Association of grandmaternal smoking during pregnancy with grandchildren DNA methylation: The Isle of Wight Study

Rui Luo1, Hongmei Zhang1, Nandini Mukherjee1, Wilfried Karmaus1, Veeresh Patil3, Hasan Arshad2,3, and Fawaz Mzayek\*,1

1 Division of Epidemiology, Biostatistics and Environmental Health, School of Public Health, University of Memphis, Memphis, USA

2 The David Hide Asthma and Allergy Research Centre, Newport, UK

3 Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK

\***Author for correspondence:** Tel: +1-(901) -678-1662; Fax: +1-(901)-678-1715; fmzayek@memphis.edu

**Abstract**

**Background:** To investigate intergenerational effects of grandmaternal smoking during pregnancy (GMSDP) on grandchildren's DNA methylation (DNAm).

**Methods:** Data from the Isle of Wight birth cohort with information of grandmaternal smoking during pregnancy and DNAm profiling at birth of grandchildren (n=161) were used. Differentially methylated cytosine-guanine dinucleotides (CpG) sites related to GMSDP were identified using a testing-training screening, ANOVA, and MANCOVA. The association between identified CpG sites and expression levels of neighboring genes was tested by linear regression.

**Results:** Twenty-three CpGs in grandchildren were differentially methylated due to GMSDP, and eight of them were associated with expression levels of 13 neighboring genes.

**Conclusions:** GMSDP has an intergenerational effect on DNAm profile of grandchildren independent from maternal smoking during pregnancy.

**Keywords:** DNA methylation; grandmaternal smoking during pregnancy; grandchildren; intergenerational effect; smoke

**Lay abstract**

How grandmaternal smoking during pregnancy affect their grandchildren? Underlying mechanisms my included epigenetic modifications. In order to address this question, we investigate intergenerational effects of grandmaternal smoking during pregnancy (GMSDP) on grandchildren’s DNA methylation at birth based on Isle of Wight birth cohort. Twenty-three cytosine-guanine dinucleotides sites (CpGs) in grandchildren were differentially methylated due to GMSDP, and eight of them were associated with expression levels of 13 neighboring genes. Thus, GMSDP has an intergenerational effect on DNAm profile of grandchildren independent from maternal smoking during pregnancy.

**INTRODUCTION**

Maternal smoking during pregnancy (MSDP) carries adverse health effects for the mother and her child and is a significant public health concern around the world [1-4]. Previous reports showed that MSDP is a risk factor for several adverse perinatal outcomes, including compromised lung function [5], low birth weight [6-8], and preterm birth [9, 10], as well as increased susceptibility to various diseases later in life, including childhood asthma [11, 12] and obesity [13, 14]. However, very few studies have examined the intergenerational effects of grandmaternal smoking during pregnancy (GMSDP) on grandchildren's health. Specifically, little is known about the epigenetic mechanisms potentially mediating these effects.

The evidence from animal and human studies points to an intergenerational effect of prenatal exposures on grandchildren's health outcomes. One study in rats reported emphysematous, and other structural changes in the lungs of the offspring after their pregnant grandmothers were exposed to nicotine [15]. Of four population-based studies that explored the relationship between GMSDP and asthma in the grandchild, three have found that GMSDP increased the risk of asthma in the grandchildren independent of the smoking status of the mother [16-18]. However, a study from the Avon Longitudinal Study of Parents and Children (ALSPAC) found no effect of GMSDP on grandchildren's respiratory outcomes [19].

The potential mechanisms underlying the observed adverse health effects of GMSDP on grandchildren's health are poorly understood. Current evidence suggests that epigenetic mechanisms, such as DNA methylation (DNAm), may mediate the effect of MSDP on children's health outcomes. MSDP-induced alterations in DNAm were observed in the placenta, umbilical cord blood, and buccal cells of offspring [20, 21]. Besides, a dose-response effect of MSDP on the methylation levels of 15 cytosine-guanine dinucleotides (CpG) sites, located in seven gene regions, has been reported before [22]. It has also been shown that prenatal exposure to tobacco smoke can affect germline reprogramming during embryonic development and result in increased susceptibility to disease during adult life [12]. Together these pieces of evidence raise the question of whether the underlying mechanisms for the effects of GMSDP on grandchildren's health may also involve epigenetic mechanisms such as DNA methylation.

In this study, we examined the associations between GMSDP and grandchildren's DNAm profile using an epigenome-wide association study (EWAS) approach and utilizing data from three generations from the Isle of Wight Birth Cohort (IOWBC).

**MATERIAL AND SUBJECTS**

**The Isle of Wight Birth Cohort and the Study Population**

Participants in this study are a subgroup of the ongoing Isle of Wight Whole Population Birth Cohort [23, 24]. This cohort was initially established in 1989 in the United Kingdom to observe the natural history of allergic conditions. 1518 pregnant women (F0 generation) were approached at the time of delivery during 1989-1990 and invited to enroll in the study, along with all their newborns (F1 generation). Of those, 1456 mother-child pairs (94.8%) consented [25]. Between 2011-2018, 231 F1 daughters became pregnant and gave birth to grandchildren (F2 generation, n = 405). This study utilizes information from 161 F2 grandchildren with complete data on their DNAm and their grandmother (F0) and mother (F1) smoking behavior. Exclusion criteria include severe clinical disease of the mother (F1), such as Type I and Type II diabetes, cancer (leukemia, lymphoma), hypertension, preeclampsia, or severe cardiovascular disease. For mothers with data available on multiple children, only the first child is included in the analysis (Figure 1).

**Exposure Assessment**

Information about grandmaternal (F0) prenatal smoking status was determined by self-report and serum cotinine levels at the end of pregnancy. A grandmother who had a serum cotinine level greater than ten µg/L was classified as a smoker [26]. Mothers' (F1) prenatal smoking was measured by self-report at 20 weeks and 28 weeks of pregnancy, and by urine cotinine measured at each of the two aforementioned times. Participants with urine cotinine levels equal to, or greater than 2.47 ng/ml were classified as smokers [27]. For both the grandmothers and the mothers non-smoker status was defined as reporting no smoking during pregnancy and had a negative serum or urine cotinine test, respectively. Unfortunately, we did not have information regarding whether grandmother or mother quitted smoking during pregnancy.

