DNA Methylation Trajectories During Pregnancy

Epigenetics Insights
Volume 12: 1–9

DOI: 10.1177/2516865719867090
SAGE

Olena Gruzieva[1,2], Simon Kebede Merid[1], Su Chen[3], Nandini Mukherjee[3], Anna M Hedman[4], Catarina Almqvist[4,5], Ellika Andolf[6], Yu Jiang[3], Juha Kere[7,8,9], Annika Scheynius[10,11,12], Cilla Söderhäll[7,13], Vilhelmina Ullemar[4], Wilfried Karmaus[3], Erik Melén[1,11], Syed Hasan Arshad[14,15] and Göran Pershagen[1,2]

[1]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [2]Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden. [3]Division of Epidemiology, Biostatistics, and Environmental Health Science, School of Public Health, The University of Memphis, Memphis, TN, USA. [4]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [5]Astrid Lindgren Children's Hospital, Karolinska University Hospital, Stockholm, Sweden. [6]Department of Clinical Sciences, Division of Obstetrics and Gynaecology, Danderyd Hospital, Karolinska Institutet, Stockholm, Sweden. [7]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden. [8]Molecular Neurology Research Program, University of Helsinki and Folkhälsan Institute of Genetics, Helsinki, Finland.[9]School of Basic & Medical Biosciences, King's College London, London, UK. [10]Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden.[11]Sachs' Children and Youth Hospital, Södersjukhuset, Stockholm, Sweden. [12]Clinical Genomics, Science for Life Laboratory, Stockholm, Sweden. [13]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden. [14]The David Hide Asthma and Allergy Research Centre, Newport, UK. [15]Clinical and Experimental Sciences, University of Southampton, Southampton, UK.

ABSTRACT: There is emerging evidence on DNA methylation (DNAm) variability over time; however, little is known about dynamics of DNAm patterns during pregnancy. We performed an epigenome-wide longitudinal DNAm study of a well-characterized sample of young women from the Swedish Born into Life study, with repeated blood sampling before, during and after pregnancy (n = 21), using the Illumina Infinium MethylationEPIC array. We conducted a replication in the Isle of Wight third-generation birth cohort (n = 27), using the Infinium HumanMethylation450k BeadChip. We identified 196 CpG sites displaying intra-individual longitudinal change in DNAm with a false discovery rate (FDR) $P < .05$. Most of these (91%) showed a decrease in average methylation levels over the studied period. We observed several genes represented by ⩾3 differentially methylated CpGs: *HOXB3, AVP, LOC100996291*, and MicroRNA 10a. Of 36 CpGs available in the replication cohort, 17 were replicated, all but 2 with the same direction of association (replication $P < .05$). Biological pathway analysis demonstrated that FDR-significant CpGs belong to genes overrepresented in metabolism-related pathways, such as adipose tissue development, regulation of insulin receptor signaling, and mammary gland fat development. These results contribute to a better understanding of the biological mechanisms underlying important physiological alterations and adaptations for pregnancy and lactation.

KEYWORDS: Cohort, DNA methylation, Illumina EPIC and Infinium chip, pregnancy

RECEIVED: June 28, 2019. **ACCEPTED:** July 10, 2019.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Swedish Research Council, the Swedish Heart Lung Foundation, the Swedish Asthma and Allergy Association's Research Foundation, the Stockholm County Council (ALF projects), FORTE, FORMAS, the Strategic Research Program in Epidemiology at Karolinska Institutet, and the National Institute of Allergy and Infectious Diseases in the United States (R01AI091905).

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Olena Gruzieva, Institute of Environmental Medicine, Karolinska Institutet, Nobels väg 13, SE-171 77 Stockholm, Sweden.
Email: olena.gruzieva@ki.se

Introduction

DNA methylation (DNAm) plays an important role in the maintenance of normal cellular function, gene expression regulation, and embryonic development.[1] There is emerging evidence that an individual's DNAm profile is subject to temporal changes on both a short- (hours to days) and long-term (months to years) timescale.[2] A few studies have examined the temporal stability of DNAm in serial peripheral blood samples from the same individuals,[3-6] reporting various estimates of proportions of stable methylation sites over time. Other studies investigated associations of methylation with age, indicating high correlations and prediction.[7,8] Most of the existing studies are limited to a few candidate sites,[9] to particular age groups (eg, childhood populations),[10-13] or to ages after the reproductive period.[14] However, little is known about dynamics of DNAm status during pregnancy. Pregnancy is characterized by substantial physiological changes, including alterations in glucose and fat metabolism as well as immune system to meet the increased metabolic demands of the mother and fetus and prevent allogenic rejection.[15,16] A better understanding of changes during pregnancy in epigenetic marks may help to explain the biological mechanisms underlying important physiological alterations and adaptations needed to allow development of the fetus as well as to prepare the mother for childbirth and the postnatal period. Furthermore, causal mechanisms behind adverse pregnancy outcomes such as preeclampsia or coexisting

conditions such as asthma may be elucidated. To the best of our knowledge, only 1 study has investigated genome-wide methylation patterns in blood samples during pregnancy, showing a very limited number of both consistent and significantly changed Cytosine-phosphate-Guanine (CpGs) during the transition from early to late pregnancy.[6] However, no detailed assessment was performed of time trends in methylation of specific CpG sites of functional importance. To address this knowledge gap, we performed an epigenome-wide longitudinal DNAm study (EWAS) of a well-characterized sample of young women with repeated blood sampling before, during and after pregnancy, using the Illumina EPIC array. Furthermore, we conducted a replication analysis in an independent cohort using the Infinium HumanMethylation450k BeadChip.

