Human Molecular Genetics, 2015, Vol. 24, No. 11 3021–3029

doi: 10.1093/hmg/ddv013
Advance Access Publication Date: 29 January 2015
Original Article

ORIGINAL ARTICLE

Maternal gestational diabetes is associated with genome-wide DNA methylation variation in placenta and cord blood of exposed offspring

Sarah Finer[1,2,*], Chris Mathews[1,†], Rob Lowe[1,†], Melissa Smart[1], Sara Hillman[3], Lin Foo[1], Ajay Sinha[1], David Williams[3], Vardhman K. Rakyan[1] and Graham A. Hitman[1]

[1]Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK, [2]Wellcome Trust-MRC Institute of Metabolic Sciences, University of Cambridge, Cambridge CB2 0QQ, UK and [3]Institute for Women's Health, University College London, London WC1E 6AU, UK

*To whom correspondence should be addressed at: University of Cambridge Metabolic Research Laboratories, Wellcome Trust-MRC Institute of Metabolic Science, Level 4, Box 289, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK, Tel: +44 1223336792; Fax: +44 1223330598; Email: sf494@medschl.cam.ac.uk

Abstract

Exposure of a developing foetus to maternal gestational diabetes (GDM) has been shown to programme future risk of diabetes and obesity. Epigenetic variation in foetal tissue may have a mechanistic role in metabolic disease programming through interaction of the pregnancy environment with gene function. We aimed to identify genome-wide DNA methylation variation in cord blood and placenta from offspring born to mothers with and without GDM. Pregnant women of South Asian origin were studied and foetal tissues sampled at term delivery. The Illumina HumanMethylation450 BeadChip was used to assay genome-wide DNA methylation in placenta and cord blood from 27 GDM exposed and 21 unexposed offspring. We identified 1485 cord blood and 1708 placenta methylation variable positions (MVPs) achieving genome-wide significance (adjusted *P*-value <0.05) with methylation differences of >5%. MVPs were disproportionately located within first exons. A bioinformatic co-methylation algorithm was used to detect consistent directionality of methylation in 1000 bp window around each MVP was observed at 74% of placenta and 59% of cord blood MVPs. KEGG pathway analysis showed enrichment of pathways involved in endocytosis, MAPK signalling and extracellular triggers to intracellular metabolic processes. Replication studies should integrate genomics and transcriptomics with longitudinal sampling to elucidate stability, determine causality for translation into biomarker and prevention studies.

Introduction

Gestational diabetes mellitus (GDM) is a common, asymptomatic metabolic disorder of pregnancy manifest by maternal impaired glucose tolerance from the late second trimester of pregnancy onwards. The prevalence of gestational diabetes has been estimated at 2–10% of pregnancies in the USA (1), and 3.5% in the UK (2). Risk factors for gestational diabetes include ethnicity (South Asian, black Caribbean, Middle-Eastern), obesity, previous gestational diabetes, first degree family history of diabetes (3) and highlight common genetic risk between type 2 diabetes and gestational diabetes (4).

Glucose intolerance in gestational diabetes is a result of peripheral insulin resistance and failure of β-cell expansion and

† These two authors contributed equally to the work.

Received: September 1, 2014. **Revised and Accepted:** January 16, 2015

Table 1. Maternal phenotypic and metabolic characteristics

	Controls ($n = 28$) Mean (SD)		GDM ($n = 21$) Mean (SD)		*t*-test
Maternal variables					
Gestation at study visit (weeks)	32 (2.8)		33 (3.0)		n.s.
Age (years)	27 (4.9)		31 (4.3)		**
Body mass index (m/kg^2)	28 (6.1)		27 (5.0)		n.s.
Parity (n)	1.3 (1.4)		1.7 (1.4)		n.s.
Systolic blood pressure at booking (mmHg)	112 (15.0)		111 (11.4)		n.s.
Diastolic blood pressure at booking (mmHg)	72 (12.7)		73 (11.3)		n.s.
Glucose tolerance test—0 min (mmol/L)	4.5 (0.5)		5.2 (1.1)		*
Glucose tolerance test—120 min (mmol/L)	5.6 (0.9)		9.5 (1.7)		***
Fasting insulin	19.1 (13.0)		15.1 (10.7)		n.s.
Offspring variables					
Birth weight (g)	3010 (400.8)		3151 (438.8)		n.s.
Gestational age at delivery (days)	272 (9.1)		272 (10.6)		n.s.
Ponderal index (weight/height^3)	28.5 (4.3)		30.1 (4.9)		n.s.
Cord blood haematocrit (L/L)	0.47 (0.07)		0.46 (0.06)		n.s.
Offspring sex (n)	15M, 13F		14M, 7F		n.s. (χ^2 test)
	Mean (SD)	Median (SEM)	Mean (SD)	Median (SEM)	
Customized birth weight centile	33.4 (27.9)	27.0 (6.1)	46.1 (27.0)	49.3 (5.1)	n.s.

Mean and standard deviation are presented, and where means are significantly different *$P < 0.05$, **$P < 0.005$, ***$P < 0.0005$, n.s. = $p \geq 0.05$. Where the mean and median are presented, data are non-normally distributed and *t*-tests are performed on log-transformed data. n.s. = $P \geq 0.05$.

maternal insulin production to cope with ensuing hyperglycaemia. Maternal glucose crosses the placental interface but insulin does not, leading to foetal hyperglycaemia and hyperinsulinaemia. Foetal hyperinsulinism results in accelerated foetal growth, leading to macrosomia and organomegaly and neonatal hypoglycaemia. Additional disruptions to materno-placento-foetal nutrient transfer in GDM include altered transit of amino acids (5), lipids (6) and fatty acids (7) and altered vascular and endothelial function in placenta (8).

The altered pregnancy environment in gestational diabetes is thought to confer future risk of diabetes and obesity to the developing foetus through 'developmental programming'. Multiple cross-sectional studies have identified an association between maternal gestational diabetes (GDM) and a phenotype of diabetes and obesity in offspring. For example, studies in the Pima Indian population of the USA showed that offspring born to women with gestational diabetes developed fasting hyperglycaemia, higher rates of abnormal glucose tolerance and obesity, compared with controls (9). In a rat model, an induced phenotype of maternal hyperglycaemia in the third trimester of pregnancy induces a programmed phenotype of glucose intolerance and impaired insulin secretion in F1 offspring and similar consequences transmitted to the F2 generation, independent of genetic risk (10). Additional human studies have employed careful study design to try and minimize genetic confounding in studies of foetal programming via GDM and their evidence also supports that non-genetic factors are involved (11,12).

