

Consistency and Variability of DNA Methylation in Women During Puberty, Young Adulthood, and Pregnancy

Genetics & Epigenetics
Volume 9: 1–9
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1179237X17721540



Su Chen¹, Nandini Mukherjee², Vimala Devi Janjanam²,
S Hasan Arshad^{3,4}, Ramesh J Kurukulaaratchy^{3,4},
John W Holloway^{3,5}, Hongmei Zhang² and Wilfried Karmaus²

¹Department of Mathematical Sciences, The University of Memphis, Memphis, TN, USA. ²Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, The University of Memphis, Memphis, TN, USA. ³Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. ⁴The David Hide Asthma and Allergy Research Centre, Newport, UK. ⁵Academic Unit of Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK.

ABSTRACT: Prior DNA methylation (DNA-m) analyses have identified cytosine-phosphate-guanine (CpG) sites, which show either a significant change or consistency during lifetime. However, the proportion of CpGs that are neither significantly different nor consistent over time (indifferent CpGs) is unknown. We investigated the methylation dynamics, both longitudinal changes and consistency, in women from preadolescence to late pregnancy using DNA-m of peripheral blood cells. Consistency of cell type-adjusted DNA-m between paired individuals was assessed by regressing CpGs of subsequent age on the prior, stability by intraclass correlation coefficients (>0.5), and changes by linear mixed models. In the first 2 transitions (10–18 years and 18 years to early pregnancy), 19.5% and 20.9% CpGs were consistent, but only 0.35% in the third transition (from early to late pregnancy). Significant changes in methylation were found in 0.7%, 5.6%, and 0% CpGs, respectively. Functional enrichment analyses of genes with significant changes in DNA-m in early pregnancy (5.6%) showed that the maternal DNA-m seems to reflect signaling pathways between the uterus and the trophoblast. The transition from early to late pregnancy showed low consistency/stability and no changes, suggesting the presence of a large proportion of indifferent CpGs in late pregnancy.

KEYWORDS: Puberty, pregnancy, DNA methylation, CpGs, consistency, stability, change

RECEIVED: March 8, 2017. **ACCEPTED:** May 3, 2017.

PEER REVIEW: Two peer reviewers contributed to the peer review report. Reviewers' reports totaled 712 words, excluding any confidential comments to the academic editor.

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 study is supported by the

grants from the National Institute of Allergy and Infectious Diseases (R01AI091905) and the National Heart, Lung, and Blood Institute (R01HL132321).

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: Wilfried Karmaus, Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, The University of Memphis, 301 Robison Hall, 3825 DeSoto Avenue, Memphis, TN 38152, USA.
Email: karmaus1@memphis.edu

Introduction

Differentially methylated cytosine-phosphate-guanine (CpG) sites, both with stable and varying levels, have been associated with multiple adverse health conditions.¹ A few studies have investigated the temporal stability of DNA methylation (DNA-m) from peripheral blood cells over time.^{2–4} One earlier study, using samples from an Icelandic cohort and from Salt Lake City, Utah, compared global DNA-m and found that only 8% to 10% of the participants showed a change of more than 20% in DNA-m over 11 to 16 years.² Focusing on 227 inter-individually variable methylated regions (VMRs) in the same Icelandic cohort, Feinberg et al³ reported that 119 (52.4%) of these represent stable VMRs; 41 (18.1%) were classified as having higher intra-individual differences. Flanagan et al⁴ focused on 92 participants of a British cohort, 35 to 84 years of age, and determined the methylation CpG dinucleotide sequences twice, 6 years apart. They found that 17% of 353 633 CpGs were stable after exclusion of CpGs that were annotated with single-nucleotide polymorphisms (SNPs) in DNA-m measurement probes, also called probe-SNPs. A Danish study investigated statistically significant changes in CpG methylation in 43 paired measurements 10 years apart (age range at first measurement:

73–82 years).⁵ The investigators removed a smaller number of CpGs affected by probe-SNPs resulting in a total of 424 706 analyzed CpGs followed by cell composition adjustment. Their results addressing epigenetic changes showed that only 0.054% (2284/424 706) showed significant changes. The number of steady CpGs was not provided in this study.

In a recently published article addressing DNA-m changes in early life, 538 paired samples (aged 0 and 4/5 years) and 726 paired samples (aged 4 and 8 years) from Spain, France, Sweden, and the Netherlands were assessed.⁶ Focusing on changes only, the authors identified that the methylation of 2.2% of 439 306 CpGs reduced significantly both from age 0 to 4/5 and 4 to 8 years, and 1% CpGs increased significantly in both periods. There were 0.08% of CpGs showing a decrease in methylation from age 0 to 4/5 years followed by an increase from 4 to 8 years and vice versa in 0.23% CpGs.

Other studies investigated associations of methylation and age reporting high correlations and prediction.^{7–9} However, most of these studies focused on ages after the reproductive lifetime which are relative steady life periods and excluded early transition periods of puberty and pregnancy.



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

In contrast, our intention is to investigate important transition periods, puberty, and pregnancy. Only 2 genome-wide studies to date have investigated the transition period of puberty, none has studied pregnancy. In a European study of 376 adult healthy women (33–75 years of age), a link between self-reported age at menarche and peripheral blood DNA-m patterns in adults was investigated.¹⁰ The results were internally inconclusive as DNA-m was not determined in the period of pubertal transition but later in adulthood. In a Danish investigation of 32 boys and 22 girls, DNA-m was measured twice, at the beginning and the end of the pubertal transition.¹¹ When boys and girls were analyzed separately, significant associations of puberty with methylation changes of CpGs were only found in boys, which may be related to the smaller sample size of girls in this study.

