DNA methylation signatures link prenatal famine exposure to growth and metabolism

Elmar W.Tobi[1], Jelle J.Goeman[3] , Ramin Monajemi[3], Hongcang Gu[8,9], Hein Putter[3], Yanju Zhang[1], Fabian Müller[4] , Erik. W. van Zwet[3], Christoph Bock[5,6], Alexander Meissner[8,9], L.H.Lumey[1,2] , P.Eline Slagboom[1,7], Bastiaan T.Heijmans[1,7]

1. Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA
3. Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands
4. Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarbrücken, Germany
5. Research Center for Molecular Medicine (CeMM), Austrian Academy of Sciences, Vienna, Austria
6. Department of Medical and Chemical Laboratory Diagnostics, University of Vienna, Vienna, Austria
7. The Netherlands consortium for Healthy Ageing, Leiden University Medical Center, Leiden, The Netherlands
8. Harvard University of Stem cell and Regenerative Biology, Cambridge, MA, USA
9. The Broad Institute of MIT and Harvard, MA, USA

Manuscript submitted

Abstract

Environmental conditions during embryonic development, such as malnutrition, have been linked to adult health. It has been suggested that epigenetic mechanisms may form the molecular link between development and later health. For instance, we have shown that prenatal famine exposure to the Dutch Hunger Winter, a severe war-time famine at the end of WW2, is associated with DNA methylation changes at candidate loci. However, a genome-wide assessment of which genomic annotations and pathways are sensitive to prenatal famine and the relation of such differential methylated regions (DMRs) with phenotypes of interest in relation to prenatal famine exposure is currently lacking.

We generated genome-scale DNA methylation data at 1.2M CpG sites for 24 individuals exposed to the Dutch Famine during early gestation and 24 same-sex unexposed sibling controls. Differential DNA methylation after famine exposure was observed at regions with a regulatory potential. Of the 181 prenatally induced differentially methylated regions (P-DMRs) 60.7% were located in gene-bodies and entailed differences of ~5% between exposed individuals and controls. Validation of six P-DMRs in all available individuals exposed during early gestation (N=60) and unexposed same-sex sibling controls (N=60) showed an attenuation of differential methylation in individuals conceived during April and May 1945, the last 2 months of the famine. The modest differences at individual regions extended to their biological pathways. Methylation differences were predominantly related to the regulation of growth and lipid and cholesterol metabolism. In our study group with early gestational famine exposure, DNA methylation of *INSR* was also associated with birth weight and *CPT1A* methylation with LDL cholesterol.

Prenatal exposure to famine may affect regulatory regions. Although the observed differences at individual regions are modest, they extend across biologically relevant pathways and could potentially explain phenotypes previously associated with prenatal famine exposure.

Introduction

It has been hypothesized that epigenetic mechanisms, including DNA methylation, underlie the link between environmental exposures during early development and adult phenotypes[1]. Experiments in animal models showed that prenatal malnutrition during early development can result in persistent changes in DNA methylation at candidate regions including promoters[2], retrotransposons[3] and imprinted regions[4]. These can subsequently be transmitted during cell differentiation and growth and explain phenotypic variation. It remains unclear whether prenatally-induced differential methylation contributes to the risk of disease in humans.

The quasi-experimental setting of well-documented historical famines represents a powerful opportunity to study this question in humans. Exposure to the Dutch Hunger Winter, a severe war-time famine at the end of WW2, is associated with an adverse metabolic profile consisting of a decreased insulin response, higher body mass index (BMI)[5], and elevated total and LDL cholesterol[6,7], but also an increased risk on schizophrenia[8]. The study of the Dutch Famine, also called the Hunger Winter, is of interest as it offers a clear developmental exposure[5]. More over, its relevance may extend to modern exposures since maternal smoking[9] and gestational diabetes[10] have been associated to similar health outcomes as prenatal famine exposure.

We previously linked prenatal famine exposure to differential DNA methylation at candidate regions, including promoters and imprinted regions[11,12]. Of interest, DNA methylation at a subset of these regions was also sensitive to maternal smoking[13] and gestational diabetes[14]. In line with the observation that the epigenome is particularly dynamic around conception and embryo implantation[15], differential methylation was predominantly observed after exposure to famine during early gestation. To further understand the potential role of prenatally-induced differential methylation in human disease, a genome-wide approach is required to uncover the characteristics of affected genomic regions and annotate the function of the genes and biological pathways involved.

We used Reduced Representation Bisulfite Sequencing[16] to generate DNA methylation data on 1.2M individual CpG dinucleotides in individuals exposed to famine in early gestation and an unexposed same-sex sibling as control. Using a step-wise analysis strategy involving genomic annotation and technical and biological validation, we identified genomic characteristics of prenatally-induced differentially methylated regions (P-DMRs)[17]. Moreover, we also identified pathways affected by prenatal famine exposure and describe individual P-DMRs that are associated with phenotypic outcomes associated with early gestational famine exposure.

Results

We studied 24 same-sex sibling pairs (N=48) from a larger group of 60 pairs of which one sibling was exposed to the Dutch Famine around conception and up to 10 weeks into development (i.e. periconceptional exposure). The subset was constrained to siblings within the ongoing Dutch Hunger Winter Families Study[18] with an age difference of less than 5 years and included an equal number of male and female pairs as well as an equal number of pairs with the control siblings conceived and born before or after the Famine (**Supplement II, Table S1**). We employed Reduced Representation Bisulfite Sequencing (RRBS) on DNA from whole blood to obtain single nucleotide high resolution DNA methylation data on a genome-wide scale[16]. Sequencing was performed on an Illumina GAIIx, utilizing one sequencing lane per individual. This resulted in 25.6 million high quality reads on average per sample (**Supplement II, Table S1 and S2**). Exclusion of CpG sites with a low or extremely high coverage (≤5-fold or >200) and those that were uninformative (mapping to a random chromosome, median methylation=0% or 100%) resulted in 1.206.161 unique CpG sites with an average median sequencing depth of 28 and the data for a CpG site was on average for 99.8% complete (**Supplement II, Table S2, Figure S1A and Figure S1B**). The data showed the classical bimodal nature of DNA methylation (**Supplement II, Figure S2**)

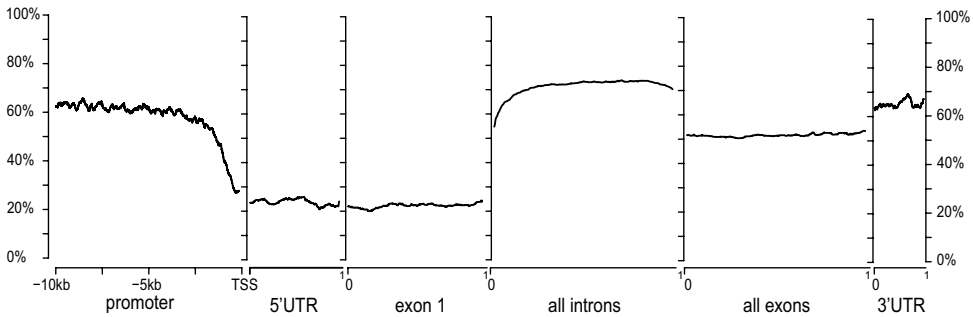

Figure 1. The methylation level of genes in the RRBS dataset

The methylation level across a gene. A lowess has been fit across all data for all entrez genes. The width of the gene elements represents the relative amount of data for such elements in the total dataset.

with promoter and intragenic methylation levels matching those observed in the comprehensive blood methylome[19] (**Figure 1**).

Genome-scale data corroborates previous findings

In previous studies of prenatal famine exposure[20], in particular during early gestation[21], we did not observe differences in global methylation between exposed individuals and sibling controls. Again no difference was found as the mean global methylation extracted from all sequenced reads was 42.55% in the exposed and 42.64% in the sibling controls (Paired t-test, P=0.91, **Supplement II, Figure S3**). Since the average methylation of retrotransposons is regularly taken as a measure of global methylation, we also tested the average DNA methylation of 993 RepBase Update database[22] repeat sequences, again finding no associations (P_{FDR}>0.26). The genome-scale data covered the 16 candidate regions we reported on previously[11,12]. The regions originally found to be differentially methylated after periconceptional famine exposure were also associated in the current data set (P=9.0x10^{-3}). Likewise, no association was observed for regions at which no differential methylation was observed previously (P=0.19), validating our genome-scale approach.