The study of epigenetics is providing insight into mechanisms by which genetic and environmental factors can interact to predispose to disease, including in animal and human models of foetal programming (13–17). Recently, studies of target gene regions have identified epigenetic modifications in foetal tissues from women with obesity and gestational diabetes (18) and according to quantitative metabolic traits in pregnancy (19). Additional evidence that epigenetic factors may play a role in the nutritional environment of pregnancy comes from a study of genes encoding high-affinity glucose transporters (GLUT1 and GLUT3) in placenta which show a relationship between DNA methylation and their gene expression that could impact glucose flux between the maternal and foetal circulation (20).

To date, most studies identifying epigenetic variation in association with adverse in utero exposures have employed targeted approaches to study DNA methylation at specific genes of interest and show small methylation differences between exposed and unexposed are small (<5% in case–control analyses and correlations of ≤0.35 with quantitative traits) (17,19,21). While these studies provide useful observations, they do not offer insights into complex processes of genomic regulation and are biased towards regions of genetic interest. A genome-wide approach to studying epigenetic variation offers an opportunity to investigate gene–environment interactions, unbiased by *a priori* hypotheses and thus may provide useful insight into the mechanisms of foetal programming. To date, one genome-wide study of DNA methylation in offspring exposed to gestational diabetes has been performed, but the differences it detected did not hold up to false discovery rate control due to presumed insufficient sample size and power to detect differences (22).

This study has been designed as a discovery-based investigation of genome-wide epigenetic variation in foetal tissues exposed to maternal GDM to provide focus for future studies of gene function and longitudinal stability of marks and association with phenotype. UK-based South Asians were studied due to their high prevalence of gestational diabetes and excess contribution to the worldwide risk of type 2 diabetes and cardiovascular disease in home and migrant populations (23).

Results

Women with gestational diabetes ($n = 28$) and controls ($n = 21$) were recruited in the third trimester of pregnancy. Women in the gestational diabetes (GDM) group were, on average, 4 years older than controls, but had no significant phenotypic differences (Table 1). Of those mothers with GDM, the majority (74%) were diet-controlled, and the remaining were treated with insulin alone (19%) or the combination of metformin and insulin. Offspring born to control and GDM mothers were no different

with respect to the mean birth weight or ponderal index (Table 1). Customized birth weight centiles, which adjust for gestational age and sex and maternal factors, based on UK population-wide data (24), show a trend towards higher customized birth weight centile in GDM offspring compared with controls (median 49.3 versus 27.0), but this difference did not reach statistical significance. Analysis of cord blood samples did not show erythrocytosis or difference in cellular composition (Supplementary Material, Table S1) (25). There was no difference in gestational age at delivery of control and GDM offspring, and they were all born at term (>255 days). Six placenta (3 GDM, 3 control) samples and one cord blood (GDM) sample were removed as they failed quality control checks after DNA extraction and bisulphite conversion.

Analysis of genome-wide methylation data was performed on placenta ($n = 25$ versus 18) and cord blood ($n = 27$ versus 21) samples from offspring born to GDM and control mothers. Of 485 512 probes present per array, the following probes were removed: 40 590 cross hybridizing, 11 648 X and Y chromosome, 1661 (placenta) and 1805 (cord blood) failing to meet the detection P-value cut-off ($P \geq 0.05$), resulting in 431 613 (placenta) and 431 469 (cord blood) probes at which CpG methylation data could be interrogated.

Significant genome-wide methylation differences were identified in placenta and cord blood. In placenta, methylation variable positions ('MVPs') were identified at 4219 probes where significant genome-wide statistical difference was achieved (FDR-corrected $P < 0.05$) and of these, 1373 MVPs represented a β-value difference of >5% (Table 2). In cord blood, 13 561 MVPs were identified ($P < 0.05$, FDR-corrected); of which, 1418 had a β-value difference of >5% (Table 2). A subset of 378 MVPs were common to both placenta and cord blood. A summary of the top MVPs, selected by FDR-adjusted P-value, in cord blood and placenta is presented in Table 3 (for full listings of MVPs, see Supplementary Material, S2).

Table 2. Summary table of number of MVPs in cord blood and placenta, and their proportional distribution in relationship to all array probes and mean methylation (β-value) density

	Mean beta value across all samples			Total
	<0.3	≥0.3 to <0.7	≥0.7	
Placenta				
All array probes	169 752	105 051	156 810	431 613
Probes as a proportion of total array (%)	39	24	36	
MVPs (*n*)	374	733	266	1373
MVPs as proportion of total array (%)	27	53	19	
Cord blood				
All array probes (*n*)	179 000	41 152	211 317	431 469
Probes as a proportion of total array (%)	41	10	49	
MVPs (*n*)	454	473	491	1418
MVPs as proportion of total array (%)	32	33	35	

MVPs are determined as probes where methylation variation between GDM-exposed and unexposed reaches FDR-corrected $P < 0.05$ and β-value difference >5%. Observed versus expected distribution of MVPs was significantly different in both placenta (χ^2 46.8, df = 2, $P < 0.0001$) and cord blood (χ^2 58.9; df = 2, $P < 0.001$).