**Assessment of DNAm in the grandchildren**

Grandchildren’s DNAm data, which were measured in umbilical cord blood (n = 106) or Guthrie card (n = 55), were utilized in this study. High agreement between CpGs DNAm from cord blood and neonatal blood on Guthrie cards has been reported before, making Guthrie cards an acceptable source of neonatal blood [28]. Umbilical cord blood samples of the grandchildren were collected after delivery and before the umbilical cord was cut. Dried blood spots on Guthrie cards were collected from heel prick within 5-8 days after birth as part of the UK Newborn Screening Program. Guthrie cards were stored in the dark at room temperature in individual paper protectors.

The umbilical cord blood samples were centrifuged at 3,000 rpm to separate blood cells from plasma and stored at -80 ºC. And then, DNA was extracted by using standard salting out procedure [29] through QIAamp DNA Blood Midi kit (QIAGEN, Crawley, UK) according to manufacturer's guidelines. The DNA from dried blood spots on Guthrie cards was extracted using the procedure described by Beyan et al [30]. The samples were purified using the Qiagen QiaAmp mini columns (QIAGEN, Crawley, UK) and eluted in 100 μl of elution buffer. After that, Qubit instrument (Invitrogen) was used to measure DNA concentration.

Samples were randomized according to their study IDs when placed on plates. One microgram of DNA from each subject underwent cytosine to thymine conversion with sodium bisulfite by EZ-96 DNAm kit (Zymo Research, Irvine, CA, USA). Genome-wide DNAm was measured using the Illumina Infinium Human Methylation450 Beadchip (Illumina, Inc., San Diego, CA, USA). Beadchips were scanned and processed using a standard protocol based on the random self-assembly of a bead pool onto a patterned substrate [31]. β-values were extracted from image data files using the methylation module of Genome Studio software (Illumina, Inc., San Diego, CA, USA). The β-values (percent methylation) represent the proportion of methylated (M) over the sum of methylated and unmethylated (U) allele intensities (β=M/[M+U+c]), where c indicates a constant to prevent dividing by zero [32]. To adjust for different cell type fractions, we estimated the proportion of cell types in cord blood and Guthrie card using the *estimateCellCounts* function in *minfi* package following the Bakulski approach using cord blood reference panel [33] and Houseman approach [34] using adult reference panel [35], respectively.

**DNA methylation data preprocessing**

To improve the DNAm data quality, the Bioconductor IMA (Illumina methylation analyzer) package was used for quantile normalization and type I and type II probe peak correction [36]. Furthermore, ComBat package was used for removing batch effects and other unwanted variations [37]. Additionally, CpG sites with probe single-nucleotide polymorphisms (SNPs) were removed from the list of all the CpG sites if minor allele frequency (MAF) of probe SNPs was larger than 1%. Probe SNPs are those SNPs located within the probes sequence and can cause false interpretation of Illumina microarray results [38].

**Gene expression measurement**

Gene expression data were available for a subsample of 87 grandchildren with complete GMSDP and DNAm data. Total RNA was extracted from umbilical cord blood according to the manufacturer's instructions (Qiagen, Valencia, CA, USA). Samples were randomized according to their study IDs when placed on plates. Total RNA yield and the absence of DNA contamination were measured by Qubit 2.0 fluorometer (Life Technologies, Grand Island, NY, USA). The RNA quality was evaluated by a Bioanalyzer 2100 via RNA 6000 Nano Chips (Agilent Technologies, Santa Clara, CA, USA). RNA was reverse-transcribed into complementary DNA (cDNA). cDNA was subsequently used as a template for DNA amplification by real-time PCR using commercially available primer and probes and following Cy5-labeled by Agilent's Single-Color Microarray-Based Gene Expression Analysis protocol version 6.0 (Agilent Technologies, Santa Clara, CA, USA). For each gene, multiple primers were designed and created, which can match sense and antisense complementary DNA. Next, sample cDNAs and reference cDNA were hybridized separately on SurePrint G3 Human GE v2 8x60K Agilent Microarray (Agilent Technologies, Santa Clara, CA, USA). After hybridization, slides were washed and scanned using Agilent's G2565AA Microarray Scanner System (Agilent Technologies, Santa Clara, CA, USA). Scanned images were subsequently analyzed based on fluorescence intensities. Quantitative expression levels were reported after correction for background signal and normalization through Agilent's Feature Extraction Software version 9.5.

**Measurement of covariates and potential confounders**

Exposure to passive smoking during pregnancy of the F1 mothers was ascertained by a questionnaire that was administered to both parents and was defined as the occurrence of one of the following: 1) an F1 father, who lived in the same residence with the mother, reported smoking during her pregnancy; 2) other people in the household smoked during the F1 mother pregnancy, or 3) the mother was regularly exposed to secondhand smoking at places other than home, such as in the workplace. The mother's SES was assessed by a composite variable based on her education level, income, and the number of rooms of the house and was categorized into four categories [39]. Information of gestational age and the gender of the grandchild were also collected. Age at the conception of the F1 mothers was not considered as a confounder because of the narrow range of F1 mothers’ age at their first pregnancy.

**Statistical Analysis**

The comparison of main characteristics between the study samples and the whole cohort was made using *t* test and chi-square test.

The grandchildren were classified into four exposure groups according to the different combinations of smoking statuses of the grandmother and the mother: both the grandmother and the mother smoked during pregnancy (yes/yes group), only the grandmother smoked during pregnancy (yes/no group), only the mother smoked during pregnancy (no/yes group), and neither smoked during pregnancy (no/no group). And the difference between four groups in response to main characteristics were determined by ANOVA and chi-square test, respectively.

We applied a training-testing screening (tt-screening) method to the cell proportion-adjusted beta-values to identify potential CpG sites in the grandchildren's blood that are differentially methylated in relation to their GMSDP/MSDP exposure status [40]. Briefly, tt-screening is an efficient algorithm that uses a randomly-selected two-thirds of the data for training and the other one-third for testing. This approach has been shown to have better control on type I and II errors compared to the commonly used approaches for multiple testing correction [40]. CpG sites detected in at least 60% of the training-testing pairs (100 randomly selected pairs in total) were considered as being potentially associated with GMSDP/MSDP and were used in subsequent analyses.

