Diabetologia (2016) 59:1714–1723
DOI 10.1007/s00125-016-3983-7

ARTICLE

DNA hypermethylation of CD3$^+$ T cells from cord blood of infants exposed to intrauterine growth restriction

Lyda Williams[1] · Yoshinori Seki[1] · Fabien Delahaye[2] · Alex Cheng[1] · Mamta Fuloria[3] · Francine Hughes Einstein[2] · Maureen J. Charron[1,2,4]

Received: 22 February 2016 / Accepted: 19 April 2016 / Published online: 17 May 2016

Abstract
Aims/hypothesis Intrauterine growth restriction (IUGR) is associated with increased susceptibility to obesity, metabolic syndrome and type 2 diabetes. Although the mechanisms underlying the developmental origins of metabolic disease are poorly understood, evidence suggests that epigenomic alterations play a critical role. We sought to identify changes in DNA methylation patterns that are associated with IUGR in CD3$^+$ T cells purified from umbilical cord blood obtained from male newborns who were appropriate for gestational age (AGA) or who had been exposed to IUGR.
Methods CD3$^+$ T cells were isolated from cord blood obtained from IUGR and AGA infants. The genome-wide methylation profile in eight AGA and seven IUGR samples was determined using the HELP tagging assay. Validation analysis using targeted bisulfite sequencing and bisulfite massARRAY was performed on the original cohort as well as biological replicates consisting of two AGA and four IUGR infants. The Segway algorithm was used to identify methylation changes within regulatory regions of the genome.
Results A global shift towards hypermethylation in IUGR was seen compared with AGA (89.8% of 4,425 differentially methylated loci), targeted to regulatory regions of the genome, specifically promoters and enhancers. Pathway analysis identified dysregulation of pathways involved in metabolic disease (type 2 diabetes mellitus, insulin signalling, mitogen-activated protein kinase signalling) and T cell development, regulation and activation (T cell receptor signalling), as well as transcription factors (*TCF3*, *LEF1* and *NFATC*) that regulate T cells. Furthermore, bump-hunting analysis revealed differentially methylated regions in *PRDM16* and *HLA-DPB1*, genes important for adipose tissue differentiation, stem cell maintenance and function and T cell activation.
Conclusions/interpretation Our findings suggest that the alterations in methylation patterns observed in IUGR CD3$^+$ T cells may have functional consequences in targeted genes, regulatory regions and transcription factors. These may serve as biomarkers to identify those at 'high risk' for diminished attainment of full health potential who can benefit from early interventions.
Access to research materials HELP tagging data: Gene Expression Omnibus database (GSE77268), scheduled to be released on 25 January 2019.

Keywords Human · Metabolic syndrome · Prediction of type 2 diabetes

Abbreviations

AGA	Appropriate for gestational age
BSMAP	Bisulfite sequencing mapping platform
DML	Differentially methylated loci

Lyda Williams, Yoshinori Seki and Fabien Delahaye contributed equally to this work and are joint first authors.

Electronic supplementary material The online version of this article (doi:10.1007/s00125-016-3983-7) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

✉ Maureen J. Charron
maureen.charron@einstein.yu.edu

[1] Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Rm F312, Bronx, NY 10461, USA

[2] Department of Obstetrics and Gynecology and Women's Health, Albert Einstein College of Medicine, Bronx, NY, USA

[3] Department of Pediatrics, Albert Einstein College of Medicine, Bronx, NY, USA

[4] Department of Medicine, Division of Endocrinology, Albert Einstein College of Medicine, Bronx, NY, USA

DMR	Differentially methylated region
FDR	False discovery rate
GSEA	Gene Set Enrichment Analysis
HH	Hedgehog
IUGR	Intrauterine growth restriction
KEGG	Kyoto Encyclopedia of Genes and Genomes
LGA	Large for gestational age
MAPK	Mitogen-activated protein kinase
PBL	Peripheral blood lymphocyte
ROS	Reactive oxygen species
Treg	Regulatory T cell
TBS	Targeted bisulfite sequencing
TSS	Transcription start site

Introduction

The developmental origins of health and disease hypothesis implicate an adverse prenatal environment in the development of chronic diseases [1, 2], a concept known as programming. Epidemiological research has demonstrated that an adverse environment during development can program metabolic pathways in the fetus and increase susceptibility to metabolic syndrome [3, 4]. This finding is supported by studies using animal models of intrauterine growth restriction (IUGR), including maternal nutrient deprivation (e.g. energy or protein restriction) or excess (e.g. high-fat diet and hyperglycaemia) [5, 6].

Annually, approximately 8% of non-anomalous infants are born with IUGR, a multifactorial disorder that likely contributes to the worldwide epidemic of metabolic disease [7]. Studies from the Dutch Famine Cohort, and other epidemiological studies, linked low birthweight with increased risk for cardiovascular disease-related death and type 2 diabetes [8]. Although the molecular mechanisms underlying this phenomenon are poorly understood, several decades later, peripheral blood leucocytes showed an association between nutrient deprivation during fetal life and lower methylation of *IGF2*, a key factor in growth and development [9]. Other studies have shown distinct methylation patterns in haematopoietic $CD34^+$ stem cells purified from cord blood of infants exposed to IUGR (referred to as 'IUGR infants') as well as those who were large for gestational age (LGA), when compared with infants who were appropriate for gestational age (AGA). Interestingly, a sexual dimorphic effect was observed in which the number of hypermethylated loci was markedly higher in IUGR male infants and LGA female infants [10]. Furthermore, methylation differences were found in genes belonging to pathways for MODY and hedgehog (HH) signalling, both of which are associated with premature glucose intolerance, type 2 diabetes and stem cell proliferation/renewal [10]. These studies highlight the potential role of dysregulated DNA methylation in key metabolic pathways in metabolic disease susceptibility.

Obesity has been associated with an increased number of macrophages, neutrophils, T cells, B cells and mast cells in adipose tissue [11]. Furthermore, obesity-induced inflammation contributes to a number of chronic illnesses, including type 2 diabetes and metabolic syndrome, in humans and experimental models [12, 13]. Recent studies have demonstrated that $CD3^+$ T cells play an important role in obesity and other inflammatory diseases by inducing inflammation and promoting insulin resistance via accumulation of inflammatory macrophages in adipose tissue [14]. IUGR is also associated with a decrease in the number of circulating regulatory T cells (Tregs) in cord blood [15] and a decrease in subpopulations responsible for regulating inflammation, including $CD3^+$ cells in uterine decidua [16]. Thus, exposure to inflammation during fetal life may link growth restriction with subsequent development of type 2 diabetes [17]. While genes displaying differential methylation associated with IUGR have been identified in peripheral blood lymphocytes (PBLs) [9], the cell heterogeneity of the samples may have introduced experimental artefacts that could alter findings [18].