Methods

Discovery study design

This study is based on the prospective birth cohort study Born into Life, following a cohort of women before, during, and after pregnancy, as well as their children in the peri- and postnatal period.[17] Born into Life used the larger LifeGene study as a sampling frame, which is described in detail elsewhere.[18] In brief, LifeGene is a prospective cohort study including persons aged 18 to 45 years, from certain areas in Sweden, as well as their household members (partner and any children). A comprehensive web-based questionnaire comprised multifaceted questions concerning health conditions and various risk factors was administered at baseline, briefly followed up annually, and readministered every 5 years thereafter.

Women in the LifeGene cohort who had responded to the baseline questionnaire and later became pregnant were invited to the Born into Life study. They were recruited both before and after gestational weeks 10 to 14, but no later than weeks 26 to 28. Pregnant women answered questionnaires regarding pregnancy, lifestyle, and health at gestational weeks 10 to 14 and 26 to 28. They also provided biological material including peripheral blood at gestational weeks 10 to 14, 26 to 28, as well as in conjunction with the newborn screening test of their child 2 to 4 days after delivery (Table 1). Ethical approval was obtained by the Regional Ethics Review Board in Stockholm, Sweden, and participation was based on written informed consent.

Replication study design

The Isle of Wight (IoW) birth cohort was established in the United Kingdom to prospectively study the natural history of asthma and allergic conditions. Details on study design, enrollment and follow-up procedures are described in detail elsewhere.[19] For analyses of DNAm, peripheral blood samples were drawn at the age of 18 years, as well as during the first (8-21 weeks) and the second (22-38 weeks) halves of pregnancy. Questionnaires were completed at age 18 years, as well as during pregnancy. Information was gathered about lifestyle factors and health, medication, asthma, eczema, and atopy status.[6] Female participants of the birth cohort have been followed through their pregnancy occurring between years 2011 and 2015. Ethics approval was given by the Isle of Wight Local Research Ethics Committee. Permission was granted for all follow-ups as well as collection of samples for genetic studies. Written informed consent was obtained from all participants before they participated in the study and at all follow-ups. For the 18-year follow-up, an approval was received from the Isle of Wight, Portsmouth, and SE Hampshire Local Research Ethics Committee. The investigations of pregnant cohort members were approved by the Isle of Wight, Portsmouth, and SE Hampshire Local Research Ethics Committee. At the University of Memphis, the internal review board approved the project.

Table 1. Overview of DNA methylation data availability from maternal blood samples in the Born into Life study.

SAMPLING TIME	FEMALE PARTICIPANTS (N)
Baseline[a] (preconception)	34
Pregnancy weeks 10 to 14	24
Pregnancy weeks 26 to 28	34
After pregnancy[b]	27
Data available from all 4 sampling occasions	21

[a]On average, 1.6 years (range: 0.86-3.35 years) before conception.
[b]In conjunction with the newborn screening test of the child 2 to 4 days after delivery.

DNAm profiling

In the Born into Life study, epigenome-wide DNAm was measured in 119 samples obtained from 34 women, using DNA extracted at the Karolinska Institutet Biobank from peripheral whole blood. Samples were drawn at 4 occasions: once before, twice during, and once after pregnancy. Of those, 21 women provided blood samples at all 4 occasions (Table 1).

An aliquot (500 ng) of DNA per sample underwent bisulfite conversion using the EZ-96 DNA Methylation Kit (Zymo Research Corporation, Irvine, CA, USA). Samples were plated onto 96-well plates in randomized order and processed with the MethylationEPIC BeadChip, using the manufacturer's standard protocol (Illumina, Inc., San Diego, CA, USA) at the Mutation Analysis Facility, Karolinska Institutet (www.maf.ki.se). This chip measures 866 836 CpG sites across the genome.

Quality control of analyzed samples was performed using standardized criteria. GenomeStudio Software was used to process the raw methylation intensities, and the detection

P value for each CpG was used as a quality control measure of probe performance. The Bioconductor Illumina Minfi package and the ComBat package[20] were used to remove background noise, adjust for inter-array variation, perform peak correction, and remove batch effects. DNA methylation levels for each CpG were estimated as the proportion of intensity of methylated (M) over the sum of methylated (M) and unmethylated (U) probes, $\beta = M/[c + M + U]$, with c being a constant to prevent dividing by zero. The CpG sites that had detection *P* values >.01 in 3% of all samples, and CpG sites with probe-SNPs (single-nucleotide polymorphisms), were excluded from all analyses. We found 23 016 probes with detection *P* values >.01 in 3% of all samples.

Samples were excluded in case of sample call rate <3. Color, low staining efficiency, poor extension efficiency, poor hybridization performance, low stripping efficiency after extension, and poor bisulfite conversion were evaluated using R package ENmix.[21] Furthermore, we applied median intensity plots for methylated and unmethylated intensity using the Minfi R package. All the above-mentioned criteria led to exclusion of 1 sample.

