The tissue-specific aspect of genome-wide DNA methylation in newborn and placental tissues: implications for epigenetic epidemiologic studies

Herzog, E.M.; Eggink, A.J.; Willemsen, S.P.; Slieker, R.C.; Felix, J.F.; Stubbs, A.P.; ... ; Steegers-Theunissen, R.P.M.

Journal of Developmental Origins of Health and Disease

www.cambridge.org/doh

The tissue-specific aspect of genome-wide DNA methylation in newborn and placental tissues: implications for epigenetic epidemiologic studies

Emilie M. Herzog[1], Alex J. Eggink[1], Sten P. Willemsen[1,2], Roderick C. Slieker[3], Janine F. Felix[4,5,6], Andrew P. Stubbs[7], Peter J. van der Spek[7], Joyce B. J. van Meurs[8], Bastiaan T. Heijmans[3] and Régine P. M. Steegers-Theunissen[1,9]

[1]Department of Obstetrics and Gynaecology, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands; [2]Department of Biostatistics, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands; [3]Department of Molecular Epidemiology, Leiden University Medical Centre, Postbus 9600, 2300 RC Leiden, The Netherlands; [4]Department of Epidemiology, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands; [5]Generation R Study Group, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands; [6]Department of Paediatrics, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands; [7]Department of Bioinformatics, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands; [8]Department of Internal Medicine, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands and [9]Department of Paediatrics, Division of Neonatology, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA Rotterdam, The Netherlands

Original Article

Cite this article: Herzog EM, Eggink AJ, Willemsen SP, Slieker RC, Felix JF, Stubbs AP, van der Spek PJ, van Meurs JBJ, Heijmans BT, and Steegers-Theunissen RPM. The tissue-specific aspect of genome-wide DNA methylation in newborn and placental tissues: implications for epigenetic epidemiologic studies. *Journal of Developmental Origins of Health and Disease* doi: 10.1017/S2040174420000136

Received: 23 June 2019
Revised: 31 December 2019
Accepted: 28 January 2020

Keywords:
EWAS; tissue specificity; HUVEC; umbilical cord blood; placenta

Address for correspondence:
Régine P. M. Steegers-Theunissen, MD, PhD, Professor in Periconception Epidemiology, Department of Obstetrics and Gynaecology, Erasmus MC, University Medical Centre Rotterdam, Postbus 2040, 3000 CA, Rotterdam, The Netherlands.
Email: r.steegers@erasmusmc.nl

Abstract

Epigenetic programming is essential for lineage differentiation, embryogenesis and placentation in early pregnancy. In epigenetic association studies, DNA methylation is often examined in DNA derived from white blood cells, although its validity to other tissues of interest remains questionable. Therefore, we investigated the tissue specificity of epigenome-wide DNA methylation in newborn and placental tissues. Umbilical cord white blood cells (UC-WBC, $n = 25$), umbilical cord blood mononuclear cells (UC-MNC, $n = 10$), human umbilical vein endothelial cells (HUVEC, $n = 25$) and placental tissue ($n = 25$) were obtained from 36 uncomplicated pregnancies. Genome-wide DNA methylation was measured by the Illumina HumanMethylation450K BeadChip. Using UC-WBC as a reference tissue, we identified 3595 HUVEC tissue-specific differentially methylated regions (tDMRs) and 11,938 placental tDMRs. Functional enrichment analysis showed that HUVEC and placental tDMRs were involved in embryogenesis, vascular development and regulation of gene expression. No tDMRs were identified in UC-MNC. In conclusion, the extensive amount of genome-wide HUVEC and placental tDMRs underlines the relevance of tissue-specific approaches in future epigenetic association studies, or the use of validated representative tissues for a certain disease of interest, if available. To this purpose, we herewith provide a relevant dataset of paired, tissue-specific, genome-wide methylation measurements in newborn tissues.

Introduction

Adverse influences during early prenatal development are associated with an increased risk of cardiovascular, metabolic and neurodevelopmental dysfunction in child- and adulthood which is in line with the 'Developmental Origins of Health and Disease' paradigm[1]. These associations are potentially mediated by epigenetic mechanisms[1]. Epigenetic reprogramming is essential during early embryonic development and placentation, when active and passive demethylation takes place immediately after fertilisation, followed by *de novo* methylation at the late morula stage[2,3].

Epigenetic epidemiological studies are essential to elucidate the role of DNA methylation of embryonic and extra-embryonic tissues, in the associations between prenatal exposures and health or disease in later life. Most studies, however, focus on DNA methylation levels in white blood cells as surrogate tissue because the target tissue of interest is not easily accessible in human studies[4]. In 1983, it was demonstrated in rodents and followed by many other studies that DNA methylation is established in a tissue-specific manner and as such accompanies lineage differentiation[5–7]. Recently, this has been addressed in epigenome-wide analyses of DNA methylation with attempts to identify differentially methylated positions (DMPs) and differentially methylated regions (DMRs) associated with tissue specificity or disease[8–14].

CAMBRIDGE UNIVERSITY PRESS

To further examine tissue specificity in this context, DNA methylation is best studied in newborn tissues such as placental tissue and umbilical cord (blood), since postnatal environmental exposures affecting DNA methylation can be excluded[15]. The placenta represents an essential extra-embryonic organ for embryonic and fetal growth and development. The epigenetic programming of the placenta is involved in the regulation of fetal demands and intrauterine conditions[16]. To examine the newborn vasculature as target tissue of linking prenatal exposure to future vascular health, human umbilical vein endothelial cells (HUVEC) are the best accessible representative tissue, and no prior genome-wide methylation studies have been performed in HUVEC before. During initial embryonic development and differentiation, the development of these cells is also highly regulated by various epigenetic mechanisms, influenced by external exposures and the local haemodynamic conditions of the pregnancy[17–19]. One of the challenges of whole blood epigenetic studies is the variation in blood cell mixtures as potential confounder for DNA methylation differences[20]. This can be investigated by comparing DNA methylation levels between umbilical cord white blood cells (UC-WBC) and a subgroup without granulocytes: umbilical cord blood mononuclear cells (UC-MNC).

Against this background, the objective of this study was to examine the tissue specificity of epigenome-wide DNA methylation of 450,000 methylation sites in placental tissue, HUVEC and UC-MNC, in comparison to the widely used UC-WBC as representative tissue, in order to ultimately increase the use of validated representative tissues for diseases of interest in future epigenome-wide association studies (EWAS). Although the genome-wide Illumina 450K array covers a relatively small fraction of the genome compared to high-resolution techniques, the regions measured are biologically informative and relevant for the surrounding genome biology, and larger sample size measurements are possible.

Methods

Study design

Pregnant women and their newborn babies were recruited between June 2011 and June 2013 in a nested case-control study embedded in the Rotterdam Periconceptional Cohort (Predict Study), at the Erasmus MC, University Medical Centre Rotterdam, The Netherlands[21].

Maternal and newborn characteristics

The case-control study aimed to examine genome-wide DNA methylation in pregnancies with various pregnancy-induced diseases. For the current study, we utilised 36 uncomplicated control pregnancies, from which data were generated for up to 25 cases per tissue type, as depicted in Fig. 1. Uncomplicated pregnancies were defined as pregnancies without the following pregnancy-specific complications: preeclampsia, gestational hypertension, fetal growth restriction or preterm birth. Women with HIV infection, those aged <18 years, those not able to read or understand the Dutch language, multiple birth pregnancies or women with pregnancies complicated by fetal congenital malformations were excluded. Due to the tertiary hospital setting where the study was carried out, chronic comorbidities, including endocrine, metabolic, autoimmune, renal or cardiovascular diseases, were not exclusion criterion. None of the woman smoked during pregnancy. Maternal and newborn characteristics were obtained from hospital medical records. All women gave written informed consent before participation, and parental informed consent was obtained for the child. All methods were performed in accordance with the principles expressed in the Declaration of Helsinki, and ethical approval for the study was given by the Erasmus MC, University Medical Centre Research Ethics Board (MEC-2004-227).