Table 3. Summary of the top 10 MVPs (selected by FDR-adjusted P-value) in cord blood and placenta

	Probe name	Chr	Position	Gene name	FDR-adjusted P-value	% Methylation difference (GDM versus controls)
Cord blood	cg06717289	15	75978121	CSPG4	1.04E−06	8.7
	cg06355908	10	134897731		2.16E−06	6.0
	cg14910854	2	132254702	LOC150776	2.16E−06	9.4
	cg09060090	9	15510150	PSIP1	2.16E−06	−9.7
	cg00820581	1	174969299	CACYBP	2.89E−06	−5.7
	cg10633906	7	1595602	TMEM184A	2.89E−06	8.1
	cg19893717	12	62653989	USP15	2.89E−06	−5.3
	cg00221035	7	7222342	C1GALT1	2.92E−06	−7.9
	cg14558566	5	141016453	RELL2;HDAC3	3.16E−06	−5.9
	cg14728024	6	90543432	CASP8AP2	3.16E−06	14.9
Placenta	cg25192619	6	37467404	C6orf129	3.21E−03	−7.3
	cg00591353	16	83670892	CDH13	3.21E−03	−6.9
	cg27098574	16	75285489	BCAR1	3.21E−03	−7.0
	cg09309286	17	32960898	TMEM132E	3.21E−03	10.7
	cg13898647	9	136275791	REXO4	3.27E−03	5.6
	cg19085699	1	231004314	C1orf198	3.82E−03	−9.0
	cg10680376	10	116391843	ABLIM1	3.82E−03	−6.7
	cg11062956	16	2723328		3.82E−03	−8.3
	cg16897005	1	27962412	FGR	3.82E−03	6.0
	cg26714514	8	81889651	PAG1	3.82E−03	6.4

These 'top hit' MVPs have the most significant FDR-adjusted P-values, selected per tissue type. The table shows the Illumina 450k probe and chromosome position as well as the gene name where relevant. The methylation differences at these top hits are shown by comparing GDM with controls.

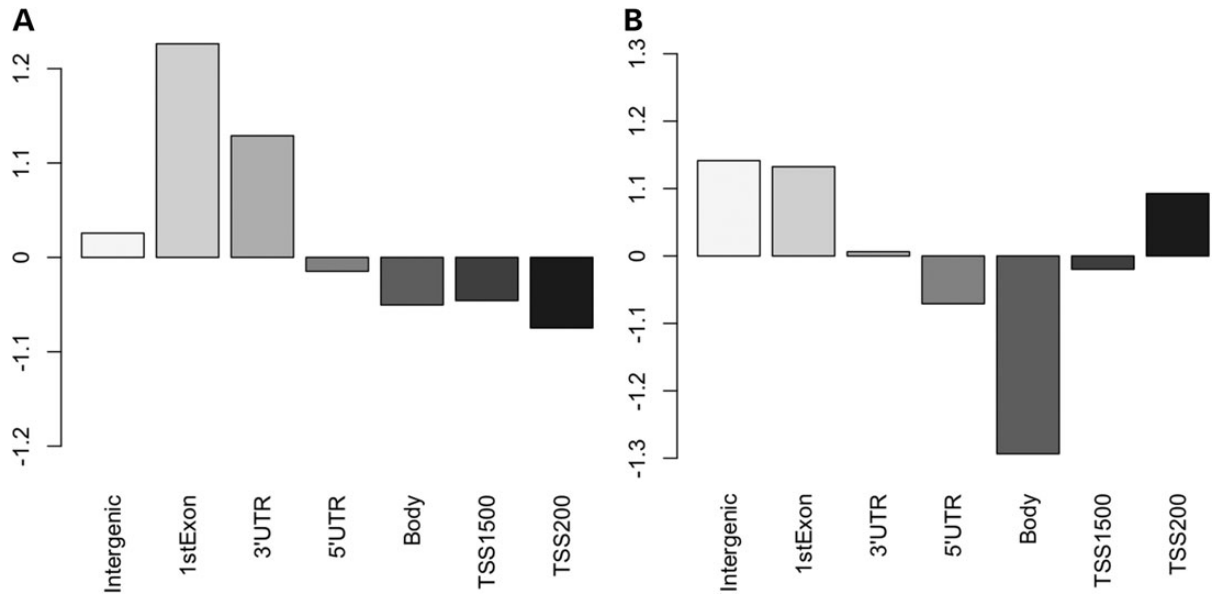

Figure 1. (**A**) Placenta and (**B**) cord blood. Enrichment of methylated variable positions against background according to genomic feature. Y-axis denotes enrichment of genomic features (+1) or depletion of genomic features (–1) in MVPs compared with background (*P-value < 0.05). X-axis denotes genomic features, defined by Illumina annotation. (TSS1500, within 1.5kB of transcriptional start site; TSS200, within 200 bp of transcriptional start site.)

To investigate the effect of cellular heterogeneity on cord blood MVPs, we used an *in silico* method (26) to calculate the predicted levels of different blood cell types from probe data and compared these with actual blood cell type levels, where available (Supplementary Material, Table S1). This comparison showed good correlation for both lymphocyte ($r^2 = 0.30$, $P = 0.0012$) and monocyte ($r^2 = 0.45$, $P = 1.7 \times 10^{-5}$) levels but not for granulocyte ($r^2 = 0.057$, $P = 0.18$) levels (Supplementary Material, Fig. S1). We therefore re-calculated MVPs for cord blood using these calculated levels of blood subtypes (Supplementary Material, Methods) and found 12 454 MVPs were identified ($P < 0.05$, FDR-corrected). Of these, 9806 were in common with the unadjusted MVPs and 1258 were in common with those which had a β-value difference of >5%. As there are no reference-based data sets available for placenta, it was not possible to use the same approach for correcting for cellular heterogeneity in the placenta samples and hence we applied a reference-free approach (27). Using this method, we identified 1 852 MVPs ($P < 0.05$, FDR) of which 1068 were in common with the unadjusted MVPs and 325 were in common with those that had a β-value difference of >5%. Interestingly applying this same approach to the cord blood samples yielded 11 690 MVPs, a reduction from the 12 454 MVPs found using the reference-based approach, although 1136 MVPs at >5% difference were still identified. While the large reduction in MVPs using the reference-free-based approach is suggestive that cellular heterogeneity is an important confounder, the number of MVPs (65) that were common with cord blood is also significantly reduced, suggesting that the method could be too conservative and hence is removing 'true' biological differences. We therefore propose to restrict all further analysis to MVPs unadjusted for cell count reaching genome-wide significance ($P < 0.05$, FDR-corrected) and with a β-value difference >5%. We report in the Supplementary Material the results from all the analysis below using the reference-free adjusted MVP with the same criteria.

To investigate whether the differential methylation identified could be driven by SNP-associated variation, we performed enrichment analysis of our MVP calls with all probes at, or adjacent to, known SNPs (218 551 probes). Of the placenta MVPs, 604 overlapped SNP-associated probes representing a small but significant depletion (1.1-fold depletion, $P < 0.001$), while for the cord blood MVPs, 514 probes overlapped SNP-associated probes also representing a significant depletion (1.3-fold depletion, $P < 0.001$). This suggests the genome-wide methylation variation persists in the absence of SNP-associated probes and we have therefore included these probes in further downstream analysis.