Here, our overall objective was to identify differentially methylated loci (DML) in purified $CD3^+$ T cells of IUGR and AGA term male newborns using a genome-wide approach. $CD3^+$ T cells are metabolically relevant cells due to their role in inflammatory responses and represent a relatively homogenous population composed of terminally differentiated $CD4^+$ and $CD8^+$ cells [19]. We elected to study male IUGR infants based on our previous studies demonstrating greater methylation changes in $CD34^+$ haematopoietic stem cells from IUGR male infants compared with female infants [10]. By identifying DML in $CD3^+$ T cells associated with IUGR using a genome-wide approach, we identify potential functional targets that may be related to inflammatory changes seen in metabolic disease pathogenesis. Our secondary objective was to map differential methylation to genomic regulatory elements of known metabolic pathways/genes of interest that could prove useful in narrowing the list of candidate genes for future research and clinical studies.

Methods

Approval from the Institutional Review Board of the Montefiore Medical Center and the Committee on Clinical Investigation at the Albert Einstein College of Medicine was obtained in accordance with the Health Insurance Portability and Accountability Act regulations. Written informed consent was obtained from all participants prior to the study. To help guide the reader, a flowchart of procedures is presented in electronic supplementary material (ESM) Fig. 1.

Sample collection Cord blood samples were obtained from 21 consenting women who delivered healthy, non-anomalous term neonates without evidence of fetal distress (normal Apgar scores and cord blood gases without acidaemia) at the Weiler Division of Montefiore Medical Center. One woman in the IUGR group had pre-eclampsia, whereas all other pregnancies were uncomplicated in both groups. Birthweight and ponderal index were used to identify cases and controls matched for gestational age at delivery, sex and self-identified race/ethnicity. IUGR was defined by birthweight and ponderal index values <tenth percentile for gestational age and sex. AGA infants had normal birthweight and ponderal index percentiles (>tenth and <90th) for both variables. Maternal and infant characteristics are shown in Table 1. Cohort 1 (with genome-wide methylation data) is composed of eight AGA and seven IUGR infants and Cohort 2 (for technical and biological validation) is composed of two AGA and four IUGR infants.

Isolation of $CD3^+$ T cells Cord blood samples were collected in a standardised protocol used routinely for public cord blood donation [20] and $CD3^+$ T cells were isolated using an immunomagnetic separation technique. Mononuclear cells were separated by Ficoll–Paque density gradient (GE Healthcare, Little Chalfont, UK) or by PrepaCyte-WBC (BioE, St Paul, MN, USA). $CD3^+$ T cells were obtained by positive immunomagnetic bead selection after $CD34^+$ cell depletion using the AutoMACS Separator (Miltenyi Biotech, Bergisch Gladbach, Germany) resulting in isolation of cells with ≥95% purity. Purified cells were cryopreserved in 10% dimethyl sulfoxide using controlled rate freezing.

HELP tagging assay The HELP tagging assay was used to determine genome-wide methylation [21] of purified $CD3^+$ T cells from 15 infants (eight AGA, seven IUGR). Briefly, genomic DNA was prepared from frozen $CD3^+$ T cells and digested to completion by either HpaII or MspI. Methylation-sensitive HpaII profiles were obtained for each sample. Methylation scores were calculated using a previously generated MspI human reference.

Table 1 Maternal and neonatal characteristics of the participants

Characteristic	AGA ($n=10$)	IUGR ($n=11$)
Neonate		
Gestational age (weeks)	39.2 ± 0.7	39.2 ± 1.9
Birthweight (g)	3,221 ± 290	2,601 ± 432***
Length (cm)	48.5 ± 1.3	47.9 ± 2.5
Ponderal index (g/cm^3)	2.8 ± 0.2	2.3 ± 0.1****
Mother		
Maternal age (years)	26.0 ± 4.8	24.5 ± 6.7
Pre-pregnancy weight (kg)	74.8 ± 20.8	68.5 ± 12.1
Pre-pregnancy BMI (kg/m^2)	28.9 ± 8.2	26.5 ± 5.5
Weight at term (kg)	86.5 ± 18.9	75.4 ± 9.5
Weight gain during pregnancy (kg)	11.7 ± 6.4	7.6 ± 5.8
BMI at term (kg/m^2)	33.4 ± 7.1	29.2 ± 4.5

***$p<0.001$ and ****$p<0.0001$ vs AGA. Data are means ± SD

Data analysis DNA methylation scores from 0 (fully methylated) to 100 (unmethylated) were filtered by confidence scores determined for each sample in Cohort 1 (eight AGA, seven IUGR) based on the total number of HpaII-generated reads as a function of the total number of MspI-generated reads. Confidence score was used to eliminate loci with insufficient representation to ensure the robustness of the data. Methylation scores were compared between AGA and IUGR infants and DML defined as those having an absolute difference of >20, and a significance <0.05. Statistical analysis was run using LIMMA (Bioconductor, https://bioconductor.org/packages/release/bioc/html/limma.html) [22] on preprocessed data using ComBat (Bioconductor, https://bioconductor.org/packages/release/bioc/html/sva.html) [23]. ComBat is a means to correct for various biological and technical confounders, such as batch effect, race, maternal BMI and other clinical information, to allow focus on our variable of interest (i.e. IUGR vs AGA). Principal component analysis was used to assess variability across samples and validate ComBat correction (ESM Fig. 2). We obtained 4,425 significant DML out of 1,077,400 testable loci after confidence score filtering. All data are available through the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo/) under the accession number GSE77268 and are scheduled to be released on 25 January 2019.

Bisulfite MassARRAY validation Technical validation of HELP tagging assay data using primers for loci with a range of methylation scores and validation of DML for *MARK2* and *CDH12* were performed using bisulfite MassARRAY (Mass array nanodispenser, Sequenom, San Diego, CA, USA) [24] on Cohort 1. Primers were designed to cover loci with varying levels of DNA methylation (ESM Table 1). Matched peak data were exported using EpiTYPER software (Sequenom, San Diego, CA, USA) and analysed for quality using analytical tools that we developed [25].