Probes with a SNP in the single base extension site with a frequency of >5% were excluded[22] (n = 11 726), as were probes with nonoptimal binding (nonmapping or mapping multiple times to either the normal or the bisulfite-converted genome) (n = 42 558), and the probes belonging to chrX (n = 18 578), resulting in addition the exclusion of 72 862 probes. Furthermore, we implemented "DASEN" recommended from wateRmelon package for signal correction and normalization.[23] The batch effect was corrected using ComBat.[20]

In the IoW study, DNA was extracted from whole blood from a subsample, collected at the age of 18 years (n = 249), first half of pregnancy (weeks 8-21, n = 131), and second half of pregnancy (after week 22, n = 130) using a standard salting-out procedure. A total of 27 women had DNAm data measured at all 3 occasions, ie, prepregnancy, first, and second half of pregnancy. DNA concentration was determined by Qubit quantitation. One microgram of DNA was bisulfite-treated for cytosine to thymine conversion using the EZ-96 DNA Methylation Kit, following the manufacturer's standard protocol. Epigenome-scale DNAm was assessed using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, Inc.), which interrogates 485 512 CpG sites.

Data quality control was undertaken and quantile-normalized beta values were processed following the CPACOR pipeline.[24] The samples from different array platforms and batches were combined and only the shared probes between EPIC and 450k arrays were chosen. ComBat[20] was used to remove batch effects in the combined data set. The CpG sites with probe-SNPs within 10 base pairs (bp) and with minor allele frequency greater than 0.007 (which represented about 10 subjects in expectation in the complete study cohort) were excluded.

Statistical analyses

To estimate changes in DNAm, we applied generalized linear mixed models with repeated measurement taking into account the within-subject time effect using lmerTest package in R.[25] The analyses were adjusted for maternal age at delivery (years) and parity (0/≥1). Replication analysis in the IoW cohort was additionally adjusted for maternal smoking during pregnancy (there were no smokers in Born into Life cohort). To account for potential differences in DNAm that may arise from variability of cell composition in whole blood,[26] we estimated 7 cell types (CD4+ T cells, CD8+ T cells, natural killer cells, B cells, monocytes, eosinophils, and granulocytes) using the R package minfi according to the algorithm developed by Houseman et al.[27] We adjusted for cell composition by including these estimated cell-type fractions as covariates in the linear regression. DNAm sites were annotated based on data provided by Illumina. We applied Bonferroni correction ($P < .05$/[the number of CpGs analyzed]), as well as the false discovery rate (FDR) procedure to account for multiple testing in EWAS results.[28] For the top 6 most significant results, we used the web-based plotting tool CoMet to graphically display additional information about all available CpGs within the region of 50 kb up- and downstream, including physical location, and statistical significance.[29] We additionally performed the diptest for unimodality for each CpG to test for non-unimodal distributions.[30] Our main analysis in Born into Life included DNAm measurements from 4 sampling occasions (ie, prepregnancy, early and late pregnancy, as well as after delivery) in 21 women. Because the replication cohort IoW (n = 27) had DNAm from only 3 occasions (ie, prepregnancy, first, and second halves of pregnancy), we reran our analysis based on 3 measurements, excluding the 2 to 4 days after delivery (n = 24) time point, and the FDR-significant results from these 3 sampling occasions were used for the replication analysis.

Biological pathway analysis

Furthermore, we performed a functional enrichment analysis to identify significant biological pathways characterized by multiple genes, which correspond to CpGs that varied significantly (FDR $P < .05$) throughout pregnancy. Gene Ontology (GO) terms and pathways in the FDR-significant CpGs were analyzed using Enrichr web tool (http://amp.pharm.mssm.edu/Enrichr/).[31]

Results

Repeated genome-wide DNAm measures were available for 34 women in the Born into Life study (Table 1). For the present analyses, we included 21 women who all had blood samples drawn before pregnancy (on average 1.6 years prior conception), twice during pregnancy (weeks 10-14 and weeks 26-28 of pregnancy), and postnatally at the time of the newborn screening test of the child 2 to 4 days after delivery. Table 2

Table 2. Descriptive statistics of the study samples of women in the Born into Life (BiL) and Isle of Wight (IoW) third-generation cohort with repeated DNA methylation measurements around pregnancy.

	BIL (N = 21) NO. (%)	IOW THIRD-GENERATION (N = 27) NO. (%)
Age at delivery[a], y	33.1 (4.1), [24.0-39.7]	23.7 (1.01), [21.7-25.6]
Parity,[b] ⩾1	5 (24)	7 (26)
Smoking during pregnancy	0	11 (42)
Maternal education		
University	16 (80)	10 (37.0)
Secondary school	3 (15)	6 (22.2)
Other	1 (5)	11 (40.7)
Pregnancy symptoms[c]	8 (38)	(Not collected)
Mode of delivery		
Vaginal	11 (52)	22 (81.5)
Caesarian section	6 (29)	5 (18.5)
Vacuum extraction	4 (19)	0
Gestational age[a], wk	39.3 (1.3), [38-42]	39.1 (1.5), [35-41]
Maternal chronic diseases	7 (33)[d]	10 (37)[e]
Season of delivery		
Winter (December-February)	5 (24)	8 (29.6)
Spring (March-May)	7 (33)	7 (25.9)
Summer (June-August)	5 (24)	7 (25.9)
Autumn (September-November)	4 (19)	5 (18.5)
BMI during early pregnancy[a]	23.2 (3.9), [19.5-36.1]	29 (7.2), [19.8-53.6]
Medication during pregnancy	8 (38)[f]	23 (92)[g]