It is known that neighbouring CpG sites show correlated methylation states and hence to determine if our MVPs were robust, we identified 'co-methylation' in a 1000 bp window around MVPs using a bioinformatic algorithm. When applied to MVPs defined by *P*-value alone, placenta co-methylation was 67%, increasing to 76% at MVPs with β-value difference >5%. For cord blood MVPs, co-methylation was 56 and 62% for MVPs unfiltered and filtered by >5% β-value difference, respectively. Increasing rates of co-methylation with MVP β-value difference suggest our filtering approach of 5% β-value difference is valid and helps determine those methylation differences that are true from those that may be due to chance or technical variation.

In both placenta and cord blood samples, a greater proportion of MVPs were hypermethylated in GDM compared with control (926 versus 447 placenta; 828 versus 590 cord blood). Compared with background (all array probes) placenta and cord blood MVPs were located disproportionately frequently at probes showing intermediate methylation (β-value ≥0.3 to <0.7) compared with those with high (≥0.7) or low (<0.3) methylation density (3 × 2 factorial design; placenta χ^2 46.8, df = 2, $P < 0.0001$; cord blood χ^2 58.9; df = 2, $P < 0.001$) (Table 2). The location of MVPs in relationship to known genomic architecture was summarized and compared with background Fig. 1A and B). Placenta MVPs were disproportionately located in the first exon of genes compared with background and this difference reached statistical significance (placenta $P = 0.0027$). Cord blood MVPs were significantly enriched at CpG islands ($P = 0.012$) (Fig. 2A and B). MVPs found

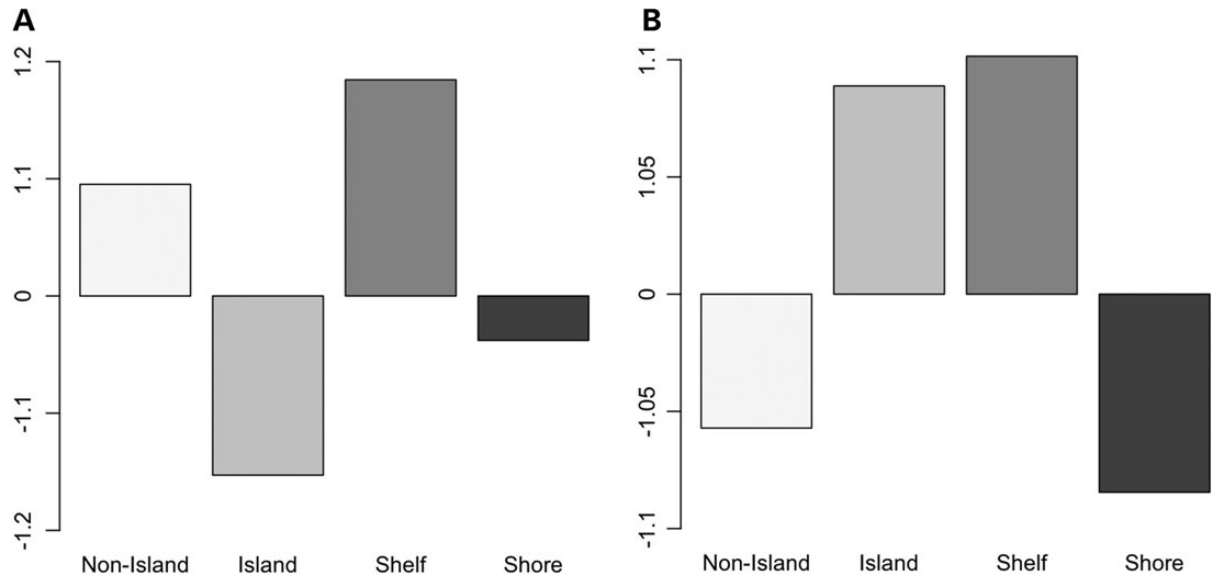

Figure 2. (**A**) Placenta and (**B**) cord blood. Enrichment of methylated variable positions against background according to their relationship with CpG islands. *Y*-axis denotes enrichment (+1) or depletion (−1) in MVPs compared with background (*P-value <0.05). CpG shores are 0–2 kB away from CpG islands, shelves 2–4 kB away from CpG islands.

Table 4. KEGG analysis of placenta methylated variable positions (MVPs)

Term	Fold enrichment	*P*-value
hsa04144: endocytosis	16.6	1.5E−12
hsa05200: pathways in cancer	11.0	3.4E−12
hsa04010: MAPK signalling pathway	12.4	8.6E−12
hsa04670: leucocyte transendothelial migration	20.8	7.7E−11
hsa04514: cell adhesion molecules (CAMs)	18.4	2.6E−10
hsa03010: ribosome	23.6	1.2E−09
hsa04510: focal adhesion	12.7	2.5E−09
hsa04020: calcium signalling pathway	13.8	4.7E−09
hsa04062: chemokine signalling pathway	12.6	1.2E−08
hsa04080: neuroactive ligand–receptor interaction	10.4	2.1E−08
hsa04910: insulin signalling pathway	15.0	5.9E−08
hsa00230: purine metabolism	13.0	2.2E−07

This table displays the gene pathways overrepresented by MVPs in placenta. The top 12 pathways (ranked by *P*-value) are shown.

Table 5. KEGG analysis of cord blood methylated variable positions (MVPs)

Term	Fold enrichment	*P*-value
hsa04144: endocytosis	22.4	1.3E−19
hsa04010: MAPK signalling pathway	15.8	9.6E−17
hsa05200: pathways in cancer	11.5	1.1E−12
hsa04080: neuroactive ligand–receptor interaction	11.6	2.8E−09
hsa04062: chemokine signalling pathway	13.1	1.4E−08
hsa04510: focal adhesion	12.3	2.4E−08
hsa04310: Wnt signalling pathway	13.8	2.2E−07
hsa05010: Alzheimer's disease	13.5	2.3E−07
hsa05016: Huntington's disease	12.2	5.4E−07
hsa04360: axon guidance	14.8	5.1E−07
hsa04114: oocyte meiosis	16.0	1.6E−06
hsa04660: T-cell receptor signalling pathway	15.7	1.3E−19

This table displays the gene pathways overrepresented by MVPs in cord blood. The top 12 pathways (ranked by *P*-value) are shown.

in both cord blood and placenta were located within the first exon of genes at a rate greater than that expected by chance ($P < 0.001$).