Targeted bisulfite sequencing After bisulfite conversion of genomic DNA (200 ng) using the EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA), separate PCR amplification of 55 individual target regions was performed (primers listed in ESM Table 1), the amplicons were pooled in equal ratios and Illumina libraries were generated using robotic automation (Tecan, Zurich, Switzerland). A total of 21 libraries (eight AGA, seven IUGR [Cohort 1]; two AGA, four IUGR [Cohort 2]) were multiplexed on the Illumina MiSeq (Illumina, San Diego, CA, USA) for 250 bp paired

end sequencing, including standards for 0% and 100% methylation. Amplicons were selected to validate biologically relevant loci as well as to represent the range of DNA methylation values from 0 to 100%.

Bisulfite sequence alignment and DNA methylation calls Sequence reads from Illumina MiSeq were trimmed for adapter sequences and aligned to the human genome using the default settings of BSMAP (bisulfite sequencing mapping platform) (version bsmap/2.89/gcc.4.4.7, http://lilab.research.bcm.edu/dldcc-web/lilab/yxi/bsmap/bsmap-2.90.tgz) [26], requiring a PHRED score of ≥37 during alignment. Bisulfite conversion efficiency was determined (C→T in CH contexts, ESM Table 2) and the methylation score for each sample, from 0 (fully methylated) to 100 (unmethylated), was quantified for every CpG in the amplicon using the methratio tool provided by BSMAP. Validation was performed on Cohort 1 and on biological replicates in Cohort 2.

Genome annotation Publicly available data from chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-seq) was obtained from the Roadmap in Epigenomics project for $CD3^+$ T cells from several donors (ESM Table 3). Annotation involved processing the raw data for chromatin accessibility (DNAse hypersensitivity) along with ChIP-seq data for six histone modifications (H3K4me1, H3K27me3, H3K27ac, H3K4me3, H3K36me3 and H3K9me3), followed by use of the Segway algorithm [27] to predict seven features as previously described [10]. Annotation was generated by analysing enrichment of each feature for the different chromatin tracks with a one-way proportion test performed with continuity correction of the observed vs expected proportion. Genome-wide repartition of each feature was further investigated in relation to transcription start site based on RefSeq gene reference data.

Based on enrichment and genomic localisation, feature 3 was annotated as candidate active promoter (H3K4me3), feature 1 as candidate inactive promoter (H3k27me3), features 0 and 6 as candidate enhancers (H3K1me3 and H3K9me3, respectively) and feature 5 as transcribed sequences (H3K36me3). Finally, features 2 and 4 did not show sufficient specific enrichment to be annotated. Every HpaII site was assigned to a genomic feature.

Functional enrichment analysis The Bioconductor package GoSeq [28] was adapted to control for bias associated with the variation in the number of HpaII sites associated with different genes as previously described [10]. The candidate DML were assigned to genes only if they overlapped with candidate promoters within ±2 kb (features 1 and 3) or candidate enhancers within ±5 kb (features 0 and 6), features with a greater likelihood of having functional consequences. KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis was performed using candidate DML and false discovery rate (FDR) <0.05 was applied to determine significantly different pathways. Permutation analysis was performed to confirm significance of these results, 1,000 iterations of the same number of randomly selected genes failing to identify the same pathways ($p < 0.001$). Transcription factor analysis was performed by Gene Set Enrichment Analysis (GSEA) [29].

Bump hunting To identify differentially methylated regions (DMRs) we used bump hunting, specifically the dmrFind algorithm within the CHARM package (Bioconductor, charm version 2.16.1, https://www.bioconductor.org/packages/release/bioc/html/charm.html) [30]. DMRs are defined by four or more consecutive DML within a 1 kb distance. DMRs 1 kb upstream of the transcription start site (TSS) or within the gene body were mapped to the corresponding gene for further analysis.

Literature search The PubMed database was searched for publications between 2000 and 2015 using the search terms IUGR and methylation, metabolic disease, T cells or T cell-associated genes. Overlaps between DML from the present study and those in the literature were identified.

Results

Genome-wide DNA methylation profiles Genome-wide DNA methylation profiles of purified $CD3^+$ T cells from the cord blood of Cohort 1 (AGA $n = 8$; IUGR $n = 7$) demonstrated a shift towards hypermethylation in IUGR (Fig. 1a). From the 4,425 loci defined as candidate DML, 3,973 loci (89.8%) showed increased DNA methylation (hypermethylation) and 452 loci (10.2%) showed decreased methylation (hypomethylation) in IUGR compared with control (Fig. 1b). This subset of loci is sufficient to dissociate AGA samples from IUGR samples as illustrated by the heatmap in Fig. 1c. Furthermore, hypermethylation in IUGR is generally seen throughout the genome (Fig. 1d).

Based on this annotation, DML were assigned to 1,807 genes. Interestingly, 318 of these genes account for two or more DML (Table 2). The gene accounting for the most DML is *PTPRN2*, with 13 DML (12 hypermethylated CpGs and one hypomethylated CpG). *PRDM16*, which is involved in differentiation of brown adipose tissue [31] and stem cell maintenance [32], was hypermethylated on six CpG sites. *NCOR2*, a co-repressor of importance for homeostatic processes, inflammation and circadian rhythms [33], had five hypermethylated CpGs and one hypomethylated CpG. We then assessed the presence of regions of vulnerability to methylation across the whole genome. Regions of vulnerability, or DMRs, are defined as regions with average methylation over several CpGs that are significantly different between AGA