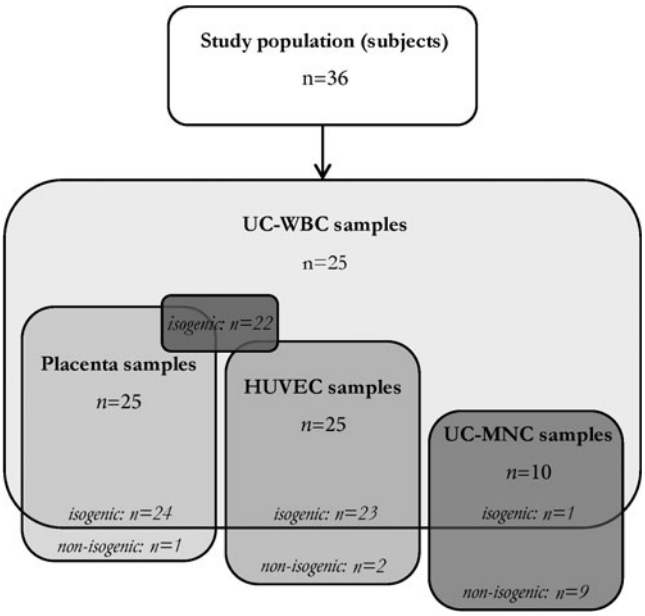

Fig. 1. Flow chart of the sample distribution in the study population. UC-WBC, umbilical cord blood white blood cells; UC-MNC, umbilical cord blood mononuclear cells; HUVEC, human umbilical vein endothelial cells.

Data collection

Umbilical cord blood (UCB) samples from the umbilical vein were collected with the placenta still *in situ*. Thereafter, placental tissue was obtained, and HUVEC were isolated and stored until DNA extraction. A detailed description of the data collection, UC-WBC and UC-MNC processing, HUVEC isolation and DNA extraction is provided in Appendix 1.

DNA methylation measurement

Isolated genomic DNA (500 ng) was treated with sodium bisulphite using the EZ-96 DNA methylation kit (Shallow) (Zymo Research, Irvine, CA, USA). Hybridisation was performed following manufacturer's instructions. DNA methylation of cytosine guanine dinucleotides (CpGs) was measured by the Illumina HumanMethylation450K BeadChip using the manufacturer's protocol (Illumina, Inc., San Diego, CA, USA)[22–24].

Data quality control and pre-processing

All data pre-processing and statistical analyses were performed in R software version 3.2.2 and Bioconductor version 3.1[25]. A quality control protocol was conducted in Illumina GenomeStudio software using the methylation module. It included a sample call rate check, colour balance check and control dashboard checks. Probes targeting a CpG with documented single-nucleotide polymorphisms in the C or G nucleotides were removed ($n = 17{,}196$) (minor allele frequency >0.05, European population, 1000 Genomes Project). Probes directed at the sex chromosomes ($n = 11{,}648$) and with a detection of p-value >0.01 in more than

1% of samples ($n = 2773$) were also excluded. Out of 485,512 probes, a total of 454,892 were left for further analysis. Normalisation was performed for all samples together by the Dasen method, which consists of background adjustment and between-array normalisation, applied to Type I and Type II probes separately (Bioconductor package watermelon version 1.80)[26]. The effect of the Dasen normalisation procedure was checked by plotting raw and normalised density plots per tissue, where the β-values with and without Dasen normalisation showed a very high correlation ($r > 0.99$, Supplementary Fig. S1).

tDMP identification

Methylation β-values were converted to M-values to achieve a normal distribution of the data using M-value = $\log_2$ (β-value/(1 − β-value))[27]. M-values were used for all statistical analyses, whereas β-values were depicted in the presentation of our data for a better biological understanding. For every CpG on the array, we estimated a linear mixed model with the M-value as response, tissue as categorical predictor and a random intercept for each subject, adjusted for bisulphite plate batch and gestational age to compare each tissue with the UC-WBC (R package 'lme4')[28]. Sensitivity analyses were performed for the following covariates: batch effect of the bisulphite plate, gestational age, birth weight, fetal sex, comorbidity and moment of inclusion for the study in or >1st trimester. Only gestational age and bisulphite plate were thereafter included in our statistical model as potential confounders. The effect of all potential technical and biological covariates on the methylation β-values is depicted in a correlation heatmap (Supplementary Fig. S2). It is demonstrated that the first three principle components are represented by the different tissues, explaining 98.7% of total variation present in our data. We additionally checked whether adjustment for variation in underlying cell populations in UC-WBC and placenta demonstrated an effect on our model. For this, we rerun our final model with additional covariates representing the first two principle components of imputed blood cell populations and imputed placental underlying cell variation using Houseman's data[20,29]. These additional covariates were not applied to our final model due to the very strong correlation that was demonstrated between the two models in HUVEC ($r = 0.996$) and placental samples ($r = 0.998$). The correlation in the UC-MNC samples is lower due to the initial very low tissue effect in this comparison, which is attenuated after correction for the different blood cell populations in UC-WBC ($r = 0.770$) (Supplementary Fig. S3). The potential effect of 'mode of delivery' was also excluded after comparing our final statistical model with and without additional covariate 'mode of delivery' (vaginal delivery versus caesarean section) and observing a very strong correlation between the two models in UC-MNC ($r = 0.95$), HUVEC ($r > 0.99$) and placental samples ($r > 0.99$) (Supplementary Fig. S4).

We used the following model:

$$M_{ij} = b_i + \beta_0 + \beta_1 T_{\text{HUVEC}} + \beta_2 T_{\text{PLACENTA}} + \beta_3 T_{\text{UC-MNC}} + \beta_4 GA + \beta_5 \text{PLATE}_2 + \beta_6 \text{PLATE}_3 + \varepsilon_{ij}.$$

M_{ij} is the logit of the methylation of individual i in tissue j, where j is UC-WBC, UC-MNC, HUVEC or placenta. b_i is the subject-specific intercept which is assumed to follow a normal distribution, T_{HUVEC}, T_{PLACENTA} and $T_{\text{UC-MNC}}$ are tissue-specific indicator variables, GA is the gestational age, and PLATE_2 and PLATE_3 are indicator variables for the sample plates. β_0 Is the intercept, and β_1 to β_7 are the regression coefficients indicating the effect of the various covariates. ϵ_{ij} is the measurement error, assumed to be independently normally distributed with a constant variance and mean of zero. All tissues were compared against one reference tissue: UC-WBC.

A false discovery rate (FDR)-adjusted p-value below 0.05 was considered significant. CpG sites were classified as tissue-specific differentially methylated position (tDMP), if statistically significant and presenting with a minimal effect size ΔM of 1.3 versus UC-WBC. We considered various M-value cut-off values for further data analysis and applied the widely used robust cut-off of 1.3 ΔM-value (~20% on β-value scale)[30] (Supplementary Table S1 and Supplementary Fig. S5). For the comparison with the smallest number of samples ($n = 10$ UC-WBC samples), we were able to measure methylation differences of at least 9.4% Δβ, with a power of 0.7 (SD of 0.05 and $p < 0.05$). The lambda-values for the different epigenome-wide analysis performed include 0.02 for UC-WBC versus UC-MNC, 6.4 for UC-WBC versus HUVEC and 76.6 for UC-WBC versus placenta. These values are to be expected looking at the amount of CpGs that were found significantly different and based on the underlying biology of our study samples.

A partial replication was performed by a correlation analysis between the estimates (β) for the tissue effect in our dataset with the Δβ of three publicly available Gene expression Omnibus (GEO) datasets for UC-WBC (GSE69176: $n = 152$ samples), HUVEC (GSE82234: $n = 6$ samples) and placental tissue (GSE75248: $n = 174$ samples with adequate gestational age range)[31,32]. Δβ was calculated as the Δβ mean methylation of HUVEC or placenta versus UC-WBC for all our tDMPs (Supplementary Fig. S7). To check the reproducibility of our tissue-specific differentially methylated regions (tDMRs), we compared our results to three previous EWAS with available tDMR datasets and similar study designs with respect to type of methylation array and tissues. We were able to compare our tDMR genes in HUVEC to tDMR genes in vascular tissue of Lokk *et al.*, our tDMR genes in placenta to amniotic tDMR genes of Slieker *et al.* 2015 (both extra-embryonic tissues), and all our tDMR genes in HUVEC and placenta to those of Slieker *et al.* 2013 using the online Jvenn tool[9,10,33,34] (Supplementary Table S2).