[a]Mean (SD), [min-max].
[b]Number of previous pregnancies resulting in a live birth or stillbirth.
[c]Pregnancy symptoms reported at the follow-up during pregnancy weeks 26 to 28 included nausea (n = 8), vomiting (n = 2).
[d]Maternal chronic diseases retrieved from maternal ward records included endocrinological thyroid disease (n = 2), gynecologic disease or operation (n = 6).
[e]Maternal chronic diseases during pregnancy included asthma (5), hay fever (2), previous depression (2), sacral nerve damage (1), supraventricular tachycardia (1), Hodgkin lymphoma (1), and hypermobility (1).
[f]Any medication during pregnancy included vitamins and supplements, ie, folic acid (n = 8), inhaled corticosteroids (1), aspirin and omeprazole (1).
[g]Medication during pregnancy included paracetamol (12), antibiotic (2), becotide (1), clexane (1),

cocodamol (1), cyclizine (1), dihydrocodiene (1), gaviscon (1), iron supplements (1), salbutamol (2), sudafed (1).

summarizes some characteristics of the study participants in the discovery cohort (Born into Life) and the replication cohort (IoW third-generation study). None of the women in the Born into Life study sample reported smoking during pregnancy, having serious pregnancy complications, or taking medication. Women in the replication cohort had on average higher body mass index, as well as higher frequency of smoking (42%), and medication use during pregnancy (92%). All participants were of European ancestry and many had a university education (80% in the Born into Life study and 37% in the IoW third-generation study).

Discovery EWAS analysis

After exclusion of CpGs with probe-SNPs and with low quality, a total of 774 408 CpGs were included in the present analyses. Because cell types are known to influence DNAm, we estimated cell proportions for each sample. Supplementary Figure S1 shows the estimated cell composition. Compatible with the expected immunomodulation during pregnancy,[32] we observed significant differences in estimated proportions of the cell types between baseline and early pregnancy (Kruskal-Wallis test $P < .05$), with the exception of monocytes

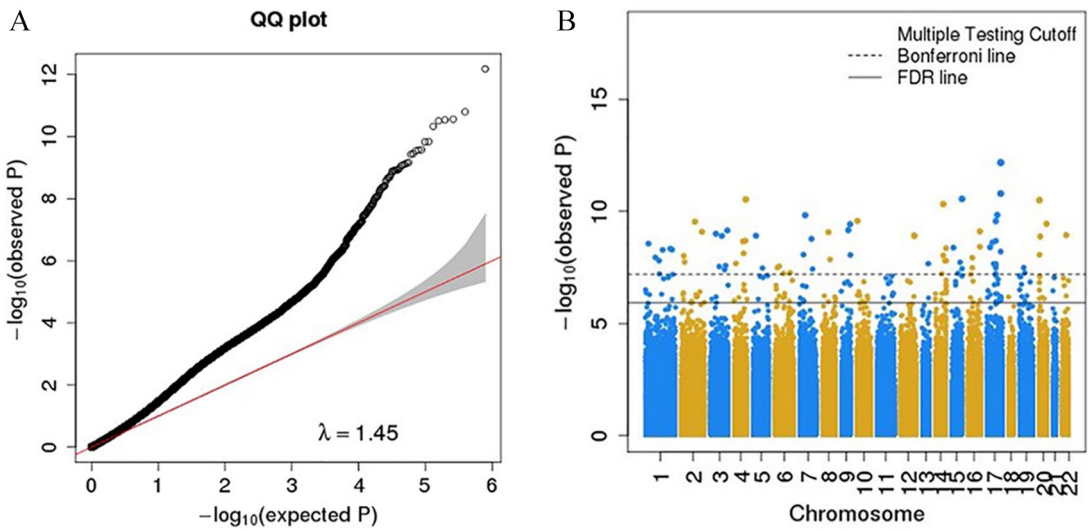

Figure 1. (A) Quantile-quantile plot and (B) Manhattan plot for epigenome-wide results testing for DNA methylation differences before, during, and after pregnancy in the Born into Life cohort (n = 21). (B) A total of 196 CpGs exhibit significant trend in DNA methylation measured in peripheral whole blood sampled at 4 occasions (once before, twice during, and once after pregnancy). The x-axis displays the chromosome on which the CpG is located, and the y-axis displays $-\log_{10}$ (*P* value). The solid and dashed horizontal lines represent the Bonferroni-adjusted and FDR *P* value thresholds, respectively. FDR indicates false discovery rate.

and eosinophils, whereas no statistically significant temporal change in any of the calculated cell proportions was found between the later time points.

We assessed inflation of test statistics by plotting the observed and expected *P* values across all CpGs on the array (Figure 1A). The inflation factor (λ) was 1.45. Figure 1B shows a Manhattan plot of *P* values for testing trends in DNAm levels throughout the 4 time points around and during pregnancy across the autosomes. Our analysis identified 196 CpGs displaying intra-individual longitudinal change with an FDR $P < .05$ (Supplementary Table S1). The vast majority of these CpGs (90.8%) demonstrated decrease in average methylation levels over the study period (178 out of 196 CpGs). Furthermore, we observed multiple differentially methylated CpGs within the same genes, namely, 5 CpGs mapping to Homeobox Protein Hox-B3 (*HOXB3* chr17), 4 CpGs to arginine vasopressin gene (*AVP* chr20); 3 CpGs to *LOC100996291* (chr 17) and to MicroRNA 10a (*MIR10A* chr 17), respectively, as well as 2 CpGs within each of the following genes: Serpin Family A Member 1 (*SERPINA1* chr14), CUE Domain Containing 1 (*CUEDC1* chr17), NACC Family Member 2 (*NACC2* chr9), Zinc Finger CCHC-Type Containing 14 (*ZCCHC14* chr16), NSE1 homolog (*NSMCE1* chr16), Strawberry Notch Homolog 2 (*SBNO2* chr19), Myosin IE (*MYO1E* chr15), Transmembrane Protein 49 (*TMEM49* chr17), NACC Family Member 2 (*NACC2* chr9), Pyrroline-5-carboxylate reductase (*PYCRL* chr8), and HOXA10 Antisense RNA *(HOXA10-AS* chr7). Figure 2 illustrates differential temporal patterns of methylation levels during pregnancy for the 6 top significant CpGs. In general, monotonous trends were indicated, sometimes with more pronounced changes from prior to pregnancy, but not 2 to 4 days after delivery. Furthermore, by examining the top 6 FDR-significant CpGs, we found some evidence for localized clustering around cg19748455 (*LOC100996291*) and cg05844798 with multiple genome-wide significant CpGs appearing within the same region (Figure S2). Also, the diptest did not reveal statistically significant deviation from unimodality for any of the 196 identified CpGs (FDR-adjusted $P > .05$).