Analysis of Variance (ANOVA) was used to assess the differences in DNAm levels across the four exposure groups (yes/yes, yes/no, no/yes, and no/no). To tease out the independent effect of GMSDP from that of MSDP, methylation levels among those who were exposed only to GMSDP (the yes/no group) were compared to the unexposed (no/no) group using Tukey's HSD test.

 Additionally, we compared the direction and magnitude of DNAm changes between the grandchildren who were exposed to both GMSDP and MSDP (the yes/yes group) and those exposed to GMSDP alone (the yes/no group) in relation to the unexposed (no/no) group. Only the CpG sites with DNAm changes in the same direction and at least the same magnitude in response to GMSDP and MSDP were selected for further analysis. The reason is because to conclude an effect of exposure to GMSDP only on DNAm of the grandchildren, the effect of exposure to both GMSDP *and* MSDP should be in the same direction and at least as big.

Multivariable analysis of covariance (MANCOVA) was used to adjust for potential confounders. Cell-type adjusted DNAm level (beta-value) was the dependent variable, and GMSDP exposure status, mother's (F1) exposure to secondhand smoking, the gender of the grandchild, gestational age of the grandchild, SES of the mother, and source of the grandchild's blood (Guthrie card or cord blood) were included as independent variables. The magnitude of DNAm changes was evaluated between GMSDP-exposure only group (the yes/no group) and the unexposed (no/no) group. FDR was used to control multiple testing with a cut-off *p*-value of 0.05.

GMSDP-associated CpG sites were explored for biological relevance by assessing their association with gene expression of the neighboring genes (250 kb upstream and downstream of each CpG) [41]. Corresponding annotated genes were identified from methylation label file (Infinium MethylationEPIC v1.0 B4 Manifest File), SNIPPER (<https://csg.sph.umich.edu/boehnke/snipper/>) [42] and the University of California Santa Cruz (UCSC) Genome Browser (<https://genome.ucsc.edu/>) [43]. Linear regression models were run with gene expression level as the dependent variables and DNAm levels (beta-value), gender, source of blood, maternal smoking during pregnancy, secondhand smoking during pregnancy, and gestational age as the independent variables.

The statistical analyses were performed by SAS 9.4 (SAS Institute, Cary NC) and R 3.4.4.

**RESULTS**

**Baseline characteristics**

Of the 161 eligible grandmother-mother-grandchild triads, 34 (21.1%) were exposed to both GMSDP and MSDP, 45 (30.0%) were exposed to GMSDP only, 29 (18.0%) were exposed to MSDP only, and 53 (32.9%) were unexposed to both (Table 1). The average age of F1 mothers at the time of conception was 23.7 years old (range: 18-28). The comparison of main characteristics between analytical samples and the whole cohort revealed that there was no significant difference in gender distribution (F2), birthweight (F2), exposure to passive smoking during pregnancy (F1) and socioeconomic status of the F1 mother. However, the proportions of participants with smoking grandmothers and smoking mothers were higher among the study sample (Supplementary table 1).

**EWAS for differential DNAm associated with grandmaternal smoking during pregnancy**

Out of 1,664 CpG sites selected by tt-screening as differentially methylated in association with GMSDP and/or MSDP, 27 were still significant after adjusting for multiple testing (Table 2). Of these 27 CpG sites, 25 were significantly differentially methylated in the same direction (i.e., hyper- or hypomethylated) in grandchildren exposed to either GMSDP only or to GMSDP + MSDP (ten CpGs were more methylated, and 15 were less methylated). Figure 2 shows two examples of such CpG sites, one with increased and one of decreased methylation due to GMSDP and/or MSDP. After adjusting for potential confounders, 23 CpG sites remained significantly associated with GMSDP (Table 2).

 **Association of GMSDP-induced DNAm alterations with gene expression**

There were 87 grandchildren with available data on DNAm and gene expression, of which 33 (37.9%) were exposed to GMSDP. The 23 differentially methylated CpG sites were mapped to 256 genes (using a proximity range of +/- 250 kb). DNAm levels among eight significantly differentially methylated CpG sites were associated with the expression levels of 13 nearby genes (Supplementary Table 2, Table 3).

**DISCUSSION**

Very little is known about the intergenerational effects of GMSDP on grandchildren's health outcomes. Using three-generation data from IOWBC, we examined whether GMSDP is associated with changes in the epigenome of the grandchildren. We identified 23 CpG sites that were differentially methylated in association with GMSDP, of which eight were associated with neighboring gene expression.

These findings indicate that GMSDP has lasting effects that persist across generations. We compared our results with studies focused on effect of MSDP and DNA methylation, which has have identified over 6,000 differentially methylated CpG sites [44, 45]. Interestingly, none of the 23 CpGs identified in this study as associated with GMSDP (Table 2) was reported in studies of MSDP. However, when we compared DNAm in grandchildren exposed to both GMSDP and MSDP with the unexposed grandchildren in our sample, we found 12 CpG sites that were differentially methylated (Supplementary table 3) that were located on three genes repeatedly reported to be associated with MSDP: *MYO1G*, *AHRR*, and *GFI1*. Together, these findings indicate that GMSDP and MSDP affect different epigenetic locations, which may translate into different epigenetic effects. The mechanism of the intergenerational epigenetic transmission remains unclear, since there is no evidence to date to support direct germline transmission of DNA methylation from mother to offspring [46]. One possible way to explain our findings is that GMSDP exerts a direct effect on F2 gametes located in the developing ovaries of F1 fetus [47, 48]. [44]

Among the 23 CpGs that are associated with GMSDP, eight were located in CpG islands, four in N shore, three in N shelf, and one in S shore (Table 2). The methylation level of these sites could have a functional impact on nearby genes, some of which are involved in important biological processes [49]. For example, *MYOM2*  and *MYOM3* genes are expressed in skeletal and cardiac muscles [50], While *IGF1R* and *IGFALS* genes are essential for normal human growth [51-53] and central nervous system development [54, 55]. Mutations in these two genes have been related to short stature, delayed puberty, and growth retardation [56, 57].

