Differential regulation of the DNA methylome in adults born during the Great Chinese Famine in 1959-1961

Li, Shuxia

2021-11

Li , S , Wang , W , Zhang , D , Li , W , Lund , J , Kruse , T , Mengel-From , J , Christensen , K & Tan , Q 2021 , ' Differential regulation of the DNA methylome in adults born during the Great Chinese Famine in 1959-1961 ' , Genomics , vol. 113 , no. 6 , pp. 3907-3918 . https://doi.org/10.1016/j.ygeno.2

http://hdl.handle.net/10138/339568
https://doi.org/10.1016/j.ygeno.2021.09.018

Genomics 113 (2021) 3907–3918

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Differential regulation of the DNA methylome in adults born during the Great Chinese Famine in 1959–1961

Shuxia Li [a,*,1], Weijing Wang [b,1], Dongfeng Zhang [b], Weilong Li [a,c], Jesper Lund [a,d], Torben Kruse [e], Jonas Mengel-From [a], Kaare Christensen [a], Qihua Tan [a,e,*]

[a] Epidemiology and Biostatistics, Department of Public Health, University of Southern Denmark, Odense, Denmark
[b] Qingdao University School of Public Health, Qingdao, China
[c] Population Research Unit, Faculty of Social Sciences, University of Helsinki, Finland
[d] Digital Health & Machine Learning Research Group, Hasso Plattner Institute for Digital Engineering, Potsdam, Germany
[e] Unit of Human Genetics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

ARTICLE INFO

Keywords:
Fetal adversity
Chinese famine
DNA methylation
Association

ABSTRACT

Background: Extensive epidemiological studies have established the association between exposure to early-life adversity and health status and diseases in adults. Epigenetic regulation is considered as a key mediator for this phenomenon but analysis on humans is sparse. The Great Chinese Famine lasting from 1958 to 1961 is a natural string of disasters offering a precious opportunity for elucidating the underlying epigenetic mechanism of the long-term effect of early adversity.

Methods: Using a high-throughput array platform for DNA methylome profiling, we conducted a case-control epigenome-wide association study on early-life exposure to Chinese famine in 79 adults born during 1959–1961 and compared to 105 unexposed subjects born 1963–1964.

Results: The single CpG site analysis of whole epigenome revealed a predominant pattern of decreased DNA methylation levels associated with fetal exposure to famine. Four CpG sites were detected with $p < 1e\text{-}06$ (linked to *EHMT1*, *CNR1*, *UBXN7* and *ESM1* genes), 16 CpGs detected with $1e\text{-}06 < p < 1e\text{-}05$ and 157 CpGs with $1e\text{-}05 < p < 1e\text{-}04$, with a predominant pattern of hypomethylation. Functional annotation to genes and their enriched biological pathways mainly involved neurodevelopment, neuropsychological disorders and metabolism. Multiple sites analysis detected two top-rank differentially methylated regions harboring *RNF39* on chromosome 6 and *PTPRN2* on chromosome 7, both showing epigenetic association with stress-related conditions.

Conclusion: Early-life exposure to famine could mediate DNA methylation regulations that persist into adulthood with broad impacts in the activities of genes and biological pathways. Results from this study provide new clues to the epigenetic embedding of early-life adversity and its impacts on adult health.

1. Introduction

Susceptibility to certain common diseases may have been embeded in the genome through epigenetic reprogramming by early-life adversity [14,16,41]. Poor fetal and early postnatal nutrition during critical growth phases may alter the structural and physiologic functional development of vital organs thus "program" the susceptibility to diseases in adulthood. Animal studies have provided strong evidence that poor fetal nutrition leads to low birth weight in offspring and increased blood pressure in adulthood [4,25] and altered glucose tolerance [26]. In humans, however, direct experimental evidence on early-life adversity and adult health and diseases is not available due to ethical and practical issues. An opportunity for conducting human studies is to assess the impact of a natural disaster, when fetus and infant suffered from the adverse effects of food shortage and undernutrition [42,44]. Famine provides quasi-experimental conditions to examine the effect of early-life adversity on epigenetic "programing" and health consequences in humans. One example is studies on the health effects of early-

* Corresponding author at: Epidemiology and Biostatistics, Dept of Public Health, Faculty of Health Science, University of Southern Denmark, J. B. Winsløws Vej 9B, DK-5000 Odense C, Denmark.
E-mail addresses: sli@health.sdu.dk (S. Li), weilong.li@helsinki.fi (W. Li), jlund@health.sdu.dk (J. Lund), torben.kruse@rsyd.dk (T. Kruse), jmengel-from@health.sdu.dk (J. Mengel-From), kchristensen@health.sdu.dk (K. Christensen), qtan@health.sdu.dk (Q. Tan).
[1] Equal contributions.

https://doi.org/10.1016/j.ygeno.2021.09.018
Received 23 March 2021; Received in revised form 24 August 2021; Accepted 25 September 2021
Available online 29 September 2021

life exposure to the Dutch Famine in World War II from November 1944 to April 1945 [10,43,45,55,64]. The Dutch famine lasted about 6 months with relatively short exposure by victims. However, the Dutch famine occurred in the context of an otherwise well-nourished population that may have been somewhat buffered. The Great Chinese Famine between the spring of 1959 and the end of 1961 was the largest in human history spanning approximately 3 years [50]. All provinces in China were affected with more devastating conditions in the rural areas. In contrast to the relatively brief Dutch famine, the exceptionally long Chinese famine provides a unique opportunity to examine the "programming" effects of poor nutrition during fetal development and infancy in modifying the risk of adult-onset diseases. With increasing available data on the cohorts affected, there has been an emerging body of epidemiological studies on the long-term health consequences of early-life exposure to the Great Chinese Famine covering overweight [67], metabolic syndrome [30], type II diabetes [31], hypertension [65], schizophrenia [53,54,70] and mortality [7,38,52]. Overall, the studies suggest that exposure to the Great Chinese Famine during pregnancy and/or infancy increased the risk of common diseases in adulthood. Unfortunately only few studies have investigated the molecular basis of the "programing" of health effects in Chinese famine [17,49].

Different from traditional epidemiologic approaches that seek observational evidence by associating environmental exposure with disease outcome, epigenetic epidemiology searches for molecular mediators between environmental factors (especially early-life events) and functional regulation of genes and diseases with epigenetics serving as the link between nature (genome) and nurture (environment) [59,60]. The unique "natural experiment" of Great Chinese Famine [50] is ideal for conducting unmatched epigenetic studies on the molecular basis for the reported association between early-life adversity and adult-age health advantaged by the large population affected and the long exposure time (about 3 years). By using a microarray technology platform for high throughput genomic DNA methylation analysis, we performed an epigenome-wide association study (EWAS) to identify and characterize the epigenetic signatures of prenatal exposure to the Chinese famine. This paper reports findings from the study to provide potential epigenetic insights for the reported epidemiological observations concerning early-life adversity and adult-age health and diseases.