Biological pathway analysis

Next, the list of the FDR-significant 196 CpGs was used to search for significantly enriched GO terms and the KEGG pathways using the Enrichr tool. Pathway analysis demonstrated that these CpGs belong to genes enriched in metabolic pathways, eg, adipose tissue development (*SH3PXD2B; RORC; ARID5B*), regulation of insulin receptor signaling (*SOCS2; KANK1; SOCS3; CISH; GRB10*), and mammary gland fat development (*SH3PXD2B; RORC; ARID5B*) (Table 3).

Replication analysis

Finally, we pursued replication analyses in an independent cohort with a similar biosampling protocol, the IoW study. However, for this replication attempt, we redid the analysis in Born into Life focusing on DNAm measurements from the 3 sampling occasions comparable with those available in IoW third-generation birth cohort (age 18, first and second halves of pregnancy). The distribution of covariates used in the analysis in the IoW study is shown in Supplementary Table S2. A total of 36 FDR-significant CpGs were available in the IoW sample based on the Infinium HumanMethylation450k BeadChip. Out of these CpGs, 17 (47%) replicated in the IoW cohort ($P < .05$), all but 2 with the same direction of effect (Supplementary Table S3).

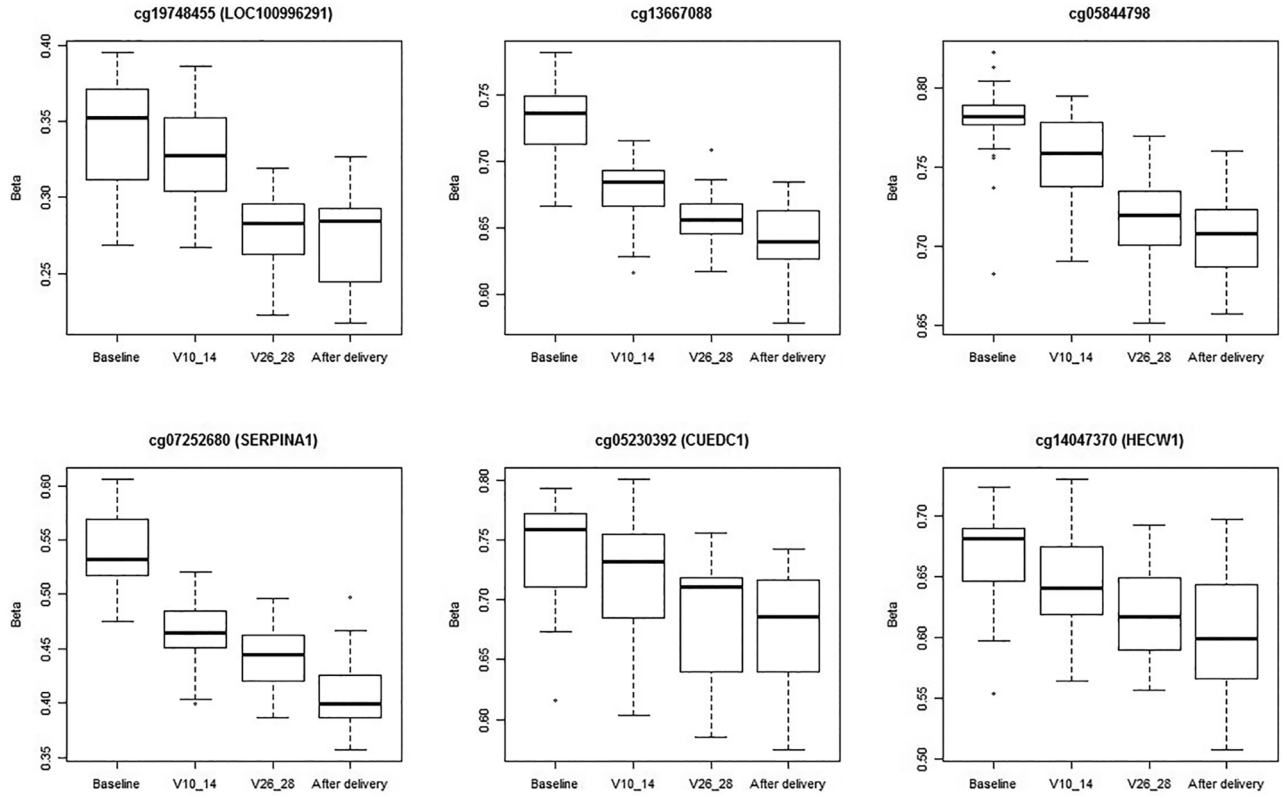

Figure 2. Boxplots of the cumulative distribution function for methylation intensities for the 6 top-ranked CpG sites throughout pregnancy in Born into Life women (n = 21).
Y-axis beta = methylation intensity; X-axis denotes biosampling occasions: baseline = preconception, V10_14 = pregnancy weeks 10 to 14, V26_28 = pregnancy weeks 26 to 28, after delivery = 2 to 4 days after delivery. Each box contains the middle 50% of the data, with the upper edge (hinge) of the box indicating the 75th percentile, and the lower one indicating the 25th percentile, the interquartile range (IQR). The line in the box represents the median. The upper and lower ends of the vertical lines ("whiskers") indicate upper quartile +1.5 IQR and lower quartile −1.5 IQR, respectively. Individual values outside this range are marked as circles.