 Moreover, methylation levels at eight CpG sites were associated with nearby gene expression levels (Table 3), which indicate a potential functional implication from GMSDP. Among these genes, *CD33* gene encodes a sialic acid-binding immunoglobulin-like lectin that is involved in the regulation of leukocyte functions during an inflammatory and immune response [58, 59]. Additionally, one study suggested that *CHD1L* gene plays a vital role in nervous system development and neuronal differentiation [60], as well as kidney and urinary tract development [61]. The *ELMOD3* gene was associated with hearing impairment [62] and autism spectrum disorder [63]. *FMO5* gene is a key regulator of metabolic aging [64], Absence of *FMO5* protein was associated with high glucose tolerance and insulin sensitivity [65].

Our results support an intergenerational effect of GMSDP on grandchildren's DNAm, which, in turn, may affect the health of the grandchildren. Of note that the novel CpG sites identified in this study were not reported in previous studies that assessed the association of MSDP with DNAm of the newborn. This implies that epigenetic mechanisms involved in MDSP and GMSDP may be different. Finally, these results provide evidence of the potential physiological effects of these changes through gene function changes.

There are several methodological strengths in our study. To directly observe the intergenerational effect of GMSDP, we utilized three-generation data from IOWBC. By comparing grandchildren exposed to different combinations of GMSDP and MSDP exposures were able to identify the pure effect of GMSDP on grandchildren's DNAm. Also, we used state-of-the-art methodology for exposure and outcome assessments resulting in high-quality data.

The study has several limitations, including the small sample size, which may have reduced the power of the study, and the failure to find another available cohort for replication. Also, only the effect of the maternal grandmother smoking on the grandchildren's DNAm was assessed. It is important to examine the effect of paternal grandmother smoking, as well, to have a complete picture of the intergenerational effect of GMSDP. Also, we analyzed DNAm from both cord blood and Guthrie cards. However, the comparability in DNAm between umbilical cord blood and Guthrie cards has been examined in our data based on 34 subjects in the Isle of Wight 3rd Generation Birth Cohort, which showed that around 70% of the CpGs agreed in DNAm mean between cord blood and Guthrie cards [28]. Additionally, the source of DNAm (cord blood or Guthrie card) was included as a covariate in the multivariable model and was not associated with DNA methylation levels. Although grandmaternal smoking behavior was assessed by self-report and validated by serum cotinine levels measured at the end of pregnancy, misclassification could still be a possibility. Smoking cessation in early pregnancy may negate effects on methylation at birth [66]. Unfortunately, we were unable to observe whether it would be the same of the effect from grandmaternal smoking cessation during pregnancy on grandchildren DNA methylation due to lack of cessation information. Finally, loss to follow-up could be another limitation considering IOWBC included three generations, but there is no evidence to suggest that the loss to follow-up was not at random.

**Conclusion**

This study identified novel loci differentially methylated in newborn grandchildren associated with GMSDP independent of MSDP. To our knowledge, this is the first study assessing the intergenerational effect of grandmaternal smoking during pregnancy on DNAm patterns of their newborn grandchildren. Also, the study provides preliminary evidence of the potential health effects of GMSDP on grandchildren's health by affecting the function of several important genes.

[https://csg.sph.umich.edu/ boehnke/snipper/](https://csg.sph.umich.edu/%20boehnke/snipper/)

**Future perspective**

Although we demonstrated a potential biological effect of GMSDP on several CpG sites in the blood of the grandchild, the nature and extent of these effects remain largely unknown. It remains unknown whether the differential levels due to GMSDP are biologically and clinically relevant. Future studies are needed to examine whether these changes are related to clinical outcomes in the grandchildren.

**Summary points**

* We identified an intergenerational effect of grandmaternal smoking during pregnancy on DNAm profile of grandchildren at birth, independent from maternal smoking during pregnancy.
* tt-screening method was performed to identify the differentially methylated CpGs in association with GMSDP and/or MSDP in the peripheral blood of the grandchildren.
* A total of 23 CpG sites were identified as differentially methylated in association with GMSDP in the newborn after adjusting for confounding and multiple testing.
* DNAm levels of eight significantly differentially methylated CpG sites were associated with the expression levels of 13 nearby genes, which suggest a functional effect of GMSDP on grandchildren’s health
* Some of those 13 genes have important functional implications, including *CD33*, *CHD1L*, and *ELMOD3* genes that are related to leukocyte function regulation, nervous system development, as well as kidney development.
* The identified CpG sites could potentially serve as blood biomarkers of intergenerational smoking exposure.

**Table 1.** Characteristics of grandmother-mother-grandchild triads in IOWBC

|  |  |
| --- | --- |
|   | GMSDP/MSDP Groups |
|   | Yes/yes (n=34) | Yes/no (n=29) | No/yes (n=45) | No/no (n=53) |
| Gender n (%) |  |  |  |  |
| Male | 17 (50.0) | 20 (44.4) | 21 (72.4) | 32 (60.4) |
| Female | 17 (50.0) | 25 (55.6) | 8 (27.6) | 21 (39.6) |
| Birth Weight (mean, SD, gram)b | 3294.8 ± 515.8 | 3315.5 ± 617.7 | 3641.5 ± 414.0 | 3427.6 ± 476.7 |
| Gestational age (week)b | 38.8 ± 2.3 | 39.0 ± 2.2 | 40.2 ± 1.2 | 39.6 ± 1.4 |
| Blood Source n (%) |  |  |  |  |
| Cord blood | 21 (61.8) | 30 (66.7) | 18 (62.1) | 37 (69.8) |
| Guthrie card | 13 (38.2) | 15 (33.3) | 11 (37.9) | 16 (30.2) |
| Exposure to passive Smoking n (%)b |  |  |
| No | 5 (15.2) | 6 (13.3) | 11 (37.9) | 37 (69.8) |
| Yes | 28 (84.9) | 39 (86.7) | 18 (62.1) | 16 (30.2) |
| Socioeconomic Status n (%) |  |  |  |
| 1 (lower) | 6 (18.8) | 8 (17.8) | 6 (20.7) | 9 (17.0) |
| 2 | 13 (40.6) | 20 (44.4) | 7 (24.1) | 10 (18.9) |
| 3 | 7 (21.9) | 6 (13.3) | 11 (37.9) | 12 (22.6) |
| 4 (higher) | 6 (18.8) | 11 (24.4) | 5 (17.2) | 22 (41.5) |