CpG density and gene-centric enrichment analysis of tDMPs

tDMPs were annotated according to their position relative to CpG islands and relative to genes using UCSC database. In relation to CpG islands, we identified CpG shores as the 2-kb CpG island flanking region and shelves as the 2-kb CpG shore flanking region. Remaining tDMPs were annotated as non-CpG island regions[9]. Enrichment analysis was applied to tDMPs. Relative to genes, tDMPs were annotated as gene body (+500 bp to 3' end of the gene), distal promotor (>10 to 1.5 kb from the nearest transcription start site (TSS)), proximal promotor (−1.5 to +500 bp from the nearest TSS), intergenic (>10 kb from the nearest TSS), and downstream regions (3' end to +5 kb from 3' end). Human genome build 37 was used for all annotations.

tDMR identification

tDMRs were generated according to a previously published algorithm as regions in which at least three tDMPs were detected with an inter-CpG distance ≤1 kb and not interrupted by more than three non-DMPs within the DMR (DMRfinder)[9].

Table 1. Maternal and newborn characteristics of uncomplicated control pregnancies ($n = 36$)

	UC-WBC ($n = 25$)	UC-MNC ($n = 10$)	HUVEC ($n = 25$)	Placenta ($n = 25$)	Overall ($n = 36$)
Maternal characteristics					
Age (years)	31.8 (5.4)	32.1 (4.7)	31.9 (5.4)	31.8 (5.3)	31.8 (5.1)
Nulliparous, *n* (%)	7 (28.0)	3 (30.0)	8 (32.0)	8 (32.0)	11 (30.6)
Ethnicity, *n* (%)					
Western geographic origin	20 (80.0)	9 (90.0)	20 (80.0)	20 (80.0)	30 (83.3)
Non-Western geographic origin	5 (20.0)	1 (10.0)	5 (20.0)	5 (20.0)	6 (16.7)
Preconceptional BMI (kg/m^2)	24.9 (4.4)	23.4 (3.1)	24.9 (4.4)	24.7 (4.5)	24.3 (4.0)
Comorbidity (yes), *n* (%)	6 (24.0)	1 (10.0)	7 (28.0)	7 (28.0)	8 (22.2)
Newborn characteristics					
Males, *n* (%)	14 (56.0)	5 (50.0)	16 (64.0)	14 (56.0)	21 (58.3)
Gestational age at birth* (weeks)	39.9 (1.9)	39.8 (1.9)	39.9 (1.7)	39.9 (1.9)	39.7 (1.1)
Birth weight (g)	3815 (396)	3352 (377)	3811 (364)	3788 (398)	3691 (418)

BMI, body mass index; UC-WBC, umbilical cord white blood cells; UC-MNC, umbilical cord blood mononuclear cells; HUVEC, human umbilical vein endothelial cells.
Data are presented as mean (standard deviation) or number (%).
*Non-parametric data are presented as median (interquartile range). No statistical testing has been performed due to overlapping pregnancies between the tissue groups.

Gene Ontology term enrichment analysis and Ingenuity Pathway Analysis of tDMR genes

HUVEC and placental tDMRs were mapped to the nearest gene based on Ensembl annotations from UCSC, also when facing multiple genes. Assigned Ensembl genes were uploaded to the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool to examine possible enrichment of corresponding gene ontology (GO) terms using the GO_BP_FAT annotation category (DAVID Bioinformatics Resources 6.7)[35,36]. Fisher Exact was applied to measure the gene enrichment of annotated GO terms of the uploaded gene list, against the whole human genome list as a background. To focus on the biology of the annotated GO terms, clusters of similar annotations were examined from the DAVID Functional Annotation Clustering table tool. The clustering algorithm is based on the hypothesis that similar annotations have similar gene members, resulting in a group enrichment score to rank their biological significance.

We additionally conducted Ingenuity Pathway Analysis (IPA) with the annotated DMP Ensembl gene lists to validate the DAVID enrichment and focused on canonical pathways and networks. Associated canonical pathways were subjected to Benjamini–Hochberg procedure for controlling FDR ($p < 0.05$). Networks were generated based on network eligible molecules, which were encoded by our DMP genes and also interact with other molecules in the Ingenuity Pathways Knowledge Base. A high score for a network indicates a more approximate fit between network eligible molecules and the molecules that constitute the network, calculated using the right-tailed Fisher's exact test.

Results

Tissues from 36 uncomplicated control pregnancies were studied, derived from a nested case-control study embedded in the Rotterdam Periconceptional Cohort[21]. From a large majority of 22 pregnancies, we obtained isogenic UC-WBC, HUVEC and placental samples, and in four pregnancies, one or two tissues were obtained. Nine pregnancies provided UC-MNC samples only. This resulted in 25 UC-WBC, 10 UC-MNC, 25 HUVEC and 25 placental samples (Fig. 1).

Maternal and newborn characteristics of the four tissue groups are shown in Table 1 and were overall comparable. The mean age of the women was 32 years and 83% were of Western geographic origin. Newborns were born at a mean gestational age of 40 weeks and 58% were male. In the UC-MNC tissue group, comorbidity was present in 10% of the pregnancies and the mean newborn birth weight was 3352 g versus the other tissue groups presenting with comorbidity in 24% to 28% and a mean birth weight around 3800 g.

tDMP identification

Genome-wide DNA methylation data revealed an overall bimodal distribution (Supplementary Fig. S6). In UC-WBC, UC-MNC and HUVEC, only a small amount of CpGs demonstrated methylation levels around 50%, in contrast to a relatively large amount of CpGs with around 50% methylation in placental tissue. Most annotated regions demonstrated median methylation levels between 50% and 85%, except for CpG islands and proximal promotor regions, which revealed lower median methylation levels of around 10%, (Supplementary Table S3).

A heatmap based on the clustering of methylation according to CpG and sample revealed clustering of the three different tissues (UCB, HUVEC and placenta) (Fig. 2). No distinctive clusters were identified for the two different UCB cell fractions. On an epigenome-wide level, we identified tissue-specific CpG methylation using UC-WBC as reference tissue. We observed 1636 (0.4%) differentially methylated CpGs between UC-WBC and UC-MNC, 193,945 (43%) between UC-WBC and HUVEC and 333,061 (73%) between UC-WBC and placental tissue (all FDR-adjusted $p < 0.05$). Those statistically differentially methylated CpGs with an additional effect size >1.3 in M-value (corresponding to logit2 of the β-value: ~Δ 20% β-value) were defined as tDMPs. This revealed 2 (4×10^{-6}%) MNC-tDMPs, 49,979 (11%) HUVEC tDMPs and 126,482 (28%) placental tDMPs, in comparison to UC-WBC (Fig. 3a, Supplementary Table S1). We provided a

