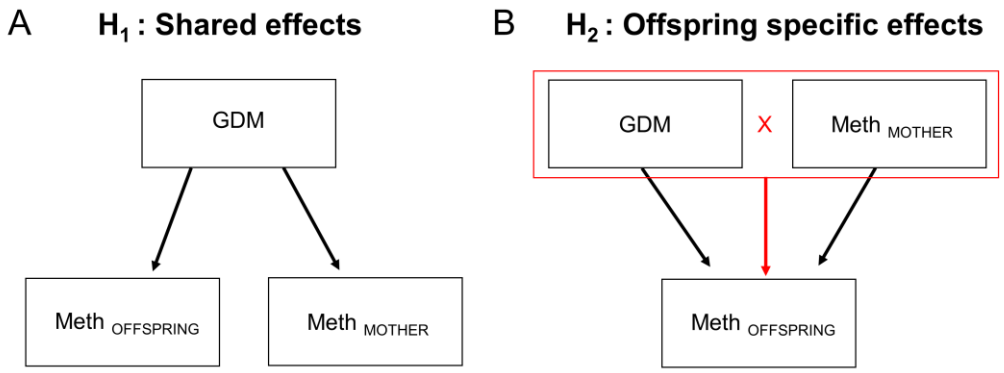
[1]Mean (SD); n (%)
[2]Welch Two Sample t-test; Fisher's exact test
[3] > + 2 SD (40)
[4] < - 2 SD (40)