Discussion

This study represents an epigenome-wide analysis evaluating temporal changes in DNAm from prepregnancy, to early and late, and postpregnancy in 21 women of the Swedish Born into Life cohort. We identified 196 differentially methylated CpGs across the genome after correction for multiple testing and blood cell type correction. Seventeen of 36 CpGs of these pregnancy-related methylation changes were significant in an independent replication cohort, IoW third-generation group of 27 women, and 15 of these 17 CpGs were also directionally consistent. Biological pathway analysis revealed that our identified CpGs belong to genes overrepresented in metabolism-related pathways. In addition, there were significant changes in the white blood cell types in the beginning of pregnancy, compatible with the expected immunomodulation of the maternal system to prevent allogenic rejection of the fetus.[32]

Several of the identified differentially methylated CpGs were located within or nearby genes previously implicated in fetal development, as well as respiratory and metabolic outcomes. We found multiple genome-wide significant CpGs annotated to the *HOX* gene family: 5 CpGs to the *HOXB3* gene and 2 CpGs to *HOXA10 Antisense RNA* gene. Hox proteins have earlier been shown to control differentiation and pattern formation throughout embryogenesis.[33] HoxB3 has been suggested as relevant for lung development by mechanistic studies demonstrating differential expression of *HOXB3* gene in embryonic lung cells.[34,35] *HOXA10* has been mainly differentially expressed in normal adult human lung, as well as with primary pulmonary hypertension.[36] We also found 4 differentially methylated CpGs mapping to the arginine vasopressin gene (*AVP*). Although AVP is well recognized as an antidiuretic hormone, increasing evidence supports its role in regulation of blood glucose levels.[37] A recent study based on a 9-year-long follow-up of more than 5000 subjects from France has reported associations between a set of SNPs in *AVP* and incidence of hyperglycemia.[38] Another 3 CpGs were localized in the RNA gene *MIR10A*. MiR-10a, along with a few other circulating microRNAs, has been suggested as potential biomarkers for diagnosis of cardio- and cerebrovascular diseases.[39,40] Differentially methylated CpGs were also identified in *SERPINA1*, which encodes for a serine protease inhibitor (alpha-1 antitrypsin, AAT), and has been shown to play a role in the development of chronic obstructive pulmonary disease[41]; *SBNO2*, previously identified as differentially methylated in association with body mass index, as well as interleukin-10 and serum C-reactive protein[42,43]; *CUEDC1*, the latter has

Table 3. Biological pathway enrichment analysis in the Born into Life cohort based on Gene Ontology biological process database and the KEGG pathway database.

TERM	*P* VALUE	ADJUSTED *P* VALUE	GENES
Gene Ontology biological process database			
Negative regulation of insulin receptor signaling pathway involved in determination of adult lifespan (GO:1903105)	2.15E−07	2.74E−04	*SOCS2; KANK1; SOCS3; CISH; GRB10*
Negative regulation of insulin receptor signaling pathway by insulin receptor internalization (GO:0038014)	2.67E−07	2.74E−04	*SOCS2; KANK1; SOCS3; CISH; GRB10*
Negative regulation of insulin receptor signaling pathway (GO:0046627)	3.50E−06	2.39E−03	*SOCS2; KANK1; SOCS3; CISH; GRB10*
Fat pad development (GO:0060613)	1.48E−05	5.04E−03	*SH3PXD2B; RORC; ARID5B*
Fat body development (GO:0007503)	9.89E−06	5.04E−03	*SH3PXD2B; RORC; ARID5B*
Adipose tissue development (GO:0060612)	1.48E−05	5.04E−03	*SH3PXD2B; RORC; ARID5B*
Mammary gland fat development (GO:0060611)	1.16E−04	3.39E−02	*SH3PXD2B; RORC; ARID5B*
The KEGG pathway database			
Jak-STAT signaling pathway_Homo sapiens_hsa04630	.002	.160	*SOCS2; SOCS3; CISH; PIM1; IL22RA2*
Notch signaling pathway_Homo sapiens_hsa04330	.003	.160	*JAG1; NOTCH1; MAML2*
Acute myeloid leukemia_Homo sapiens_hsa05221	.004	.174	*ZBTB16; PIM1; RUNX1*
Prolactin signaling pathway_Homo sapiens_hsa04917	.008	.237	*SOCS2; SOCS3; CISH*
Glycosphingolipid biosynthesis—lacto and neolacto series_Homo sapiens_hsa00601	.010	.237	*GCNT2; ST3GAL3*
Transcriptional misregulation in cancer_Homo sapiens_hsa05202	.020	.407	*ZBTB16; MLLT1; SUPT3H; RUNX1*
Vasopressin-regulated water reabsorption_Homo sapiens_hsa04962	.026	.455	*ADCY3; AVP*
Type II diabetes mellitus_Homo sapiens_hsa04930	.031	.455	*SOCS2; SOCS3*
Arginine and proline metabolism_Homo sapiens_hsa00330	.033	.455	*PYCRL; SMOX*
Cyanoamino acid metabolism_Homo sapiens_hsa00460	.039	.486	*GGT6*
Metabolic pathways_Homo sapiens_hsa01100	.050	.559	*GALC; GGT6; PYCRL; CHSY1; NDUFA10; MGAT3; GCNT2; PHGDH; NADK; PLB1; COASY; ST3GAL3*

been shown to have a high methylation level in the placental tissue in women with preeclampsia.[44] Furthermore, 2 of the identified genes have been demonstrated to affect female reproductive system carcinogenesis, namely, *CUEDC1* has been associated with early stage cervical[45] and *TMEM49* with ovarian cancer.[46]