We investigated 3 transition periods in the life course of women: (1) childhood to adolescence (puberty: between 10 and 18 years), (2) young adulthood to the first half of pregnancy (18 years to early pregnancy), and (3) early to late pregnancy (within pregnancy dynamics (up to week 21 of gestation compared with week 22 and later). In addition, we addressed the 2 possible poles in DNA-m pattern over the ages: (1) stability/consistency and (2) change of DNA-m over time. To investigate this polarity, using paired samples, we will distinguish between *consistency* over time, defined as significant predictability of the second measurement of DNA-m by the first, and *stability* evaluated by intraclass correlation coefficients (ICCs) (>0.5). In addition, change of DNA-m, the other end in the polarity in contrast to stability/consistency, will be assessed using paired *t* test without cell type adjustment. Linear mixed model with repeated measurements will be applied to additionally adjust for cell proportions.

Materials and Methods

DNA-m analysis

DNA was extracted from peripheral blood sampled at age 10 years ($n=34$), at age 18 years ($n=245$), in early pregnancy ($n=139$), and in late pregnancy ($n=127$). For peripheral blood, standard salting out procedure was used to isolate DNA and its concentration was determined by Qubit quantitation. About 1 μg of DNA was bisulfite treated for cytosine to thymine conversion using the EZ 96-DNA methylation kit (Zymo Research, Irvine, CA, USA), following the manufacturer's standard protocol. The Infinium HumanMethylation450 BeadChip from Illumina (Illumina, San Diego, CA, USA) was used to obtain methylation levels, using the manufacturer's standard protocol. This chip measures more than 450 000 CpG sites. Arrays were processed at the Wellcome Trust Centre for Human Genetics (Oxford, UK). Multiple identical control samples were assigned to each bisulfite conversion batch to assess assay variability. Batch effects were controlled by randomly distributing samples on beadchips.

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 Methylation Analyzer¹² package and the ComBat¹³ package 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.¹⁴ The CpG sites that had detection *P* values > .01 in >10% of all samples (age 10, age 18, early, and late pregnancy), and CpG sites with probe-SNPs, were excluded from all analyses. We found 89 892 sites in the Illumina annotation file affected by probe-SNPs defined as having a SNP in the probe 10 and less base pairs apart from the CpG site (http://support.illumina.com/downloads/humanmethylation450_15017482_v1-2_product_files.html). In addition, we focused on the 22 autosomes excluding all CpGs on sex chromosomes, which resulted in 218 222 CpGs used in the assessments. M values (logit-transformed β values) were used in all analyses because β values are often heteroscedastic (unequal variance related to predictors).¹⁵

Analytical design

To assess linear associations (i.e., consistency), we applied linear regressions and compared genome-wide pairs of individual CpG methylation measurements determined 8 years apart (1: age 10 to age 18 years, 34 pairs), on average 5.3 years (2: age 18 to early pregnancy, 41 pairs) and few (2.3) months apart (3: early to late pregnancy, 63 pairs). Intraclass correlation coefficients were used to compare the stability of the methylation of 218 222 CpG pairs.

To adjust for cell composition in each period (age 10, age 18, early, and late pregnancy) using cell-specific CpGs, we estimated 7 cell types, CD4T, CD8T, natural killer cells, B cells, monocytes, granulocytes, and eosinophils using the R-package *minfi* according to the approach by Houseman et al¹⁶ and Jaffe and Irizarry.¹⁷ Then, the methylation of 218 222 CpGs was regressed on the proportion of the cell types, and the residuals not explained by cell types were estimated for each time period. These estimated residuals were used to test the consistency and stability between DNA-m measured in 2 adjacent time periods.

First, we tested linear associations between the paired CpG measurements with the residuals from the first methylation measurement as predictor and the second measurement as response as follows: Cell-adjusted residuals of the methylation_{*t*+1} = $\alpha + \beta_1$ * cell-adjusted residuals of the methylation_{*t*}.

Because we conducted a large number of 218 222 tests, we controlled for false discovery rate (FDR) to adjust for multiple testing preventing a large proportion of false positives.¹⁸ The variation of statistically significant associations in these 3 periods is demonstrated by Manhattan plots. We used volcano plots to illustrate whether the statistically significant predictions are positive or negative. To check the stability independent of consistency,

the ICCs of all the 218222 CpGs were assessed for each of the 3 transition periods.

To estimate changes in DNA-m, we applied linear mixed model with repeated measurement taking into account the within-subject time effect as well as adjusting for cell type effects:

$$y_{ij} = \beta_{0i} + \beta_1 t_j + \sum_{k=1}^6 \beta_{2k} z_{ijk} + e_{ij}$$

where y_{ij} denotes the DNA-m of subject i measured at time j ; $\beta_{0i} = \beta_{00} + U_i$ is the random intercept for each subject with $U_i \sim N(0, \sigma_0^2)$; β_1 represents time effect; z_{ijk} is the k th cell type proportion for i th subject measured at time j , $j=1, 2$; and $e_{ij} \sim N(0, \sigma^2)$ is the random error. If the cell type effects were removed from the model, the mixed model is equivalent to the paired t test. In this model, we assume that cell types have the same effect on DNA-m measured at the 2 time points included in these comparisons.