2. Methods

2.1. Study samples

Participants of the study were randomly recruited from a large project on diabetes prevention jointly at Qingdao Center for Disease Control and Prevention and Qingdao University Medical College, China. The population-based cross-sectional surveys on diabetes prevention were conducted in 2001–2002 and 2006 in Qingdao, China [40]. We first randomly recruited 186 subjects born during famine (born from Jan. 1, 1959 to Dec. 31, 1961) as cases and then recruited 401 subjects born after famine (born from Jan. 1, 1963 to Dec. 31, 1964) as controls, with date of birth based on civil registration (Supplementary Fig. S1). The controls were matched with cases by village, residential, and economic factors. Two-hundred seventy-eight participants with a history of hypertension, diabetes, chronic obstructive pulmonary disease, cancer, stroke, severe mental disorders, tuberculosis, hepatitis, other infectious diseases, and occupational diseases were excluded. Fifteen participants having myocardial infarction, stroke, chronic obstructive pulmonary disease, and cancer during the survey were further dropped. Based on the clinical indicators from examination, 110 subjects with diabetes or hypertension were also excluded. The remaining samples were interviewed in 2017 when blood samples were taken for blood testing and stored under −80 °C for DNA methylation analysis. In total, 79 participants who were born from January 3, 1959 to December 21, 1961 during the famine and 105 participants who were born after the famine (at least one year later than the exposure group) from January 18, 1963 to May 29, 1964 were collected for the study. In addition to exposure to famine, multiple anthropometric including age, sex, weight, height, BMI, marriage, education, and health variables including blood pressure, blood glucose, triglycerides, total cholesterol, high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C) were collected. This study was approved by the Ethics Committee of Qingdao CDC with approval number WDF-07-308. Written informed consent was obtained from each participant. The study was conducted in accordance with the principles of the Helsinki Declaration.

2.2. Quantification of DNA methylation

Genomic DNA (0.2–1.0 μg) was bisulfite converted using an EZ DNA Methylation-Direct Kit (Zymo Research). DNA samples were bisulfite converted by incubation with the CT conversion reagent for 8 min at 98 °C, 3.5 h at 64 °C, followed by 18 h at 4 °C in a thermocycler. The treated DNA was added to a Zymo-Spin IC Column, desulfonated using M-desulphonation buffer, and then eluted from the column in 12 μl of M-elution buffer.

Methylation profiling of the bisulfite-treated DNA was performed using Illumina Infinium MethylationEPIC BeadChip (Illumina) according to standard protocol. In brief, 4 μl of bisulfite-treated DNA was denatured, neutralized and amplified with an overnight whole-genome amplification reaction. The amplified DNA was then enzymatically fragmented, precipitated and re-suspended in hybridization buffer before being dispensed onto the MethylationEPIC BeadChips for hybridization. After hybridization, the BeadChips were processed through a primer-extension protocol and subsequently stained. Finally, the BeadChips were coated and imaged using Illumina's HiScan System. DNA methylation analysis was performed by BioMiao Biological Technology (Beijing) Co., Ltd., 100044, Beijing, China.

2.3. Data preprocessing

The raw methylation data cover 866,091 CpG sites across the genome of each individual. The R package *minfi* [1] was used for data preprocessing including quality control (QC) and normalization, with the pipeline for raw data processing illustrated in Supplementary Fig. S2. For each CpG site of a sample, a detection *p*-value was first calculated as 1 − *p*-value computed from the background model characterizing the chance that the signal was distinguishable from negative controls. We filtered out 2386 CpGs with detection *p*-value >0.01 in more than 5% of the overall samples (i.e. 9 samples). After QC, a total of 863,705 CpG sites remained. We further removed all CpGs on the sex chromosomes and CpGs physically overlapping with SNPs leaving 695,536 CpGs. QC at sample level was done by plotting the log median intensity in the methylated (M) against that of the unmethylated (U) channels using getQC and plotQC functions in *minfi*, with no bad sample found using a default cut-off. Data normalization was performed using subset-quantile within array normalization (SWAN) [33] implemented in R package *minfi*. At each CpG site, DNA methylation level was summarized by calculating a methylation "beta" value defined by the Illumina's formula as $\beta = M/(M + U + 100)$. Before statistical testing, we filtered out CpGs of very low methylation variation by calculating standard deviations (SD) based on methylation beta values [66] and dropping CpGs with SD < 0.02 with 487,229 CpGs remaining for subsequent analysis. Before statistical analysis, the β-values were converted to methylation M-values for better statistical properties by logit transformation with $M = \log_2(\beta/(1- \beta))$ [11].

2.4. Estimating cell-type composition

Since the target tissue is whole blood comprising multiple cell types, cellular heterogeneity among samples can be an important confounding factor in epigenetic association analysis due to cell specificity of DNA methylation. To control for cell-type composition effect, the proportions

of major leukocytes were estimated using the Houseman's method [21] implemented in the R package *minfi*. Based on the DNA methylation data, the method estimated blood cell composition in each individual for 6 blood cell types: CD8T, CD4T, natural killer cell, B cell, monocyte, and granulocyte. The estimated cell type proportions were included in downstream statistical analysis to adjust for cell composition effects.

2.5. Handling age-affected methylation sites

In a systematic literature review on epidemiological studies of the Chinese famine, Li and Lumey [29] demonstrated a common bias in health-related risk estimates for famine exposure using post-famine born controls. The mean age difference of 4 years between the exposed and unexposed groups could also affect DNA methylation levels of age-related methylation sites. To handle this problem, we drop the age-affected methylation sites by.

(1) extracting all 177 samples with ages from 51 to 60 years (about the same sample size and same age range of the studied samples 52–59 years) from a large in-house cohort with DNA methylation data measured by the Illumina EPIC (850 K) methylation array;
(2) performing an epigenome-wide association analysis to find significant age-dependent methylation sites (after correcting for multiple testing);
(3) removing the significant age-affected sites from the famine study and report only results without the age-affected sites.

To detect age-affected CpGs, we apply an assumption-free generalized correlation coefficient to test the correlation between age and methylation [37]. The method is able to capture all patterns of age-dependent DNA methylation unlimited to linear correlation. A total of 18,023 CpGs were found to associate with age after correcting for multiple testing using FDR < 0.05. These CpGs were then removed leaving a total of 469,206 CpGs for subsequent statistical analysis.

2.6. Statistical analysis

For each CpG site, a linear regression model was fitted regressing DNA methylation M-value on famine exposure adjusting for age, sex and cell proportions. Correction for multiple testing on the 469, 206 CpGs was done by calculating the false discovery rate (FDR) [3] and define FDR < 0.05 as genome-wide significant. For CpGs with FDR ≥ 0.05, we define $p <$ 1e-06 for suggestive significance and 1e-06 $\leq p <$ 1e-05 for weaker-than-suggestive significance. In addition to single site testing, we also performed multiple site analysis by introducing the bumphunting approach [22] implemented in R package *minfi*. The bumphunting approach is a region-based analysis that detects and tests differentially methylated regions (DMRs) enriched by multiple CpGs exhibiting same direction of effects. The method starts with regressing the DNA methylation M-value at each CpG site on exposure variables. It assumes that at the population level, the locus-specific slope estimates of the exposure variable are smooth along the strand of a chromosome and applies a smoothing technique to smooth estimated slope coefficients for CpGs within a pre-defined region. We then calculated the 99th percentile of the smoothed coefficients across all regions to obtain upper and lower thresholds. The thresholds were used to define hyper- or hypo-methylated DMRs with smoothed peaks above or below the thresholds. For each identified DMR, the method took the sum of the absolute values of all the smoothed coefficients within that region. This sum statistic was then used to rank all DMRs (high to low). Statistical significance of each DMR was determined by assessing random DMRs from each of 1000 permutations. Adjusting for multiple testing was done by computing the family-wise error rate (FWER) for each DMR area as the proportion of maximum area values per permutation larger than the observed area. A significant DMR was defined as FWER$<$0.1. Empirical uncorrected *p*-value for a single DMR was computed as the proportion of all random DMRs from 1000 permutations that are larger than the area of the observed DMR. The R packages *chromoMap* was used for plotting the differentially methylated CpGs and *Gviz* for plotting the DMRs [15] along the chromosomes.

2.7. EWAS power analysis

We have recently published a computer simulation study on power of EWAS using different study designs [32]. According to the study, even for the lowest simulated effect size of 10% methylation variance explained by the environmental exposure, a sample size of about 80 exposed (total sample with equal number of unexposed $2 \times 80 = 160$, about the same sample size as our famine study) would have a statistical power of 80%. Another power study for EWAS [62] estimated that a sample size of 50 cases (plus equal number of controls $2 \times 50 = 100$) would have a statistical power of 80% in capturing a percentage of methylation difference of 6%. In sum, the power studies on EWAS estimated sufficient power for detecting even low effect size of prenatal famine exposure.