GMSDP = Grandmaternal smoking during pregnancy, MSDP = Maternal smoking during pregnancy

a displayed in mean ± standard deviation

b indicated significant association between variables and GMSDP and/or MSDP

**Table 2.** Differentially-methylated CpG sites associated with GMSDP (n = 161).

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CpG site | Chr | Position | Nearest Annotated genea | UCSC RefGene Group | Relation to CpG Island | Unadjusted model effect sizeb | *P*b | FDRb | Adjusted model effect sizec | *P*c | FDRc |
| cg01718065 | 19 | 51774429 | *SIGLECLd* |  | Island | 0.247 | <0.001 | 0.003 | 0.338 | <0.001 | 0.001 |
| cg02296145 | 16 | 1840711 | *IGFALS* | Body | Island | 0.290 | 0.001 | 0.005 | 0.354 | 0.001 | 0.004 |
| cg02399044 | 12 | 2500229 | *CACNA1C* | Body |  | -0.263 | <0.001 | 0.003 | -0.295 | 0.001 | 0.004 |
| cg02473254 | 8 | 2000469 | *MYOM2* | Body | N Shelf | -0.180 | <0.001 | 0.002 | -0.177 | 0.001 | 0.003 |
| cg02615735 | 6 | 31648276 | *LY6G5C* | TSS200 | N Shore | -0.260 | 0.003 | 0.007 | -0.244 | 0.014 | 0.020 |
| cg03807330 | 11 | 76327232 | *LRRC32 d* |  |  | -0.111 | 0.001 | 0.004 | -0.078 | 0.045 | 0.049 |
| cg06293294 | 11 | 100557954 | *FLJ32810* | TSS1500 | Island | 0.177 | 0.001 | 0.005 | 0.159 | 0.123 | 0.123 |
| cg08859138 | 5 | 74532894 | *ANKRD31* | TSS200 | Island | 0.460 | <0.001 | 0.003 | 0.455 | 0.005 | 0.010 |
| cg11596397 | 11 | 75526167 | *UVRAG* | TSS200 | N Shore | 0.138 | 0.002 | 0.005 | 0.142 | 0.038 | 0.045 |
| cg12732436 | 18 | 77752726 | *TXNL4A d* |  | S Shelf | -0.229 | 0.004 | 0.007 | -0.287 | 0.007 | 0.012 |
| cg12799739 | 3 | 97595056 | *CRYBG3* | TSS1500 |  | -0.093 | <0.001 | 0.003 | -0.112 | <0.001 | 0.001 |
| cg13749070 | 15 | 99451935 | *IGF1R* | Body |  | -0.235 | 0.001 | 0.005 | -0.199 | 0.025 | 0.033 |
| cg14284055 | 1 | 24439399 | *MYOM3* | TSS1500 |  | -0.233 | <0.001 | 0.003 | -0.193 | 0.002 | 0.004 |
| cg15497724 | 19 | 51774377 | *SIGLECL1 d* |  | N Shore | 0.176 | 0.001 | 0.005 | 0.218 | 0.001 | 0.004 |
| cg15517847 | 13 | 41241076 | *FOXO1* | TSS1500 | Island | -0.190 | 0.003 | 0.007 | -0.228 | 0.011 | 0.018 |
| cg16409505 | 14 | 73425357 | *DCAF4* | Body |  | -0.126 | <0.001 | 0.003 | -0.098 | 0.071 | 0.074 |
| cg20056908 | 2 | 85808945 | *VAMP8* | 3'UTR | N Shelf | -0.105 | 0.001 | 0.005 | -0.104 | 0.009 | 0.016 |
| cg23072559 | 19 | 55944231 | *SHISA7* | 3'UTR | N Shore | 0.104 | <0.001 | 0.003 | 0.100 | 0.014 | 0.020 |
| cg23243463 | 18 | 35146454 | *BRUNOL4* | TSS1500 | Island | -0.352 | 0.001 | 0.005 | -0.420 | 0.005 | 0.010 |
| cg23350744 | 5 | 122422098 | *PRDM6 d* |  | N Shelf | 0.279 | <0.001 | 0.003 | 0.272 | 0.006 | 0.012 |
| cg23670101 | 10 | 46168396 | *ANUBL1* | TSS200 | Island | 0.141 | <0.001 | 0.004 | 0.124 | 0.043 | 0.049 |
| cg24735611 | 1 | 146494928 | *LOC728989* | Body |  | -0.107 | <0.001 | 0.004 | -0.081 | 0.032 | 0.040 |
| cg25256538 | 6 | 159343074 | *C6orf99 d* |  |  | -0.188 | 0.001 | 0.004 | -0.201 | 0.002 | 0.004 |
| cg26961622 | 2 | 20426146 | *SDC1* | TSS1500 | S Shore | -0.155 | <0.001 | 0.002 | -0.159 | 0.002 | 0.004 |
| cg27106230 | 13 | 99229068 | *STK24* | 1stExon;5'UTR | Island | 0.184 | 0.002 | 0.005 | 0.292 | <0.001 | 0.002 |

Chr: Chromosome; FDR: False Discovery Rate; RefGene Group = Gene region feature category describing the CpG position

Effect size = difference in DNAm (cell-proportion adjusted beta-value) between grandchildren among those who were exposed to GMSDP only and those who were unexposed to neither GMSDP nor MSDP.

a only displayed information of the closest corresponding gene

b unadjusted *p*-value

c adjusted for maternal passive smoking during pregnancy, maternal SES, source of grandchild's blood, gender, and gestational age

d based on UCSC Genome Browser (hg 19)