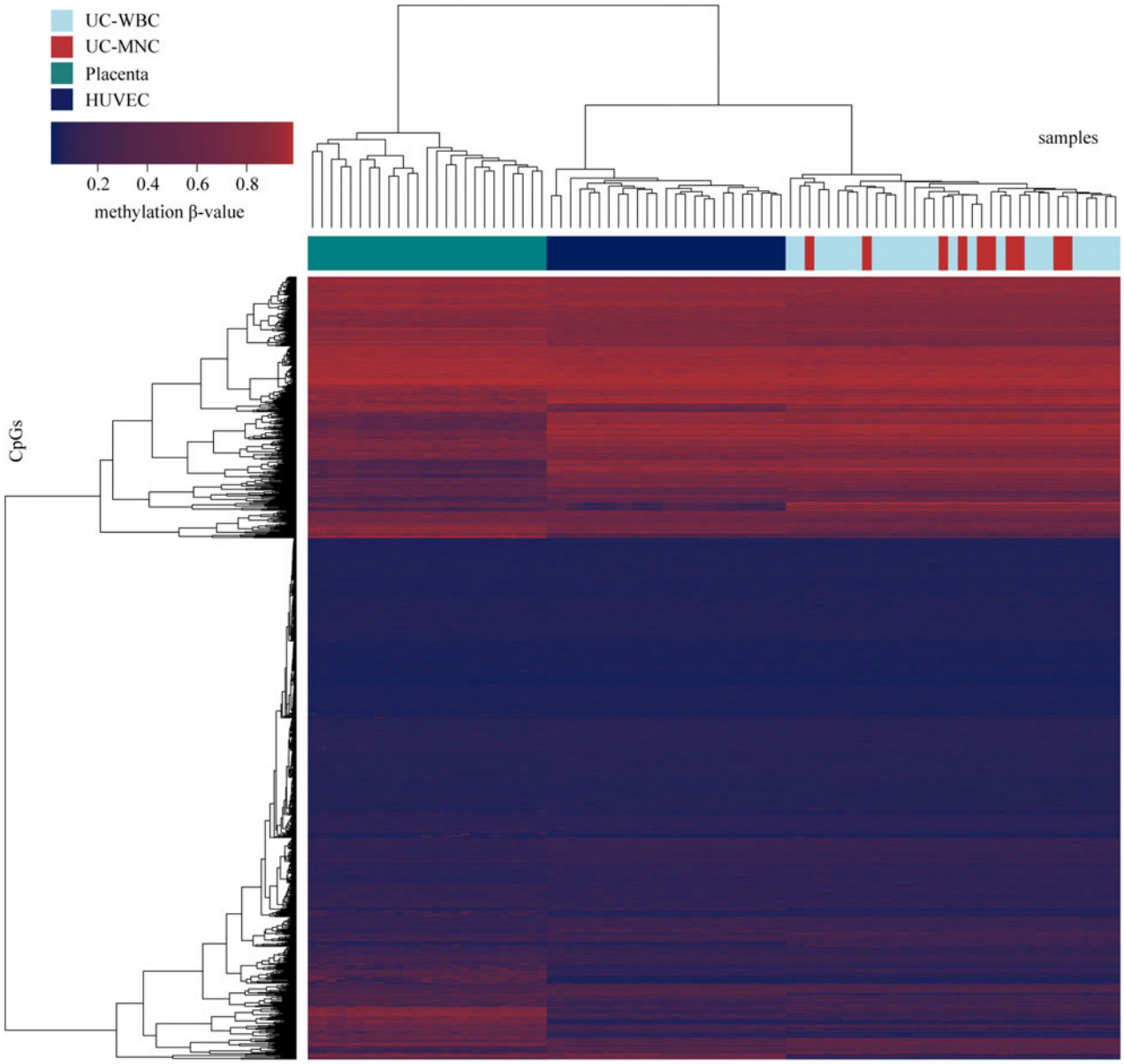

Fig. 2. Heatmap based on clustering of methylation (β-value) of all differentially methylated CpGs. Samples are depicted on the horizontal axis and CpGs on the vertical axis. Samples cluster by tissue type, without distinctive clustering for the two different UCB cell fractions. β-Values are depicted for a better biological understanding of the figure. UC-WBC, umbilical cord white blood cells; UC-MNC, umbilical cord blood mononuclear cells; HUVEC, human umbilical vein endothelial cells.

Supplementary Table with all tDMPs observed in UC-MNC, HUVEC and placenta, including CpG identifiers, β-values and *p*-values (Supplementary Table S2). A partial replication demonstrated a strong correlation between the estimates (β) for the tissue effect in our own dataset and the Δβ methylation differences of HUVEC and placenta versus UC-WBC of three independent datasets for all our tDMPs ($r = 0.94$ in HUVEC samples and $r = 0.98$ in placental samples, $p < 0.0001$) (Supplementary Fig. S7).

CpG density and gene-centric enrichment of tDMPs

To evaluate whether tDMPs were enriched in certain genomic annotations, CpG island and gene-centric annotations of HUVEC and placental tDMR genes were examined (Fig. 4, Supplementary Table S4). tDMP annotation relative to CpG islands showed that tDMPs were significantly enriched in CpG shores, shelves and especially in non-CpG island regions (non-CPG island odds ratio (OR)^{HUVEC} 2.02, 95% confidence interval (CI) 1.98–2.06, OR^{Placenta} 1.54, 95% CI 1.52–1.56) and strongly depleted in CpG islands (CpG island OR^{H} 0.47, 95% CI 0.47–0.48, OR^{P} 0.27, 95% CI 0.26–0.27). Both annotation patterns were concordant for HUVEC and placental tDMPs.

tDMPs were significantly enriched in all gene-centric regions except for a strong depletion of tDMPs in proximal promotors (OR^{HUVEC} 0.55, 95% CI 0.54–0.57, OR^{Placenta} 0.49, 95% CI 0.48–0.50). The strongest placental enrichment was observed in intergenic regions. HUVEC's strongest enrichment was observed in gene bodies (Fig. 4, Supplementary Table S4).

A combined gene and CpG island annotation of tDMPs is depicted in Supplementary Fig. S8. A prominent depletion of tDMPs in proximal promotors due to a strong underrepresentation of tDMPs in CpG island proximal promotors was observed

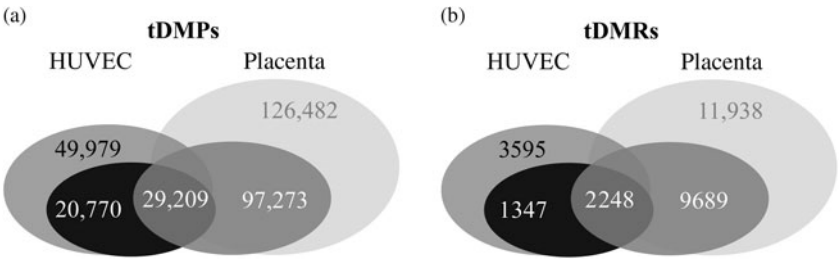

Fig. 3. Venn diagram of the number of unique and overlapping HUVEC and placental (a) tDMPs and (b) tDMRs, relative to UC-WBC. tDMPs, tissue-specific differentially methylated positions; tDMRs, tissue-specific differentially methylated regions; HUVEC, human umbilical vein endothelial cells.

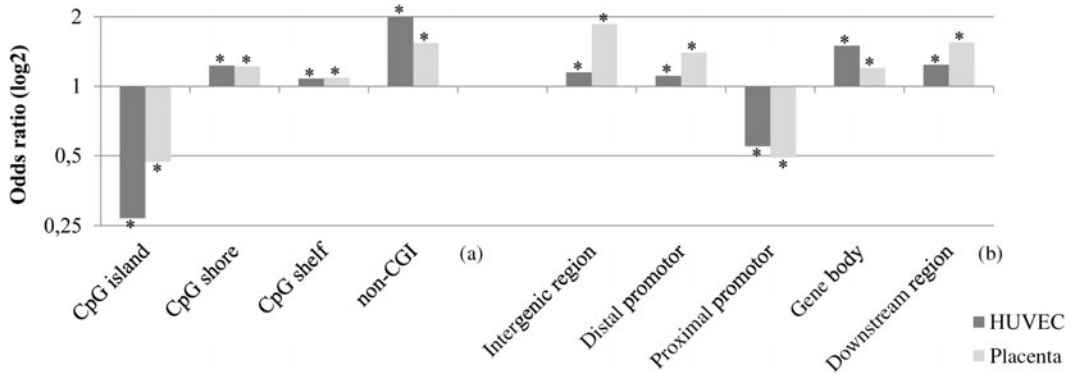

Fig. 4. HUVEC and placental tDMP enrichment in the (a) CpG density and (b) gene-centric annotation, in comparison to UC-WBC. (a) tDMP annotation relative to CpG islands showed that tDMPs were significantly enriched in CpG shores, shelves and especially in non-CpG island regions and strongly depleted in CpG islands. Annotation patterns in relation to CpG islands were concordant for HUVEC and placental tDMPs. (b) tDMP annotation relative to genes demonstrated that tDMPs were significantly enriched in all gene-centric regions except for a strong depletion of tDMPs in proximal promotors. The strongest placental enrichment was observed in intergenic regions. HUVEC's strongest enrichment was observed in gene bodies. *$p < 0.05$. Non-CGI, non-CpG island; HUVEC, human umbilical vein endothelial cells.