2.8. Functional annotations

The identified CpGs were linked to nearest genes and tested for over-representation of gene-sets (pathways) using the Molecular Signatures Database through Gene Set Enrichment Analysis (GSEA) [56] based on canonical pathways at https://www.gsea-msigdb.org/gsea/index.jsp. The over-representation analysis calculates a probability from the hypergeometric distribution for testing if the submitted list of genes contains more genes from a pathway or gene set than would be expected by chance, i.e.

$$p(X \geq k) = 1 - \sum_{r=0}^{k} \binom{m}{r} \binom{N-m}{n-r} \Big/ \binom{N}{n}$$

where N is the number of genes annotated to all CpGs on the Epic array, m is the number of genes linked to the detected CpGs, n is the number of genes in a particular biological pathway, k is the number of genes belonging to both the pathway under testing and the list of genes linked to the detected CpGs. The test produces a probability score for each pathway or gene set, which is corrected by calculating FDR using the Benjamini-Hochberg method [3].

In addition to the above gene-based over-representation analysis, functional annotation was also conducted by submitting the chromosomal coordinates of the detected DMRs to the Genomic Regions Enrichment of Annotations Tool (GREAT) [36] at (http://bejerano.stanford.edu/great/public/html/) to analyze the functional significance of cis-regulatory regions identified by localized measurements of DNA binding events across an entire genome [36] using the Genome Reference Consortium Human Build 37 (GRCh37) as the RefSeq database. GREAT incorporates annotations from 20 ontologies and associates genomic regions with genes by defining a 'regulatory domain' for each gene such that all non-coding sequences that lie within the regulatory domain are assumed to regulate that gene. The 'two nearest genes' was assigned as the association rule from genomic regions to genes, which extends each gene's regulatory domain from its transcription start site (TSS) to the nearest upstream and downstream TSS, up to 1 MB in each direction. GREAT uses a binomial test over genomic regions to provide a more accurate picture of annotation enrichments [36].

2.9. Replication analysis

For the detected top significant methylation sites, we performed a replication analysis using the Danish birth-weight discordant twin pairs [58] representing an independent adult cohort exposed to neonatal stress from a different population. The sample consisted of 78 pairs of male and 72 pairs of female monozygotic twins with a median age of 57

years, a total of 300 twins. The average percentage of birth-weight discordance (defined as $\frac{\Delta(birthweight)}{max(birthweight)*100}$ for a twin pair) was 17.73% (range: 5.26%–47.06%). DNA methylation data was measured using the Illumina Infinium HumanMethylation450 BeadChip. Raw data pre-processing was described in detail in Tan et al. [58] which used a similar pipeline and criteria as described in the current study. The association of DNA methylation with birthweight discordance was assessed by the mixed effect model assigning birth-weight and discordance (larger or smaller) together with their interaction as fixed effect variables while twin pairing and experimental factors as random effect variables using the R package *lme4*.

3. Results

Table 1 presents the basic description of the study samples and their general health-related parameters from anthropometry, blood pressure to clinical biochemical tests. The exposure and the control groups did not show statistically significant difference in most of the variables. The mean age of unexposed group was 4 years younger than the exposed group. Although weight of the unexposed group was slightly heavier than the exposed group ($p = 0.02$), the BMI did not significantly differ ($p = 0.08$). To minimize the effect age difference between the exposure and control groups, we remove the 18,023 age-associated CpGs and focused on the 469, 206 CpGs in subsequent analysis.

3.1. Single CpG site EWAS

Each CpG was tested by fitting a linear regression model on the DNA methylation M-value and famine exposure adjusting for age, sex and cell proportions. The Manhattan plot and the Q-Q plot for the single CpG site EWAS results are shown in Fig. 1 (1a, 1b) with detailed test statistics showing in Supplementary Table S1 ($p < 0.05$) and Table 2 ($p < 1e-05$, 20 CpGs). After correcting for multiple testing, no CpG reached genome-wide significance as defined by FDR < 0.05. Fig. 2 is a volcano plot for the estimated coefficients for famine exposure against their inverse-log transformed *p*-values. The 4 CpGs (marked red) on the very top of the volcano are cg16187328 ($p = 1.74e-07$, FDR = 0.08), cg06077226 ($p = 3.37e-07$, FDR = 0.08), cg24875889 ($p = 5.43e-07$, FDR = 0.09) and cg20451680 ($p = 8.96e-07$, FDR = 0.11), all show suggestive significance with $p < 1e-06$. There are 16 CpGs (marked purple) detected as showing weaker-than-suggestive significance ($1e-06 \leq p < 1e-05$). One important phenomenon in the estimates was that the CpGs on top of the volcano (colored dots for 177 CpGs with $p < 1e-04$) in Fig. 2 are predominantly hypomethylated (only 8 CpGs hypermethylated) by famine exposure. This is also illustrated by the chromosome map in Fig. 1c displaying the locations of the 177 CpGs with $p < 1e-04$ (blue color for hypomethylated, red color for hypermethylated sites). In Supplementary Fig. S3, we plot the distribution of CpGs hypermethylated (red line) or hypomethylated (blue line) with famine exposure ($p < 1e-04$) over gene regions (S3a for relative proportions, S3b for absolute proportions) or positions relative to CpG island (CGI) (S3c for relative proportion, S3d for absolute proportion). The hypermethylated sites are more allocated to 5′UTR, while hypomethylated sites to gene body and intergenic regions (S3a, S3b). In Supplementary Fig. S3c and S3d, the hypermethylated sites are distributed to CGIs and shelf, but the hypomethylated site to the open sea and other positions. The absolute proportions in Supplementary Fig. S3 b and d show predominant hypomethylation over gene regions and positions.

Table 1
Sample description (count, mean and standard deviation) and comparison of health variables.

Variables	Unexposed ($n = 105$)	Exposed ($n = 79$)	Z score	P value
Age, year	53	57	4.17	3.08E-05
Sex (male)	48	30	0.74	4.59E-01
Height, cm	165 (6.80)	162.05 (6.87)	−1.51	1.30E-01
Weight, kg	68.55 (12.97)	65 (9.28)	−2.32	2.03E-02
BMI, kg/m^2	24.91 (4.55)	24.63 (3.02)	−1.73	8.32E-02
SBP, mmHg	125 (8.82)	127.5 (8.64)	1.08	2.80E-01
DBP, mmHg	79 (5.95)	80 (6.08)	1.29	1.98E-01
Pulse, mmHg	73 (6.14)	72.25 (7.63)	−0.73	4.66E-01
FBG, mmol/L	5.06 (0.68)	5.24 (0.74)	1.25	2.12E-01
TC, mmol/L	5.125 (1.01)	5.35 (0.89)	0.29	7.73E-01
TG, mmol/L	1.48 (1.17)	1.31 (0.87)	−1.07	2.84E-01
HDL-C, mmol/L	1.51 (0.39)	1.53 (0.34)	0.11	9.15E-01
LDL-C, mmol/L	2.11 (0.53)	2.28 (0.56)	0.61	5.43E-01

FBG: fasting blood glucose; TC: total cholesterol; TG: triglyceride; HDL-C: high-density lipoprotein cholesterol; LDL-C: low-density lipoprotein cholesterol; BMI: body mass index.