**Table 3.** Association between DNAm of GMSDP-associated CpG sites and annotated gene expression level (n=87)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CpG sites | Nearest Annotated gene | UCSC RefGene Group | Relation to CpG Island | Gene Symbol | Probe Name | Estimate | Chi-square | p-value |
| cg02296145 | *IGFALS* | Body | Island | *GFER* | A\_24\_P31003 | 0.29 | 4.22 | 0.040 |
| cg02296145 | *IGFALS* | Body | Island | *IFT140* | A\_23\_P140725 | 0.28 | 4.80 | 0.029 |
| cg02615735 | *LY6G5C* | TSS200 | N Shore | *LY6G6F* | A\_33\_P3214334 | 0.88 | 7.20 | 0.007 |
| cg15497724 | *SIGLECL1* |  | N Shore | *CD33* | A\_21\_P0000094 | 0.73 | 6.11 | 0.013 |
| cg15497724 | *SIGLECL1* |  | N Shore | *SIGLEC12* | A\_23\_P164596 | 0.55 | 3.90 | 0.048 |
| cg20056908 | *VAMP8* | 3'UTR | N Shelf | *ELMOD3* | A\_33\_P3235204 | 1.36 | 7.07 | 0.008 |
| cg23072559 | *SHISA7* | 3'UTR | N Shore | *PPP6R1* | A\_23\_P119448 | -1.60 | 3.87 | 0.049 |
| cg23072559 | *SHISA7* | 3'UTR | N Shore | *SSC5D* | A\_21\_P0000127 | -1.26 | 5.80 | 0.016 |
| cg23350744 | *PRDM6* |  | N Shelf | *SNX24* | A\_33\_P3702364 | -0.40 | 5.52 | 0.019 |
| cg24735611 | *LOC728989* | Body |  | *CHD1L* | A\_23\_P45831 | 1.04 | 8.14 | 0.004 |
| cg24735611 | *LOC728989* | Body |  | *FMO5* | A\_33\_P3376214 | 1.27 | 5.46 | 0.019 |
| cg27106230 | *STK24* | 1stExon;5'UTR | Island | *FARP1* | A\_33\_P3238856 | -0.57 | 5.55 | 0.019 |
| cg27106230 | *STK24* | 1stExon;5'UTR | Island | *SLC15A1* | A\_33\_P3373273 | -0.75 | 6.58 | 0.010 |

RefGene Group = Gene region feature category describing the CpG position

**n= 161**

Participants from F2 generation with valid data on DNAm, grandmaternal and maternal smoking during pregnancy

**n= 299**

Participants from F2 generation with valid DNAm data

**n= 405**

Participants from F2 generation

**n= 106**

- Missing data on DNAm (n=106)

**n= 138**

- Siblings or a twin member (n=50).

- Missing data on grandmaternal and maternal smoking during pregnancy (n=72).

- Participants with severe maternal chronic disease (n=16): leukemia (n=2), diabetes (n=4) severe heart disease (n=2), Hopkins lymphoma (n=1), hypertension (n=8), and preeclampsia (n=2)\*.

**Figure 1.** Flowchart of participants.

\*Some participants may have more than one maternal chronic diseases



**Figure 2.** Relationship between GMSDP and/or MSDP and cell-proportion adjusted methylation levels (β-values) at cg24735611 and cg02296145.

**Author Contributions**

Luo R tested hypotheses, conducted the analysis, and drafted the manuscript. Zhang H contributed to the conception and interpretation of the data. Mukherjee N helped the data analysis and interpretation of results. Karmaus W contributed data interpretation and manuscript revision. Patil V was involved in manuscript revision. Arshad H was involved in sample collection, data acquisition, DNA methylation measurement in IOW cohort, and manuscript revision. Mzayek F designed the study, guided the analysis, and involved in drafting and revision of the manuscript prior to submission. The manuscript has been read and approved by all authors.

**Acknowledgments**

The authors gratefully acknowledge the valuable effort of all the staff at the David Hide Asthma and Allergy Research Centre in maintaining the Isle of Wight birth cohort. And we would like to thank the help of the team at the ITS Research Computing office at the University of Memphis to support high-performance computing. The authors are thankful to the participants and their families, who helped us over the last two decades.

**Financial and competing interest disclosure**

This study was funded by the National Institutes of Health, USA (grant number R01 AI091905-01). The authors have no financial conflict with any organization or entity.

**Information pertaining to writing assistance**

No writing assistance was utilized during the writing process of this manuscript

**Ethical conduct of research**

The authors state that they have obtained appropriate approval by local research ethics committee of the Isle of Wight Asthma and Allergy Center Committee (reference number: 09/H0504/129), as well as the internal review board of the University of Memphis (FWA00006815, 7 December 2012). In addition, informed consent has been obtained from the participants involved.

**Data sharing statement**

NA

**Reference**

**Papers of special not have been highlighted as: ⃰ of interest; ⃰ ⃰ of considerable interest**

1. Gaysina D, Fergusson DM, Leve LD *et al*. Maternal smoking during pregnancy and offspring conduct problems: evidence from 3 independent genetically sensitive research designs*.* *JAMA Psychiatry* 70(9), 956-963 (2013).

2. Wang X, Zuckerman B, Pearson C *et al*. Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight*.* *JAMA* 287(2), 195-202 (2002).

3. Stick SM, Burton PR, Gurrin L, Sly PD, Lesouef PN. Effects of maternal smoking during pregnancy and a family history of asthma on respiratory function in newborn infants*.* *Lancet* 348(9034), 1060-1064 (1996).

4. Cnattingius S. The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes*.* *Nicotine Tob Res* 6 Suppl 2 S125-140 (2004).

5. Hanrahan JP, Tager IB, Segal MR *et al*. The effect of maternal smoking during pregnancy on early infant lung function*.* *Am Rev Respir Dis* 145(5), 1129-1135 (1992).

6. Horta BL, Victora CG, Menezes AM, Halpern R, Barros FC. Low birthweight, preterm births and intrauterine growth retardation in relation to maternal smoking*.* *Paediatr Perinat Epidemiol* 11(2), 140-151 (1997).

7. Chiolero A, Bovet P, Paccaud F. Association between maternal smoking and low birth weight in Switzerland: the EDEN study*.* *Swiss Med Wkly* 135(35-36), 525-530 (2005).

8. Pereira PP, Da Mata FA, Figueiredo AC, De Andrade KR, Pereira MG. Maternal Active Smoking During Pregnancy and Low Birth Weight in the Americas: A Systematic Review and Meta-analysis*.* *Nicotine Tob Res* 19(5), 497-505 (2017).