Differential methylation occurs at open chromatin regions

We selected 28 genomic annotations from the literature, mapped the CpG dinucleotides to these annotations and evaluated at which of these annotations DNA methylation was susceptible to early gestational famine exposure. DNA methylation at 5 annotations was associated with famine exposure, namely non-CGI promoters with an open chromatin conformation ('bonafide non-CGI promoters')[23], enhancers[24], exons, DNaseI/FAIRE-seq open chromatin regions, and enhancers active during the pre- and peri-implantation period[43] (P_{FDR}<0.05; **Table 1**). Apart from exons these regions represent potential regulatory regions. Annotations not associated with famine exposure included CGIs, putative meta-stable epialleles[25] and highly variable DNA methylation

Table 1. Outcome of genome-scale testing of genomic annotations

Genomic annotations*	Regions in genome	Covered	$P_{nominal}$	P_{FDR}
Non-CGI, non 'bonafide' promoters[1]	7,014	2,024 (28%)	$9.1x10^{-4}$	0.026
Enhancers[2]	59,466	6,207	$1.9x10^{-3}$	0.026
DNaseI/FAIRE-seq regions[3]	590,252	79,728	$4.4x10^{-3}$	0.036
Middle exons	17,848	1,570	$5.8x10^{-3}$	0.036
Developmental enhancers type I[4]	5,118	922	$6.5x10^{-3}$	0.036
'bonafide' CGI – shores[5]	88,871	27,688	0.012	0.053
Non-coding RNA[6]	718	59	0.015	0.053
Conserved regions[7]	165,937	1,386	0.016	0.053
CGI –shores	319,509	67,811	0.017	0.053
3'UTR	21,004	2,909	0.035	0.085
Non genic CGI[8]	129,049	41,023	0.036	0.085
'Bonafide' CGI – border	88,074	22,777	0.036	0.085
Developmental enhancer type II	2,287	320	0.078	0.15
CGI	343,925	113,673	0.078	0.15
Introns	201,640	61,816	0.080	0.15
hESC bivalent chromatin domains	1,797	1,741	0.16	0.28
Bonafide CGI	44,439	35,271	0.20	0.32
cell-type specific gene promoters	2,372	2,106	0.21	0.32
First exons	51,497	13,507	0.25	0.36
Promoters	23,689	16,904	0.26	0.36
HSC bivalent chromatin domains	2,910	2,779	0.28	0.36
Imprinted promoters	46	42	0.29	0.36
'Bonafide' CGI promoter	16,674	14,880	0.32	0.37
CTCF insulators from CD4+ cells	28,661	4,396	0.32	0.37
Imprinted DMRs	14	6	0.33	0.37
Putative Metastable epialles	38	29	0.43	0.47
Variably methylated regions	227	56	0.55	0.57
Promoters – cancer genes	888	795	0.63	0.63

1. Promoters without CGIs but with a relatively open chromatin state[23]
2. Enhancers characterized by H3K4me1, non-overlapping with promoters[24]
3. Regions with an open Chromatin state as defined by DNaseI and FAIRE-seq signals (UCSC track ENCODE)
4. Enhancers active during first stages of blastocyst development[44]
5. Shores of so-called bonafide CGI, CGI island with an ubiquitously open chromatin structure; oe>0.6, GC%>50% & length >700bp[23]
6. Body of various type of non-coding RNAs
7. Conserved regions outside promoters, CGIs, exons and UTR
8. CGI >10kb from gene

* Details on the genomic annotations can be found in the materials and methods.

regions (VMR)[26]. There was no difference between the average GC content of the 5 significant annotations and the non-significant genomic annotations (t-test, P=0.79).

Characteristics of individual DMRs associated with prenatal famine

Next, the 90,451 individual regions within the 5 associated genomic annotations were tested (median size= 1.4 kb; mean number of CpG dinucleotides= 6.7). This analysis identified 181 regions as prenatally induced differentially methylated regions (P-DMRs, P_{FDR}<0.05). The difference in DNA methylation between exposed individuals and their same-sex sibling was variable at these regions (up to >10%), but generally moderate (median 4.6%; **Figure 2**). DNA methylation at P-DMRs was more commonly higher (60.8%) than lower among exposed individuals which suggests that differential methylation observed is not a simple one-to-one consequence of a possible methyl donor deficiency in the rations during the Famine.

Of the 181 P-DMRs, 60.7% was located in gene bodies, 11.6% in promoters, 10.5% in upstream genic regions, 10.0% in downstream genic regions and 7.2% were intergenic. Further characterization with EPIGRAPH[27] showed an increased co-occurrence with histone marks associated with active

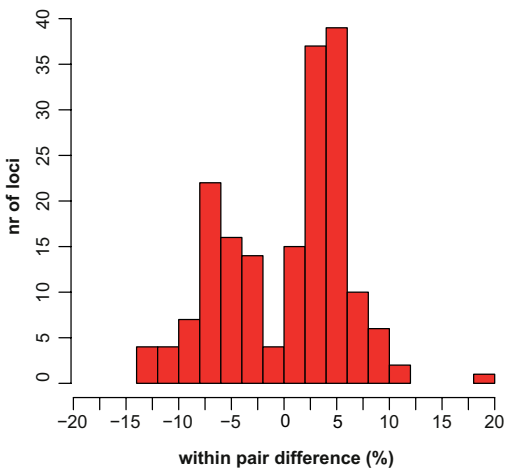

Figure 2. The average within pair difference for the 181 loci surviving multiple testing

enhancers, transcribed gene bodies, active regulatory sites and expressed exons (H3K4me1-3, H2AZ, H3K9me1, PolII, H3K79me1, H3K27me1)[28-30], while a lower co-occurrence with SINES and a lower overall repeat score was observed (P_{FDR}<0.05). Interestingly, although the putative P-DMRs were identified on the basis of whole blood samples, the genes nearest to the P-DMRs were not enriched for tissue-specific gene expression, including blood and bone marrow (OR=0.98 [95% CI:0.59-1.64] and OR=1.01 [95% CI:0.55-1.87], respectively). In contrast, the P-DMRs were enriched for genes differentially expressed during the pre-implantation stage of development (OR=4.95 [95% CI:3.53-6.49])[31] and organogenesis (OR=4.76 [95% CI: 3.11- 7.30])[32] in humans.

Validation of genome-scale measurements

Findings from genome-scale DNA methylation studies require validation with an independent technology. From the 181 putative P-DMRs, we prioritized 11 regions with the lowest p-values and 9 regions according to consistency of DNA methylation differences across CpG dinucleotides and mapping to genes with known functions. For 19/20 regions a working assay could be designed using the mass spectrometry based method Epityper (**Supplement II, Table S5**)[33]. The overall correlation between the average DNA methylation measured using the genome-wide method RRBS and the locus-specific method Epityper was good (r=0.81) (**Figure 3**), despite the fact that the length of RRBS regions defined by genome annotations were larger (>=1 kb) than the regions targeted by Epityper assays (<560bp [as 600bp is the technical limit]). Of the 19 regions, 13 were again associated with famine (P<0.05, **Supplement II, Table S3**). A correspondence between the RRBS and the Epityper measurements of individual regions was defined as a Pearson's correlation greater than 0.7 and was found for 6 loci mapping to *CDH23*, *SMAD7*, *INSR*, *CPT1A*, *KLF13* and *RFTN1* (**Figure 3**). These 6 Epityper assays were taken further for validation in the complete set of 60 periconceptionally exposed individuals and their same-sex sibling controls.

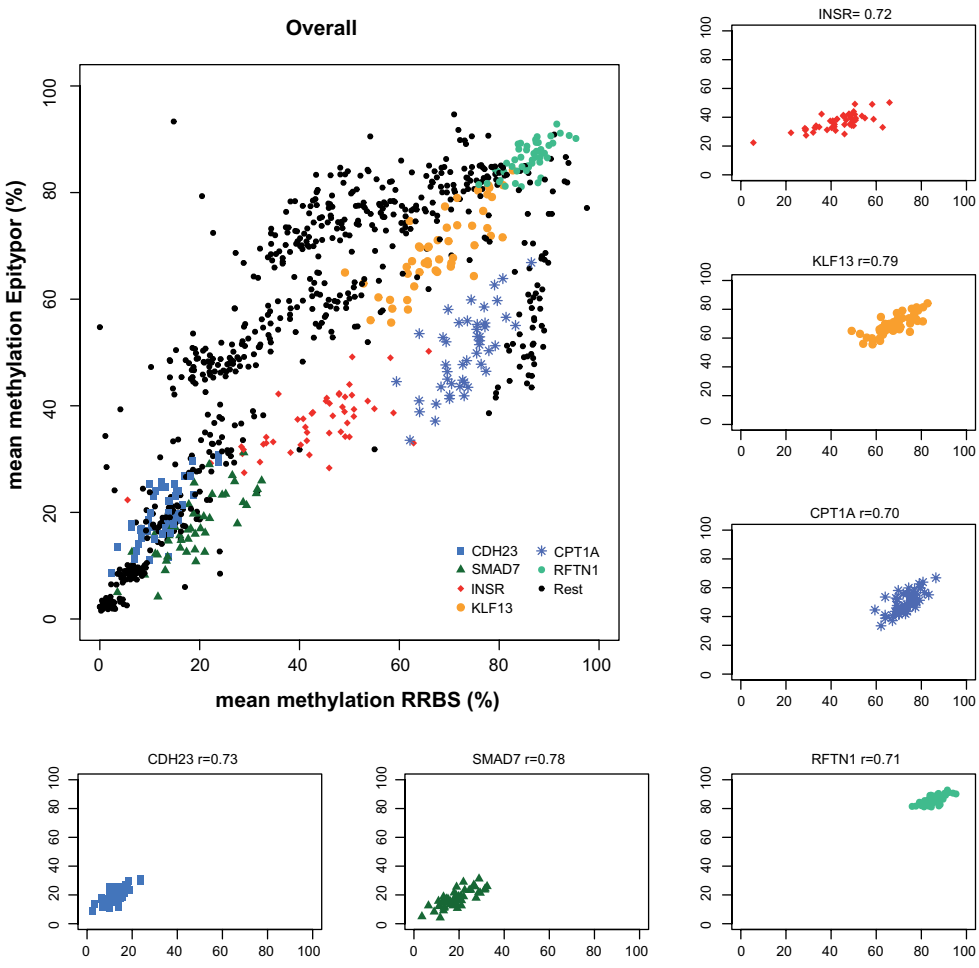

Figure 3. The relationship between the average methylation of the RRBS regions and the Epityper measurements.