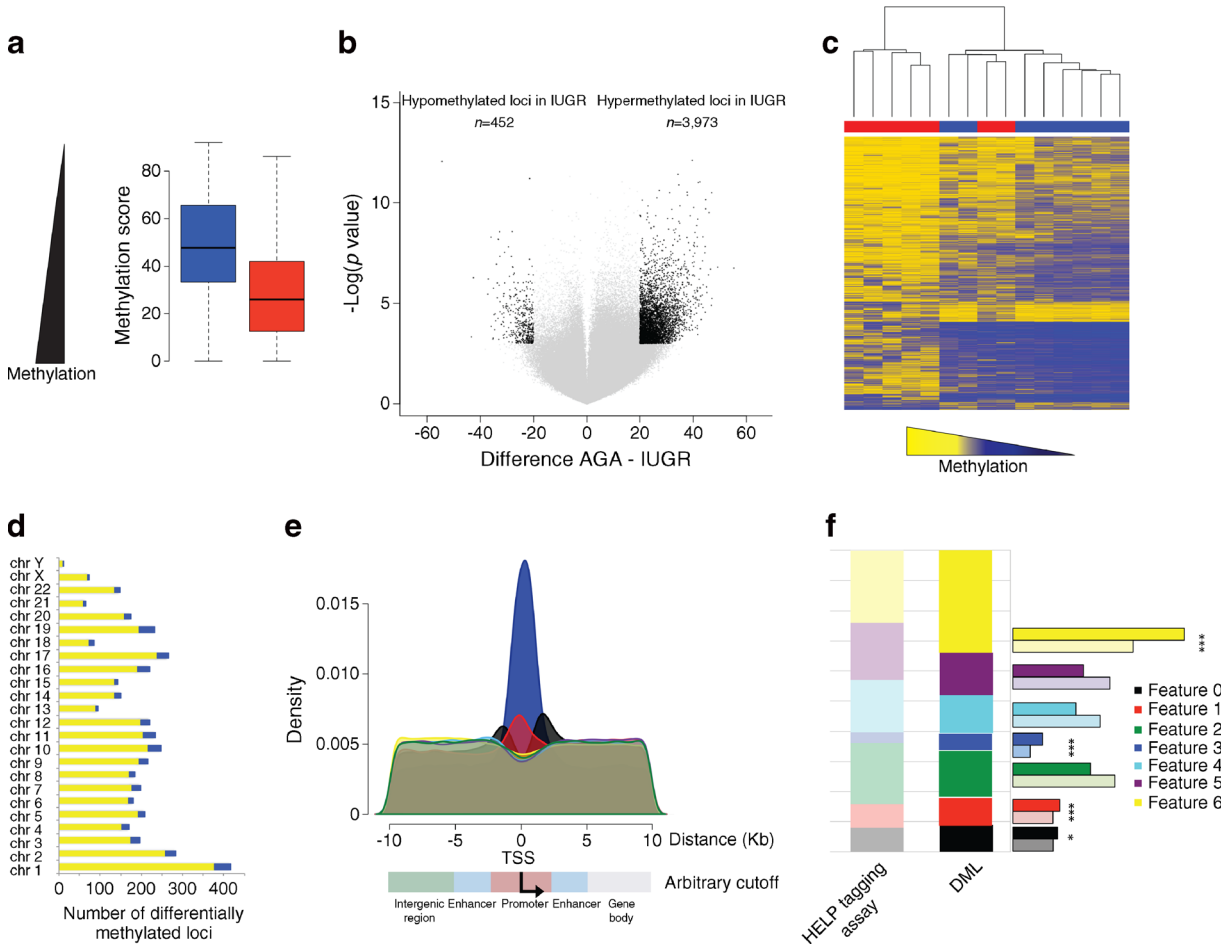

Fig. 1 Dysregulated methylation in CD3$^+$ T cells from umbilical cord blood of IUGR neonates. (**a**) Box plot of methylation scores (0, fully methylated; 100, unmethylated) for IUGR (red) and AGA (blue) infants. The thick horizontal line represents the median value, the whiskers represent the distribution of values and the edges of the rectangle represent the interquartile range. (**b**) Volcano plot of DNA methylation score differences between AGA and IUGR infants based on 1,077,400 loci throughout the genome. DML with $p < 0.05$ and methylation difference |>20| are shown in black. (**c**) Self-organising heatmap of candidate DML showing clustering by sample. (**d**) Chromosomal distribution of DML; yellow, hypermethylated DML; blue, hypomethylated DML. (**e**) Density plot of Segway features within a 10 kb window from the TSS. (**f**) Bar plot representing the proportional distribution of each feature in terms of loci tested by HELP tagging vs the proportions of features at DML. The histogram to the right compares the proportion of loci within each feature between the tested loci vs the DML. Significant enrichment are as follows: *$p < 0.05$ and ***$p < 0.001$

and IUGR infants. Bump-hunting analysis identified 45 DMRs. Of these, *PRDM16* [31, 32] and *HLA-DPB1*, which is important in regulating activation of CD8$^+$ cells [34], were identified (Table 3).

Functional analysis Using a more precise analysis to assess the functional relevance of the DML, we first assigned each locus to a specific feature using our generated CD3$^+$ T cell-specific genome annotation (Fig. 1e). Of the 4,425 DML, 4,125 loci were assigned to genomic features 0–6 and, when compared with the distribution of tested loci by feature, DML were significantly enriched in Segway features 6 ($p < 0.001$) and 0 ($p < 0.05$) (corresponding to candidate inactive and active enhancer, respectively), feature 3 ($p < 0.001$) (candidate active promoter) and feature 1 ($p < 0.001$) (candidate inactive promoter), suggesting that the epigenetic dysregulation of CD3$^+$ T cells in IUGR is targeted towards regulatory transcriptional elements (Fig. 1f).

Pathways analysis was then performed using a modified GoSeq approach. The top 50 significant pathways using KEGG database are listed in ESM Table 4. The primary immunodeficiency pathway contained T cell-related genes that were differentially methylated between AGA and IUGR, as well as genes involved in the differentiation of helper T cells. Other significant pathways demonstrating differential methylation between AGA and IUGR infants include type 2 diabetes and mitogen-activated protein kinase (MAPK), phosphatidylinositol and Wnt signalling.

Transcription factor analysis was performed to determine whether DML localised to transcription factor binding sites. As shown in Table 4, DML are enriched in binding sites for

Table 2 Number of genes with DML and the number of DMLs in the gene

Number of DMLs in gene	Number of genes	Gene name
13CpG	1	*PTPRN2*
12CpG	1	*RBFOX3*
11CpG	0	
10CpG	0	
9CpG	0	
8CpG	1	*CDH4*
7CpG	4	*CSMD1*, *GRID1*, *OPCML*, *SORCS2*
6CpG	5	*CAMTA1*, *NCOR2*, *PRDM16*, *RPTOR*, *SHANK2*
5CpG	12	*ANKRD20A4*, *CDH23*, *COL18A1*, *DOK7*, *DUX4L4*, *IQSEC3*, *PDE4D*, *PTPRT*, *SRCIN1*, *SYT7*, *TSNARE1*, *XKR6*
4CpG	18	*ABLIM2*, *CACNA1C*, *COL22A1*, *DIP2C*, *DNAH17*, *DSCAML1*, *ERBB4*, *FAM19A5*, *FRMD4A*, *LMX1A*, *MCF2L*, *MEGF11*, *OLFM1*, *PAX3*, *RBFOX1*, *SLIT1*, *TBCD*, *UNC5B*
3CpG	47	–
2CpG	229	–
1CpG	1,490	–

TCF3, *LEF1* and *NFATC*, transcription factors involved in T cell development and T cell receptor signalling pathways.