Analysis of gene pathway enrichment at MVPs identified was performed in placenta and cord blood MVPs and those present in both tissues (FDR-corrected *P*-value <0.05, β-value difference >5%). The enriched KEGG pathways were ranked according to pathway fold-enrichment differences reaching statistical significance (Tables 4–6). Variable methylation in placenta showed enrichment in pathways with potentially important effects on placental function, e.g. endocytosis, and those involved in growth, e.g. MAPK signalling. In cord blood, enrichment was observed in multiple pathways relating to insulin signalling, and growth factor signalling. Analysis of MVPs found in placenta, cord blood and those overlapping both revealed pathway enrichment in endocytosis, focal adhesion, chemokine signalling and ligand–receptor interactions and could suggest common susceptible pathways affected by gestational diabetes exposure that are not tissue-specific. These pathways display methylation differences at a range of genes involved with the extracellular triggers to a range of intracellular signalling processes, including MAPK, PI3k-AKT, JAK-STAT and Wnt, which have multiple functions in embryonic development and metabolic pathways.

To investigate whether the treatment has biased the identification of MVPs, we repeated the MVP analysis in both placenta

Table 6. KEGG analysis of methylated variable positions (MVPs) common to both placenta and cord blood

Term	Fold enrichment	*P*-value
hsa04144: endocytosis	23.6	5.8E–06
hsa04670: leucocyte transendothelial migration	24.1	3.2E–03
hsa04360: axon guidance	21.4	3.4E–03
hsa04062: chemokine signalling pathway	14.7	1.1E–02
hsa04510: focal adhesion	13.8	1.1E–02
hsa03010: ribosome	25.9	9.4E–03
hsa00360: phenylalanine metabolism	83.8	9.3E–03
hsa03040: spliceosome	17.5	2.2E–02
hsa00983: drug metabolism	40.8	3.0E–02
hsa00350: tyrosine metabolism	38.8	3.0E–02
hsa05200: pathways in cancer	8.3	3.1E–02
hsa04740: olfactory transduction	7.9	3.4E–02

This table displays the gene pathways overrepresented by MVPs identified in both placenta and cord blood. The top 12 pathways (ranked by *P*-value) are shown.

and cord blood for metformin and/or insulin-treated GDM against control and for diet-treated GDM against control samples. We combined those treated with metformin and/or insulin due to the small numbers receiving metformin, insulin or both. In placenta, we found no probes showing methylation difference that reached genome-wide significance in the metformin and/or insulin-treated GDM ($n = 7$) compared with controls. When comparing diet-controlled GDM ($n = 18$) with controls, differential methylation was identified at 5498 probes ($P < 0.05$, FDR-corrected). We hypothesized that the lack of methylation differences in the metformin- and/or insulin-treated comparison was due to the small sample size rather than effect of treatment. In support of this hypothesis, we have plotted β-value differences of GDM versus control for these 5498 probes for the two different treatment groups and show a highly significant correlation ($\rho = 0.88$, $P < 2.2 \times 10^{-16}$) that would not be expected with strong treatment-specific differences. We performed the same comparison in cord blood samples, which included a larger sample number: 7 metformin and/or insulin-treated and 20 diet-treated GDM samples. We found methylation differences at 2138 probes when comparing metformin- and/or insulin-treated GDM with controls ($P < 0.05$, FDR-corrected), and 9304 probes comparing diet-treated GDM with controls ($P < 0.05$, FDR-corrected). Again, we found strong correlation between the β-value differences from the two treatment groups ($\rho = 0.97$, $P < 2.2 \times 10^{-16}$). The findings of these analyses suggest that there are no strong treatment-driven effects on methylation in cord blood and placenta, and that we are confident in our approach of comparing controls with a combined group of diet-treated and metformin and/or insulin-treated GDM. The limited power in our study limits the potential to identify whether a small subset of the MVPs do show treatment-specific effects and this could be of great interest in future work.

Discussion

To date, while clinical and epidemiological evidence of foetal programming is convincing, there is little molecular understanding how it may occur, but epigenetic studies are beginning to elucidate mechanisms intergenerational transmission of disease risk. In this study, we identify multiple genome-wide differences in DNA methylation in foetal tissues exposed to GDM, compared with controls, using a discovery-based experimental approach. Our cross-sectional study design using cord blood and placenta sampled during term delivery offers a window with which to understand the direct effects of pregnancy and may help to define an epigenetic signature of future disease risk, induced through 'foetal programming'.

Our sampling and analysis pipeline are consistent with consensus approaches (28,29) and we have established that there is no significant difference in cellular composition of cord blood between cases and controls. To mitigate the potential for false discovery from multiple hypothesis testing, we have used stringent Benjamini and Hochberg 'false discovery rate' (FDR) control to define methylation differences. Previously reported genome-wide studies have been unable to detect DNA methylation differences reaching this level of significance due to an assumed lack of power from smaller sample sizes (22). We have applied a further filter to define MVPs as those with methylation differences of ≥5% and provide evidence of co-methylation to further support our assertion that we have identified true variation over background noise. Although many targeted epigenetic studies report methylation differences of <5% as significant (17), the functional effects of small methylation differences in understanding complex disease pathogenesis are, as yet, undetermined. The MVPs identified are located within regulatory regions of the genome (to which the array is enriched), in or around CpG islands and with preponderance to the first exon, greater than that which would be expected by array design. It has been suggested that methylation of the first exon may have a role in transcriptional silencing (30) and may be the target site of action of Dnmt3a, a DNA methyltransferase with a role in embryonic development (31). CpG islands are well known to have an important interaction with transcriptional regulation, and many are located at transcriptional start sites and are relatively hypomethylated to enable transcription (32).

The identification of placental hypomethylation, relative to other tissues, is consistent with that of other studies of the placenta methylome (33). The predominance of hypermethylation we identify at MVPs in both foetal tissues is consistent with previous reported findings of greater mean methylation in GDM-exposed placenta and cord blood compared with unexposed tissues (22). Most GDM-associated MVPs were located in regions of intermediate methylation (β-value 0.3–0.7), despite array design representing the bimodal distribution of β-values towards hyper- and hypomethylation. Existing studies have shown epigenetic variation at regions of intermediate methylation, such as placental partially methylated domains (33), but others suggest greater technical variation in the intermediate methylation range with the 450k array (34). Technological advances in whole-genome DNA methylation profiling will help interpret this finding in future studies.