The correlation between the average methyation of the RRBS regions and the Epityper measurements of smaller subsets of these regions. Individual regions with a Pearson correlation >0.7 are denoted in color and plotted separately along the main figure. The correlation of the other loci can be found in Supplement II, Table S3.

The associations were consistent for individual CpG dinucleotides across the P-DMRs (**Supplement II, Table S4**), although the within pair differences observed in these 60 sibling pairs using the Epityper assay were somewhat attenuated as compared with those observed in the genome-scale data (**Table 2**). The associations were not affected by age, smoking (neither

current smoking nor package years), socio-economic status (SES) and current diet (kcal/day nor the percentage of fat, carbohydrates or protein in the diet), indicating that the DMRs were independent of current, post-natal environmental exposures of the individuals studied. Lastly, we excluded cellular heterogeneity of whole blood samples as a confounding factor by exploring the potential association of DNA methylation at the 6 regions with blood cell counts in an independent sample set (**Supplement II, Table S7**).

Table 2. Validation of RRBS associations with Epityper

Annotation		Genome-wide (N=48)				Epityper Validation (N=120)	
Type*	Nearest Gene (kb)	Meth. Controls (%)	Within pair diff. (%)	P	P_{FDR}	Within pair diff. (%)	P
1 & 2	*SMAD7* (+25)	21.3	4.2	$1.0x10^{-7}$	$1.1x10^{-3}$	3.2	$2.5x10^{-3}$
1 & 2	*CDH23* (0)	12.4	4.0	$1.3x10^{-7}$	$1.1x10^{-3}$	2.2	$6.3x10^{-3}$
1	*INSR* (0)	43.3	8.1	$3.9x10^{-6}$	0.010	2.0	0.031
1 & 3	*RFTN1* (0)	86.3	-2.3	$3.2x10^{-5}$	0.030	-0.9	0.09
1	*CPT1A* (0)	67.0	4.5	$4.0x10^{-5}$	0.031	2.0	0.05
1	*KLF13* (0)	67.1	-7.9	$6.1x10^{-5}$	0.042	-3.1	0.014

* Type of genomic annotation: 1 = Open chromatin, 2= enhancer, 3 = exon.

Critical window of exposure

It is generally assumed that the influence of a prenatal exposure on DNA methylation depends on its exact timing during development[1]. The gestational timing of the famine exposure is based on the mother's last menstrual period (LMP), a commonly used proxy for the start of pregnancy. To test for an influence of timing on differential methylation we tested for an interaction between the month of the start of pregnancy (month of LMP) and famine exposure.

There was an influence of timing ($P_{interaction}$=0.016) therefore the average within sibling pair difference in DNA methylation was plotted for the 6 P-DMRs against the start of the mother's pregnancy of the prenatally exposed (**Figure 4A**). DNA methylation differences were present throughout the first 4 months

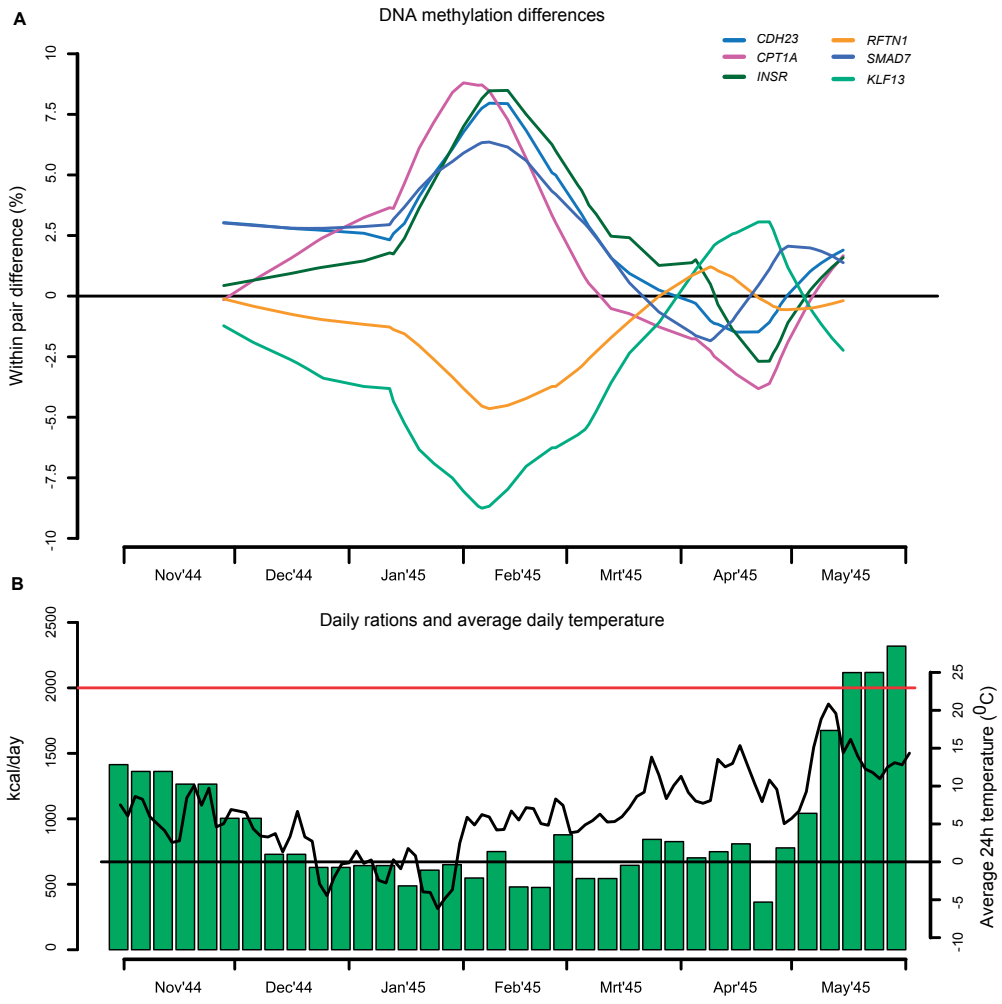

Figure 4. Results across the Famine period

A. A lowess depicting the average within pair difference (y-axis) ass stratified by the estimate of the moment of conception (last menstrual period; x-axis). Each colored line represents a locus.

B. The official daily rations (kcal/day) for each calendar week (green bars) and a lowess depicting the average 24h temperature (black line; source KNMI; DeBilt weather station). The daily requirement of non-pregnant women of 2000kcal/day is denoted in red

of the Famine, but waned towards the end to become virtually absent by the start of April 1945 (test for interaction pre-April and later pregnancies, $P_{interaction}=4.6x10^{-3}$). Re-examination of candidate loci identified as P-DMRs in earlier studies of the same individuals[11,12,21], showed the same attenuation

of DNA methylation differences in pregnancies starting after April 1945 ($P_{interaction}$=1.8x10^{-3}).

These associations with famine exposure for pairs with an exposed sibling conceived before April was consistent. Inspection of the RRBS data in exposed individuals from pre-April pregnancies and sibling controls (n=36 individuals) showed that the DNA methylation differences were virtually identical to those in the complete set of discovery pairs (n=48). These pre-April associations were reliably validated using an independent technology (Epityper). Crucially, the effect sizes in the pre-April exposed individuals and controls not used for discovery (n=36) were remarkably similar to those in the discovery set obtained using the same technology (**Table 3**). Indeed, the replication was high (95% CI:3/6-6/6), confirming the consistency of the data. Conversely, no DNA methylation difference was observed at these regions in exposed individuals conceived later during the Famine as compared to their sibling controls (N=48, P>0.20).

This attenuation may be explained by changes in the conditions towards the end of the Famine, but the rations remained low during the entire famine period and also the daily temperature was not very different in April and May from March (**Figure 4B**). Alternatively, the data may be compatible with intrinsic features of early development and may suggest that DMRs are induced after conception during periods at which large scale DNA methylation changes occur[15] or slowly accumulate during a more prolonged exposure.