($OR^{Placenta}$ 0.21, 95% CI 0.21–0.22, OR^{HUVEC} 0.12, 95% CI 0.11–0.12). The combined annotation enrichment analysis also revealed opposite annotation patterns for HUVEC and placental tDMPs, mainly in various gene-centric regions situated in CpG islands and CpG shelves.

tDMR identification

To obtain a more robust measure of genomic regions with a large proportion of tissue-specific sites, we used the tDMPs to generate tDMRs with DMRfinder. This revealed 3595 HUVEC tDMRs with a mean length of 639 basepairs (bp) and 11,938 placental tDMRs with a mean length of 894 bp. The numbers of overlapping HUVEC and placental tDMRs are presented in Fig. 3b.

GO term enrichment of tDMR genes

To gain insight in tDMR functional categories, tDMRs were mapped to their nearest gene. This yielded 2882 unique (without duplicate genes) HUVEC genes and 7629 unique placental genes. The 2882 HUVEC tDMR genes mapped to 2296 genes in DAVID because the remaining annotations were mainly transcripts. The 2296 genes were enriched for involvement in embryogenesis, regulation of gene expression, cell motility and vascular development. The top 3000 placental tDMR genes were selected based on the regions with the largest absolute mean value of each tDMP M-value difference within the DMR, to meet the maximum number of genes for the Functional Annotation Clustering tool in DAVID. After excluding transcripts, DAVID mapped 2208 placental tDMR genes and revealed enriched GO term involvement in embryogenesis, regionalisation and regulation of gene expression. HUVEC and placental highest significantly enriched GO terms are presented in Table 2.

Qiagen IPA of tDMR genes

Ingenuity software mapped 2523 of the 2881 HUVEC tDMR genes and identified enriched canonical pathways. Top-ranked pathways were mainly involved in immune response processes. We further conducted network analyses, using Fisher's exact test, revealing that the top networks were mainly associated with cellular function, movement and signalling, immune response and embryonic development.

Out of 7629 placental tDMR genes, 6333 were mapped by Ingenuity, demonstrating top five enriched pathways involved in a broad spectrum of cell signalling processes and embryonic stem cell pluripotency. The top networks were mainly associated with cardiovascular disease, embryonic development and cellular movement and function. The highest significantly enriched HUVEC and placental pathways and networks are presented in Table 3.

Discussion

This epigenome-wide DNA methylation study shows 43% HUVEC- and 73% placental tissue-specific DNA methylation out of all measured CpGs, in comparison to UC-WBC. Enrichment of tDMPs was demonstrated in gene bodies and non-CpG islands. No tDMRs were observed in UC-MNC compared to UC-WBC.

The most pronounced identification of placental compared to HUVEC tDMRs might be explained by the time point of placental-specific development. After fertilisation and global demethylation

Table 2. Top 10 DAVID GO Functional annotation clusters of HUVEC and placental tDMR nearest genes

HUVEC	Annotation Cluster Enrichment score
1. Embryonic morphogenesis	8.4
2. Negative regulation of gene expression	8.4
3. Positive regulation of gene expression	8.3
4. Cell motility	7.6
5. Vascular development and morphogenesis	7.1
6. Embryonic development	7.1
7. Haematopoiesis	7.0
8. Inflammatory response	6.7
9. Bone development	5.7
10. Cell adhesion	5.1
Placenta	**Annotation cluster enrichment score**
1. Embryonic development	13.9
2. Embryonic morphogenesis	12.9
3. Regionalisation	11.7
4. Negative regulation of gene expression	10.7
5. Embryonic organ development	9.3
6. Regulation of transcription	8.9
7. Embryonic development and morphogenesis	8.5
8. Positive regulation of gene expression	8.3
9. Cellular morphogenesis	8.3
10. Cell motility	7.2

tDMRs, tissue-specific differentially methylated regions; DAVID, Database for Annotation, Visualization and Integrated Discovery; HUVEC, Human umbilical vein endothelial cell. HUVEC and placental tDMRs were mapped to the nearest gene using Ensembl annotations from UCSC database. Assigned Ensembl genes were uploaded to the DAVID tool to examine possible enrichment of corresponding GO terms using the GO_BP_FAT annotation category. Fisher's exact test was applied to measure the gene enrichment of annotated GO terms of the uploaded gene list, against the whole human genome list as a background. To focus on the biology of the annotated GO terms, clusters of similar annotations were examined from the DAVID Functional Annotation Clustering table tool. The clustering algorithm is based on the hypothesis that similar annotations have similar gene members, resulting in a Group Enrichment Score to rank their biological significance, which is based on the Fisher's exact *p*-values of each GO term within the cluster. The higher the enrichment score, the more the enrichment is.

of the paternally and maternally derived genome, the very first *de novo* DNA methylation at the late morula stage determines the initial developmental lineage differentiation of the inner cell mass that will give rise to the embryo and the trophectoderm that will evolve into placental tissue. The inner cell mass methylation is more pronounced than the relatively hypomethylated trophectoderm, resulting in an asymmetrical methylation status from this developmental stage onwards, which is consistent with the observed extent of genome-wide placental-specific methylation and in line with previous studies[2,3,33]. The relatively hypomethylated state of the trophectoderm is necessary for the highly proliferative and invasive process of first trimester placentation requiring an active transcription of genes[3]. Schroeder *et al.* described partially methylated regions (PMRs) in the placenta, characterised by methylation levels around 45%, in contrast to most human tissues demonstrating mainly a high methylation (>70%) of the majority of the genome[37]. This is in line with the relatively abundant placenta-specific intermediate methylation levels around 50% in our data and that of others, especially since the PMRs were annotated to genes with tissue-specific functions[12].

HUVEC displayed another substantial tissue-specific methylation pattern compared to UC-WBC, but less extreme than placental tissue. This may be explained by the fact that HUVEC and the surrounding umbilical cord originate from the inner cell mass[38]. UCB originates from endothelial cells in the ventral aorta of the developing embryo during initial haematopoiesis and thus displays a more common background with HUVEC, substantiating the less diverging methylation profiles of UC-WBC and HUVEC[39]. These observations in HUVEC suggest that carefully selected tDMR-poor genomic regions may be represented adequately by UC-WBC in terms of DNA methylation in candidate gene studies, but this needs further functional validation.

In comparison to peripheral blood, Slieker *et al.* identified 3500 DMRs in peripheral tissues and 5400 DMRs in post-mortem internal tissues, mainly situated in CpG-poor regions, which is in agreement with our data considering HUVEC tDMR quantity and genomic occurrence[9]. Also in line with our observations are the recently published data of Lowe *et al.* showing highly distinctive epigenome-wide methylation levels in whole blood DNA versus several somatic tissues using publicly available Illumina 450K databases[8,40]. Although other studies have identified tDMRs to examine tissue specificity, these studies are limited in terms of comparability, due to the use of post-mortem tissues or cultured cells and a wide variety of statistical approaches[10,41,42]. There exists partial overlap between tDMRs identified in our study and in previous EWAS. These potentially represent highly tissue-specific sites throughout several different tissues and suggest a certain degree of reproducibility of our results (Supplementary Table S2).