To address the biological importance of the findings, we performed a functional enrichment and identified the significant biological pathways characterized by multiple genes, which correspond to CpGs that varied significantly in the transition periods. ToppFun, a subfunction of the ToppGene Suite (<https://toppgene.cchmc.org/>), was used to identify functional pathways. For genes that were not identified in the databases, gene names or synonyms were imputed before the conduction of gene pathway enrichment analysis.¹⁹ Controlling for multiple testing, we used a FDR (FDR B&Y) of 0.05.²⁰

Results

In the 3 transition periods (age 10 to age 18 years, age 18 to early pregnancy, and early to late pregnancy), different numbers of participant pairs with repeated methylation measurements in peripheral blood cells at both time points were available. From the same cohort of women, the number of participants was as follows: $n=34$ for the transition from 10 to 18 years, $n=41$ for the transition from 18 years to early pregnancy, and $n=63$ for the transition from early to late pregnancy. The time span for the first transition was on average 8 years, 5.3 years (range: 3–6.7 years apart) for the second transition, and 2.3 months (range: 1.2 to 4.1 months apart) for the third transition. After exclusion of CpGs with probe-SNPs and with low quality at one time point and those with probe-SNPs, we analyzed 218222 CpGs on the 22 autosomes at each time point. In addition, because cell types are known to influence DNA-m, we estimated cell composition. Supplemental Table 1 shows the estimated cell composition based on cell-specific CpGs. With the exception of monocytes, proportions of all cell types (B cells, CD4 cells, and CD8T cells, granulocytes, and eosinophilic granulocytes) were statistically significantly different in the transition from age 10 to age 18 and from age 18 to early pregnancy. Interestingly, paired estimations of cell types in early and late pregnancy ($n=63$) suggest a similar cell composition at both time points.

Figure 1 shows Manhattan plots of FDR-adjusted P values for testing the consistency of 218222 CpG sites used in this study across the autosomes (22 chromosomes) for each of the 3 transition periods. The dashed lines indicate threshold of $-\log_{10}(0.05) \approx 1.3$ corresponding to an FDR-adjusted P value of .05. The dots above the dashed lines represent significant CpGs. The number of CpGs with statistically significant regression coefficients (both positive and negative) of subsequent CpG methylations is drastically reduced in the transition from early to late pregnancy (Figure 1C) compared with the transition from age 10 to age 18 years (Figure 1A) and to that from age 18 years to early pregnancy (Figure 1B). These Manhattan plots do not display the direction of the effects.

To inspect the directions of the regression coefficients (positive or negative), we used volcano plots, 1 for each of the 3 transition periods on the 218222 CpG pairs (Figure 2). We additionally show these plots with and without adjustment for cell types (Figure 2A to C and Figure 2D to F, respectively). The CpGs with statistically significant intra-individual regression coefficients with an FDR-adjusted P value of $<.05$ are colored in red. Those with positive coefficients are plotted to the right in the volcano plot and those with negative coefficients to the left. The nonsymmetric pattern indicates a predominant trend of CpGs showing a positive regression coefficient in the 3 transition periods. A positive coefficient results when the methylation value of a CpG at 2 time points shows a similar direction; e.g., the methylation of a CpG is low at age 10 years as well as at age 18 years, whereas if the methylation of specific CpGs at age 10 years is relatively high and the methylation at age 18 years becomes low, this would be indicated by a negative coefficient. Fewer CpGs show a negative regression coefficient in the 3 transition periods (left sites of the plots), which, however, rarely achieves a coefficient of -1 . A regression coefficient of 1 implies no change (not only being consistent but also being stable) in DNA-m between 2 measurements. After the first transition, the number of CpGs with coefficients >1 (increase in methylation from the first time point to the second) reduces to nearly 0 in the 2 subsequent transition periods (Figure 2).

The ICC measures absolute agreement and an ICC of 1 shows perfect accordance. The ICC plots (Figure 3) suggest that CpGs with an ICC >0.50 , implying a higher tendency of having more stable methylation levels in 2 subsequent DNA-m measurements, are reduced in the transition from early to late pregnancy, irrespective of adjustment for cell type. Table 1 shows that the estimated stability of DNA-m (percentage of CpGs with an ICC >0.5) in the first 2 transitions is 16.1% and 14.5% with cell type adjustment but only 0.13% in the repeated measurements in early and late pregnancies. The proportion with consistency (percentage of CpGs with positive significant [FDR P value $<.05$] regression coefficients) is higher: 19.5%, 20.9%, and 0.35% in the 3 transition periods (and includes the proportion with stability; Table 1). The proportion of CpGs with FDR-adjusted P values ($\leq.05$) decreases when adjusted for cell composition.