The P-DMRs

In all pre-April exposed individuals and sibling controls (n=72), the average within pair differences were 4.2% for *SMAD7* (P=6.8x10^{-4}), 3.6% for *CDH23* (P=3.3x10^{-4}), 3.2% for *INSR* (P=7.1x10^{-3}), -2.1% for *RFTN1* (P=3.6x10^{-3}), 4.1% for *CPT1A* (P=1.9x10^{-3}) and -6.0% for *KLF13* (P=1.7 x10^{-4}). The P-DMRs were located in regions defined by their chromatin state to be an enhancer or isolator and all were located in open chromatin regions according to ENCODE data, indicating these loci may have a regulatory function (**Supplement II, Figure S4A-D**). For instance, the *INSR* P-DMR is located

Table 3. The outcome of a stratified analysis for timing

	Genome-wide <April '45 (N=36)[1]		Technical validation (idem to RRBS)[2]		Replication (N=36)[3]		Epityper April & May '45 (N=48)[4]	
P-DMR	Diff	P	Diff	P	Diff	P	Diff	P
SMAD7	4.3	$2.2x10^{-7}$	3.8	0.010	4.4	0.027	1.6	0.39
CDH23	4.2	$4.5x10^{-6}$	3.6	$1.4x10^{-3}$	3.6	0.034	0.1	0.91
INSR	9.0	$7.4x10^{-7}$	2.5	0.16	3.8	0.016	0.5	0.77
RFTN1	-3.7	$1.1x10^{-7}$	-2.3	$3.5x10^{-3}$	-1.9	0.13	0.8	0.21
CPT1A	4.9	$1.1x10^{-5}$	4.4	0.041	3.8	0.021	-1.1	0.49
KLF13	-8.6	$1.6x10^{-5}$	-5.5	$7.8x10^{-3}$	-6.0	$7.5x10^{-3}$	1.0	0.57

1. Outcome of the famine association for the pairs with one sibling conceived from Nov'44 to Mar'45 (18 pairs, 36 individuals) in the RRBS data.
2. Outcome of the famine associations for the Epityper measurements on the same pairs as measured by RRBS and with a sibling conceived between Nov'44 and Mar'45.
3. Outcome of the famine association for the Epityper measurements for pairs *not* measured by RRBS but with one sibling conceived between Nov'44 and Mar'45 (18 pairs, 36 individuals).
4. Outcome the famine association for the Epityper measurements for all pairs with one sibling conceived in April and May '45 (24 pairs, 48 individuals).

in an intronic enhancer of *INSR* according to ENCODE data and marks a DNaseI hypersensitivity site in 31 different tissues and cell lines (**Figure 6B**) and the P-DMR at *CPT1A* overlapped a weak enhancer in the blood derived GM12787 line and H1 embryonic stem cells. Moreover, it overlaps a strong binding site of the BAF155 transcription factor which is an repressor of self renewal[34] and vital for early liver development[35] (**Figure 6D**).
Indeed the genes in which the P-DMRs are located or are closest have been implicated in development, like eye development (*CDH23* and *RFTN1*), forebrain formation (*SMAD7*[36]), growth (*INSR*) and sustaining early pregnancy (*KLF13*[37]). But beside roles in development most genes also had metabolic functions, including insulin signaling (*INSR*), pancreatic beta cell functioning (*SMAD7*[38]), fatty acid oxidation (*CPTIA*[39]) and cholesterol metabolism (*KLF13*[40]). All in all the P-DMRs are interesting subjects for study.

Famine associations extend into pathways

However, the average within pair differences were modest. It is hypothesized that epigenetic modulation may occur through multiple smaller changes across a gene network[41], a hypothesis that recently gained empirical support[21,42]. So we tested if the association with prenatal famine exposure and DNA methylation extended to the pathways of which the genes belonging to the P-DMRs are a part of. So we re-visited the genome-wide DNA methylation dataset for the pre-April sibling pairs (N=36) and tested for an additional association along the MSigDB pathways and GO Biology terms to which the 6 genes belong to by removing all methylation data mapping to these 6 genes, including the P-DMRs.
DNA methylation of 25 out of 101 gene-sets was associated with famine exposure ($P_{FDR}<0.05$). The three most significant pathways were the GO terms *positive regulation of growth* ($P_{FDR}=5.5x10^{-3}$), *response to activity* ($P_{FDR}=5.6x10^{-3}$) and *regulation of embryonic growth* ($P_{FDR}=0.021$). Four MSigDB terms were significant, namely the KEGG insulin signaling pathway ($P_{FDR}=0.029$), the BIOCARTA *HDAC* pathway ($P_{FDR}=0.032$; involved in myogenesis and cardiac development), the REACTOME *IRS related events*

(P_{FDR}=0.032; encompassing the insulin signaling cascade) and REACTOME *metabolism of lipids and lipoproteins* (P_{FDR}=0.039).

The significant GO terms were clustered based on their relatedness and redundant terms were removed by REVIGO[43]. Visualization of these results showed that the largest cluster was formed from the GO term *regulation of growth* (**Figure 5).** Multiple clusters contained pathways related to lipid and cholesterol metabolism. Indeed, beside the REACTOME term *metabolism of lipids and lipoproteins*, multiple GO terms related to lipid metabolism showed significantly different DNA methylation between exposed individuals and sibling controls (GO: *positive regulation of lipid metabolic process*, P_{FDR}=0.028; *lipid homeostasis*, P_{FDR}=0.042; *triglyceride metabolic process*, P_{FDR}=0.049).

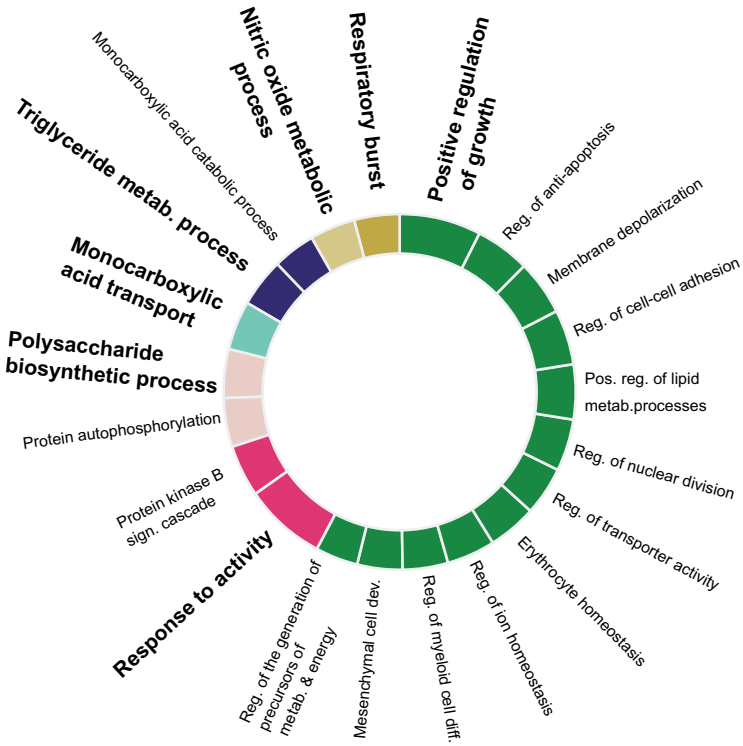

Figure 5. Revigo analysis of the significant pathways

A sunburst graph of the non-redundant clustered FDR significant GO terms associated with prenatal famine exposure. The size of the circular boxes represents the statistical evidence. In bold are the dominant terms of the clusters, which are denoted in different colors.

P-DMRs and phenotypic outcomes

The pathways and P-DMRs are of interest to the reported phenotypic associations with early developmental famine exposure and development in general. Individuals exposed early in gestation were reported to have a higher birth weight than controls[11] and replicated later-life phenotypic outcomes include a higher body mass index (BMI) and an altered glucose response[5] and elevated LDL and total cholesterol levels[6,7]. Phenotypic data for all outcomes except schizophrenia were available and associations between these five phenotypes and methylation at the 6 P-DMRs were explored.

Birth weight data were available for exposed individuals (N=60), but not their sibling controls. In this study group with early gestation exposure, *INSR* methylation was positively correlated with birth weight (R-squared=0.111, $P=9.0x10^{-3}$; **Figure 6A**). The association remained statistically significant after multivariate analysis and after accounting for multiple testing (6 P-DMRs and 5 phenotypic outcomes), gestational age, age at examination, and adult characteristics as SES, diet and smoking ($\beta_{exp\text{-}birth\ weight}$=3.9%/1kg, P_{FDR}=0.033).

The later-life phenotypic outcomes tested were BMI, plasma glucose 120 minutes after an oral glucose tolerance test and LDL and total cholesterol levels. This information was available in both the exposed individuals and the unexposed same-sex siblings (N=120; 60 pairs). *CPT1A* methylation was positively correlated with LDL cholesterol levels (R-squared=0.077, $P=3.5x10^{-3}$, **Figure 6C**), in line with the association of early gestational famine with *CPT1A* methylation and LDL. The statistical association with *CPT1A* methylation and LDL cholesterol remained after correction for multiple testing and further adjustments for age, sex, BMI, SES, smoking and current diet (β_{LDL}=2.4%/mmol*l^{-1}, P_{FDR}=0.033). The observed effects were almost identical in the prenatally exposed individuals and their unexposed same-sex siblings (β_{exp}=2.4%/mmol*l^{-1}; β_{sibs}=2.3%/mmol*l^{-1}), suggesting that *CPT1A* may be a quantitative trait locus for LDL. Exclusion of individuals (N=10) using lipid lowering medication did not affect the association (β_{LDL}=2.5%/mmol*l^{-1}, $P=2.5x10^{-3}$). Similar results were obtained for total cholesterol, but

could be attributed to LDL cholesterol, which was highly correlated with total cholesterol ($r=0.92$, $P=2.2x10^{-16}$).