We expected to find substantial differences in DNA methylation of the two blood cell fractions since DNA methylation is important in haematopoiesis and blood cell differentiation, and MNC represent a strongly different subgroup as they lack the largest proportion of total WBC: granulocytes[43]. A genome-wide profiling study using Illumina 27K Methylation arrays compared haematopoietic pluripotent cells (HPC) with granulocytes and monocytes from the same UCB and demonstrated that further differentiation of HPC is associated with demethylation of certain epigenetic programmes, depending on the cell type[43]. In our data, however, only 0.4% of all epigenome-wide measured CpG sites were statistically differentially methylated, resulting in only two tDMPs and no tDMRs. This is partly explained by the decreased power as a result of the 2.5 times lower UC-MNC sample size, but mainly by the smaller effect sizes we find when studying the UC-MNC group and the biological origin being much more similar to the UC-WBC than the other tissues. Moreover, instead of the HUVEC and placental samples that were genetically almost identical, the MNC samples were in 90% non-genetically matched to the UC-MNC, resulting in genetically driven differences that overwhelm cell-specific differences when closely related cells are compared. This, combined with the fact that not all probes associated with single nucleotide polymorphisms were removed, and the contribution of methylation quantitative trait loci – single nucleotide polymorphisms, may explain the UC-MNC result. It seems therefore equally adequate to perform DNA methylation studies in DNA derived from UC-WBC as from UC-MNC, if the use of homogeneous blood cells is not feasible. Nevertheless, adjustment for WBC mixtures should always be applied in whole blood EWAS to avoid confounding of underlying WBC mixtures[20].

Table 3. Top 10 Ingenuity networks and pathways of HUVEC and placental tDMR nearest genes

HUVEC canonical pathways	–log *p*-value
1. Dendritic cell maturation	4.48
2. Antigen presentation pathway	4.48
3. T helper cell differentiation	4.48
4. Role of NFAT in regulation of the immune response	4.48
5. PKCθ signalling in T lymphocytes	4.48
6. iCOS-iCOSL signalling in T helper cells	4.48
7. OX40 signalling pathway	4.05
8. Hepatic fibrosis/hepatic stellate cell activation	4.04
9. Altered T cell and B cell signalling in rheumatoid arthritis	3.23
10. Allograft rejection signalling	3.11
HUVEC networks	**Score**
1. Behaviour, reproductive system development and function, cellular function and maintenance	31/35
2. Cell-to-cell signalling and interaction, inflammatory response, lipid metabolism	31/35
3. Cellular function and maintenance, inflammatory response, digestive system development and function	29/34
4. Cellular movement, immune cell trafficking, cell morphology	27/33
5. Skeletal and muscular system development and function, embryonic development, organismal development	25/32
6. Infectious diseases, DNA replication, recombination, and repair, gene expression	25/32
7. Glomerular injury, organismal injury and abnormalities, renal fibrosis	25/32
8. Cell-to-cell signalling and interaction, haematological system development and function, immune cell trafficking	23/31
9. Cell signalling, molecular transport, vitamin and mineral metabolism	23/31
10. Connective tissue development and function, skeletal and muscular system development and function, tissue development	23/31
Placenta canonical pathways	**–log *p*-value**
1. G protein-coupled receptor signalling	5.18
2. Glutamate receptor signalling	5.14
3. Neuropathic pain signalling in dorsal horn neurons	5.02
4. Human embryonic stem cell pluripotency	5.02
5. cAMP-mediated signalling	5.01
6. Regulation of the epithelial–mesenchymal transition pathway	5.01
7. Axonal guidance signalling	4.78
8. Transcriptional regulatory network in embryonic stem cells	4.43
9. Hepatic fibrosis/hepatic stellate cell activation	4.26
10. Antigen presentation pathway	3.38
Placenta networks	**Score**
1. Cardiovascular disease, organismal injury and abnormalities, reproductive system disease	21/35
2. Cellular movement, cell morphology, cellular assembly and organisation	21/35
3. Cardiovascular disease, cellular movement, developmental disorder	21/35
4. Cell-to-cell signalling and interaction, cellular function and maintenance, embryonic development	21/35
5. Cell death and survival, cellular compromise, neurological disease	21/35
6. Cardiovascular system development and function, small molecule biochemistry, connective tissue development and function	19/34
7. Cell death and survival, infectious diseases, inflammatory disease	19/34
8. Haematological system development and function, organismal functions, cell-to-cell signalling and interaction	19/34
9. Cellular movement, cell death and survival, cardiovascular system development and function	19/34
10. Connective tissue disorders, cancer, gastrointestinal disease	19/34

tDMRs, tissue-specific differentially methylated regions; HUVEC, Human umbilical vein endothelial cells.
Ingenuity Pathway Analysis was performed with the annotated DMP Ensembl gene lists with focus on canonical pathways and networks. Enriched canonical pathways were calculated using the Fisher's exact test and subjected to Benjamini–Hochberg procedure for controlling FDR ($p < 0.05$). Networks were generated based on network eligible molecules, which were encoded by our DMP genes and also interact with other molecules in the Ingenuity Pathways Knowledge Base. A high score for a network indicates a more approximate fit between network eligible molecules and the molecules that constitute the network, calculated using the right-tailed Fisher's exact test.

To this purpose, a cord blood-specific algorithm was recently developed to estimate cell proportions[44].

Our observations concerning UC-WBC seem in agreement with the study of Wu *et al.*, demonstrating significantly correlated methylation levels within the same individuals for MNC and WBC in adult peripheral blood, although these were measured by repetitive element methylation levels in different assays[45].

Enrichment analysis of the identified HUVEC and placental tDMPs showed that tDMPs in general occur quite randomly relative to CpG density and genes. However, a strong depletion of tDMPs in CpG islands and proximal promotors and a strong enrichment in non-CpG islands was evident. Additionally, we demonstrated HUVEC tDMP enrichment in gene body regions, in agreement with Lokk *et al.*[10]. Gene bodies and non-CpG island regions seem to be more tissue-specific and susceptible to variation in DNA methylation and may therefore represent interesting genomic regions for future epidemiologic epigenetic association studies. This also suggests that we should focus less on the concept of DNA methylation resulting in complete silencing of transcription which is related to DNA methylation of promotor regions. Methylation of gene bodies does in fact not block transcription and is associated with transcription stimulation[46]. Evidence emerges that CpG island methylation of promoter regions, the main region of interest in previous work, is mostly associated with long-term repression of gene expression (Imprinting, X-chromosome inactivation i.a.), which is in line with the shown depletion of tDMPs in CpG island proximal promotors[9,41,46].

GO functional annotations of the tDMRs demonstrated that tissue-specific DNA methylation is involved in specific tissue-related biological functions such as vascular development in HUVEC, suggesting that DNA derived from HUVEC might be more informative than from UC-WBC if one would be interested in epigenetic involvement during prenatal vascular development. Moreover, general biological functions and early developmental processes were also observed in association with tDMR genes, such as the regulation of gene expression, cellular function and signalling and embryonic development. Both categories underline that the tissue-specific aspect should be considered in future epigenetic association studies.

The Ingenuity Knowledge database validated most GO functional annotations in revealing a broad spectrum of housekeeping and embryonic development pathways. In addition, HUVEC tDMR genes were annotated to immune response processes which fit to the known immune cell adhesion and migration function of endothelial cells[17]. This, however, also demonstrates that a certain reference tissue might reveal corresponding tDMR genes, and one could argue that the annotated immune response genes might in fact be more attributable to our reference tissue UC-WBC than to HUVEC-specific methylation. It should also be noted that pathway enrichment analyses may be susceptible to bias due to the design of the array, but we believe this is unlikely given the strength of the enrichments consistent between DAVID and IPA[47].

These data provide novel insights into the tissue specificity of epigenome-wide DNA methylation in newborn and placental tissues without confounding of postnatal exposures. The study was performed in human tissues with a much larger sample size than used in previous studies. A major strength of our study is the fact that our samples are largely matched and therefore not confounded by the potential influence of genetic variation on DNA methylation. Moreover, we standardised the tissue sampling and applied an optimised tDMR identification technique. We therefore believe that our findings support and contribute to the existing evidence demonstrating the importance of tissue specificity in genome-wide DNA methylation measurements.