3.2. Analysis of differentially methylated regions (DMRs)

We performed DMR analysis on 4482 CpGs with $p < 0.005$. A total of 58 regions were enriched with $p < 0.05$. In Supplementary Table S2, we show the list of 887 detected regions ranked by corresponding *p*-values. After controlling for multiple testing, we found only one DMR reached predefined statistical significance (FWER<0.1), but at borderline level (FWER<0.101, $p < 1.24e-04$). For comparison purpose, we also calculated Šidák corrected *p* values for the DMRs which are interestingly very close to our estimated FWERs (Supplementary Table S2). We obtained a Šidák p value for this region as 0.104. The region is located on chromosome 6 with base pair (bp) position 30,039,374-30,039,476. The position falls into the middle of the human MHC region (28,477,797 bp-33,448,354 bp. assembly: GRCh37). Interestingly, there is another detected region nearby the aforementioned DMR on chromosome 6 positioned at 30,039,142 bp-30,039,206 bp (FDR < 0.57, $p < 1.09e-03$). Fig. 3 displays the genomic region harboring the 2 DMRs on chromosome 6. They are hosted by a coding region (exon) of the *RNF39* gene, and also in the vicinity of the *PPP1R11* gene. The rank two DMR in Supplementary Table S2 is located at 157,369,895 bp-157,369,960 bp on chromosome 7 in the intron region of *PTPRN2* gene (FDR < 0.28, $p < 3.99e-04$) (Supplementary Fig. S4).

3.3. Functional annotations

From GREAT analysis, genomic regions of the transcriptional domain covered by CpGs with $p < 0.005$ were enriched by 12 GO biological processes with FDR < 0.05, as shown in Table 3. The GO terms included negative regulation of leukocyte proliferation, ventral spinal cord interneuron fate commitment, spinal cord dorsal/ventral patterning, dendritic cell chemotaxis, etc. Supplementary Fig. S5 is a hierarchical visualization of the GO terms from GREAT in Table 3. As an alternative, 1223 genes were linked to CpGs with $p < 0.005$ and which were submitted to GSEA for over-representation analysis. This resulted in 10 gene sets enriched with FDR < 0.05 (Table 4). The top significant gene sets were dominated by pathways involved in nervous system development, neuron differentiation and progression, etc.

3.4. DNA methylation of the IGF2 and INSR genes

DNA methylation levels of the *IGF2* and *INSR* have been associated with fetus exposure to the Dutch [18,61]. In the current study, we perform an independent replication analysis using a total of 122 CpGs mapped to the *IGF2* gene and 75 CpGs mapped to the *INSR* gene. Fig. 4 plots the distribution of the 122 CpGs in the genomic region of *IGF2* on chromosome 11. Most of them (94 CpGs) are located in the promoter region and less than half (54 CpGs) reside in the gene body, with 41 CpGs sitting in both promoter and gene body of the isoforms of *IGF2* (Fig. 4). Fig. 5 is a volcano plot displaying the coefficients for famine exposure against statistical significance. Because of the 41 CpGs overlapping between gene body and promoter, we tested the methylation patterns associated with famine exposure at the promoter and at the

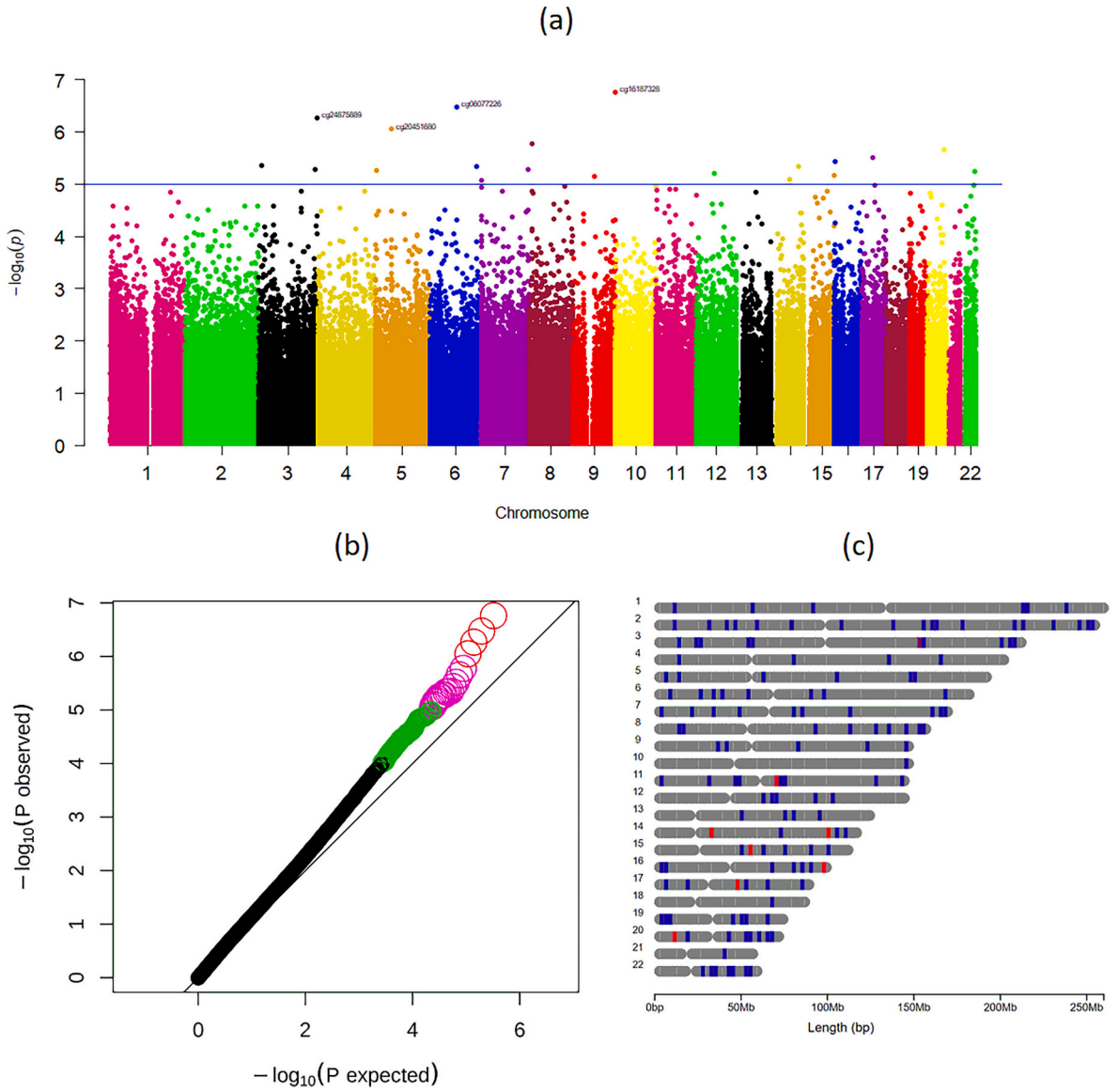

Fig. 1. EWAS results for fetal exposure to famine shown by a Manhattan plot plotting log-scaled *p* value of CpGs against their chromosome locations (1a), by a QQ-plot plotting observed *p*-values (top 177 CpGs with $p <$ 1e-06 colored red, 1e-06 < p < 1e-05 colored pink, 1e-05 < p < 1e-04 colored green) against a random distribution (1b), and by a chromosome location map for the 177 CpGs with blue color indicating hypomethylation, red color indicating hypermethylation (1c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

gene body separately. In the Fig. 5a, the 94 warm colored dots (red for CpGs only from promoter, pink for CpGs from both promoter and gene body) are mostly hypermethylated (65 CpGs, 69.15%) which are in contrast to the 28 cold colored dots (cyan for CpGs only from gene body, black for CpGs from other regions) (10 CpGs, 35.71%). A chi-squared test showed a highly significant difference in the proportions of hyper-methylation at promoter and non-promoter regions ($\chi^2 = 8.82$, $p <$ 2.98e-03). A similar test for gene body versus outside gene body showed no difference in their exposure associated methylation patterns ($\chi^2 =$ 2.6, $p < 0.107$). Fig. 5b displays the results on CpGs of *INSR* gene. The CpGs are mostly from the gene body both hyper- or hypo-methylated with low statistical significance except cg10473041 (hypomethylated at gene body with $p <$ 1.39e-04).