Validation To be assured of the significance of our findings, several validation steps were performed. First, using single locus quantitative validation studies, reproducibility of the DNA methylation differences at candidate DML using two different techniques, bisulfite MassARRAY and targeted bisulfite sequencing (TBS), was tested. MassARRAY analyses confirmed a strong correlation between MassARRAY and HELP-tagging data ($R^2 = 0.921$) (data not shown).

Biological validation was then performed in an independent cohort consisting of two AGA and four IUGR infants. We performed TBS at 55 loci in all 21 infants (Cohort 1 and Cohort 2). The methylation scores obtained from TBS and HELP tagging were strongly inversely correlated ($R^2 = 0.703$) (Fig. 2a) demonstrating the technical robustness of the HELP tagging assay (ESM Table 5). Several genes playing key roles in T cell function, metabolic disease pathogenesis and DNA methylation, including *NFATC2*, *GCK*, *PRDM16*, *MMP9*, *LDLR*, *KCNQ1* and *DNMT3B*, were validated (Figs 2b and 3).

Further, comparison of DML from $CD3^+$ T cells with those previously found in $CD34^+$ haematopoietic progenitor stem cells purified from IUGR infants [10] revealed a direct overlap of 35 precise loci between the two cell types (ESM Table 6). Validation of *MARK2* and *CDH12* was performed using bisulfite MassARRAY as Wnt signalling genes were previously identified as hypermethylated in $CD34^+$ cells from IUGR [10] (Fig. 4).

Finally, we surveyed the list of DML for genes previously shown to play a role in T cell development as well as for those shown to be differentially methylated in other IUGR cohorts (ESM Table 7). Overlapping DML were identified in 13 genes that play a role in T cell development (e.g. *NFATC2* and *TCF3*) and 15 genes previously shown to be linked to IUGR (e.g. *KCNQ1*, *INSR* and *RXRA*).

Discussion

We found an overall shift towards hypermethylation in $CD3^+$ T cells of IUGR compared with AGA male infants. Furthermore, the DML were targeted specifically to regulatory regions of genes in key pathways known to have downstream metabolic and inflammatory effects.

$CD3^+$ cell differentiation plays a key role in the inflammatory state seen in metabolic disease, and alteration of the $CD4^+$:$CD8^+$ cell ratio may have downstream effects in the development of age-associated diseases [14]. Interestingly, an analysis of transcription factors revealed that several involved in the regulation of $CD4^+$ and $CD8^+$ cells were differentially methylated in IUGR infants. These include *NFATC2*, which is critical in the development and function of the immune system, T cell differentiation and thymocyte development [35], and *TCF3*, which functions in T cell development [36] and the regulatory circuitry of embryonic stem cells to influence self-renewal and pluripotency [37]. Using $CD34^+$ haematopoietic stem cells, Delahaye et al identified several genes, also involved in stem cell differentiation, as being differentially methylated in IUGR offspring [10]. From these, *RXRA* (a binding partner for *PPARA*), *CREBBP* (involved in the transcriptional co-activation of numerous transcription

Table 3 Genes identified by bump hunting

Chr	Start	End	Refseq	Gene
Chr1	1229386	1230044	NM_030649	*ACAP3*
Chr1	3105013	3105616	NM_022114	*PRDM16*[a]
Chr1	17053828	17054531		
Chr1	32231409	32232907	NM_001703	*BAI2*
Chr1	42609618	42611568		
Chr2	1215843	1216892	NM_018968	*SNTG2*[a]
Chr2	122659926	122660898		
Chr2	131046362	131046475		
Chr2	239036191	239036813	NM_194312	*ESPNL*
Chr3	58621118	58622080	NM_138805	*FAM3D*[a]
Chr4	7124404	7124611		
Chr5	1856656	1857343		
Chr5	77142738	77144046		
Chr5	139057708	139058750	NM_016463	*CXXC5*
Chr5	149485756	149487572	NM_005211	*CSF1R*
Chr6	33048530	33048938	NM_002121	*HLA-DPB1*[a]
Chr8	1320665	1321638		
Chr8	21540181	21540918		
Chr8	58055166	58056159		
Chr9	36980179	36981484	NM_016734	*PAX5*[a]
Chr9	90622042	90622857		
Chr9	124988609	124989739	NM_199160	*LHX6*
Chr10	413784	414154	NM_014974	*DIP2C*[a]
Chr10	80940939	80941651	NM_020338	*ZMIZ1*[a]
Chr10	81053183	81053987	NM_020338	*ZMIZ1*[a]
Chr10	132847006	132847505		
Chr10	133048757	133049574	NM_174937	*TCERG1L*[a]
Chr10	134778649	134778988		
Chr12	132973731	132974247		
Chr13	28395096	28397352		
Chr14	24631391	24631505	NM_006084	*IRF9*[a]
Chr14	106347613	106349218		
Chr15	27087673	27088015		
Chr15	74520957	74521366		
Chr16	32199346	32200057		
Chr16	32264750	32265760		
Chr16	32752014	32752724		
Chr16	33140329	33141040		
Chr16	33205690	33206692	NM_016212	*TP53TG3*
Chr16	33262225	33263227	NM_001099687	*TP53TG3B*
Chr17	80449470	80449822		
Chr21	41282715	41283895	NM_006198	*PCP4*
Chr21	46850587	46851389	NM_130445	*COL18A1*
Chr22	37732941	37734380		
Chr22	49075586	49075908	NM_015381	*FAM19A5*[a]

Genes were identified by bump hunting using a 1 kb window

[a] Genes that also have DML

Table 4 Transcription factor binding sites enriched in DML in $CD3^+$ T cells from cord blood