9. Kyrklund-Blomberg NB, Cnattingius S. Preterm birth and maternal smoking: risks related to gestational age and onset of delivery*.* *Am J Obstet Gynecol* 179(4), 1051-1055 (1998).

10. Kyrklund-Blomberg NB, Granath F, Cnattingius S. Maternal smoking and causes of very preterm birth*.* *Acta Obstet Gynecol Scand* 84(6), 572-577 (2005).

11. Jaakkola JJ, Gissler M. Maternal smoking in pregnancy, fetal development, and childhood asthma*.* *Am J Public Health* 94(1), 136-140 (2004).

12. Burke H, Leonardi-Bee J, Hashim A *et al*. Prenatal and passive smoke exposure and incidence of asthma and wheeze: systematic review and meta-analysis*.* *Pediatrics* 129(4), 735-744 (2012).

13. Power C, Jefferis BJ. Fetal environment and subsequent obesity: a study of maternal smoking*.* *Int J Epidemiol* 31(2), 413-419 (2002).

14. Ino T. Maternal smoking during pregnancy and offspring obesity: meta-analysis*.* *Pediatr Int* 52(1), 94-99 (2010).

**\*Descibe the risks of grandmaternal nicotine exposure during gestation**

\*15. Maritz GS, Mutemwa M. The effect of grand maternal nicotine exposure during gestation and lactation on lung integrity of the F2 generation*.* *Pediatr Pulmonol* 49(1), 67-75 (2014).

16. Magnus MC, Haberg SE, Karlstad O, Nafstad P, London SJ, Nystad W. Grandmother's smoking when pregnant with the mother and asthma in the grandchild: the Norwegian Mother and Child Cohort Study*.* *Thorax* 70(3), 237-243 (2015).

17. Lodge CJ, Braback L, Lowe AJ, Dharmage SC, Olsson D, Forsberg B. Grandmaternal smoking increases asthma risk in grandchildren: A nationwide Swedish cohort*.* *Clin Exp Allergy* 48(2), 167-174 (2018).

18. Li YF, Langholz B, Salam MT, Gilliland FD. Maternal and grandmaternal smoking patterns are associated with early childhood asthma*.* *Chest* 127(4), 1232-1241 (2005).

19. Miller LL, Henderson J, Northstone K, Pembrey M, Golding J. Do grandmaternal smoking patterns influence the etiology of childhood asthma? *Chest* 145(6), 1213-1218 (2014).

20. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation*.* *Am J Respir Crit Care Med* 180(5), 462-467 (2009).

21. Maccani JZJ, Koestler DC, Houseman EA, Marsit CJ, Kelsey KT. Placental DNA methylation alterations associated with maternal tobacco smoking at the RUNX3 gene are also associated with gestational age*.* *Epigenomics* 5(6), 619-630 (2013).

22. Richmond RC, Simpkin AJ, Woodward G *et al*. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC)*.* *Hum Mol Genet* 24(8), 2201-2217 (2015).

23. Arshad SH, Hide DW. Effect of environmental factors on the development of allergic disorders in infancy*.* *J Allergy Clin Immunol* 90(2), 235-241 (1992).

24. Soto-Ramirez N, Ziyab AH, Karmaus W *et al*. Epidemiologic methods of assessing asthma and wheezing episodes in longitudinal studies: measures of change and stability*.* *J Epidemiol* 23(6), 399-410 (2013).

25. Arshad SH, Holloway JW, Karmaus W *et al*. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC)*.* *Int J Epidemiol* 47(4), 1043-1044i (2018).

26. Verification SSOB. Biochemical verification of tobacco use and cessation*.* *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco* 4(2), 149-159 (2002).

27. Balhara YPS, Jain R. A receiver operated curve-based evaluation of change in sensitivity and specificity of cotinine urinalysis for detecting active tobacco use*.* *J Cancer Res Ther* 9(1), 84-89 (2013).

**\*\*Provides justifications for using DNAm profile from umbilical cord and Guthrie cards**

\*\*28. Jiang Y, Wei JF, Zhang HM *et al*. Epigenome wide comparison of DNA methylation profile between paired umbilical cord blood and neonatal blood on Guthrie cards*.* *Epigenetics-Us* 15(5), 454-461 (2020).

29. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells*.* *Nucleic Acids Res* 16(3), 1215 (1988).

30. Beyan H, Down TA, Ramagopalan SV *et al*. Guthrie card methylomics identifies temporally stable epialleles that are present at birth in humans*.* *Genome Res* 22(11), 2138-2145 (2012).

31. Bibikova M, Fan J-B. GoldenGate assay for DNA methylation profiling*.* *Methods Mol Biol* 507 149-163 (2009).

32. Du P, Zhang X, Huang C-C *et al*. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis*.* *BMC bioinformatics* 11 587 (2010).

33. Bakulski KM, Feinberg JI, Andrews SV *et al*. DNA methylation of cord blood cell types: Applications for mixed cell birth studies*.* *Epigenetics-Us* 11(5), 354-362 (2016).

34. Houseman EA, Accomando WP, Koestler DC *et al*. DNA methylation arrays as surrogate measures of cell mixture distribution*.* *Bmc Bioinformatics* 13 (2012).

35. 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* 7(7), (2012).

36. Wang D, Yan L, Hu Q *et al*. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data*.* *Bioinformatics* 28(5), 729-730 (2012).

37. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods*.* *Biostatistics* 8(1), 118-127 (2007).

38. 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* 16 37 (2015).

39. Ogbuanu IU, Karmaus W, Arshad SH, Kurukulaaratchy RJ, Ewart S. Effect of breastfeeding duration on lung function at age 10 years: a prospective birth cohort study*.* *Thorax* 64(1), 62-66 (2009).

**\*\*Provides the statistical model for screening in epigenome-wide association study**

\*\*40. Ray MA, Tong X, Lockett GA, Zhang H, Karmaus WJ. An Efficient Approach to Screening Epigenome-Wide Data*.* *Biomed Res Int* 2016 2615348 (2016).

41. Zhang H, Wang F, Kranzler HR *et al*. Identification of methylation quantitative trait loci (mQTLs) influencing promoter DNA methylation of alcohol dependence risk genes*.* *Hum Genet* 133(9), 1093-1104 (2014).