3.5. Replication using Danish birth-weight discordant twins

As an effort for replication, we used birth-weight discordant twin pairs described in Methods to verify the identified sites in Table 2. Among the top 4 CpGs with p < 1e-06, 2 (cg20451680, cg24875889) were matched to the 450 k array. As shown in Supplementary Table S3, both CpGs were hypomethylated showing same direction as in Table 2 with $p =$ 4.16e-03 for cg20451680 and $p = 0.12$ for cg24875889. Among the 20 CpGs with $p <$ 1e-05 (Table 2), 10 were matched to the 450 k array data for Danish twins with eight hypomethylated. Seven out of the 10 CpGs were replicated in the same direction as in Table 2, with 4 of them replicated with $p < 0.09$ (Supplementary Table S3). One CpG, cg15821562, was replicated with $p <$ 5.32e-03 but showing an opposite direction of effect. Overall, the replication showed clear consistency of hypomethylation associated with prenatal adversity in the two independent samples from different populations.

4. Discussion

We have performed a genome-wide analysis of DNA methylation patterns in association with early-life exposure to the largest famine in human history [50] using the Illumina Epic 850 K array. The results

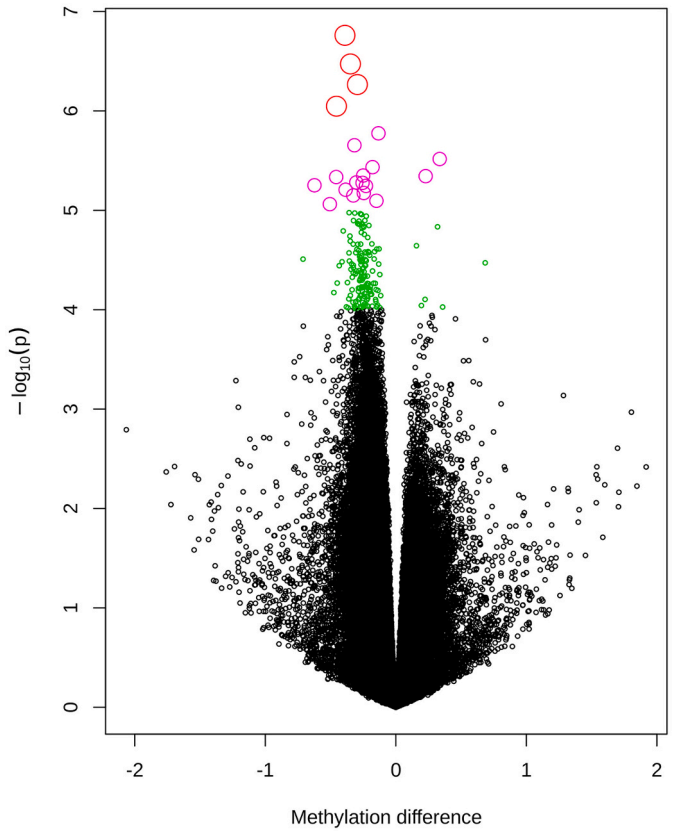

Fig. 2. A volcano plot displaying log-scaled p value of each CpG against difference in M value between exposed and unexposed groups. CpGs with p < 1e-06 colored red, 1e-06 < p < 1e-05 colored pink, 1e-05 < p < 1e-04 colored green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from our epigenetic analysis of this unique natural experiment may shed insight into epigenetic sites which are altered in response to sustained adversity as a result of famine.

According to EWAS results from this study, deprived nutritional environment early in life as induced by fetal exposure to famine is associated with regulatory modification of the epigenome characterized by predominant hypomethylation of the DNA methylome (Fig. 1). As most of the CpG sites are methylated except at CpG islands of promoters in the general population, the pattern of extensive hypomethylation by famine exposure could indicate a systematic epigenetic remodeling in response to early-life stress due to fetal programming. Interestingly, Krishna et al. [24] very recently reported a global hypomethylation pattern in their EWAS on intrauterine growth restriction (IUGR) which is a disease characterized by retarded fetus growth.

The predominant pattern of hypomethylation for the 20 CpGs in Table 2 was replicated by the 10 CpGs mapped to the 450 k array for the Danish twin data, with 7 CpGs consistently replicated for their direction of effect (hypomethylation) (Supplementary Table S3). Some of the CpGs were replicated with low *p* values, for example, cg20451680 ($p <$ 4.16e-03) and cg20083839 ($p <$ 1.74e-2). The consistent pattern of hypomethylation suggests both conditions could have introduced similar epigenetic responses associated with neonatal stress. Note that the replication results were obtained from a different population (Danish) for a less severe condition (birth-weight discordance in twin pairs). Besides, for the replication samples e.g. birth weight discordant twins, the nutritional or blood supply deprivation for the smaller twin terminated after birth while exposure to famine covered both prenatal and postnatal periods. All the differences could result in the less pronounced effect size and significance in the replication analysis. Future EWASs on early-life adversity (famine or the like) should help to further verify our results.

Among the top sites in Table 2, cg16187328 ($p <$ 1.74e-07) is located in the body of *EHMT1*. Heterozygous mutations or deletions of the gene are the main causes of Kleefstra syndrome, a neurodevelopmental disorder characterized by impaired memory, autistic features and mostly severe intellectual disability [23]. de Boer et al. [9] found that individuals with *EHMT1* mosaicism seem to have increased vulnerability for developing severe psychopathology, especially ASD and mood disorders. The decreased methylation of cg16187328 in the body of *EHMT1* could suppress expression of the gene and increase the susceptibility of neurodevelopmental disorders. The second site cg06077226 ($p <$ 3.37e-07) is at 3′UTR of *CNR1* gene. As an endocannabinoid system gene, genetic variation in *CNR1* has been associated with neurological phenotypes in humans [51], and to Tourette syndrome [57]. The third site, cg24875889 ($p <$ 5.43e-07) is at 3′UTR of *UBXN7* gene. This gene has been shown to be expressed in brain, however its role is poorly understood at the moment [39]. Other interesting genes linked to the top CpGs include *BLK* (hosting cg13440894 at TSS1500) with mutation linked to maturity onset diabetes of the young and β-cell dysfunction [5]; *GNAS* (cg20083839 at 3′UTR) shown to have important roles in the regulation of energy metabolism [69]. The fourth CpG in Table 2 is located at the first exon of *ESM1* gene. In a EWAS on offspring of women with gestational diabetes mellitus (GDM) using the Illumina 450 k array, Hjort et al. [19] reported 2 hypomethylated CpGs (cg09452568 at gene body; cg00992687 at 3′UTR) comparing offspring of GDM mothers with controls. Considering that DNA methylation in the region of the first exon is reportedly much more tightly linked to transcriptional silencing than is methylation in the upstream promoter region [6], the overlap with their findings could suggest implication of early-life adversity on epigenetic modifications concerning the metabolic domain.