We investigated whether methylation differences were driven by genetic variation and found that epigenomic differences were not dependent on SNP-associated methylation variation, although the latter did contribute to some variation. We have therefore included SNP-containing probes in our analysis, understanding that they may be an important contributory mechanism by which gene and environment interacts in early developmental life and that our study was not adequately powered to examine this. Future studies should incorporate allele-specific methylation techniques to understand how they affect gene function, and the combined interrogation of parental and foetal tissue may yield significant insight into the role of genetic and epigenetic factors on parent–child transmission of disease risk.

Genome-wide methylation differences associated with GDM exposure highlight putative effects on gene pathways with functional implications placental function and foetal development. KEGG pathway analysis suggests that the methylation variation in placenta exposed to gestational diabetes may have a role in endocytic processes. The role of endocytosis (including processes such as autophagy) in placental function is important as a responsive mechanism to variable pregnancy environments; two recent studies in term human placenta and cell lines suggest an inverse relationship between glucose and placental autophagy (35,36). The mitogen-activated protein kinase (MAPK) pathway, also enriched in placenta MVPs, is also thought to regulate autophagy and its effects (along with that of ERK) are known to be disrupted in gestational diabetes (37). Pathway analysis of methylation variants in cord blood suggests that exposure to gestational diabetes impacts multiple pathways including cell adhesion, chemokine signalling and ligand–receptor interactions; these incorporate key extracellular triggers to a myriad of intracellular signalling pathways involved in growth and metabolism, including PI3k-Akt, MAPK, JAK-STAT and Wnt signalling. Enrichment of pathways incorporating these common processes, and highlighting differential methylation at genes such as growth factors and key regulators of insulin signalling pathways is reassuring and suggests that this epigenomic data set does offer a window into true pathophysiological processes. We, and others (22), have identified MVPs that were common to both placenta and cord blood in association with gestational diabetes exposure. Pathway analysis was performed on MVPs common to both cord blood and placenta and shows enrichment of several pathways identified in either tissue alone; this may support a role for common susceptible pathways across these tissues. Although pathway analysis offers a means by which to explore putative biological mechanisms affected by gestational diabetes-associated epigenetic variation, it does not provide any additional proof that these epigenetic variants have a role in either the cause or consequence of gestational diabetes. Longitudinal and intervention-based studies need to be performed to establish whether there is a causal relationship between these epigenetic variants and GDM and/or future programmed disease in offspring. If epigenetic variants are shown to have a causal relationship with future disease, they might provide a useful 'molecular signature' and/or biomarker to improve current prediction tools of disease risk and response to intervention.

There are a number of limitations to this study. DNA from whole blood or placental samples consists of more than one cell type, nonetheless when corrected for cell heterogeneity, there are no substantial changes to our findings. The variation in GDM treatments and varied size of the treatment groups studied is a limitation of our study, although our analysis suggests that this does not drive methylation differences. This is a pilot study in small numbers, and therefore, it is difficult to exclude our results are not driven by treatment effects of GDM, maternal factors or foetal phenotype.

In summary, we have identified epigenomic variation in offspring exposed to GDM *in utero* through the study of DNA methylation profiles in cord blood and placenta. The detection of DNA methylation differences withstand correction for false discovery, show characteristics that distinguish them from technical variation and are not purely SNP-driven. Characterization of these epigenetic differences suggest they exist preferentially at CpG islands and within the first exon of genes and this offers a route by which to determine their role in transcriptional regulation in future studies. Pathway analysis suggests that gestational diabetes exposure may have functional impact on placental endocytic processes, increasingly understood to play a role in foetal growth and metabolism, and multiple other extra- and intracellular signalling pathways involved in growth and metabolism. While cross-sectional studies such as this one offer important insights into the effect of adverse pregnancy environments on foetal tissues, future studies must include: larger sample sizes to improve power to detect differences and control for false discovery, replication across cohorts, parallel genetic studies to establish their combined and/or independent role in parent–child transmission of disease risk, longitudinal sampling and intervention to determine causation over association and functional studies to determine the mechanism by which epigenetic variation can exert its effects.

Materials and Methods

Women of South Asian origin receiving antenatal care at the Royal London Hospital, UK, were invited to participate in the study during the third trimester of pregnancy, under institutional and ethics committee approval. All women had received screening for gestational diabetes (GDM) at 28 weeks gestation with a 2 h 75 g oral glucose tolerance test. Local diagnostic thresholds for GDM were used (0 min serum glucose ≥5.8 and/or 120 min ≥7.8 mmol/L). Women with multiple pregnancies or pre-existing diabetes (including that suggested by an abnormal glucose tolerance test at 16 weeks) were excluded. Recruited women underwent standard antenatal care and additional metabolic phenotyping during the third trimester of pregnancy.

Standardized protocols were used to sample cord blood and placenta at delivery. Cord blood was obtained from the umbilical cord vessels within 30 min of delivery and collected into EDTA tubes for full blood count measurement and storage (at −80°C within 12 h of collection). Placenta samples (2 × 2 × 2 cm) were taken adjacent to the umbilical cord insertion on the foetal surface of the placenta, membranes (chorion) were removed and the sample was rinsed thoroughly in sterile phosphate-buffered saline, placed in RNA-later solution at 4°C for 24 h and then frozen at −80°C.

DNA extraction and purification was performed from whole blood (DNeasy Blood and Tissue, QIAquick PCR purification kits, Qiagen) with standard quality control checks. Bisulphite conversion of DNA was performed using the EZ-methylation kit (Zymo) and samples with conversion efficiency (determined by qPCR) >95% were used for array hybridization. Batch effects were minimized using sample randomization per study group and tissue type prior to bisulphite conversion and hybridization.