Transcription factor	Motif	FDR q value
TCF3	CAGGTG	2.41×10^{-35}
No match to any known transcription factor	AACTTT	7.69×10^{-19}
LEF1	CTTTGT	4.61×10^{-18}
NFATC	TGGAAA	8.18×10^{-16}
MLLT7	TTGTTT	1.91×10^{-15}

factors) and *PRR5* (plays a role in platelet-derived growth factor signalling) were also identified in $CD3^+$ T cells from IUGR offspring. In addition, *MARK2*, *CDH12* and *TCF3*, all Wnt signalling-related genes, were hypermethylated in IUGR offspring in both studies [10]. Wnt signalling has previously been implicated in T cell development, activation and differentiation [38]. *MARK2* is downstream of *WNT5A*, which is involved in inflammatory diseases, metabolic disorders and embryonic development [39, 40]. *CDH12* is associated with BMI [41] and also mediates calcium-dependent cell–cell adhesion, which is regulated by the canonical Wnt signalling pathway [42]. Our results further reinforce the possibility that differential methylation of key genes in the Wnt signalling pathway may be important in programming adult disease and may be conserved through differentiation of stem cells. Furthermore, our analysis also showed methylation changes in genes related to the MAPK signalling pathway. Dysregulation of MAPK signalling has been linked to obesity and diabetes and plays a role in insulin signalling [43]. Additionally, *PRDM16*, a transcription cofactor with an important role in the development of brown adipose tissue in mice [31] and later identified as important for establishment and maintenance of the haematopoietic stem cell pool balance through the regulation of genes and transcription regulators [44] and by regulating reactive oxygen species levels (ROS) [32], was identified as both a DMR and a differentially methylated locus in IUGR. This finding strongly suggests that the epigenetic regulation of *PRDM16* may play a role in IUGR-associated metabolic disease susceptibility, including alteration of stem cell populations in multiple organs as well as the level of ROS, known to be elevated in metabolic disease. Therefore the altered methylation status of these genes provides a strong potential for dysregulation of T cell function and activation and may alter stem cell differentiation in individuals subject to IUGR.

We observed alteration in methylation patterns in genes previously associated with obesity and cardiovascular disease pathogenesis: stable methylation of *MMP9* in PBLs, which has been associated with increased BMI in obese adults [45]; promoter hypermethylation of *LDLR* observed in patients with atherosclerosis [46] and for which genome-wide association studies have shown polymorphisms in individuals with coronary artery disease [47] and hypermethylation of *GCK* [48], *FOXP1*, *MACROD2*, *JARID2*, *ZMIZ1* and the

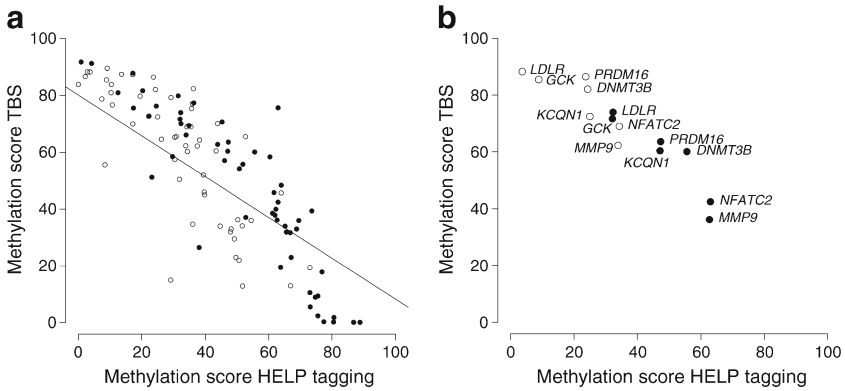

Fig. 2 Technical validation. (**a**) Correlation between HELP tagging and TBS ($R^2 = 0.70285$). (**b**) Genes discussed in the manuscript. Black circles, AGA; white circles, IUGR

maternally imprinted gene, *KCNQ1* [49], as a predictor for type 2 diabetes susceptibility. The agreement between our results and those of previous studies suggests that altered methylation of metabolically relevant genes may be one mechanism by which IUGR leads to increased susceptibility to metabolic disease.

Finally, *DNMT3B* (DNA methyltransferase 3B) was also hypermethylated in IUGR. *DNMT3B* is responsible for establishing de novo DNA methylation. Previous studies have shown that altered *DNMT3B* activity is associated with hypermethylation of a key regulator of oxidative capacity, *PGC1α* (also known as *PPARGC1A*), in the skeletal muscle of individuals with type 2 diabetes [50]. Thus, epigenetic dysregulation of a key gene involved in DNA methylation patterns identified here in T cells of IUGR neonates as well as skeletal muscle of type 2 diabetic adults reveals *DNMT3B* to be a potential biomarker for metabolic disease pathogenesis.

Our results demonstrate that poor fetal growth is associated with alterations in DNA methylation profiles in $CD3^+$ T cells. Furthermore, our results are consistent with findings from other cell types in earlier IUGR studies [10, 49]. Using our genome annotation to identify candidate promoters and enhancers in regions of interest, we can narrow the search to relevant changes in methylation that are more likely to have biologically significant downstream consequences.

While our study focused on the associations between IUGR and changes in methylation levels in $CD3^+$ T cells at birth, a key question that remains unaddressed is whether epigenetic dysregulation seen at birth persists into adulthood and plays a causative role in disease pathogenesis. We acknowledge that changes in methylation may not correlate with changes in gene expression and thus we cannot comment on the significance of the DML with respect to their potential to alter $CD3^+$ T cell gene expression. Future studies assessing the transcriptome as well as methylation profiles in specific subpopulations or ratio of $CD4^+$:$CD8^+$ cells present in cord blood will be useful in determining whether abnormal growth in early life can skew differentiation toward specific cell lineages with less favourable