Furthermore, genes with significantly changed CpG methylation during the course of pregnancy appeared to be overrepresented in pathways which control various metabolic processes such as insulin receptor signaling, mammary gland fat development, and adipose tissue development. It remains to be investigated if the observed changes in DNAm during pregnancy result from hormonal alterations, preparation for breastfeeding, or adjustment in lifestyle occurring in the course of pregnancy. Importantly, as few of these genes are expressed in the blood cells, it is possible that the presence of small numbers of fetal cells in the mother's blood during pregnancy may have contributed to the observed methylation changes.

Only 1 study has evaluated temporal trends in genome-wide DNAm during pregnancy, by contrasting early and late pregnancy.[6] The authors did not find any strong evidence for either consistent or significantly changed CpGs (0.35% and

0% of all tested CpGs, respectively), thus concluding that DNAm of most of the CpGs, 99.65%, is in an indifferent state. In contrast to our study, they compared DNAm at 2 time points during pregnancy, and their analysis was limited to 218 222 CpGs due to stringent quality control procedures applied to the methylation data measured with Illumina 450k array.

The strengths of our study include a relatively homogeneous and well-characterized study population of healthy women with repeated DNAm assessment before, during, and after pregnancy using the new Illumina EPIC array. Moreover, we adjusted for potential confounders and/or effect modifiers such as maternal age at delivery, parity, as well as cell composition. Although our sample was relatively small, we identified a number of significant CpG sites after adjustment for multiple testing and potential confounding and replicated them in an independent cohort. Furthermore, several genes were represented by multiple differentially methylated CpGs.

One challenge of genome-wide DNAm analyses in blood samples with a mixed cell composition is the difference in methylation patterns between different cell types. In the present analysis, we used the available reference data for adult peripheral blood cells to correct for cell-type proportions.[26] One should bear in mind that this reference is based on the DNAm data on healthy male blood donors derived from the Illumina 450k array. Furthermore, the maternal circulating blood may contain fetal cells with different methylation patterns that are not corrected for. We should also acknowledge that in the present analyses we were primarily interested in identifying methylation sites with a strong linear change from preconception to postpartum, potentially relevant for the mechanisms underlying maternal physiologic adaptations to pregnancy and the cause of pregnancy complications. Therefore, a generalized mixed linear modeling approach was applied. Nonlinear patterns of change in DNAm across the study period were not evaluated specifically.

Conclusions

This study adds to the sparse literature examining DNAm trends over the course of pregnancy. Significant changes in DNAm in peripheral blood during pregnancy were identified, related to genes with a possible role in adaptation for pregnancy and lactation.

Acknowledgements

The authors acknowledge study participants, and nurses, as well as Karolinska Institutet Biobank and the Department of Clinical Sciences, Danderyd University Hospital, Karolinska Institutet, Stockholm, Sweden.

Author Contributions

GP, JK, AS, CA, SC, and WK contributed to conceptualization; GP, EM, OG, AMH, and SKM to methodology; GP, CS, AS, CA, HSA, SC, NM, and SKM to investigation and formal analysis; AMH, VU, CS, and HSA to resources; OG and GP to writing—original draft; all coauthors to writing—review and editing; OG, SKM, YJ, and NM to visualization; GP to supervision; GP, CA, EA, JK, AS, and WK to funding acquisition. All authors contributed to the final version of the manuscript.