42. Welch RP, Willer CJ, J. SL, Boehnke M. Snipper: a research tool for extracting and searching biological annotations on genes near SNPs*.* (2013).

43. Karolchik D, Baertsch R, Diekhans M *et al*. The UCSC Genome Browser Database*.* *Nucleic Acids Res* 31(1), 51-54 (2003).

44. Joubert BR, Felix JF, Yousefi P *et al*. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis*.* *Am J Hum Genet* 98(4), 680-696 (2016).

45. Joubert BR, Haberg SE, Nilsen RM *et al*. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy*.* *Environ Health Persp* 120(10), 1425-1431 (2012).

46. Byrnes JJ, Babb JA, Scanlan VF, Byrnes EM. Adolescent opioid exposure in female rats: Transgenerational effects on morphine analgesia and anxiety-like behavior in adult offspring*.* *Behav Brain Res* 218(1), 200-205 (2011).

**\*Offer potential mechanisms of the intergenerational epigenetic inheritance**

\*47. Barouki R, Melen E, Herceg Z *et al*. Epigenetics as a mechanism linking developmental exposures to long-term toxicity*.* *Environ Int* 114 77-86 (2018).

48. Van Otterdijk SD, Michels KB. Transgenerational epigenetic inheritance in mammals: how good is the evidence? *Faseb J* 30(7), 2457-2465 (2016).

49. Ziller MJ, Gu HC, Muller F *et al*. Charting a dynamic DNA methylation landscape of the human genome*.* *Nature* 500(7463), 477-481 (2013).

50. Schoenauer R, Lange S, Hirschy A, Ehler E, Perriard JC, Agarkova I. Myomesin 3, a novel structural component of the M-band in striated muscle*.* *J Mol Biol* 376(2), 338-351 (2008).

51. Guler HP, Zapf J, Schmid C, Froesch ER. Insulin-like growth factors I and II in healthy man. Estimations of half-lives and production rates*.* *Acta Endocrinol (Copenh)* 121(6), 753-758 (1989).

52. Iniguez G, Argandona F, Medina P *et al*. Acid-labile subunit (ALS) gene expression and protein content in human placentas: differences according to birth weight*.* *J Clin Endocrinol Metab* 96(1), 187-191 (2011).

53. Ester WA, Hokken-Koelega AC. Polymorphisms in the IGF1 and IGF1R genes and children born small for gestational age: results of large population studies*.* *Best Pract Res Clin Endocrinol Metab* 22(3), 415-431 (2008).

54. Beck KD, Powell-Braxton L, Widmer HR, Valverde J, Hefti F. Igf1 gene disruption results in reduced brain size, CNS hypomyelination, and loss of hippocampal granule and striatal parvalbumin-containing neurons*.* *Neuron* 14(4), 717-730 (1995).

55. D'ercole AJ, Ye P, O'kusky JR. Mutant mouse models of insulin-like growth factor actions in the central nervous system*.* *Neuropeptides* 36(2-3), 209-220 (2002).

56. Hess O, Khayat M, Hwa V *et al*. A novel mutation in IGFALS, c.380T > C (p.L127P), associated with short stature, delayed puberty, osteopenia and hyperinsulinaemia in two siblings: insights into the roles of insulin growth factor-1 (IGF1)*.* *Clin Endocrinol* 79(6), 838-844 (2013).

57. Choi J-H, Kang M, Kim G-H *et al*. Clinical and functional characteristics of a novel heterozygous mutation of the IGF1R gene and IGF1R haploinsufficiency due to terminal 15q26.2->qter deletion in patients with intrauterine growth retardation and postnatal catch-up growth failure*.* *The Journal of clinical endocrinology and metabolism* 96(1), E130-134 (2011).

**\*\*Reports the role of *CD33* on immune system**

\*\*58. Lock K, Zhang J, Lu J, Lee SH, Crocker PR. Expression of CD33-related siglecs on human mononuclear phagocytes, monocyte-derived dendritic cells and plasmacytoid dendritic cells*.* *Immunobiology* 209(1-2), 199-207 (2004).

**\*\*Reports the role of *CD33* on nervous system**

\*\*59. Griciuc A, Serrano-Pozo A, Parrado AR *et al*. Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta*.* *Neuron* 78(4), 631-643 (2013).

**\*\*Reports the role of *CHD1L* in neuronal differentiation in embryonic stem cells**

\*\*60. Dou D, Zhao H, Li Z *et al*. CHD1L Promotes Neuronal Differentiation in Human Embryonic Stem Cells by Upregulating PAX6*.* *Stem Cells Dev* 26(22), 1626-1636 (2017).

**\*\*Reports the role of *CD33* on cogenital anomalities**

\*\*61. Brockschmidt A, Chung B, Weber S *et al*. CHD1L: a new candidate gene for congenital anomalies of the kidneys and urinary tract (CAKUT)*.* *Nephrol Dial Transplant* 27(6), 2355-2364 (2012).

62. Li W, Feng Y, Chen A *et al*. Elmod3 knockout leads to progressive hearing loss and abnormalities in cochlear hair cell stereocilia*.* *Hum Mol Genet* 28(24), 4103-4112 (2019).

63. Loi E, Moi L, Blois S *et al*. ELMOD3-SH2D6 gene fusion as a possible co-star actor in autism spectrum disorder scenario*.* *J Cell Mol Med* 24(2), 2064-2069 (2020).

64. Malagon SGG, Melidoni AN, Hernandez D *et al*. The phenotype of a knockout mouse identifies flavin-containing monooxygenase 5 (FMO5) as a regulator of metabolic ageing*.* *Biochem Pharmacol* 96(3), 267-277 (2015).

65. Scott F, Malagon SGG, O'brien BA *et al*. Identification of Flavin-Containing Monooxygenase 5 (FMO5) as a Regulator of Glucose Homeostasis and a Potential Sensor of Gut Bacteria*.* *Drug Metab Dispos* 45(9), 982-989 (2017).

66. Joubert BR, Haberg SE, Bell DA *et al*. Maternal smoking and DNA methylation in newborns: in utero effect or epigenetic inheritance? *Cancer Epidemiol Biomarkers Prev* 23(6), 1007-1017 (2014).