Likewise, gene-set enrichment analysis of genes linked to CpGs of low *p*-values revealed significant gene sets concerning neuron differentiation and development [34] (Table 4). The long-lasting effect of early-life adversity on the risk for developing various psychiatric illnesses has been addressed by a rich body of literature throughout the 20th century, which has become a new field of neuroepigenetics emerging as one possible mechanism for the far reaching effects of early-life adversity on adult phenotypes, behavior [48] and risk for psychiatric illness [28]. It is assumed that maternal famine during pregnancy could lead to prenatal folate deficiency which influences risk of schizophrenia in offspring [35] through dysregulation of DNA methylation given the critical role of folate as a key source of the one carbon group used to methylate DNA [8]. Neonatal exposure to famine could influence the epigenome of the fetus and modulate neuronal gene transcription to functionally alter multiple neuroendocrine and neurotransmitter systems. Consequently, these changes could impact behavior, responses to diverse environments, as well as predisposition to developing various psychiatric illnesses under certain environmental circumstances [63]. As an alternative, functional annotation based on the transcriptional domain using GREAT revealed significantly enriched GO biological processes reportedly to involve again in early adversity and development (Table 3), for example, negative regulation of leukocyte proliferation. Early-life adversity could impact immune phenotypes characterized by inflammation, impaired cellular immunity, and immunosenescence [12]. The high enrichment of the GO terms appears sensible as the immune system is a common denominator for a wide range of mental and physical symptoms overreached by early-life adversity. Similar to Table 4, the rest of the significant GO terms in Table 3 represent the neurodevelopmental consequences of maternal distress (famine) [47].

The top DMR found by our region-based association analysis is hypomethylated with famine exposure and is located in the exons of *RNF39* gene as shown in Fig. 3. It is interesting that decreased methylation in the same region has been associated with increased susceptibility for post-traumatic stress disorder in military servicemen [46]. Our finding reconfirms the involvement of this region in mediating the

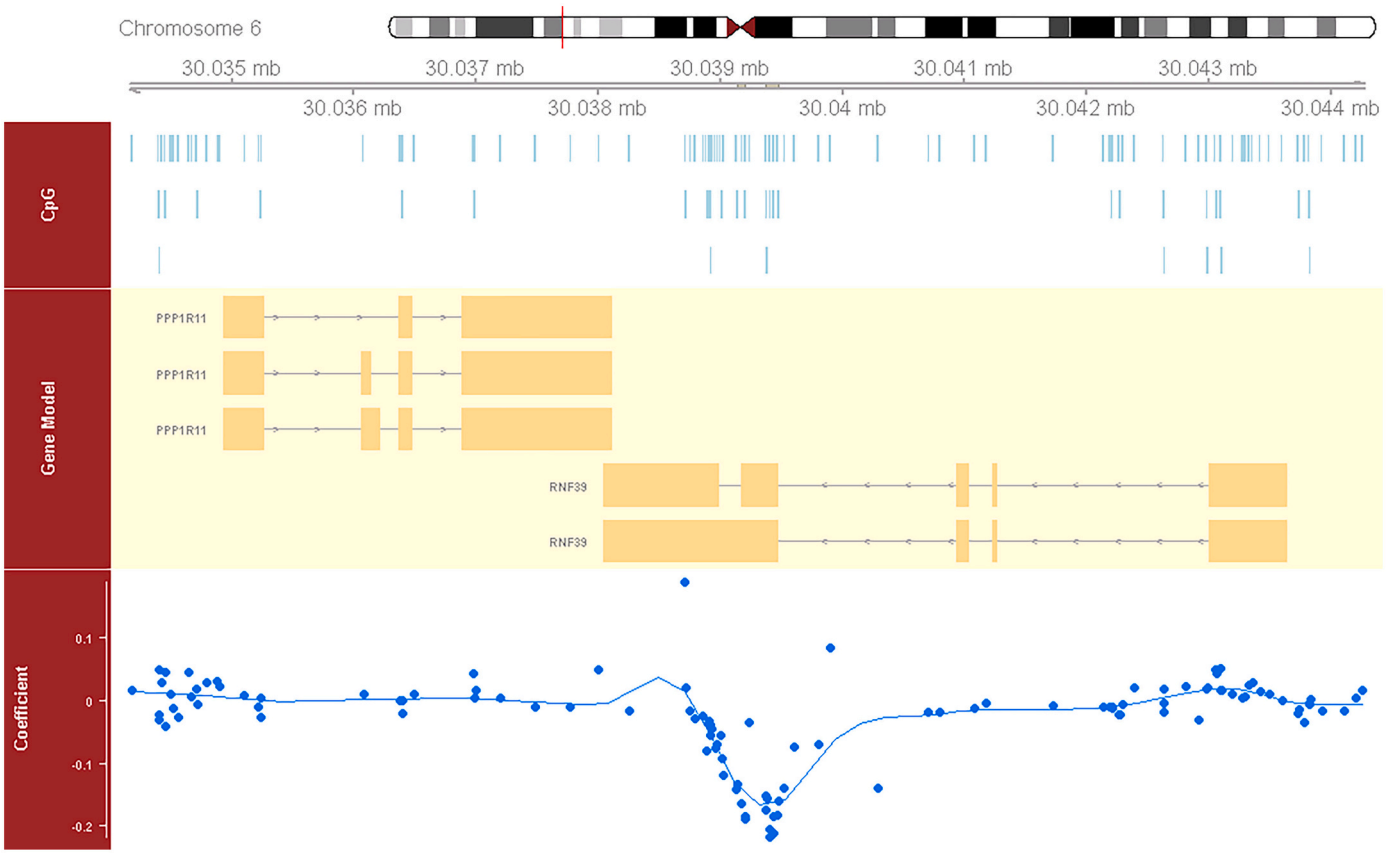

Fig. 3. Structured visualization of integrated genomic features for the top significant DMR on chromosome 6. A pattern of hypomethylation in the region of the DMR is shown at the bottom.

Table 2
List of DNA methylation sites with p < 1e-05.

CpGs	Coef.	SE	p value	FDR	Chr.	Position	Gene names	UCSC RefGene Group
cg16187328	−0.390	0.072	1.74E-07	0.082	9	140639428	EHMT1	Body
cg06077226	−0.348	0.066	3.37E-07	0.082	6	88851970	CNR1	3′UTR
cg24875889	−0.294	0.057	5.43E-07	0.088	3	196082185	UBXN7	3′UTR
cg20451680	−0.455	0.089	8.96E-07	0.109	5	54281336	ESM1	1stExon
cg13440894	−0.133	0.027	1.68E-06	0.164	8	11351135	BLK	TSS1500
cg20083839	−0.317	0.065	2.21E-06	0.180	20	57485940	GNAS	3′UTR
cg13672106	0.336	0.070	3.05E-06	0.184	17	38984421	TMEM99	5′UTR
cg11231670	−0.178	0.037	3.68E-06	0.184	16	2947301	FLYWCH2	Body
cg24042313	−0.250	0.053	4.47E-06	0.184	3	11079491	SLC6A1	3′UTR
cg15821562	0.229	0.048	4.53E-06	0.184	14	90083275	FOXN3	5′UTR
cg26859716	−0.457	0.097	4.62E-06	0.184	6	157365102	ARID1B	Body
cg15958366	−0.303	0.064	5.24E-06	0.184	3	187957027	LPP	5′UTR
cg26487422	−0.255	0.054	5.29E-06	0.184	7	155549270	RBM33	Body
cg04243827	−0.623	0.133	5.60E-06	0.184	5	3103718		
cg20376123	−0.229	0.049	5.68E-06	0.184	22	46760595	CELSR1	Body
cg06018748	−0.385	0.083	6.21E-06	0.189	12	59315217	LRIG3	TSS1500
cg24558480	−0.246	0.053	6.69E-06	0.192	15	101467705	LRRK1	Body
cg05083630	−0.326	0.070	7.09E-06	0.192	9	72169851	APBA1	5′UTR
cg18477248	−0.149	0.032	8.02E-06	0.192	14	62547992	SYT16	ExonBnd; Body
cg13727277	−0.505	0.110	8.67E-06	0.192	7	1062871	MIR339; C7orf50	TSS1500; Body

epigenetic response to stress conditions but further reveals the long-lasting effect of early-life adversity that persists into adulthood. The second top rank DMR is hypomethylated in the intron of *PTPRN2* gene on chromosome 7 (Supplementary Fig. S4). A recent study reported altered methylation of the gene in neonates of IUGR (intrauterine growth restriction) which is a pregnancy-associated disease manifested by decreased growth rate of fetus than the normal genetic growth potential [24]. The consistent finding is a reflection that prenatal exposure to famine could have introduced a stressful condition that mimics IUGR. Furthermore, the study also reported a global shift toward reduced methylation in IUGR neonates providing independent support to the pattern in Fig. 2. The *PTPRN2* gene encodes a major autoantigen and is usually considered as a metabolic gene seen in type 1 diabetes mellitus with genetically controlled CpG methylation of the gene

Table 3
GREAT bionomial test results for enrichment of GO terms[#].