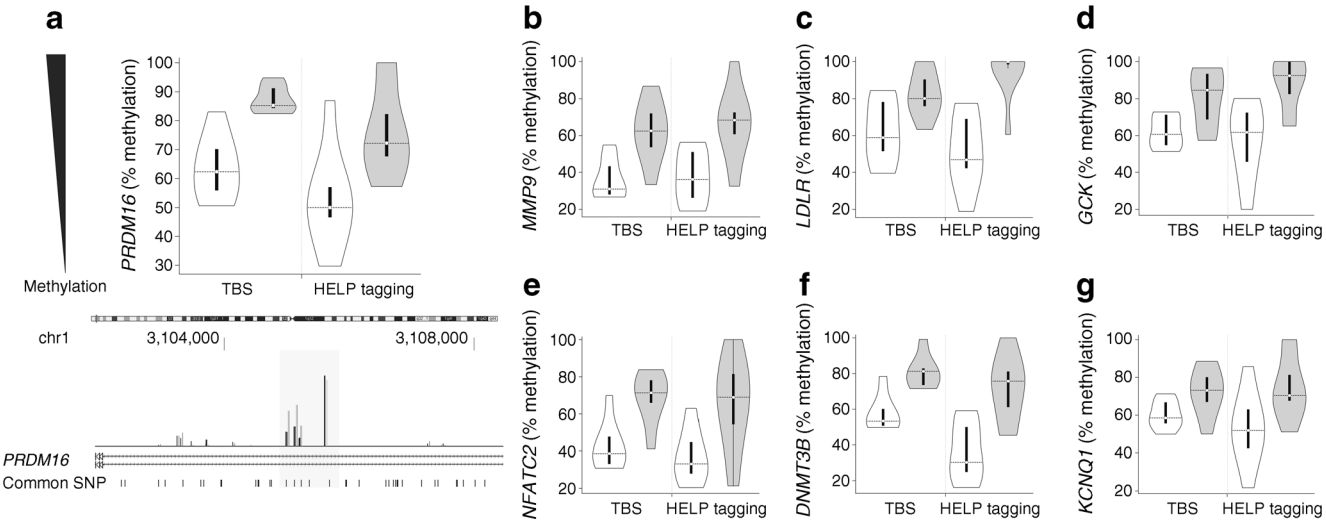

Fig. 3 Biological validation of HELP tagging data. (**a**) Violin plot representing DNA methylation repartition in AGA (white) and IUGR (grey) from TBS and HELP tagging for *PDRM16*. Per cent methylation measured by HELP tagging for several consecutive loci identified as DMRs by bump hunting within *PRDM16* is also shown. (**b**–**g**) Violin plots showing DNA methylation repartition from TBS and HELP tagging for *MMP9* (**b**), *LDLR* (**c**) and *GCK* (**d**), *NFATC2* (**e**), *DNMT3B* (**f**) and *KCNQ1* (**g**). AGA $n = 10$, IUGR $n = 11$; chr, chromosome. The horizontal line represents the median value, and the interquartile range is depicted by the thick black bar

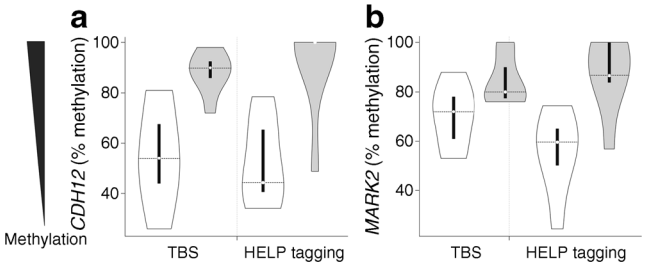

Fig. 4 Genes involved in Wnt pathways showing hypermethylation in both $CD3^+$ and $CD34^+$ cells from IUGR. Validation of DMLs with bisulfite MassARRAY. Violin plot representing DNA methylation repartition for AGA (white) and IUGR (grey) from MassARRAY and HELP tagging for *CDH12* (**a**) and *MARK2* (**b**). AGA $n = 8$, IUGR $n = 7$. The horizontal line represents the median value, and the interquartile range is depicted by the thick black bar

sequelae (i.e. greater inflammatory response). In addition, there are no studies to date that longitudinally examine the methylation changes that occur normally from birth through to adolescence and adulthood. Establishing 'normal' age-related methylation profiles will enable the identification of perturbed loci during the life course. Finally, although we have focused our attention on methylation changes, clearly the iteration of methylation and DNA sequence variation with other epigenetic regulators contribute to individual phenotype. Indeed, the role of genotype in disease susceptibility is a critical factor. A more comprehensive investigation using a multi-layered approach in a larger number of study participants would be ideal. However, these types of studies are prohibited by cost and are not likely to occur without a concerted effort from large collaborative groups. As such, the DNA methylation signatures described here provide an important framework for investigating candidate loci for more targeted experiments, perhaps to include genotypic variations and/or histone modifications that may contribute to the long-term disease susceptibility seen in adults who had been exposed to IUGR.

Funding This study was supported by the American Diabetes Association (1-13-CE-06, MJC) and the National Institutes of Health (HD063791, FHE). Additional support was generously provided by the Institute of Clinical and Translational Research of AECOM (to MJC). LW was the recipient of a National Institutes of Health, Ruth Kirschstein predoctoral fellowship (F31 DK093332).

Duality of interest The authors declare that there is no duality of interest associated with this manuscript

Contribution statement MJC, FHE and MF contributed to the conception and study design, interpretation of data and drafting and revision of the manuscript. LW and YS contributed to the acquisition, analysis and interpretation of data, and drafting and revision of the manuscript. FD contributed to the analysis and interpretation of data and revision of the manuscript. AC contributed to the analysis of the data and revision of the manuscript. All authors approved the final version to be published. MJC is the guarantor of this work.