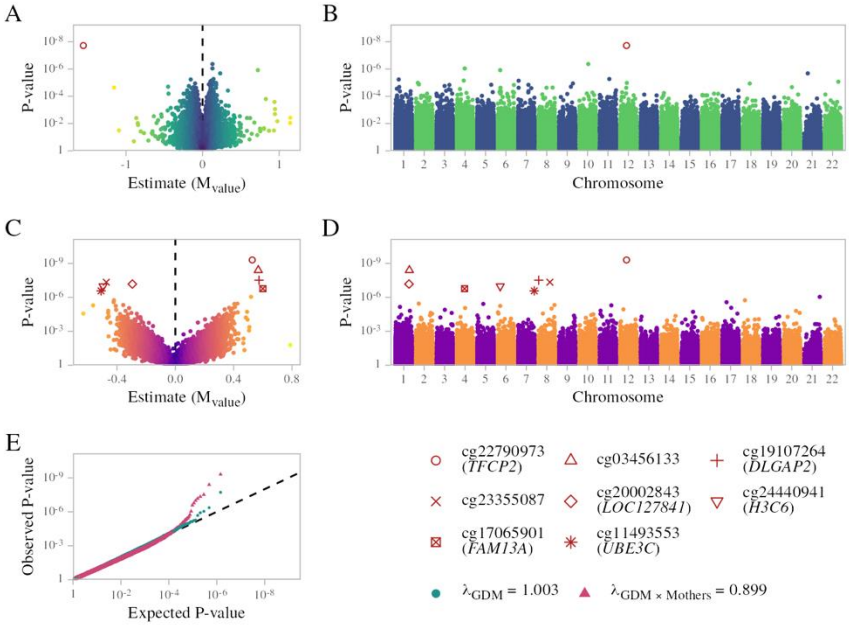
**Figure Legends:**

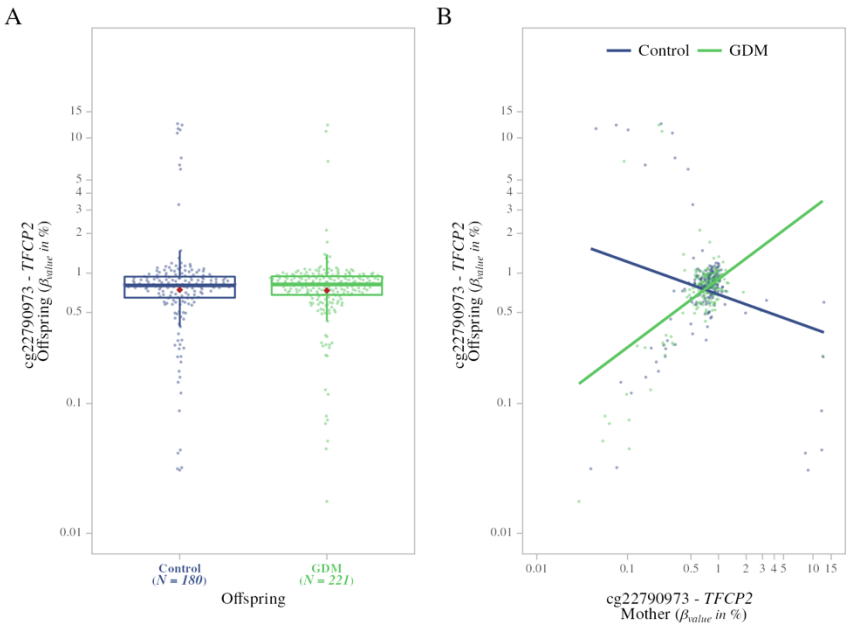
**Figure 1. The investigated study hypotheses. A) Hypothesis 1:** Assess whether there are epigenetic effects that are shared by both offspring and their mother as a result of shared exposure to GDM (offspring and maternal methylation regressed upon GDM status; see Methods); **B) Hypothesis 2:** Test if there are epigenetics effects as a result of exposure to GDM that are specific to the offspring (offspring DNA methylation regressed upon GDM status, maternal methylation for the same probe and interaction between GDM and maternal methylation status; see Methods).

**Figure 2. Summary results for offspring EWAS associations.** Linear model between offspring methylation and exposure to maternal GDM, including the methylation of mothers in the model, and the interaction between GDM exposure and maternal methylation, adjusted for offspring sex, gestational week, birthweight and cell composition. **A)** A volcano plot for offspring probe differential methylation by GDM exposure. cg22790973 has an estimate = -1.56 (FDR = $1.38 \times 10^{-2}$), equivalent to a $\beta_{value}$ of 1.03 %. **B)** Manhattan plot for the GDM exposure main effect, showing the genome-wide results for all the CpGs. **C)** A volcano plot of the GDM exposure interaction effect. **D)** Manhattan plot of the GDM exposure interaction effect, showing the genome-wide results for all the CpGs. **E)** Probability-Probability plot of the GDM exposure main effect (Green) and interaction term (Red) on the methylation of the offspring, with the black line indicating the expected distribution.

**Figure 3: Differential methylation observed for offspring at probe cg22790973 (*TFCP2*).** A) Box plot showing the positive methylation differences between offspring exposed to maternal GDM compared to non-exposed controls. B) Scatterplot showing increased methylation at this probe for offspring exposed to maternal GDM along with increased maternal methylation at the same probe. For non-exposed offspring, methylation decreases with increased maternal methylation at this probe.


A
H1 : Shared effects
GDM
Meth OFFSPRING
Meth MOTHER
B
H2 : Offspring specific effects
GDM
X
Meth MOTHER
Meth OFFSPRING


A
B
C
D
E
P-value
Estimate (M_value)
Chromosome
Observed P-value
Expected P-value
cg22790973 (TFCP2)
cg03456133
cg19107264 (DLGAP2)
cg23355087
cg20002843 (LOC127841)
cg24440941 (H3C6)
cg17065901 (FAM13A)
cg11493553 (UBE3C)
λ-GDM = 1.003
λ-GDM × Mothers = 0.899


A
B
Control
GDM
cg22790973 - TFCP2
Offspring ($\beta_{value}$ in %)
15
10
5
4
3
2
1
0.5
0.1
0.01
Control
(N = 180)
GDM
(N = 221)
Offspring
0.01
0.1
0.5
1
2
3
4
5
10
15
cg22790973 - TFCP2
Mother ($\beta_{value}$ in %)