Term Name	P-Value	*FDR Q-Val*	Fold Enrichment	Observed region hits	Region set coverage
Negative regulation of leukocyte proliferation	**4.86e-8**	**8.08e-6**	2.0446	70	1.54%
Ventral spinal cord interneuron fate commitment	**2.53e-6**	**2.27e-4**	2.5697	32	0.71%
Spinal cord dorsal/ventral patterning	**1.12e-5**	**7.59e-4**	2.1949	38	0.84%
Ventral spinal cord interneuron specification	**1.44e-5**	**9.41e-4**	2.5727	27	0.60%
Dendritic cell chemotaxis	**1.56e-5**	**9.99e-4**	3.1911	19	0.42%
Regulation of pri-miRNA transcription from RNA polymerase II promoter	**1.84e-5**	**1.15e-3**	2.1205	39	0.86%
Brush border assembly	**2.21e-5**	**1.37e-3**	11.0681	6	0.13%
Spinal cord patterning	**2.43e-5**	**1.46e-3**	2.0928	39	0.86%
Ventral spinal cord interneuron differentiation	**3.72e-5**	**2.04e-3**	2.2047	33	0.73%
Cerebellar granule cell precursor proliferation	**5.11e-4**	**1.50e-2**	2.5720	17	0.37%
Cell proliferation in hindbrain	**5.84e-4**	**1.67e-2**	2.5407	17	0.37%
Carbohydrate phosphorylation	**1.09e-3**	**2.70e-2**	2.3967	17	0.37%

The test set of 4535 genomic regions picked 5188 (28%) of all 18,549 genes.GO Biological Process has 13,145 terms covering 16,621 (90%) of all 18,549 genes, and 1,251,831 term - gene associations.13,145 ontology terms (100%) were tested using an annotation count range of [1, Inf].

associated with childhood obesity [27].

It is interesting that the identified differentially regulated genomic sites, DMRs and enriched functional pathways are commonly involved in early neurodevelopment and neuropsychological disorders. Li and Lumey [29] reported in a systematic review that schizophrenia was the only significant disorder stably more common among adults born during the famine compared with post-famine or post- plus pre-famine born controls. Their finding is in agreement with the functional interpretations of our EWAS results. We can postulate that prenatal exposure to famine could interrupt normal brain development early in life which is a sensitive period when the brain underwent its most rapid development. The identified epigenetic biomarkers could serve as molecular targets for early intervention to prevent the long-term consequences in mental health of the exposed individuals.

IGF2 is an imprinted gene that is expressed from the paternally-derived chromosome [13]. Methylation level of the gene has been shown to associate with fetus exposure to Dutch [18] and Chinese [49] famine. The CpGs analyzed by their studies were from a region located upstream of the three major imprinted promoters of *IGF2* in humans. While the Dutch study found reduced methylation specific for periconceptional exposure, the Chinese study reported elevated methylation in the same region. Different from their candidate gene analysis, the CpGs annotated to *IGF2* on the genome-wide array of this study do not overlap with their analyzed region converted to GRCh37/hg19 assembly (chr11:2169459-2169796). However, we observed a clear trend of increased methylation (coefficients for exposure >0) over the promoter

Table 4
List of gene sets over-represented by genes linked to CpGs with $p < 0.005$.

Gene Set [# Genes (K)]	Description	Genes in Overlap (k)	p-value	FDR
NEUROGENESIS [1674]	Generation of cells within the nervous system	144	$2.52 e^{-29}$	$2.59 e^{-25}$
REGULATION OF NERVOUS SYSTEM DEVELOPMENT [957]	Any process that modulates the frequency, rate or extent of nervous system development, the origin and formation of nervous tissue	98	$1.23 e^{-25}$	$6.3 e^{-22}$
CELL MORPHOGENESIS [1041]	The developmental process in which the size or shape of a cell is generated and organized	100	$5 e^{-24}$	$1.71 e^{-20}$
NEURON DIFFERENTIATION [1406]	The process in which a relatively unspecialized cell acquires specialized features of a neuron	118	$3.43 e^{-23}$	$8.81 e^{-20}$
REGULATION OF CELL DEVELOPMENT [982]	Any process that modulates the rate, frequency or extent of the progression of the cell over time, from its formation to the mature structure. Cell development does not include the steps involved in committing a cell to a specific fate	93	$5.62 e^{-22}$	$1.15 e^{-18}$
CELL PROJECTION ORGANIZATION [1588]	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of a prolongation or process extending from a cell, e.g. a flagellum or axon	124	$1.15 e^{-21}$	$1.97 e^{-18}$
NEURON DEVELOPMENT [1143]	The process whose specific outcome is the progression of a neuron over time, from initial commitment of the cell to a specific fate, to the fully functional differentiated cell	101	$1.46 e^{-21}$	$2.14 e^{-18}$
INTRINSIC COMPONENT OF PLASMA MEMBRANE [1731]	The component of the plasma membrane consisting of the gene products and protein complexes having either part of their peptide sequence embedded in the hydrophobic region of the membrane or some other covalently attached group such as a GPI anchor that is similarly embedded in the membrane	130	$3.21 e^{-21}$	$4.12 e^{-18}$
CELL MORPHOGENESIS INVOLVED IN DIFFERENTIATION [753]	The change in form (cell shape and size) that occurs when relatively unspecialized cells, e.g. embryonic or regenerative cells, acquire specialized structural and/or	78	$6.63 e^{-21}$	$7.57 e^{-18}$

(*continued on next page*)

Table 4 (*continued*)

Gene Set [# Genes (K)]	Description	Genes in Overlap (k)	p-value	FDR
	functional features that characterize the cells, tissues, or organs of the mature organism or some other relatively stable phase of the organism's life history			
REGULATION OF CELL DIFFERENTIATION [1945]	Any process that modulates the frequency, rate or extent of cell differentiation, the process in which relatively unspecialized cells acquire specialized structural and functional features	139	$1.03 e^{-20}$	1.05

region of *IGF2*, mixed also with some hypomethylated CpGs (Fig. 4). As shown in Fig. 5a, the elevated methylation with famine exposure predominates the CpGs from the promoter region ($p < 2.98e-03$) suggesting reduced activity of the gene which is in support to result from the Chinese study by Shen et al. [49].

A candidate gene study on *INSR* gene found increased methylation level of the gene in adults born during the Chinese famine [68]. Similar finding was reported by a genome-scale methylation sequencing study on the Dutch famine [61]. The region analyzed by Wang et al. [68] is the intragenic enhancer region of *INSR* where high level of methylation is associated with repression of transcription of the gene. Although no significant association was found in the region in our analysis, as shown in Fig. 5b, there are 2 CpGs outside the region in the gene body that are hypomethylated with low *p*-values (cg10473041, $p < 1.39e-04$; cg09684021, $p < 3.09e-03$). Their reduced methylation within the gene body could suggest decreased activity [71] of *INSR* related to famine exposure in early-life which is in agreement with the direction of effect found by Wang et al. [68] and Tobi et al. [61]. *INSR* mediates *IGF2* signaling to initiate intracellular signaling for metabolic regulation in response to insulin in adults and to implicate development of diabetes, obesity, and cancer [20]. More researches are needed to clarify and to characterize the effect of early-life adversity on the activities of the two genes in modifying individual's epigenetic predisposition to metabolic disorders (type 2 diabetes, obesity) and other complex diseases.