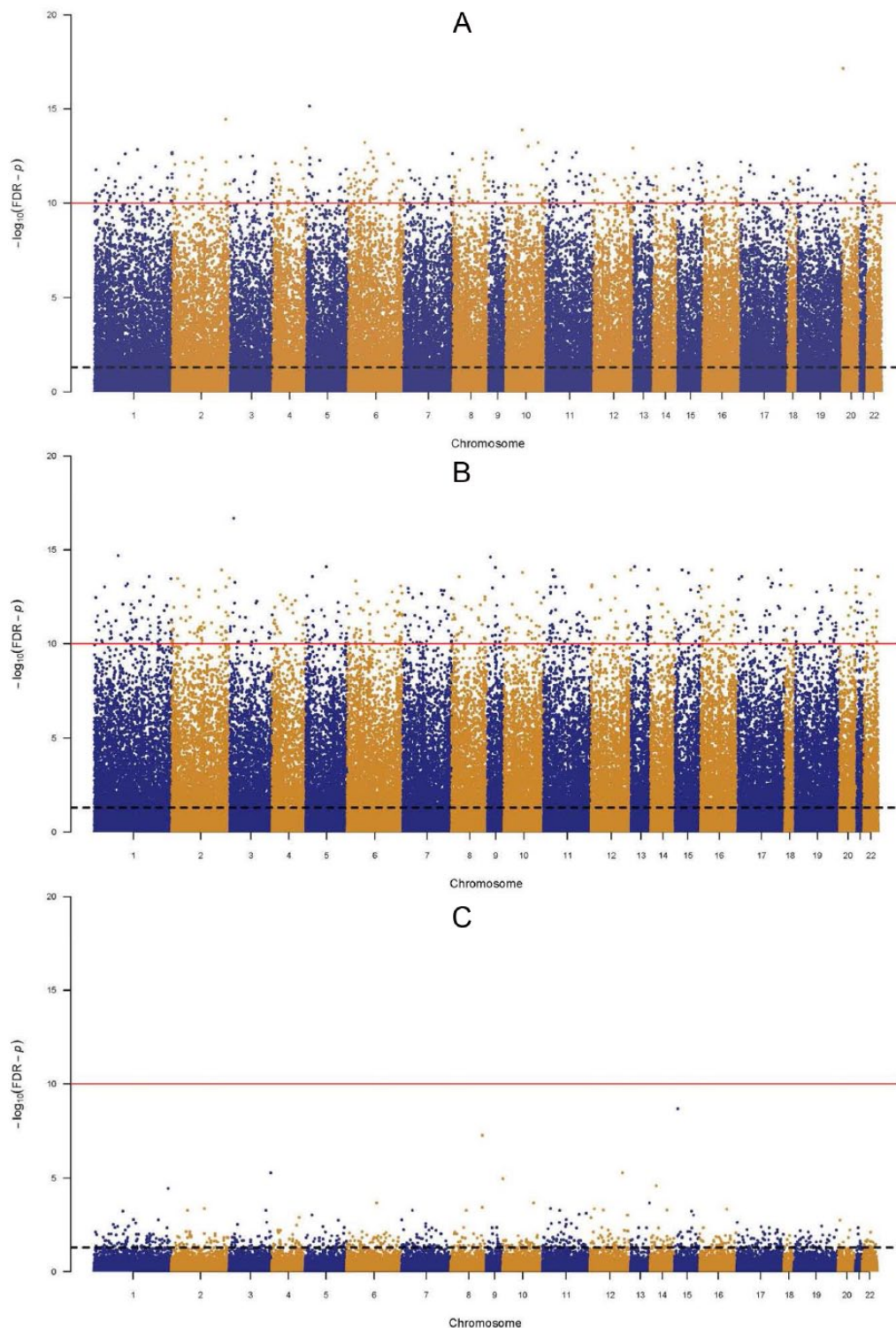


Figure 1. Manhattan plot for consistency of DNA methylation between different time periods in women (after cell type adjustment) between (A) age 10 and 18 years, (B) age 18 years and early pregnancy, and (C) early (weeks 20–21) and late pregnancy (>week 22). The black dashed line is a reference for $\text{FDR}=0.05$.

A significant difference in DNA-m of the same subject between 2 time points characterizes a change of DNA-m in each transition. In all 3 transitions, the proportion of CpGs with significantly changed DNA-m ($\text{FDR-adjusted } P < .05$) is smaller when adjusted for cell composition and larger when nonadjusted (Table 1). Between age 10 and 18 years, 0.7% of

the CpGs show a significant change in methylation and 5.6% between 18 years and early pregnancy (Table 1). From early to late pregnancy, none of the CpGs are statistically significantly changed with cell type adjustment.

In addition to assess the biological importance of the methylation changes, we investigated whether their corresponding