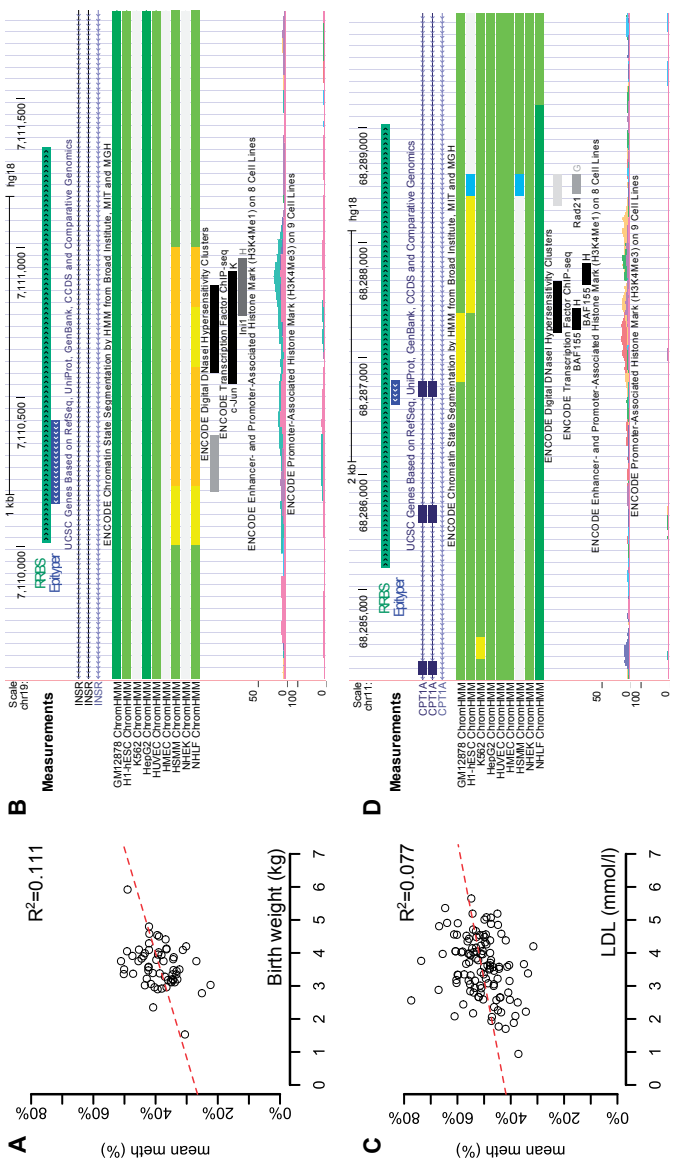

Figure 6. The INSR and CPT1A P-DMRs

A. Scatterplot between birth weight (x-axis) and the average DNA methylation of *INSR* (y-axis) in the 60 prenatally exposed individuals.

B. The genomic annotation of *INSR*. The P-DMR overlaps an enhancer in the HSMM and NHLF cell lines and a DNaseI hypersensitivity cluster in over 30 cell lines.

C. Scatterplot between LDL (x-axis) and the average DNA methylation of *CPT1A* (y-axis) in all 120 siblings.

D. The genomic annotations of *CPT1A*. The P-DMR overlaps an enhancer in the blood derived GM12878 and embryonic stem cell line H1 and a DNaseI hypersensitive cluster in over 30 cell lines. Furthermore, the BAF155 transcription factor binds in this region.

Discussion

We studied the association between famine exposure from peri-conception and into the first trimester and DNA methylation at middle age on a genome-wide scale using reduced representation bisulfite sequencing (RRBS)[16]. DNA methylation at non-CGI promoters[23], DNaseI/FAIRE-seq regions, exons, enhancers, and those enhancers active during the pre- and peri-implantation period[44] was associated with prenatal famine exposure. From these 5 annotations 181 regions were individually associated with famine. These regions were mainly located in gene bodies, co-occurred with histone marks of active chromatin, and were not enriched for tissue specific genes while being enriched for genes active during development. We found evidence that the exact timing of the famine exposure during this developmental time-frame has bearing on DNA methylation levels. We validated P-DMRs mapping to *SMAD7*, *CDH23*, *INSR*, *RFTN1*, *CPT1A* and *KLF13* by an independent methodology and found that associations for individual regions may extend to the pathways they belong to. Indeed, pathways involved in growth and lipid and cholesterol metabolism were associated with famine exposure. We found tentative associations between DNA methylation at an intronic enhancer of *INSR* with birth weight, and intragenic methylation of *CPT1A* with LDL cholesterol.

Genomic annotations affected were not the expected ones

We applied next generation sequencing of bisulfite converted DNA, generating data for 1.2M CpG dinucleotides in 24 exposed individuals and 24 unexposed same-sex sibling controls. This has allowed us to comprehensively study a large array of genomic annotations. Our approach led to discovery that regions with a regulatory potential were especially sensitive to early prenatal famine exposure. The associated regions were mainly located within gene bodies and enhancers, thus located outside the regions and annotations generally queried in epigenetic studies and previously hypothesized to be especially sensitive to prenatal environmental conditions[45].

The nature of the P-DMRs

It is unclear what mechanisms underlie the observed differences in DNA methylation after prenatal famine exposure. We think it is reasonable to entertain a biological explanation. While the prenatal environment may induce P-DMRs that in turn persistently affect gene expression[3], P-DMRs may also reflect altered transcriptional activity during development as it was shown that differential expression during development may influence DNA methylation levels[46,47]. This explanation is favored by the observation that the majority of P-DMRs were intragenic and enriched for genes changing expression during early development. Another question is whether P-DMRs merely represent genomic scars of prenatal adversity or mark an adaptive response to cope with adversity. We found that DNA methylation differences at individual regions extended towards pathways involved and growth and lipid metabolism. This finding is compatible with optimization for energy production and growth under nutritional constraint through epigenetic fine-tuning[41].

As an alternative explanation, we should consider that the observed differences in DNA methylation may have no functional meaning but merely reflect specific characteristics of the selected individuals who survived the famine. Even under non-famine conditions, early pregnancy loss may be as high as 70%[48]. During the Hunger Winter, the number of births decreased by 50%[49], presumably in part because of increased levels of early pregnancy loss. Furthermore, undernutrition during pre-implantation can cause reduced cell numbers in pre- and post implantation rat embryos[50]. Embryos or subsets of cells with an increased growth potential may have had a higher chance of survival and this may explain why we observed a P-DMR at *INSR*, a key regulator of prenatal growth. Even among individuals with exposure in early gestation, we observed an association with birth weight, and methylation changes in multiple pathways related to growth. If our findings are based on selective survival, the P-DMRs would not reflect environmentally-induced changes but characteristics that follow selection on epigenetic variation between embryos.

LDL and CPT1A

Individuals with early gestational famine exposure have a less favorable metabolic profile[5], including a higher LDL cholesterol[6]. For LDL cholesterol, data were available for both the famine exposed individuals and their unexposed sibling controls. In these groups, we found a tentative association between *CPT1A* gene body methylation and LDL. The association between *CPT1A* methylation and LDL was almost identical in the prenatally exposed individuals and their unexposed same-sex siblings. If replicated, *CPT1A* methylation may thus prove to be a LDL quantitative trait locus contributing to the higher LDL levels in the exposed. *CPT1A* is involved in shuttling fatty acids to the mitochondria and as such is the rate limiting enzyme in fatty acid oxidation[39]. The association may with LDL may be related to the observation that free fatty acids drive hepatic LDL production[51], therefore it may be of interesting to relate *CPT1A* methylation with free fatty acids levels, for which we do not have data, rather than LDL.

Tissue specificity

The famine associations along pathways and with birth weight and LDL all relate to processes taking place in other tissues than blood. However, we have not investigated whether these associations extend to other relevant tissues than blood for such samples are simply not readily available in cohorts of generally healthy individuals[17]. We do not exclude the possibility that some of the 181 regions identified entail tissue specific associations, but we did not find enrichment for P-DMRs mapping to tissue specific genes. Since we find associations arising in early gestation the DNA methylation differences in blood may reflect those in other tissues due to mitotic inheritance[52], as was shown in animal studies[3,53].

Concluding

We identified regions with a regulatory potential that may be especially sensitive to early famine exposure. We observed DNA methylation changes in middle age among individuals exposed to environmental disturbances during early human development that extend along pathways. To avoid unmeasured sources of confounding at the family level, we used unexposed same sex siblings as study controls. Our study thereby makes a significant contribution to the further systematic exploration of the relation between prenatal adversity and adult health.

Materials and Methods

Subjects

The recruitment process of the Hunger Winter Families study has been described in detail elsewhere[18]. Study subjects were selected from births between 1943-1947 at three institutions in famine-exposed cities (the midwifery training schools in Amsterdam and Rotterdam and the University Medical Center in Leiden). The series includes singletons born between February 1945-March 1946 who had been exposed to famine in utero, births in 1943 without in-utero exposure serving as pre-famine time controls, and births in 1947 without such exposure serving as post-famine time controls. Whenever possible, we recruited an unexposed same-sex sibling of each individual in this birth series to serve as a family control. This provides sibling pairs, some of which include one famine exposed individual and an unexposed control. Ethical approval for the study was obtained from the participating institutions, and all participants provided written informed consent and additional explicit approval for the current set of genome scale measurements was given by the Leiden University Medical Center medical ethics committee.