The Illumina HumanMethylation 450BeadChip (Illumina, USA) ('450k array') was used, providing genome-wide single base resolution of DNA methylation at pre-defined CpG dinucleotides, targeted to all designable RefSeq genes, CpG islands and regulatory regions. Bisulphite-converted DNA was hybridized using standard protocols to the array which uses two types of Infinium bead chemistry to determine the methylation state (β-value) of individual CpG sites (38). β-Values were generated from dye intensity [methylated probe intensity/(methylated probe intensity + unmethylated probe intensity + 100)] and detected via laser excitement and converted into *M*-values ($\log_2$ ratio of β-value) to account for heteroscedasticity and allow analysis assuming a Gaussian distribution (39). Data were filtered to remove cross-hybridizing probes, X and Y chromosome probes and those with a detection *P*–value of ≥0.05. Data were quantile normalized per colour channel, per bead and per tissue type; a method shown to reduce standard error with high performance (28). Multiple dimensional scaling analysis was performed to

exclude positional or batch effects. BMIQ was used to correct for the two different probe types contained on the array (40).

'MVPs' were identified, per probe, between offspring born to mothers with versus without gestational diabetes using Limma (41). Benjamini and Hochberg false-discovery rate control was applied (42). MVPs were selected first by P-value of <0.05 and second by β-value difference >5%. Methylation patterns at probes adjacent to MVPs (±500 bp) were characterized to identify 'co-methylation' that may indicate biological plausibility and a putative role in transcriptional regulation (43) and using methods described by others (44). We surveyed the KEGG Pathway Database was surveyed (45) to identify gene pathways overrepresented by regions of DNA methylation differences (MVPs) and provide a hypothetical functional correlate of our data.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank the participants of this research study and the staff of the Royal London Hospital for helping with this study. The clinical study was designed and performed by S.F., L.F., S.H., A.S. and G.A.H. Epigenomic lab work was designed and undertaken by S.F., M.S., L.F. and V.K.R. Bioinformatic analysis was done by C.M., R.L. and V.K.R. S.F. and G.A.H. wrote the manuscript with contribution and discussion from A.S., L.F., R.L., V.K.R. and D.W. We also thank the referees for making important suggestions for analysis of our data.

Conflict of Interest statement. None declared.

Funding

This work was funded by an MRC Clinical Research Training Fellowship (G0800441, awarded to S.F. while at QMUL) and supported by the North-East London Diabetes Research Network. S.F. 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.

References

1. Centers for Disease Control and Prevention. (2011) *National Diabetes Fact Sheet: National Estimates and General Information on Diabetes and Prediabetes in the United States*. Atlanta, GA.
2. CG63, N. Clinical Guidelines (2008) *Diabetes in Pregnancy: Management of Diabetes and Its Complications from pre-Conception to the Postnatal Period*.
3. Diagnosis and classification of diabetes mellitus. (2011) *Diabetes Care*, **34** (Suppl 1), S62–S69.
4. Kwak, S.H., Kim, S.-H., Cho, Y.M., Go, M.J., Cho, Y.S., Choi, S.H., Moon, M.K., Jung, H.S., Shin, H.D., Kang, H.M. *et al.* (2012) A genome-wide association study of gestational diabetes mellitus in Korean women. *Diabetes*, **61**, 531–541.
5. Jansson, T., Ekstrand, Y., Björn, C., Wennergren, M. and Powell, T.L. (2002) Alterations in the activity of placental amino acid transporters in pregnancies complicated by diabetes. *Diabetes*, **51**, 2214–2219.
6. Scholler, M., Wadsack, C., Lang, I., Etschmaier, K., Schweinzer, C., Marsche, G., Dieber-Rotheneder, M., Desoye, G. and Panzenboeck, U. (2012) Phospholipid transfer protein in the placental endothelium is affected by gestational diabetes mellitus. *J. Clin. Endocrinol. Metab.*, **97**, 437–445.
7. Visiedo, F., Bugatto, F., Sánchez, V., Cózar-Castellano, I., Bartha, J.L. and Perdomo, G. (2013) High glucose levels reduce fatty acid oxidation and increase triglyceride accumulation in human placenta. *Am. J. Physiol. Endocrinol. Metab.*, **305**, E205–E212.
8. Sobrevia, L., Abarzúa, F., Nien, J.K., Salomón, C., Westermeier, F., Puebla, C., Cifuentes, F., Guzmán-Gutiérrez, E., Leiva, A. and Casanello, P. (2011) Review: Differential placental macrovascular and microvascular endothelial dysfunction in gestational diabetes. *Placenta*, **32**(Suppl 2), S159–S164.
9. Fetita, L.-S., Sobngwi, E., Serradas, P., Calvo, F. and Gautier, J.-F. (2006) Consequences of fetal exposure to maternal diabetes in offspring. *J. Clin. Endocrinol. Metab.*, **91**, 3718–3724.
10. Gauguier, D., Bihoreau, M.T., Ktorza, A., Berthault, M.F. and Picon, L. (1990) Inheritance of diabetes mellitus as consequence of gestational hyperglycemia in rats. *Diabetes*, **39**, 734–739.
11. Dabelea, D., Hanson, R.L., Lindsay, R.S., Pettitt, D.J., Imperatore, G., Gabir, M.M., Roumain, J., Bennett, P.H. and Knowler, W.C. (2000) Intrauterine exposure to diabetes conveys risks for type 2 diabetes and obesity: a study of discordant sibships. *Diabetes*, **49**, 2208–2211.
12. Lindsay, R.S., Nelson, S.M., Walker, J.D., Greene, S.A., Milne, G., Sattar, N. and Pearson, D.W. (2010) Programming of adiposity in offspring of mothers with type 1 diabetes at age 7 years. *Diabetes Care*, **33**, 1080–1085.
13. Lillycrop, K.A., Phillips, E.S., Jackson, A.A., Hanson, M.A. and Burdge, G.C. (2005) Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *J. Nutr.*, **135**, 1382–1386.
14. Lillycrop, K.A., Phillips, E.S., Torrens, C., Hanson, M.A., Jackson, A.A. and Burdge, G.C. (2008) Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPAR alpha promoter of the offspring. *Br. J. Nutr.*, **100**, 278–282.
15. Waterland, R.A., Kellermayer, R., Laritsky, E., Rayco-Solon, P., Harris, R.A., Travisano, M., Zhang, W., Torskaya, M.S., Zhang, J., Shen, L. *et al.* (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.*, **6**, e1001252.
16. Tobi, E.W., Heijmans, B.T., Kremer, D., Putter, H., Delemarre-van de Waal, H.A., Finken, M.J.J., Wit, J.M. and Slagboom, P.E. (2011) DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. *Epigenetics*, **6**, 171–176.
17. Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E. and Lumey, L.H. (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl Acad. Sci. USA*, **105**, 17046–17049.
18. El Hajj, N., Pliushch, G., Schneider, E., Dittrich, M., Müller, T., Korenkov, M., Aretz, M., Zechner, U., Lehnen, H. and Haaf, T. (2013) Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes*, **62**, 1320–1328.
19. Houde, A.-A., Guay, S.-P., Desgagné, V., Hivert, M.-F., Baillargeon, J.-P., St-Pierre, J., Perron, P., Gaudet, D., Brisson, D. and Bouchard, L. (2013) Adaptations of placental and cord blood ABCA1 DNA methylation profile to maternal metabolic status. *Epigenetics*, **8**, 1289–1303.
20. Novakovic, B., Gordon, L., Robinson, W.P., Desoye, G. and Saffery, R. (2013) Glucose as a fetal nutrient: dynamic regulation of several glucose transporter genes by DNA methylation in the human placenta across gestation. *J. Nutr. Biochem.*, **24**, 282–288.