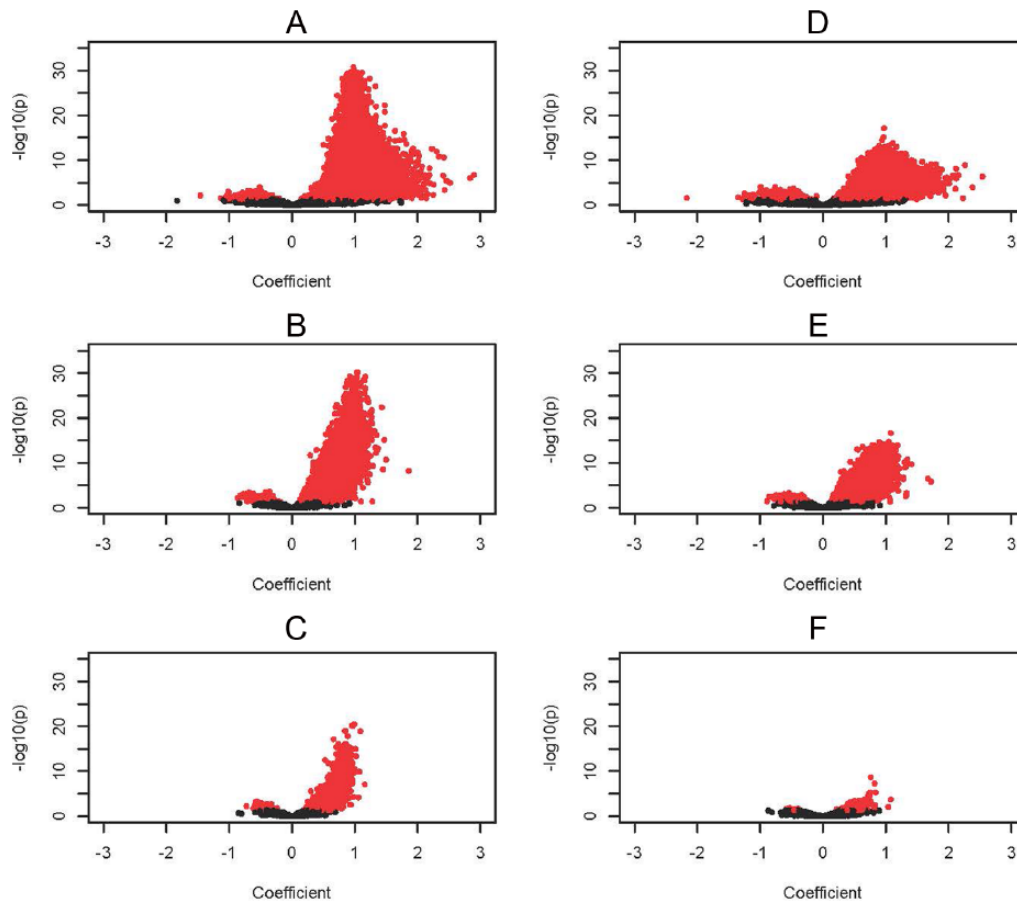


Figure 2. Volcano plots showing the regression coefficient of the paired DNA methylation measured at the latter time period on the level of the earlier measurement. The CpGs with statistically significant ($FDR < 0.05$) intra-individual regression coefficients are colored in red. Those with positive coefficients are plotted to the right and those with negative coefficients to the left. (A) Age 10 to age 18 years without adjustment for cell composition, (B) age 18 years to early pregnancy without adjustment for cell composition, (C) early (weeks 20-21) to late pregnancy (>week 22) without adjustment for cell composition, (D) age 10 to age 18 years with adjustment for cell composition, (E) age 18 years to early pregnancy with adjustment for cell composition, and (F) early (weeks 20-21) to late pregnancy (>week 22) with adjustment for cell composition. CpG indicates cytosine-phosphate-guanine; FDR, false discovery rate.

genes were suggestive of biological pathways. For the 2 transition periods of age 10 to 18 years, and 18 years to early pregnancy, the significantly changed CpGs corresponded to 1126 and 6628 unique genes, respectively. After imputation of genes with synonyms and updated gene names, ToppFun identified 1100 and 6432 genes, respectively, for the 2 transition periods. For the first transition, only 1 pathway was identified with $FDR P$ value $\leq .05$, which, however, could not be interpreted as being related to this adolescent period. For the transition between age 18 and pregnancy, we found 8 significant pathways based on the KEGG database (Supplemental Table 2). Five of these constitute signaling pathways (Rap1 signaling pathway, MAPK signaling pathway, Ras signaling pathway, calcium signaling pathway, and Hippo signaling pathway). The other genes with significantly changed CpGs belong to the pathways “Endocytosis,” “Proteoglycans in cancer,” and “Pathways in cancer.”

Discussion

Focused on 3 periods of transition in young women, this is the first study that has analyzed and compared changes in DNA-m

in peripheral blood cells from age 10 to 18 years (puberty), age 18 to early pregnancy (conception), and early to late pregnancy (gestation). Stability (high intraclass correlation) and consistency (significant regression coefficient) were estimated in paired samples of the same women by assessing whether the subsequent methylation levels (M values) were explained by the preceding levels. The results show limited consistency and stability of the methylation in 218222 analyzed CpG sites. Of importance is that we excluded 89892 CpG sites via the Illumina annotation file, which may be affected by probe-SNPs. Hence, we did not consider those CpGs that are possibly affected by genetic polymorphisms. Surprisingly, the major loss of stability and consistency of the methylation is during the third transition period of gestation even though the time gap is much closer compared with the first 2 transition periods. After adjustment for cell composition, only 0.13% and 0.35% are stable or consistent from early to late pregnancy, respectively. The predominant positive coefficients in volcano plots show that most of the significantly linearly correlated CpGs are consistent between 2 transition time points (i.e., individuals with low/high DNA-m