The following concerns, however, need to be addressed. The absence of expression data to validate our findings and the lack of information on external validity are limitations of our study. Furthermore, full replication of our analysis was not possible because a comparable independent dataset with similar patients, tissues and an almost full within-subject statistical study design is not available. A strong correlation was however observed between the estimates of the tissue effect in our own dataset and Δβ methylation differences of HUVEC and placenta versus UC-WBC of three independent datasets, suggesting partial replication of our tDMPs (Supplementary Fig. S7). Moreover, our findings could have been more relevant with the use of homogeneous cell types of UCB and placental tissue, thereby avoiding any confounding of methylation differences due to underlying cell mixture differences[48,49]. However, upon examination of the potential influence of underlying cell variation in UC-WBC and placental tissue on our model, we did see a high correlation with and without adjusting for cell heterogeneity, suggesting that the effect was marginal. A minor limitation is that we did not exclude the published list of potential cross-reactive probes that are present on the array, prior to further analysis[50]. However, given the number of cross-reactive probes in comparison to the number of identified loci, the effect will be very limited. Although we initially intended to obtain all tissues from each patient, due to logistic reasons, it appeared not feasible to perform a full within-subject analysis. The collection of sufficient blood in cord blood collection bags for UC-MNC isolation appeared extremely challenging. This seemed a result of random factors, such as stressful deliveries, caesarean sections or night shifts, but did not seem systematic. The majority of study subjects were eventually overlapping in the comparisons between UC-WBC versus placenta (24 out of 25 subjects in each group) and UC-WBC versus HUVEC (23 out of 25 subjects in each group), but the UC-MNC group only shares one patient with the UC-WBC. Fortunately, this resulted in maternal and fetal characteristics that were almost isogenic and thus highly comparable between all different tissue groups (Table 1). However, this does not exclude some residual confounding based on genetic variation, especially in the UC-MNC analysis.

Further research with repeated measurements is needed to gain insights in the maintenance and stability of tissue-specific methylation differences throughout pregnancy and postnatal life, which may increase or decrease due to accumulating environmental exposures, ageing in general and stochastically[51]. Also, the degree of correlation between tissue-specific epigenetic programming and the possibility to observe effects of environmental exposures on epigenetic features remains to be elucidated.

Conclusions

We demonstrated an extensive amount of genome-wide HUVEC and placental tDMRs in comparison to UC-WBC. This underlines the relevance of tissue-specific approaches in future epigenetic association studies or the use of a profound and validated representative tissue for a certain disease of interest, if available. Especially, the potentially limited predictive value of extra-embryonic placental tissue on epigenetic programming in fetal tissues should be considered.

We herewith provide a relevant dataset of paired, tissue-specific, genome-wide methylation measurements in newborn

tissues. These, especially our novel HUVEC data, can serve as additional reference methylomes for other investigators. Further validation in functional studies is needed to establish the usability of this data in future epigenetic epidemiological studies.

Supplementary material. For supplementary material for this article, please visit https://doi.org/10.1017/S2040174420000136

Availability of data and material. The datasets generated during the current study are available in the GEO repository with accession number GSE103253. Any additional patient characteristics will be made available for others, upon reasonable request, addressing the corresponding author.

Acknowledgements. The authors thank all medical students and staff of the Obstetrics department, in particular the Predict team, for their help with the tissue sampling, Mark F. Wildhagen for the digitalisation of the study data, Mr. M.M.P.J. Verbiest and Mrs. P. Jhamai for their support in the laboratory processing of the methylation analysis and Ms. L. Duinhouwer for her support during the haematologic laboratory work of the UC-MNC. Moreover, we thank Prof. Dr. med. Dr. h.c. J. Sehouli of the department of Gynaecology and Gynaecologic Oncology, Charité University Medicine Campus Virchow, Berlin, Germany, for his generous support during the realisation of this manuscript.

Author contributions. EH contributed to the study design, data collection, laboratory work, statistical analysis and writing of the first draft and all revisions of the manuscript. AE contributed to the study design and writing of the manuscript. SW supervised and performed the statistical analysis. RS was involved in the statistical analysis and writing of the manuscript. JF was involved in the study design, statistical analysis and writing of the manuscript. AS and PS were involved in the bioinformatic analysis and writing of the manuscript. JM was involved in writing of the manuscript and supervised the laboratory work. BH was involved in the statistical analysis and writing of the manuscript. RS initiated the study and supervised all aspects of the study and contributed to all versions of the manuscript. All authors contributed to the writing and revision of the manuscript and approved the final version.

Financial support. This work was funded by the Department of Obstetrics and Gynaecology, Erasmus MC, University Medical Centre Rotterdam, The Netherlands. J.F.F. has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 633595 (DynaHEALTH). Bioinformatic data analyses were partly funded by the Centre for Translational Molecular Medicine (CTMM) Trait and the Department of Bioinformatics.

Conflicts of interest. None.

Ethical standards. Ethical approval for the study was given by the Erasmus MC, University Medical Centre Research Ethics Board (MEC-2004-227). The research has been carried out in accordance with the Declaration of Helsinki (2013) of the World Medical Association. All women gave written informed consent before participation and parental informed consent was obtained for the child.

References

1. Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of *in utero* and early-life conditions on adult health and disease. *N Engl J Med.* 2008; 359(1), 61–73.
2. Senner CE. The role of DNA methylation in mammalian development. *Reprod Biomed Online.* 2011; 22(6), 529–535.
3. Koukoura O, Sifakis S, Spandidos DA. DNA methylation in the human placenta and fetal growth (review). *Mol Med Rep.* 2012; 5(4), 883–889.
4. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nature Rev Genet.* 2013; 14(8), 585–594.
5. Gama-Sosa MA, Midgett RM, Slagel VA, *et al.* Tissue-specific differences in DNA methylation in various mammals. *Biochimica et Biophysica Acta.* 1983; 740(2), 212–219.
6. Herzog E, Galvez J, Roks A, *et al.* Tissue-specific DNA methylation profiles in newborns. *Clin Epigenetics.* 2013; 5(1), 8.
7. Ollikainen M, Smith KR, Joo EJ, *et al.* DNA methylation analysis of multiple tissues from newborn twins reveals both genetic and intrauterine components to variation in the human neonatal epigenome. *Hum Mol Genet.* 2010;19(21), 4176–4188.
8. Lowe R, Slodkowicz G, Goldman N, Rakyan VK. The human blood DNA methylome displays a highly distinctive profile compared with other somatic tissues. *Epigenetics.* 2015; 10(4), 274–281.
9. Slieker RC, Bos SD, Goeman JJ, *et al.* Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450K array. *Epigenetics Chromatin.* 2013; 6(1), 26.
10. Lokk K, Modhukur V, Rajashekar B, *et al.* DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biology.* 2014; 15(4), r54.
11. Ma B, Allard C, Bouchard L, *et al.* Locus-specific DNA methylation prediction in cord blood and placenta. *Epigenetics.* 2019; 14(4), 405–420.
12. De Carli MM, Baccarelli AA, Trevisi L, *et al.* Epigenome-wide cross-tissue predictive modeling and comparison of cord blood and placental methylation in a birth cohort. *Epigenomics.* 2017; 9(3), 231–240.
13. de Goede OM, Lavoie PM, Robinson WP. Characterizing the hypomethylated DNA methylation profile of nucleated red blood cells from cord blood. *Epigenomics.* 2016; 8(11), 1481–1494.
14. Roadmap Epigenomics C, Kundaje A, Meuleman W, *et al.* Integrative analysis of 111 reference human epigenomes. *Nature.* 2015; 518(7539), 317–330.
15. Gomes MV, Pelosi GG. Epigenetic vulnerability and the environmental influence on health. *Exp Biol Med (Maywood).* 2013; 238(8), 859–865.
16. Sandovici I, Hoelle K, Angiolini E, Constancia M. Placental adaptations to the maternal-fetal environment: implications for fetal growth and developmental programming. *Reprod Biomed Online.* 2012; 25(1), 68–89. doi: 10.1016/j.rbmo.2012.03.017.
17. Casanello P, Schneider D, Herrera EA, Uauy R, Krause BJ. Endothelial heterogeneity in the umbilico-placental unit: DNA methylation as an innuendo of epigenetic diversity. *Front Pharmacol.* 2014; 5, 49.
18. Jin SW, Patterson C. The opening act: vasculogenesis and the origins of circulation. *Arteriosclerosis, Thrombosis, and Vascular Biology.* 2009; 29(5), 623–629.
19. Krause B, Sobrevia L, Casanello P. Epigenetics: new concepts of old phenomena in vascular physiology. *Current Vascular Pharmacology.* 2009; 7(4), 513–520.
20. Koestler DC, Christensen B, Karagas MR, *et al.* Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. *Epigenetics.* 2013; 8(8), 816–826.
21. Steegers-Theunissen RP, Verheijden-Paulissen JJ, van Uitert EM, *et al.* Cohort profile: the Rotterdam periconceptional cohort (predict study). *Int J Epidemiol.* 2016; 45(2), 374–381.
22. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. *Epigenomics.* 2011; 3(6), 771–784.
23. Sandoval J, Heyn H, Moran S, *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics.* 2011; 6(6), 692–702.
24. Bibikova M, Barnes B, Tsan C, *et al.* High density DNA methylation array with single CpG site resolution. *Genomics.* 2011; 98(4), 288–295.
25. Huber W, Carey VJ, Gentleman R, *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods.* 2015; 12(2), 115–121.
26. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics.* 2013; 14, 293.
27. Du P, Zhang X, Huang CC, *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics.* 2010; 11, 587.
28. Novakovic B, Yuen RK, Gordon L, *et al.* Evidence for widespread changes in promoter methylation profile in human placenta in response to increasing gestational age and environmental/stochastic factors. *BMC Genomics.* 2011; 12, 529.