The food supply and nutrition had been generally adequate during WWII in the Netherlands[54]. However, after October 1944, supplies became increasingly

scarce and the energy content of official rations fell below 1,000 kcal/day by the end of November 1944. By April 1945, just before liberation, the official rations were as low as 500 kcal/day[55] and the daily rations did not reach 2,000 kcal/day until May 20th. Since the famine period was shorter than the nine months of human gestation, and the population was well fed just before and directly after the famine, individuals can be identified that were exposed during specific periods of their development *in uterus*.

Within the Hunger Winter Families Study there are 313 same-sex sibling pairs who completed clinical examination. Sixty pairs included a sibling exposed to the famine around conception for up to 10 weeks into development (e.g. "periconceptional" exposure), as defined by a mothers' estimated last menstrual period between November 28, 1944 and May 15, 1945 (exposed: age 58.1y, SD 0.35; unexposed same-sex siblings: age 57.1y, SD 5.50). We included for analysis all 38 pairs of this group who had an age difference of less than 5 years at the time of examination. From these pairs we randomly selected 12 male and 12 female sibling pairs. Half of the female and half of the male pairs consisted of a sibling conceived and born after the famine as to prevent a consistent age difference between the prenatally exposed and controls and minimize a possible influence of early childhood famine exposure on the analyses (**Supplement II, Table S1**).

Birth weights were only available for the individuals in the hospital series and were taken from the birth records from the three institutions at which these individuals were born, meaning that we only had birth weight for the sixty individuals conceived during the famine and not their same-sex sibling controls. A telephone interview was performed with all participants and included questions on socio-demographic characteristics such as education, health history, health behaviors such as smoking and drinking, and medications for diabetes, cholesterol and blood pressure. The medical examinations were scheduled early in the morning and included the measurement of height (to the nearest 1mm by portable stadiometer [Seca, Hamburg, Germany]) and body weight (to the nearest 100g by portable scale [Seca, Hamburg, Germany]). The body mass index was calculated from these measurements. Participants had been asked to fast overnight before the clinic visit and were

offered conventional glucose tolerance testing with a 75-mg oral glucose challenge provided in the fasting condition. Glucose was assayed in promptly separated serum by the hexokinase reaction on a Modular P800 (Roche, Boehringer Mannheim, Germany); the interassay coefficient of variation was 1.3-1.8%. The lowest detectable amount was 3mU/L. Type 2 diabetes was defined by a diabetes history with insulin treatment or a fasting glucose value of ≥126 mg/dl (7 mmol/l) or a 2-h post-challenge glucose value ≥ 200mg/dl (11.1 mmol/l). Serum total cholesterol, HDL cholesterol and triglycerides were collected from the baseline fasting blood draw and immediately measured by standard methods[56-58] and LDL cholesterol was calculated for individuals with a triglyceride concentration lower than 400mg/dL using the Friedewald formula[59].

Library generation

Genomic DNA was isolated using the salting out method. We used Reduced Representation Bisulfite Sequencing (RRBS) of which the characteristics and a detailed protocol was published[60,61]. In short, half a microgram of genomic DNA was digested with MspI (NEB). The digestion was end-repaired using Klenow 3'-5' exo$^-$ (NEB) with a mix of 10mM methylated dCTP (Trilink), 10mM dGTP and 80mM dATP (Invitrogen). Sequencing adapters (ATDbio) with all cytosines replaced by methylcytosines were ligated to the end-repaired digestion with T4 ligase (NEB). Next, the adapter ligated digestion was run on a 2.5% 3:1 Nusieve gel (Lonza) and all fragments between 150 and 220 bp were cut out and cleaned (~40-120bp MspI fragment size). Two rounds of bisulfite conversion were performed using the Epitect 5h FFPE protocol (Qiagen). After the consecutive bisulfite treatments the 150-220 bp library was PCR amplified using Pfu Turbo Cx polymerase (Agilent Technologies) using Illumina GAIIx PCR primers (5 min at 95^{o}C, *n* x [30sec. at 95 ^{o}C, 20sec. at 65 ^{o}C, 30sec at 72 ^{o}C], 7 min at 72 ^{o}C, with n ranging from 11-19 cycles) and the resulting product was cleaned and then again put on a 2.5% 3:1 Nusieve gel (Lonza) for a final cleaning and the necessary removal of the PCR primers.

Sequencing and DNA methylation calling

Each of the 48 individual libraries was sequenced on a single Illumina GAIIx lane in the Broad Institute. An average of 25.6 million high quality 36bp single end reads were obtained for each individual for which on average 74.1% (SD 10.4%) could be uniquely aligned to the bisulfite converted human genome (hg19) using custom scripts described previously[61]. Cytosines outside the CpG dinucleotide context were used to assess the bisulfite conversion rate. The average bisulfite conversion rate was 98.9% (SD 0.68%). The number of unmethylated and methylated cytosines per measured CpG dinucleotide was determined from the .BAM alignment files using a custom python script[61]. Global mean methylation was assessed by calculating the average DNA methylation of all CpGs in all accepted reads. We also aligned all high quality reads to the prototypic repeat sequences in the RepBase Update database and again determined the amount of methylated and total number of reads per consensus sequence CpG using custom software[62]. For each repeat type the number of methylated and unmethylated reads were summed, repeats with a median coverage higher than 5 reads over all the measured sibling pairs were included in our analysis.

Data grouping to genomic annotations

Particularly when smaller effect sizes are anticipated and the study size is limited, the focus on annotated regions encompassing multiple CpG dinucleotides is superior to testing single CpG dinucleotides. For it enhances the biological relevance and interpretability of results and is statistically more efficient by reducing the number of tests and increasing the precision of DNA methylation estimates[63]. The location of CpG dinucleotides was first lifted to hg18 (NCBI36) since at the start of the study most annotations were available for hg18. All CpG dinucleotides were then mapped for an overlap with genomic features such as promoters and 'bonafide' CpG islands[23], which are CpG islands (CGIs) with a ubiquitously open chromatin conformation. Most of the genomic feature annotations were taken from *Gu et al.*[16] (available

from http://rrbs-techdev.computational-epigenetics.org). These annotations were supplemented with genome-wide annotations of H3K4me1 and H3K4me3 associated enhancers[24], CTCF binding sites[29] and some additions of particular interest for early development or prenatal environmental exposure. We included the annotations of bivalent chromatin domains from human embryonic stem cells (hESCs)[64] and human hematopoietic stem cells (HSCs)[65], highly variable regions[26], loci hypothesized to be sensitive to early nutrition (putative metastable epialleles)[25] and enhancer regions shown to be associated with genes involved in pre- and peri-implantation development (dev. Enhancers Type I) or early differentiation stages (dev. Enhancers type II)[44].

Individual CpG sites were mapped to a specific genomic locus contained within the annotations of genomic features (e.g. mapped to a particular promoter or other feature) when they had an overlap in terms of their genomic location. Each locus was subsequently mapped to the nearest entrez gene identifier within 100kb. Regions were denoted as intergenic if the distance the nearest gene was larger than 100kb.

GlobalTest for genomic annotation and pathway analysis

We used the R package GlobalTest[66] to test genomic annotation as a whole and to test groups of individual regions mapped to gene-sets. For this test DNA methylation values within a given genomic region were transformed to account for missing values and to account for the differences in the total coverage (thus accuracy of the measurement) and the number of CpG sites per locus. The transformation shrinks the fraction of methylated reads towards the average methylation fraction in the population of 48 subjects. This shrinkage is especially pronounced in subjects with few reads. The transformation was performed as follows:

$M_{i,j}$ Denotes the number of methylated reads for individual *i*, CpG dinucleotide *j* for a particular region.

$T_{i,j}$ Total number of reads for individual *i*, CpG dinucleotide *j* for a particular region.

k The number of CpG sites for a given region (=j_{max})

$M_{i\bullet} = \sum_{j=1}^{k} M_{i,j}$ Denotes the sum of all methylated reads for individual *i* for a particular region.

$T_{i\bullet} = \sum_{j=1}^{k} T_{i,j}$ Denotes the sum of all reads for individual *i* for a particular region.

$$region = \frac{M_{i\bullet} + 5 \times \left(\frac{\sum_{i=1}^{i=48} M_{i\bullet}}{\sum_{i=1}^{i=48} T_{i\bullet}} \right)}{T_{i\bullet} + 5} \times \sqrt{\frac{median(T_{1\bullet},..,T_{48\bullet})}{k}}$$

As part of the validation of this approach we tested the transformed values of the regions that were associated to prenatal famine exposure and a set of loci that were not or only nominally associated in our earlier candidate studies in this population[11,12] by GlobalTest. The positive set was again associated with prenatal famine exposure (P=9x10^{-3}) and the negative set was not (P=0.19).