21. Tobi, E.W., Lumey, L.H., Talens, R.P., Kremer, D., Putter, H., Stein, A.D., Slagboom, P.E. and Heijmans, B.T. (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.*, **18**, 4046–4053.
22. Ruchat, S.-M., Houde, A.-A., Voisin, G., St-Pierre, J., Perron, P., Baillargeon, J.-P., Gaudet, D., Hivert, M.-F., Brisson, D. and Bouchard, L. (2013) Gestational diabetes mellitus epigenetically affects genes predominantly involved in metabolic diseases. *Epigenetics*, **8**, 935–943.
23. Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J., Adair, T., Aggarwal, R., Ahn, S.Y. *et al.* (2013) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, **380**, 2095–2128.
24. Gardosi, J., Chang, A., Kalyan, B., Sahota, D. and Symonds, E.M. (1992) Customised antenatal growth charts. *Lancet*, **339**, 283–287.
25. Salvesen, D.R., Brudenell, M.J. and Nicolaides, K.H. (1992) Fetal polycythemia and thrombocytopenia in pregnancies complicated by maternal diabetes mellitus. *Am. J. Obstet. Gynecol.*, **166**, 1287–1293.
26. Hansen, K., Ayree, M., Irizarry, R.A., Jaffe, A., Maksimovic, J., Houseman, A., Fortin, J. and Triche, T. (2014) Package 'minfi' version 1.12.0.
27. Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K. and Kelsey, K.T. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform.*, **13**, 86.
28. Pidsley, R., Wong, C.C.Y., Volta, M., Lunnon, K., Mill, J. and Schalkwyk, L.C. (2013) A data-driven approach to preprocessing Illumina 450 K methylation array data. *BMC Genomics*, **14**, 293.
29. Hogg, K., Price, E.M. and Robinson, W.P. (2014) Improved reporting of DNA methylation data derived from studies of the human placenta. *Epigenetics*, **9**, 333–337.
30. Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A.J., Socci, N.D. and Scandura, J.M. (2011) DNA methylation of the first exon is tightly linked to transcriptional silencing. *PloS One*, **6**, e14524.
31. Oka, M., Rodi?, N., Graddy, J., Chang, L.-J. and Terada, N. (2006) CpG sites preferentially methylated by Dnmt3a in vivo. *J. Biol. Chem.*, **281**, 9901–9908.
32. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.*, **13**, 484–492.
33. Schroeder, D.I., Blair, J.D., Lott, P., Yu, H.O.K., Hong, D., Crary, F., Ashwood, P., Walker, C., Korf, I., Robinson, W.P. *et al.* (2013) The human placenta methylome. *Proc. Natl Acad. Sci. USA*, **110**, 6037–6042.
34. Pan, H., Chen, L., Dogra, S., Teh, A.L., Tan, J.H., Lim, Y.I., Lim, Y.C., Jin, S., Lee, Y.K., Ng, P.Y. *et al.* (2012) Measuring the methylome in clinical samples: improved processing of the Infinium Human Methylation450 BeadChip Array. *Epigenetics*, **7**, 1173–1187.
35. Signorelli, P., Avagliano, L., Virgili, E., Gagliostro, V., Doi, P., Braidotti, P., Bulfamante, G.P., Ghidoni, R. and Marconi, A.M. (2011) Autophagy in term normal human placentas. *Placenta*, **32**, 482–485.
36. Hung, T.-H., Hsieh, T.-T., Chen, S.-F., Li, M.-J. and Yeh, Y.-L. (2013) Autophagy in the human placenta throughout gestation. *PloS One*, **8**, e83475.
37. Salomón, C., Westermeier, F., Puebla, C., Arroyo, P., Guzmán-Gutiérrez, E., Pardo, F., Leiva, A., Casanello, P. and Sobrevia, L. (2012) Gestational diabetes reduces adenosine transport in human placental microvascular endothelium, an effect reversed by insulin. *PloS One*, **7**, e40578.
38. Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L. *et al.* (2011) High density DNA methylation array with single CpG site resolution. *Genomics*, **98**, 288–295.
39. Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W.A., Hou, L. and Lin, S.M. (2010) Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.*, **11**, 587.
40. Schalkwyk, L.C., Pidsley, R., Wong, C., Touleimat, N., Defrance, M., Teschendorff, A. and Maksimovic, J. (2014) Package 'wateRmelon'.
41. Hansen, K. and Aryee, M. *Minfi: Analyze Illumina's 450k Methylation Arrays*. R package version 1.2.0.
42. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. and Golani, I. (2001) Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.*, **125**, 279–284.
43. Zhang, Y., Shu, J., Si, J., Shen, L., Estecio, M.R.H. and Issa, J.-P.J. (2012) Repetitive elements and enforced transcriptional repression co-operate to enhance DNA methylation spreading into a promoter CpG-island. *Nucleic Acids Res.*, **40**, 7257–7268.
44. Bell, J.T., Tsai, P.-C., Yang, T.-P., Pidsley, R., Nisbet, J., Glass, D., Mangino, M., Zhai, G., Zhang, F., Valdes, A. *et al.* (2012) Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.*, **8**, e1002629.
45. Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**, 27–30.