29. Houseman EA, Kile ML, Christiani DC, Ince TA, Kelsey KT, Marsit CJ. Reference-free deconvolution of DNA methylation data and mediation by cell composition effects. *BMC Bioinformatics*. 2016; 17, 259.
30. Marabita F, Almgren M, Lindholm ME, *et al.* An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. *Epigenetics.* 2013; 8(3), 333–346.
31. Franzen J, Zirkel A, Blake J, *et al.* Senescence-associated DNA methylation is stochastically acquired in subpopulations of mesenchymal stem cells. *Aging Cell.* 2017; 16(1), 183–191.
32. Paquette AG, Houseman EA, Green BB, *et al.* Regions of variable DNA methylation in human placenta associated with newborn neurobehavior. *Epigenetics.* 2016; 11(8), 603–613.
33. Slieker RC, Roost MS, van Iperen L, *et al.* DNA Methylation Landscapes of Human Fetal Development. *PLoS genetics.* 2015; 11(10), e1005583.
34. Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics. 2014; 15, 293.
35. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009; 37(1), 1–13.
36. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols.* 2009; 4(1), 44–57.
37. Schroeder DI, Blair JD, Lott P, *et al.* The human placenta methylome. *Proc Natl Acad Sci U S A.* 2013; 110(15), 6037–6042.
38. Bruce M. Carlson BC. Human Embryology and Developmental Biology, 2009. MOSBY.
39. Adamo L, Garcia-Cardena G. The vascular origin of hematopoietic cells. *Dev Biol.* 2012; 362(1), 1–10.
40. Lowe R, Gemma C, Beyan H, *et al.* Buccals are likely to be a more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics.* 2013; 8(4), 445–454.
41. Byun HM, Siegmund KD, Pan F, *et al.* Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Hum Mol Genet.* 2009; 18(24), 4808–4817.
42. Rakyan VK, Down TA, Thorne NP, *et al.* An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* 2008; 18(9), 1518–1529.
43. Bocker MT, Hellwig I, Breiling A, Eckstein V, Ho AD, Lyko F. Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *Blood.* 2011; 117(19), e182–e189.
44. Bakulski KM, Feinberg JI, Andrews SV, *et al.* DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics.* 2016; 11(5), 354–362.
45. Wu HC, Delgado-Cruzata L, Flom JD, *et al.* Global methylation profiles in DNA from different blood cell types. *Epigenetics.* 2011; 6(1), 76–85.
46. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Rev Genet.* 2012; 13(7), 484–492.
47. Harper KN, Peters BA, Gamble MV. Batch effects and pathway analysis: two potential perils in cancer studies involving DNA methylation array analysis. *Cancer Epidemiol Biomarkers Prev.: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology.* 2013; 22(6), 1052–1060.
48. Grigoriu A, Ferreira JC, Choufani S, Baczyk D, Kingdom J, Weksberg R. Cell specific patterns of methylation in the human placenta. *Epigenetics.* 2011; 6(3), 368–379.
49. Joo JE, Hiden U, Lassance L, *et al.* Variable promoter methylation contributes to differential expression of key genes in human placenta-derived venous and arterial endothelial cells. *BMC Genomics.* 2013; 14, 475.
50. Chen YA, Lemire M, Choufani S, *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013; 8(2), 203–209.
51. Cencioni C, Spallotta F, Martelli F, *et al.* Oxidative stress and epigenetic regulation in ageing and age-related diseases. *International Journal of Molecular Sciences.* 2013; 14(9), 17643–17663.

Appendix 1

Data collection

Immediately after delivery of the newborn, UCB samples from the umbilical vein were collected in anticoagulant vacutainer tubes (ethylenediaminetetraacetic acid) and cord blood collection bags containing 21-ml anticoagulant citrate-phosphate-dextrose solution with the placenta still *in situ*. The complete umbilical cord was cut at the placental insertion and immediately stored in umbilical cord buffer (HBSS with 1% penicillin/streptomycin) at 4–8 °C until further HUVEC isolation within 24 h after delivery. Umbilical cord and placental samples were collected within 10 min after delivery of the placenta. Placental samples of 0.5 cm^3 were taken from the fetal side of the villi at four different sites in a 3-cm radius around the umbilical cord insertion, after carefully removing the membranes and 2 mm of the top placental layer. After washing in phosphate-buffered saline (PBS) solution to remove maternal blood, placental samples were snap-frozen in liquid nitrogen and stored at −80 °C until DNA extraction. All samples were collected by trained researchers.

UCB processing

UCB vacutainer tubes and cord blood collection bags were stored at 4–8 °C and processed within 48 h after delivery. Total UC-WBC were isolated after centrifugation of the vacutainer tubes. UC-MNC represent a subgroup of UC-WBC and were obtained from the cord blood collection bags using Ficoll gradient centrifugation to remove granulocytes from the total WBC fraction. The UC-WBC and MNC pellets were stored at −80 °C until DNA extraction.

HUVEC isolation

The umbilical cord vein was connected to infusion tubes on both extremes and rinsed with cord buffer. Once all remaining UCB was removed, the umbilical vein was filled with collagenase solution (1 mg/ml) and incubated for 15 min in a PBS water bath at 37 °C. Detached HUVEC were obtained in suspension and further purified by magnetic-activated cell separation (MACS) with CD146 MACS MicroBeads (Miltenyi Biotec GmbH, Bergisch Gladbach, Germany). PBS-washed HUVEC pellets were snap-frozen in liquid nitrogen and stored at −80 °C until DNA extraction. We assessed the cell purity of the first obtained HUVEC by counting them microscopically. It was observed that a cell purity of >95% was achieved, which we considered as sufficient for further analyses.

DNA extraction

Thawed UC-WBC and UC-MNC pellets were subjected to erythrocyte lysis by use of an Erythrocyte Lysis Buffer (Qiagen, Hilden, Germany), following manufacturer's protocols. Thirty milligrams of frozen placental tissue were grinded manually on dry ice using a tissue grinder. The placental powder was immediately added to a cell lysis buffer and stored at −80 °C until further processing. Subsequently, genomic DNA was extracted from all tissues using the Allprep DNA/RNA isolation mini kit (Qiagen, Hilden, Germany), according to manufacturer's instructions.