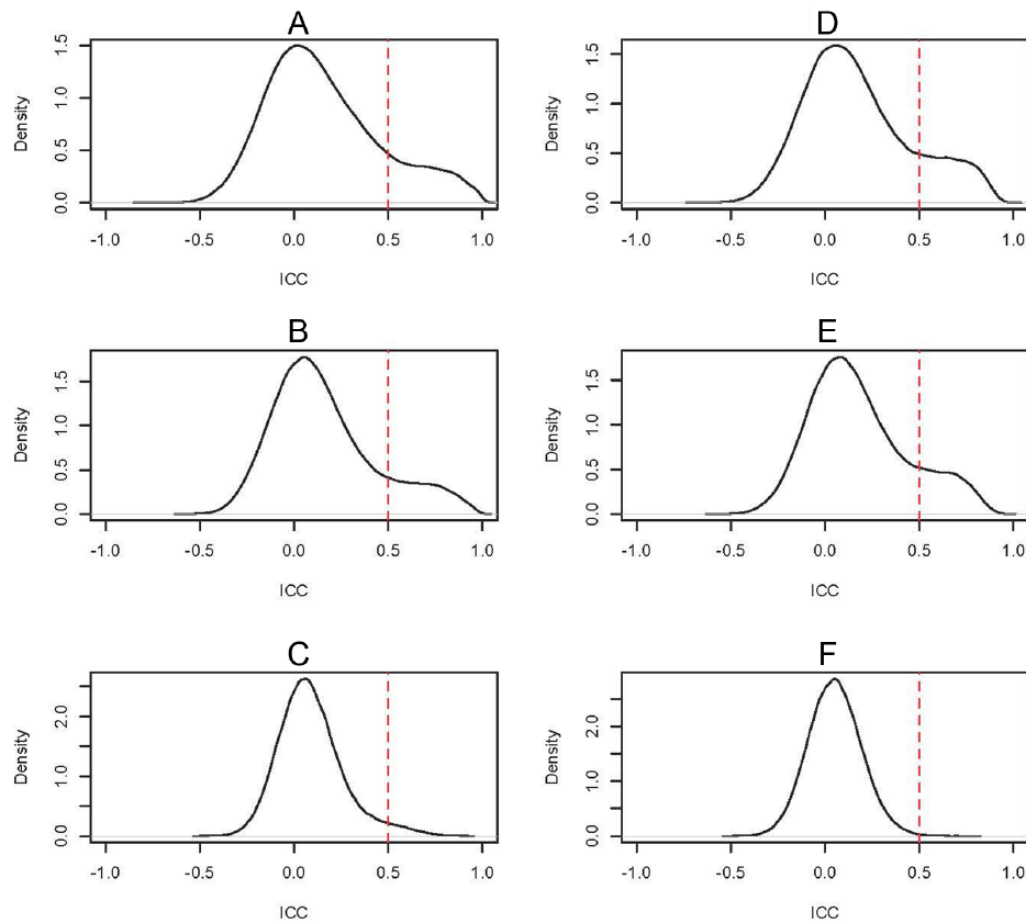


Figure 3. Analysis of the stability of 218222 CpGs using ICCs with and without prior adjustment for cell composition. (A) Age 10 to age 18 years without adjustment for cell composition—32916 CpGs are stable with a threshold of $ICC > 0.5$. (B) Age 18 years to early pregnancy without adjustment for cell composition—29832 CpGs are stable with a threshold of $ICC > 0.5$. (C) Early (weeks 20-21) to late pregnancy (>week 22) without adjustment for cell composition—6626 CpGs are stable with a threshold of $ICC > 0.5$. (D) Age 10 to age 18 years with adjustment for cell composition—35 134 CpGs are stable with a threshold of $ICC > 0.5$. (E) Age 18 to early pregnancy with adjustment for cell composition—31 748 CpGs are stable with a threshold of $ICC > 0.5$. (F) Early (weeks 20-21) to late pregnancy (>week 22) with adjustment for cell composition—290 CpGs are stable with a threshold of $ICC > 0.5$. CpG indicates cytosine-phosphate-guanine; ICC, intraclass correlation coefficient.

in the prior time period tend to have low/high value in the subsequent time point as well). However, as the volcano plots indicate, the number of CpGs with coefficient > 1 (increase in methylation from the first time point to the second) after the puberty transition is reduced to nearly 0 in the 2 subsequent transition periods. Hence, an increase in DNA-m can be observed in puberty transition but not later in life. However, in late pregnancy, fewer CpGs also have a significantly reduced methylation when comparing 2 paired levels.

Both, low consistency and low number of CpGs with significant changes, resulting in a large proportion of indifferent CpGs in late pregnancy, raised our concern that fetal DNA could be present in maternal blood, a phenomenon called microchimerism.²¹ Thus, we additionally investigated 7 Y chromosome-related CpGs out of 6 genes (excluding cross-reactive probes and polymorphic CpGs, intergenic CpGs, and genes with pseudoautosomal regions).²² We compared the levels of these Y chromosome-related CpGs in women with male and female offspring. However, we did not find any increased

methylation levels in late pregnancy for cg02233183 ($P = .65$), cg03278611 ($P = .10$), cg05782707 ($P = .67$), cg17741448 ($P = .50$), cg20474581 ($P = .64$), cg26251715 ($P = .65$), and cg26475999 ($P = .30$) (representing the following genes *NLGN4Y* (twice), *TTTY2*, *RPS4Y2*, *USP9Y*, *TTTY14*, *TSPY4*). This suggests that the indifference of CpGs in late pregnancy (neither consistent nor changeable) cannot be explained by contamination of maternal DNA due to fetal-maternal trafficking.

Regarding the proportion of CpG with significant changes, Table 1 provides information using 2 different assessments: with and without cell type adjustment. A possible limitation is that we used estimated cell type techniques in our analyses because cell sorting or single-cell methylation was not conducted in this epidemiologic cohort. However, a recent investigation has demonstrated—even using simulated cell-specific data—that with appropriate cell type references, a cell type adjustment can provide unbiased estimates of DNA-m.²³ Thus, to overcome this limitation, we focused on cell type-adjusted estimates.

Table 1. Number of CpGs with significant regression coefficients (consistency) and ICC > 0.5 (stability) in 2 subsequent methylation measurements consistency.