Epityper data generation and pre-processing

Primers were designed using Methprimer[67]. The resulting primer and amplicons locations were checked against the latest version of dbSNP for SNPs. Their spectrum characteristics were checked with the R package RSeqMeth[68]. The sequences of the primers used in our study and the genomic locations they amplify are given in **Supplement II, Table S5**. The regions measured are necessarily smaller than the regions identified by RRBS, the

genomic overlap and the number of identical CpG dinucleotides measured is given in **Supplement II, Table S6.**

One microgram of genomic DNA isolated using the salting-out method from whole blood was bisulfite treatment EZ 96-DNA methylation kit (Zymo Research) with overnight bisulfite incubation according to the supplier's protocol. The 60 sibling pairs were randomly distributed over two 96 well plates with similar proportions of male and female pairs on each plate and in similar proportions for the pairs also measured with RRBS and those who were not. DNA methylation was quantitatively assessed for each locus using the mass spectrometry based Epityper assay (Sequenom, USA) in triplicate using the manufacturers' protocol on one 384 well plate. PCR was performed with the following cycling protocol: 15 minutes at 95^{o}C, 4 rounds of 20 seconds at 95^{o}C, 30 seconds at 65^{o}C, 1 minute at 72^{o}C; followed by 40 rounds, 20 seconds at 95^{o}C, 30 seconds at (see **Supplement II, Table S5**) ^{o}C and 1 minute at 72^{o}C; ending with 3 minutes at 72^{o}C. Processing of the Epityper data has been described in detail previously[11,12,69,70]. A bisulfite converted DNA is processed on the same 384 plate in triplicate. From the bisulfite conversion onward we perform all steps inherent to the Epityper protocol 3 times. Fragments containing CpG dinucleotides ("CpG units") that have a mass within the mass range that do not overlap other CpG units were considered. Fragments were also discarded if dbSNP indicated the possible presence of a SNP in individuals of European decent with a minor allele frequency higher than 5%. In short, measurements for CpG dinucleotides containing fragments for which 2 out of 3 measurements were successful, the standard deviation of these multiple measurements was smaller than 10% and for which the overall measurement success rate in the 60 pairs was higher than 75% were included in the final analyses. Averages for these triplicate measurements were used for the analyses. For each measurement non-bisulfite converted genomic DNA and negative controls were incorporated to check for a-specific amplification and PCR artifacts, none were found. Bisulfite conversion was assessed using direct bisulfite sequencing on an ABI sequencer and was ≥98% for both 96 well plates.

Statistical tests for individual regions

We tested for within-pair differences in DNA methylation between exposed individuals and their same-sex siblings for individual loci by applying generalized linear mixed models on the sequencing data and linear mixed models for the Epityper data. With these models the correlation between adjacent CpG sites in an individual can be taken into account and all available raw but incomplete data can be used for modeling and control for possible confounders. The R programming environment was used for all analyses[71].
For the bisulfite sequencing data we used logistic mixed models using the glmer() function from the lme4 package[72] with a binomial distribution, weighing by the sequencing depth per individual observation. The dependent variable was the DNA methylation fraction. This in effect is the same as modeling each individual read as either 0 or 1 (unmethylated or methylated) in the dependent without weighing for coverage depth, hence the application of a logistic model. Exposure status (exposed vs. unexposed) and a unique identifier for each CpG dinucleotide were entered as fixed effects. To specify a within-sib-pair design, the (family) pair identifier was included as a random effect with intercept. To model the correlation in DNA methylation within an individual we make use of the fact that each family consists of an exposed individual and a same-sex sibling, therefore adding the exposure status to the model as a random slope, possibly correlated to the random intercept. This model is equivalent to one in which the individual identifier was added as a random effect. This model option allows us to use the same model for both multiple CpG sites and single CpG sites, allowing a unified analysis pipeline.
For the continuous Epityper data we used a linear mixed model based on the lmer() function from the same lme4 package, applying the same model used for the RRBS data, but now without the necessity to weigh for coverage and using a normal distribution. The models were fitted by REML or, when model fits were compared, by ML. The outcome of this model is identical to a paired t-test when an individual CpG site is assessed, there is no missing data and no corrections are applied.

We added bisulfite batch as fixed effect, since the 60 pairs were distributed across 2 96 well plates for bisulfite treatment (keeping pairs on the same plate). Since the age difference is larger for some of the pairs in this set, age at blood drawn was also entered as fixed effect. We are able to effectively correct for age because the siblings without prenatal famine exposure are equally distributed between being born before and after the war period. Additional corrections were performed by adding these respective variables as fixed effects to the model. For current diet we had data on the amount of kcal/day consumed and the percentage of fat, carbohydrates and protein in the diet. Interactions were tested by adding an interaction term as a fixed effect, always including the main terms. Models incorporating DNA methylation data from multiple loci were extended by removing exposure status as intercept and instead adding a random effect for individual with a nested random effect denoting the region. Normality of the Epityper data was checked by histograms of the raw data and the lmer() model residuals. Model fits were diagnosed by plotting the residuals against the fitted values and comparing the variance of the residuals across the exposed and unexposed and other factorial covariates.

Multiple testing correction was performed according to the method developed by Bejamini and Hochberg, better known as 'FDR' (false discovery rate) correction using the R base 'p.adjust()' function. All p-values reported are two-sided. Reported confidence intervals are at 95%, without adjustment for multiple testing. The replication rate was calculated using the closed testing procedure based on the Simes inequality as described by Goeman and Solari[73]. The validation rate is an estimate of π_1, the number of correctly rejected null hypotheses, and since its confidence interval is the most informative only this was reported in the Results section. The π_1 itself was 6/6.

Acknowledgements

We express our gratitude to the participants of the Dutch Hunger Winter Families study and the staff of TNO Quality of Life for contact tracing. At the Leiden University Medical Center we wish to acknowledge the staff of the department of Gerontology and Geriatrics Study Center for the physical examinations and the Central Clinical Chemical Laboratory for extracting DNA.

This study was supported by a Horizon Breakthrough grant in the framework of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research [93518027 to BTH], the European Union-funded Network of Excellence LifeSpan [FP6 036894 from which support was given to PES, BTH and EWT], the U.S.A.'s National Institute of Aging [R01AG042190 to LHL and BTH], The European Union's Seventh Framework Program IDEAL [259679 to PES], by the U.S. National Institutes of Health [R01-HL067914 to LHL] and the Netherlands Consortium for Healthy Ageing [#05060810 to PES]. The funders had no role in study design, data collection, analysis, decision to publish or preparation of the manuscript.

References

1. Gluckman PD, Hanson MA, Bateson P, Beedle AS, Law CM, Bhutta ZA, et al. Towards a new developmental synthesis: adaptive developmental plasticity and human disease. Lancet 2009; 373: 1654-7
2. Bogdarina I, Haase A, Langley-Evans S, Clark AJ. Glucocorticoid effects on the programming of AT1b angiotensin receptor gene methylation and expression in the rat. PLoS.One. 2010; 5: e9237
3. Morgan HD, Jin XL, Li A, Whitelaw E, O'Neill C. The culture of zygotes to the blastocyst stage changes the postnatal expression of an epigentically labile allele, agouti viable yellow, in mice. Biol.Reprod. 2008; 79: 618-23
4. Zhang S, Rattanatray L, MacLaughlin SM, Cropley JE, Suter CM, Molloy L, et al. Periconceptional undernutrition in normal and overweight ewes leads to increased adrenal growth and epigenetic changes in adrenal IGF2/H19 gene in offspring. FASEB J. 2010; 24: 2772-82
5. Lumey LH, Stein AD, Susser E. Prenatal famine and adult health. Annu.Rev.Public Health 2011; 32: 237-62
6. Lumey LH, Stein AD, Kahn HS, Romijn JA. Lipid profiles in middle-aged men and women after famine exposure during gestation: the Dutch Hunger Winter Families Study. Am.J.Clin.Nutr. 2009; 89: 1737-43
7. Lussana F, Painter RC, Ocke MC, Buller HR, Bossuyt PM, Roseboom TJ. Prenatal exposure to the Dutch famine is associated with a preference for fatty foods and a more atherogenic lipid profile. Am.J.Clin.Nutr. 2008; 88: 1648-52
8. Susser E, Neugebauer R, Hoek HW, Brown AS, Lin S, Labovitz D, et al. Schizophrenia after prenatal famine. Further evidence. Arch. Gen.Psychiatry 1996; 53: 25-31
9. Mamun AA, O'Callaghan MJ, Williams GM, Najman JM. Maternal smoking during pregnancy predicts adult offspring cardiovascular risk factors - evidence from a community-based large birth cohort study. PLoS.One. 2012; 7: e41106
10. Dabelea D, Pettitt DJ. Intrauterine diabetic environment confers risks for type 2 diabetes mellitus and obesity in the offspring, in addition to genetic susceptibility. J.Pediatr.Endocrinol. Metab 2001; 14: 1085-91
11. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl.Acad.Sci.U.S.A 2008; 105: 17046-9
12. Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, et al. DNA Methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum.Mol.Genet. 2009; 18: 4046-53
13. Soubry A, Murphy S, Huang Z, Murtha A, Schildkraut J, Jirtle R, et al. The effects of depression and use of antidepressive medicines during pregnancy on the methylation status of the IGF2 imprinted control regions in the offspring. Clin.Epigenetics. 2011; 3: 2
14. Bouchard L, Thibault S, Guay SP, Santure M, Monpetit A, St-Pierre J, et al. Leptin gene epigenetic adaptation to impaired glucose metabolism during pregnancy. Diabetes Care 2010; 33: 2436-41
15. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 2012; 484: 339-44
16. Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, et al. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. Nat.Methods 2010; 7: 133-6
17. Heijmans BT, Tobi EW, Lumey LH, Slagboom PE. The epigenome: archive of the prenatal environment. Epigenetics. 2009; 4: 526-31
18. Lumey LH, Stein AD, Kahn HS, van der Pal-de Bruin KM, Blauw GJ, Zybert PA, et al. Cohort profile: the Dutch Hunger Winter families study. Int.J.Epidemiol. 2007; 36: 1196-204
19. Li Y, Zhu J, Tian G, Li N, Li Q, Ye M, et al. The DNA methylome of human peripheral blood mononuclear cells. PLoS.Biol. 2010; 8: e1000533
20. Lumey LH, Terry MB, Delgado-Cruzata L, Liao Y, Wang Q, Susser E, et al. Adult global DNA