ORCID iD

Olena Gruzieva https://orcid.org/0000-0003-4147-5654

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Calicchio R, Doridot L, Miralles F, Mehats C, Vaiman D. DNA methylation, an epigenetic mode of gene expression regulation in reproductive science. *Curr Pharm Des*. 2014;20:1726-1750.
2. Ciccarone F, Tagliatesta S, Caiafa P, Zampieri M. DNA methylation dynamics in aging: how far are we from understanding the mechanisms? *Mech Age Develop*. 2017;174:3-17.
3. Bjornsson HT, Sigurdsson MI, Fallin MD, et al. Intra-individual change over time in DNA methylation with familial clustering. *JAMA*. 2008;299:2877-2883.
4. Feinberg AP, Irizarry RA, Fradin D, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med*. 2010;2:49ra67.
5. Flanagan JM, Brook MN, Orr N, et al. Temporal stability and determinants of white blood cell DNA methylation in the breakthrough generations study. *Cancer Epidemiol Biomarkers Prev*. 2015;24:221-229.
6. Chen S, Mukherjee N, Janjanam VD, et al. Consistency and variability of DNA methylation in women during puberty, young adulthood, and pregnancy. *Genet Epigenet*. 2017;9:1179237X17721540.
7. Florath I, Butterbach K, Muller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Hum Mol Genet*. 2014;23:1186-1201.
8. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14:R115.
9. Talens RP, Christensen K, Putter H, et al. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell*. 2012;11:694-703.
10. Wong CC, Caspi A, Williams B, et al. A longitudinal study of epigenetic variation in twins. *Epigenetics*. 2010;5:516-526.
11. Xu CJ, Bonder MJ, Soderhall C, et al. The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics*. 2017;18:25.
12. Martino D, Loke YJ, Gordon L, et al. Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biol*. 2013;14:R42.
13. Acevedo N, Reinius LE, Vitezic M, et al. Age-associated DNA methylation changes in immune genes, histone modifiers and chromatin remodeling factors within 5 years after birth in human blood leukocytes. *Clin Epigenetics*. 2015;7:34.
14. Shah S, McRae AF, Marioni RE, et al. Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Res*. 2014;24:1725-1733.
15. Lain KY, Catalano PM. Metabolic changes in pregnancy. *Clin Obstet Gynecol*. 2007;50:938-948.
16. Tan EK, Tan EL. Alterations in physiology and anatomy during pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2013;27:791-802.
17. Smew AI, Hedman AM, Chiesa F, et al. Limited association between markers of stress during pregnancy and fetal growth in "Born into Life," a new prospective birth cohort. *Acta Paediatr*. 2018;107:1003-1010.
18. Almqvist C, Adami HO, Franks PW, et al. LifeGene: a large prospective population-based study of global relevance. *Eur J Epidemiol*. 2011;26:67-77.
19. Arshad SH, Holloway JW, Karmaus W, et al. Cohort profile: the Isle of Wight whole population birth cohort (IOWBC). *Int J Epidemiol*. 2018;47:1043-1044i.
20. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8:118-127.

21. Xu ZL, Niu L, Li LP, Taylor JA. ENmix: a novel background correction method for illumina humanmethylation450 BeadChip. *Nucleic Acids Res*. 2016;44:e20.
22. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the illumina infinium methylationEPIC BeadChip. *Genom Data*. 2016;9:22-24.
23. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*. 2013;14:293.
24. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol*. 2015;16:37.
25. Kuznetsova A, Brockhoff PB, Christensen RHB. lmertest package: tests in linear mixed effects models. *J Stat Softw*. 2017;82:1-26.
26. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE*. 2012;7:e41361.
27. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86.
28. Strimmer K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics*. 2008;24:1461-1462.
29. Martin TC, Yet I, Tsai PC, Bell JT. coMET: visualisation of regional epigenome-wide association scan results and DNA co-methylation patterns. *BMC Bioinform*. 2015;16:131.
30. Hartigan JA, Hartigan PM. The dip test of unimodality. *Ann Stat*. 1985;13:70-84.
31. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44: W90-W97.
32. Sanci M, Toz E, Ince O, et al. Reference values for maternal total and differential leukocyte counts in different trimesters of pregnancy and the initial postpartum period in western Turkey. *J Obstet Gynaecol*. 2017;37:571-575.
33. Jones PL. Homeobox genes in pulmonary vascular development and disease. *Trends Cardiovasc Med*. 2003;13:336-345.
34. Nakamura N, Yoshimi T, Miura T. Increased gene expression of lung marker proteins in the homeobox B3-overexpressed fetal lung cell line M3E3/C3. *Cell Growth Differ*. 2002;13:195-203.
35. Yoshimi T, Nakamura N, Shimada S, et al. Homeobox B3, FoxA1 and FoxA2 interactions in epithelial lung cell differentiation of the multipotent M3E3/C3 cell line. *Eur J Cell Biol*. 2005;84:555-566.
36. Golpon HA, Geraci MW, Moore MD, et al. HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. *Am J Pathol*. 2001;158:955-966.
37. Nakamura K, Velho G, Bouby N. Vasopressin and metabolic disorders: translation from experimental models to clinical use. *J Intern Med*. 2017;282:298-309.
38. Roussel R, El Boustany R, Bouby N, et al. Plasma copeptin, AVP gene variants, and incidence of type 2 diabetes in a cohort from the community. *J Clin Endocrinol Metab*. 2016;101:2432-2439.
39. Luo L, Chen B, Li S, et al. Plasma miR-10a: a potential biomarker for coronary artery disease. *Dis Markers*. 2016;2016:3841927.
40. Li M, Zhang J. Circulating MicroRNAs: potential and emerging biomarkers for diagnosis of cardiovascular and cerebrovascular diseases. *Biomed Res Int*. 2015;2015:730535.
41. Lomas DA, Silverman EK. The genetics of chronic obstructive pulmonary disease. *Respir Res*. 2001;2:20-26.
42. Fernandez-Sanles A, Sayols-Baixeras S, Curcio S, Subirana I, Marrugat J, Elosua R. DNA methylation and age-independent cardiovascular risk, an epigenome-wide approach: the REGICOR study (REgistre GIroni del COR). *Arterioscler Thromb Vasc Biol*. 2018;38:645-652.
43. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature*. 2017;541:81-86.
44. Yan YH, Yi P, Zheng YR, et al. Screening for preeclampsia pathogenesis related genes. *Eur Rev Med Pharmacol Sci*. 2013;17:3083-3094.
45. Biewenga P, Buist MR, Moerland PD, et al. Gene expression in early stage cervical cancer. *Gynecol Oncol*. 2008;108:520-526.
46. Zheng LL, Chen LL, Zhang X, Zhan JF, Chen J. TMEM49-related apoptosis and metastasis in ovarian cancer and regulated cell death. *Mol Cell Biochem*. 2016;416:1-9.