	Transition From Age 10 to Age 18 Years			Transition From Age 18 to Early Pregnancy (≤ 21 Weeks)			Transition From Early Pregnancy to Late Pregnancy (> 21 Weeks)			
	Regression coefficients: FDR $P < .05$	Total no. of CpGs	Regression coefficients: FDR $P \geq .05$	Regression coefficients: FDR $P < .05$	Total no. of CpGs	Regression coefficients: FDR $P \geq .05$	Regression coefficients: FDR $P < .05$	Total no. of CpGs	Regression coefficients: FDR $P \geq .05$	Total no. of CpGs
<i>With cell type adjustment</i>										
Stability and consistency of methylation over time										
ICC > 0.5 (No. of CpGs)	35134	0	31748	31748	0	31748	290	290	0	290
ICC ≤ 0.5 (No. of CpGs)	8285	174803	14419	172055	186474	487	217445	217932	217445	217932
Total No. of CpGs	43419	174803	46167	172055	218222	777	217445	218222	217445	218222
No. of positive coefficients	42630	—	45660	—	—	766	—	—	—	—
Consistency ^a	19.5% (42630/218222)		20.9% (45660/218222)		0.35% (766/218222)					
Stability ^b	16.1% (35134/218222)		14.5% (31748/218222)		0.13% (290/218222)					
Changes of methylation over time										
CpGs with (FDR P value < .05)		1433		12308		0				
Percentage of change in CpGs, %		0.7 (1433/218222)		5.6 (12308/218222)		0				
<i>Without cell type adjustment</i>										
Stability and consistency of methylation over time										
ICC > 0.5 (No. of CpGs)	32916	0	29832	29832	0	29832	6626	6626	0	6626
ICC ≤ 0.5 (No. of CpGs)	19679	165627	21845	166545	188390	10711	200885	211596	200885	211596
Total No. of CpGs	52595	165627	51677	166545	218222	17337	200885	218222	200885	218222
No. of positive coefficients	52051	—	51194	—	—	17150	—	—	—	—
Consistency ^a	23.9% (52051/218222)		23.5% (51194/218222)		7.9% (17150/218222)					
Stability ^b	15.1% (32916/218222)		13.7% (29832/218222)		3% (6626/218222)					
Changes of methylation over time										
CpGs with (FDR P value < .05)		59483		38996		4				
Percentage of change in CpGs, %		27.3 (59483/218222)		17.9 (38996/218222)		0.002 (4/218222)				

Abbreviations: CpG, cytosine-phosphate-guanine; FDR, false discovery rate; ICC, intraclass correlation coefficient.

^aConsistency is measured by the percentage of CpGs with positive significant (FDR $P < .05$) regression coefficient out of 218222 CpGs.^bStability is measured by the percentage of CpGs with ICC > 0.5 out of 218222 CpGs.

Although the period between 10 and 18 years includes pubertal transition, the proportion of CpGs with significant changes in DNA-m between age 10 and 18 years in women is smaller than the proportion of consistent CpGs (0.7% vs 19.5%). Between age 18 and early pregnancy, the proportion of CpG with consistent methylation is similar (20.9%), but the proportion of CpGs with significant methylation changes increased to 5.6%. During gestation, contrasting early and late pregnancy, both the number of consistent and significantly changed CpGs is very small (0.35% and 0%, respectively). Focusing on methylation after adjustment for cell type, in all periods, the methylation of most of the CpGs (79.8% from 10 to 18 years, 73.5% from age 18 to early pregnancy, and 99.65% during gestation) is neither consistent nor significantly changed over time indicating an indifferent state. The dramatically increased number of these indifferent CpGs in late pregnancy may be due to mechanisms that are unique to the pregnant status and needs further investigations.

In a prior investigation of 92 participants of a British cohort, 35 to 84 years of age, the methylation of 353 633 CpGs were measured 6 years apart and adjusted for cell type.⁴ Stability was found in 17% of the CpGs. This proportion is low but still comparable with the number of stable (ICC > 0.5) CpGs of 16.1% between 10 and 18 years of age and also comparable with the transition between 18 years and early pregnancy (14.5%) in our study. Nevertheless, the methylation of specific CpGs can change during puberty and due to conception, which requires additional analyses. However, in general, gestational transition more severely suffers from low stability problem in contrast to other periods.

These findings cannot be explained by differences in the paired sample size in the 3 transition periods because the gestational transition has the largest sample size ($n = 63$ compared with $n = 34$ and $n = 41$) and thus provides statistically more powerful tests. Interestingly, on comparing the stability and consistency with and without adjustment for the changes in cell composition, we found that between age 10 and 18 years and between 18 years and early pregnancy (Supplemental Table 1), adjustment resulted only in minor improved methylation stability during these transitions and minimally reduced consistency. In contrast, in the gestational transition (early to late pregnancy), for which the cell type composition was not significantly different, adjustment for cell type reduced both stability and consistency.

Results from a Danish study identified statistically significant changes in CpG methylation in 43 paired measurements 10 years apart (age range at first measurement: 73–82 years).⁵ The investigators removed a smaller number of CpGs affected by probe-SNPs resulting in a total of 424 706 CpGs that were analyzed. Their results showed that 0.054% (2284/424 706) had significant changes. The proportion of change in our sample of women is larger in the transition from 10 to 18 years and from 18 years to early pregnancy, but comparable with changes in the few months of pregnancy.

Interestingly, genes of significantly changed CpG methylation in the transition between age 18 and pregnancy belong to 5 signaling pathways which control various processes such as cell adhesion and cell division that are crucial events occurring in the trophoblast and uterus (Rap1 signaling pathway, MAPK signaling pathway, Ras signaling pathway, calcium signaling pathway, and Hippo signaling pathway; Supplemental Table 2).^{24–27} In addition, the “Endocytosis” pathway seems also to be linked to pregnancy indicating a survival phenomenon in response to environmental stressors in the placenta.²⁸ We likewise assume that the identification of 2 cancer pathways in this transition is due to a misclassification because tumorigenesis and pregnancy share characteristics of cell survival, cell proliferation, and cell migration. It is surprising that CpGs of the genes belonging to signaling pathways involved in communication between uterus, placenta, and the trophoblast can be detected to be differentially methylated in maternal peripheral blood cells once pregnancy was established.