methylation in relation to pre-natal nutrition. Int.J.Epidemiol. 2012; 41: 116-23

21. Tobi EW, Slagboom PE, van DJ, Kremer D, Stein AD, Putter H, et al. Prenatal Famine and Genetic Variation Are Independently and Additively Associated with DNA Methylation at Regulatory Loci within IGF2/H19. PLoS.One. 2012; 7: e37933
22. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet.Genome Res. 2005; 110: 462-7
23. Bock C, Walter J, Paulsen M, Lengauer T. CpG island mapping by epigenome prediction. PLoS.Comput.Biol. 2007; 3: e110
24. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 2009; 459: 108-12
25. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. PLoS.Genet. 2010; 6: e1001252
26. Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci. Transl.Med. 2010; 2: 49ra67
27. Bock C, Halachev K, Buch J, Lengauer T. EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi) genomic data. Genome Biol. 2009; 10: R14
28. Andersson R, Enroth S, Rada-Iglesias A, Wadelius C, Komorowski J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. 2009; 19: 1732-41
29. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell 2007; 129: 823-37
30. Pekowska A, Benoukraf T, Ferrier P, Spicuglia S. A unique H3K4me2 profile marks tissue-specific gene regulation. Genome Res. 2010; 20: 1493-502
31. Zhang P, Zucchelli M, Bruce S, Hambiliki F, Stavreus-Evers A, Levkov L, et al. Transcriptome profiling of human pre-implantation development. PLoS.One. 2009; 4: e7844
32. Yi H, Xue L, Guo MX, Ma J, Zeng Y, Wang W, et al. Gene expression atlas for human embryogenesis. FASEB J. 2010; 24: 3341-50
33. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. Proc.Natl.Acad.Sci.U.S.A 2005; 102: 15785-90
34. Schaniel C, Ang YS, Ratnakumar K, Cormier C, James T, Bernstein E, et al. Smarcc1/Baf155 couples self-renewal gene repression with changes in chromatin structure in mouse embryonic stem cells. Stem Cells 2009; 27: 2979-91
35. Kleger A, Mahaddalkar PU, Katz SF, Lechel A, Joo JY, Loya K, et al. Increased reprogramming capacity of mouse liver progenitor cells, compared with differentiated liver cells, requires the BAF complex. Gastroenterology 2012; 142: 907-17
36. Ozair MZ, Noggle S, Warmflash A, Krzyspiak JE, Brivanlou AH. SMAD7 Directly Converts Human Embryonic Stem Cells to Telencephalic Fate by a Default Mechanism. Stem Cells 2012; 1: 35-47
37. Heard ME, Pabona JM, Clayberger C, Krensky AM, Simmen FA, Simmen RC. The reproductive phenotype of mice null for transcription factor kruppel-like factor 13 suggests compensatory function of family member kruppel-like factor 9 in the peri-implantation uterus. Biol.Reprod. 2012; 87: 115
38. Smart NG, Apelqvist AA, Gu X, Harmon EB, Topper JN, MacDonald RJ, et al. Conditional expression of Smad7 in pancreatic beta cells disrupts TGF-beta signaling and induces reversible diabetes mellitus. PLoS.Biol. 2006; 4: e39
39. Zammit VA, Ramsay RR, Bonomini M, Arduini A. Carnitine, mitochondrial function and therapy. Adv.Drug Deliv.Rev. 2009; 61: 1353-62
40. Natesampillai S, Fernandez-Zapico ME, Urrutia R, Veldhuis JD. A novel functional interaction between the Sp1-like protein

KLF13 and SREBP-Sp1 activation complex underlies regulation of low density lipoprotein receptor promoter function. J.Biol.Chem. 2006; 281: 3040-7

41. Stoger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? Bioessays 2008; 30: 156-66

42. Suter M, Ma J, Harris AS, Patterson L, Brown KA, Shope C, et al. Maternal tobacco use modestly alters correlated epigenome-wide placental DNA methylation and gene expression. Epigenetics. 2011; 6: 1284-94

43. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS.One. 2011; 6: e21800

44. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature 2011; 470: 279-83

45. Waterland RA, Jirtle RL. Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. Nutrition 2004; 20: 63-8

46. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, et al. Transcription is required for establishment of germline methylation marks at imprinted genes. Genes Dev. 2009; 23: 105-17

47. Waterland RA, Kellermayer R, Rached MT, Tatevian N, Gomes MV, Zhang J, et al. Epigenomic profiling indicates a role for DNA methylation in early postnatal liver development. Hum.Mol.Genet. 2009; 18: 3026-38

48. Macklon NS, Geraedts JP, Fauser BC. Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. Hum.Reprod. Update. 2002; 8: 333-43

49. Stein ZA, Susser M, Saenger G, Marolla F. Famine and Human Development: The Dutch Hunger Winter of 1944-1945. New York: Oxford University Press, 1975

50. Kwong WY, Wild AE, Roberts P, Willis AC, Fleming TP. Maternal undernutrition during the preimplantation period of rat development causes blastocyst abnormalities and programming of postnatal hypertension. Development 2000; 127: 4195-202

51. Julius U. Influence of plasma free fatty acids on lipoprotein synthesis and diabetic dyslipidemia. Exp.Clin.Endocrinol.Diabetes 2003; 111: 246-50

52. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat.Rev.Genet. 2011; 12: 529-41

53. McKay JA, Xie L, Harris S, Wong YK, Ford D, Mathers JC. Blood as a surrogate marker for tissue-specific DNA methylation and changes due to folate depletion in post-partum female mice. Mol.Nutr.Food Res. 2011; 55: 1026-35

54. Trienekens G. The food supply in the Netherlands during the Second World War, In: Smith DF, Phillips, eds. Food, science, policy and regulation in the twentieth century. International and comparative perspectives. London: Routledge, 2000: 117-33

55. Burger GCE, Drummond JC, Sandstead HR. Malnutrition and Starvation in Western Netherlands, September 1944-July 1945. The Hague: General State Printing Office, 1948

56. Rifai N, Cole TG, Iannotti E, Law T, Macke M, Miller R, et al. Assessment of interlaboratory performance in external proficiency testing programs with a direct HDL-cholesterol assay. Clin.Chem. 1998; 44: 1452-8

57. Stinshoff K, Weisshaar D, Staehler F, Hesse D, Gruber W, Steier E. Relation between concentrations of free glycerol and triglycerides in human sera. Clin.Chem. 1977; 23: 1029-32

58. Sugiuchi H, Uji Y, Okabe H, Irie T, Uekama K, Kayahara N, et al. Direct measurement of high-density lipoprotein cholesterol in serum with polyethylene glycol-modified enzymes and sulfated alpha-cyclodextrin. Clin.Chem. 1995; 41: 717-23

59. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin.Chem. 1972; 18: 499-502

60. Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat.Protoc. 2011; 6: 468-81

61. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale

DNA methylation maps of pluripotent and differentiated cells. Nature 2008; 454: 766-70

62. Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, et al. Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat.Biotechnol. 2010; 28: 1106-14

63. Bock C. Analysing and interpreting DNA methylation data. Nat.Rev.Genet. 2012; 13: 705-19

64. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, et al. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 2007; 1: 286-98

65. Cui K, Zang C, Roh TY, Schones DE, Childs RW, Peng W, et al. Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. Cell Stem Cell 2009; 4: 80-93

66. Goeman JJ, van de Geer SA, de KF, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics. 2004; 20: 93-9

67. Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002; 18: 1427-31

68. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ. Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Res. 2007; 35: e119

69. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J. 2010; 24: 3135-44

70. Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, Finken MJ, et al. DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. Epigenetics. 2011; 6: 171-6

71. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2012

72. Bates D, Maechler M, Bolker B. lme4: Linear mixed-effects models using S4 classes, 2012

73. Goeman JJ, Solari A. Multiple testing for exploratory research. Statistical Science 2011; 26: 584-97