References

1. Barker DJ (1992) The fetal origins of adult hypertension. J Hypertens Suppl 10:S39–S44
2. Law CM, Shiell AW, Newsome CA et al (2002) Fetal, infant, and childhood growth and adult blood pressure: a longitudinal study from birth to 22 years of age. Circulation 105:1088–1092
3. Hales CN, Barker DJ (1992) Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. Diabetologia 35:595–601
4. Boney CM, Verma A, Tucker R, Vohr BR (2005) Metabolic syndrome in childhood: association with birth weight, maternal obesity, and gestational diabetes mellitus. Pediatrics 115:290–296
5. Williams L, Seki Y, Vuguin PM, Charron MJ (2014) Animal models of in utero exposure to a high fat diet: a review. Biochim Biophys Acta 1842:507–519
6. Martin-Gronert MS, Ozanne SE (2010) Mechanisms linking suboptimal early nutrition and increased risk of type 2 diabetes and obesity. J Nutr 140:662–666
7. Mandruzzato G, Antsaklis A, Botet F et al (2008) Intrauterine restriction (IUGR). J Perinat Med 36:277–281
8. Roseboom T, de Rooij S, Painter R (2006) The Dutch famine and its long-term consequences for adult health. Early Hum Dev 82:485–491
9. Heijmans BT, Tobi EW, Stein AD et al (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A 105:17046–17049
10. Delahaye F, Wijetunga NA, Heo HJ et al (2014) Sexual dimorphism in epigenomic responses of stem cells to extreme fetal growth. Nat Commun 5:5187
11. Kanneganti TD, Dixit VD (2012) Immunological complications of obesity. Nat Immunol 13:707–712
12. Kohlgruber A, Lynch L (2015) Adipose tissue inflammation in the pathogenesis of type 2 diabetes. Curr Diab Rep 15:92
13. Richardson VR, Smith KA, Carter AM (2013) Adipose tissue inflammation: feeding the development of type 2 diabetes mellitus. Immunobiology 218:1497–1504
14. McLaughlin T, Liu LF, Lamendola C et al (2014) T cell profile in adipose tissue is associated with insulin resistance and systemic inflammation in humans. Arterioscler Thromb Vasc Biol 34: 2637–2643
15. Mukhopadhyay D, Weaver L, Tobin R et al (2014) Intrauterine growth restriction and prematurity influence regulatory T cell development in newborns. J Pediatr Surg 49:727–732
16. Williams PJ, Bulmer JN, Searle RF, Innes BA, Robson SC (2009) Altered decidual leucocyte populations in the placental bed in preeclampsia and foetal growth restriction: a comparison with late normal pregnancy. Reproduction 138:177–184
17. Jaeckle Santos LJ, Li C, Doulias PT, Ischiropoulos H, Worthen GS, Simmons RA (2014) Neutralizing Th2 inflammation in neonatal islets prevents beta-cell failure in adult IUGR rats. Diabetes 63: 1672–1684
18. Michels KB, Binder AM, Dedeurwaerder S et al (2013) Recommendations for the design and analysis of epigenome-wide association studies. Nat Methods 10:949–955
19. Luckheeram RV, Zhou R, Verma AD, Xia B (2012) $CD4^+$ T cells: differentiation and functions. Clin Dev Immunol 2012:925135
20. National cord blood program (2015) www.nationalcordbloodprogram.org, accessed 23 March 2016
21. Suzuki M, Jing Q, Lia D, Pascual M, McLellan A, Greally JM (2010) Optimized design and data analysis of tag-based cytosine methylation assays. Genome Biol 11:R36
22. Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article 3

23. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8:118–127
24. Ehrich M, Nelson MR, Stanssens P et al (2005) Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc Natl Acad Sci U S A 102: 15785–15790
25. Thompson RF, Suzuki M, Lau KW, Greally JM (2009) A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. Bioinformatics 25:2164–2170
26. Xi Y, Li W (2009) BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinf 10:232
27. Hoffman MM, Buske OJ, Wang J, Weng Z, Bilmes JA, Noble WS (2012) Unsupervised pattern discovery in human chromatin structure through genomic segmentation. Nat Methods 9:473–476
28. Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 11:R14
29. Subramanian A, Tamayo P, Mootha VK et al (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102: 15545–15550
30. Jaffe AE, Murakami P, Lee H et al (2012) Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. Int J Epidemiol 41:200–209
31. Seale P, Bjork B, Yang W et al (2008) PRDM16 controls a brown fat/skeletal muscle switch. Nature 454:961–967
32. Chuikov S, Levi BP, Smith ML, Morrison SJ (2010) Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. Nat Cell Biol 12:999–1006
33. Mottis A, Mouchiroud L, Auwerx J (2013) Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. Genes Dev 27: 819–835
34. Weng NP, Araki Y, Subedi K (2012) The molecular basis of the memory T cell response: differential gene expression and its epigenetic regulation. Nat Rev Immunol 12:306–315
35. Macian F (2005) NFAT proteins: key regulators of T cell development and function. Nat Rev Immunol 5:472–484
36. Yui MA, Rothenberg EV (2014) Developmental gene networks: a triathlon on the course to T cell identity. Nat Rev Immunol 14:529–545
37. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA (2008) Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev 22:746–755
38. Wong C, Chen C, Wu Q, Liu Y, Zheng P (2015) A critical role for the regulated Wnt-Myc pathway in naive T cell survival. J Immunol 194:158–167
39. Mamidi A, Inui M, Manfrin A et al (2012) Signaling crosstalk between TGFβ and Dishevelled/Par1b. Cell Death Differ 19: 1689–1697
40. Saadat I, Higashi H, Obuse C et al (2007) Helicobacter pylori CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity. Nature 447:330–333
41. Ng MC, Hester JM, Wing MR et al (2012) Genome-wide association of BMI in African Americans. Obesity (Silver Spring) 20:622–627
42. Heuberger J, Birchmeier W (2010) Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. Cold Spring Harb Perspect Biol 2:a002915
43. Gehart H, Kumpf S, Ittner A, Ricci R (2010) MAPK signalling in cellular metabolism: stress or wellness? EMBO Rep 11:834–840
44. Aguilo F, Avagyan S, Labar A et al (2011) Prdm16 is a physiologic regulator of hematopoietic stem cells. Blood 117:5057–5066
45. Feinberg AP, Irizarry RA, Fradin D et al (2010) Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci Transl Med 2:49ra67
46. Zhi YF, Huang YS, Li ZH, Zhang RM, Wang SR (2007) Hypermethylation in promoter area of LDLR gene in atherosclerosis patients. Fen Zi Xi Bao Sheng Wu Xue Bao 40:419–427 [article in Chinese]
47. Martinelli N, Girelli D, Lunghi B et al (2010) Polymorphisms at LDLR locus may be associated with coronary artery disease through modulation of coagulation factor VIII activity and independently from lipid profile. Blood 116:5688–5697
48. Tang L, Ye H, Hong Q et al (2014) Elevated CpG island methylation of GCK gene predicts the risk of type 2 diabetes in Chinese males. Gene 547:329–333
49. Hillman SL, Finer S, Smart MC et al (2015) Novel DNA methylation profiles associated with key gene regulation and transcription pathways in blood and placenta of growth-restricted neonates. Epigenetics 10:50–61
50. Barres R, Osler ME, Yan J et al (2009) Non-CpG methylation of the PGC-1alpha promoter through DNMT3B controls mitochondrial density. Cell Metab 10:189–198