The extreme reduction in both stability and consistency of DNA-m during gestation may result from endocrine effects, adjustment in lifestyle, or late environmental exposures occurring in the course of pregnancy. Because this is the first investigation of these transition periods in women, there is a necessity that future studies better characterize these periods, particularly gestation. However, independently, these results raise some doubts about whether DNA-m in late pregnancy should be used to estimate inheritance of DNA-m between mother and offspring.

Finally, there is a need to analyze whether the methylation of CpGs in the postpregnancy period returns to prepregnancy levels or partially remains in an altered state after pregnancy. This is of importance because in women, pregnancy is a transition period that is related to multiple changes of health risks. Pregnancy and parity seem to change the risk of a number of diseases including autoimmunity, multiple sclerosis, ovarian, breast, and liver cancer in women,^{29–33} which may be explained and prevented by altering DNA-m and pregnancy-related conditions.

Author Contributions

Conceived the study: SC and WK. Designed the original cohort from which the follow-up data were collected: SHA, RJK, JWH and WK. Directed the methods for recruitment and collection of cord blood samples: SHA, RJK, JWH, and WK. Directed the DNA methylation assays: JWH and HZ. Prepared data sets and performed preliminary analyses: SC, NM and VDJ. Performed final analyses and led the writing of the paper with input from SC, NM, VDJ and WK. All authors reviewed and approved of the final manuscript.

REFERENCES

1. Bakulski KM, Fallin MD. Epigenetic epidemiology: promises for public health research. *Environ Mol Mutagen*. 2014;55:171–183.

2. Bjornsson HT, Sigurdsson MI, Fallin MD, et al. Intra-individual change over time in DNA methylation with familial clustering. *JAMA*. 2008;299:2877–2883.
3. 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.
4. 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.
5. Tan Q, Heijmans BT, Hjelmborg JV, Soerensen M, Christensen K, Christiansen L. Epigenetic drift in the aging genome: a ten-year follow-up in an elderly twin cohort. *Int J Epidemiol*. 2016;45:1146–1158.
6. Xu CJ, Bonder MJ, Soderhall C, et al. The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics*. 2017;18:25.
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. Freire-Aradas A, Phillips C, Mosquera-Miguel A, et al. Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. *Forensic Sci Int Genet*. 2016;24:65–74.
9. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14:R115.
10. Demetriou CA, Chen J, Polidoro S, et al. Methylome analysis and epigenetic changes associated with menarcheal age. *PLoS ONE*. 2013;8:e79391.
11. Almstrup K, Lindhardt Johansen M, Busch AS, et al. Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood. *Sci Rep*. 2016;6:28657.
12. Wang D, Yan L, Hu Q, et al. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. *Bioinformatics*. 2012;28:729–730.
13. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8:118–127.
14. Kuan PF, Wang S, Zhou X, Chu H. A statistical framework for Illumina DNA methylation arrays. *Bioinformatics*. 2010;26:2849–2855.
15. 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.
16. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86.
17. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol*. 2014;15:R31.
18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B Meth*. 1995;57:289–300.
19. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;37:W305–W311.
20. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001;29:1165–1188.
21. Jeanty C, Derderian SC, Mackenzie TC. Maternal-fetal cellular trafficking: clinical implications and consequences. *Curr Opin Pediatr*. 2014;26:377–382.
22. Chen YA, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8:203–209.
23. Kaushal A, Zhang H, Karmaus WJJ, et al. Comparison of different cell type correction methods for genome-scale epigenetics studies. *BMC Bioinformatics*. 2017;18:216.
24. Fritz R, Jain C, Armant DR. Cell signaling in trophoblast-uterine communication. *Int J Dev Biol*. 2014;58:261–271.
25. O'Neill C, Li Y, Jin XL. Survival signaling in the preimplantation embryo. *Theriogenology*. 2012;77:773–784.
26. Lorthongpanich C, Issaragrisil S. Emerging role of the hippo signaling pathway in position sensing and lineage specification in mammalian preimplantation embryos. *Biol Reprod*. 2015;92:143.
27. Kusama K, Yoshie M, Tamura K, Daikoku T, Takarada T, Tachikawa E. Possible roles of the cAMP-mediators EPAC and RAP1 in decidualization of rat uterus. *Reproduction*. 2014;147:897–906.
28. Bildirici I, Longtine MS, Chen B, Nelson DM. Survival by self-destruction: a role for autophagy in the placenta? *Placenta*. 2012;33:591–598.
29. Parsa P, Parsa B. Effects of reproductive factors on risk of breast cancer: a literature review. *Asian Pac J Cancer Prev*. 2009;10:545–550.
30. Borchers AT, Naguwa SM, Keen CL, Gershwin ME. The implications of autoimmunity and pregnancy. *J Autoimmun*. 2010;34:J287–J299.
31. D'Hooghe MB, D'Hooghe T, De Keyser J. Female gender and reproductive factors affecting risk, relapses and progression in multiple sclerosis. *Gynecol Obstet Invest*. 2013;75:73–84.
32. Zhong GC, Liu Y, Chen N, et al. Reproductive factors, menopausal hormone therapies and primary liver cancer risk: a systematic review and dose-response meta-analysis of observational studies. *Hum Reprod Update*. 2016;23:126–138.
33. Salehi F, Dunfield L, Phillips KP, Krewski D, Vanderhyden BC. Risk factors for ovarian cancer: an overview with emphasis on hormonal factors. *J Toxicol Environ Health B Crit Rev*. 2008;11:301–321.