In the replication cohort of birth-weight discordant twin pairs, the smaller twin in a twin pair at birth is assumed to have been exposed to prenatal stressful factors as compared to the bigger twin of the pair; and this situation could mimic prenatal exposure to famine as both are early-life adversities that created stressful conditions during prenatal development. Most importantly, among the top differentially methylated sites by famine that are matched to the replication data, many of them showed similar directions of effect as in the discovery cohort and some even with very low *p* values (Results section). Indeed, the replication samples are of a different population. But the fact that replication can be done in different populations could indicate that the discovery results represent more biological effects independent of populations.

This study applied different filtering schemes to remove poorly detected CpGs using detection p value, CpGs on the sex chromosomes and CpGs physically overlapping with SNPs (CpG-SNPs). Although these are routine steps in DNA methylation microarray data analysis, dropping the sex chromosomes to avoid analytical difficulties due to their sex differences could miss important sites with sex-specific effects especially on the relatively large X-chromosome. Strategic modeling of the sex-chromosome data is needed to help with solving the issue. Moreover, dropping CpG-SNP sites could also remove methylation sites with allele-specific effects on gene expression or the methylation quantitative loci

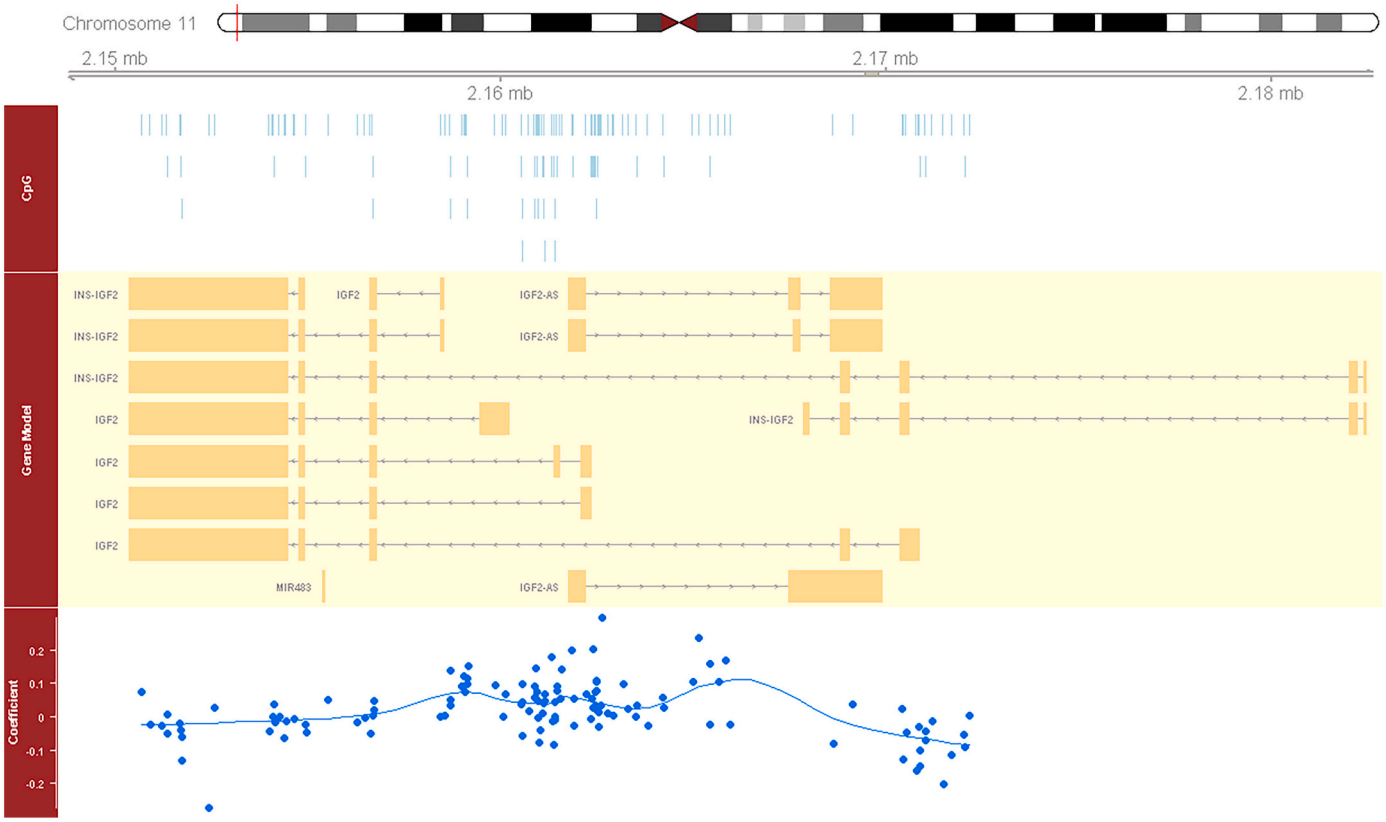

Fig. 4. Structured visualization of integrated genomic features for CpGs annotated to *IGF2* gene on chromosome 11.

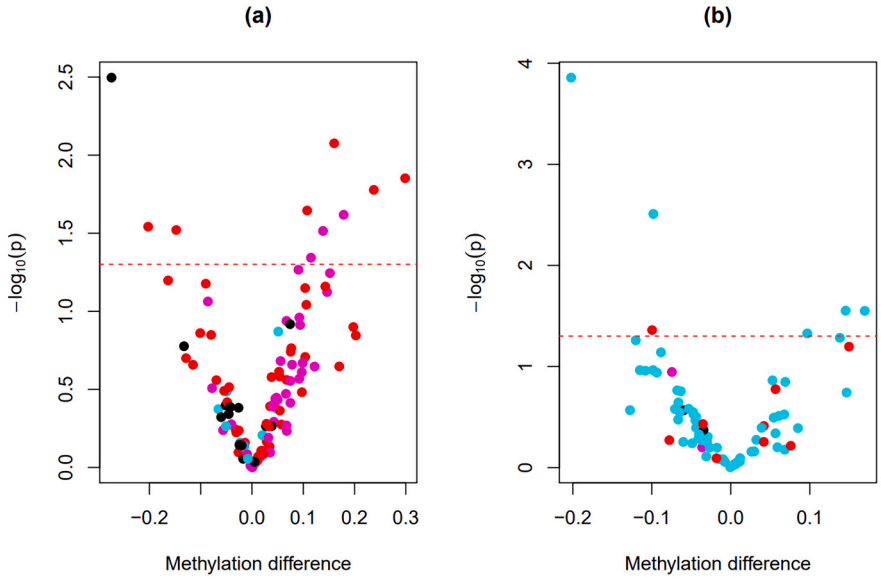

Fig. 5. Volcano plots for CpGs annotated to *IGF2* (5a) and *INSR* (5b) genes with red dots for CpGs only from promoter, pink dots for CpGs from both promoter and gene body, cyan dots for CpGs only from gene body, black dots for CpGs from other regions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(meQTLs). It would be interesting to explore the regulatory meQTLs showing interaction with prenatal famine exposure by joint analysis of the methylation data on CpG-SNPs and SNP genotype data to be collected.

In conclusion, our genome-wide DNA methylation profiling on fetal exposure to Chinese famine identified a predominant pattern of hypomethylation in the differentially regulated genomic sites with top CpGs, DMRs and enriched functional pathways functionally implicated in early neurodevelopment, neuropsychological disorders and metabolism. The results from our epigenetic analysis of the "natural experiment" provided new clues to the epigenetic embedding of early-life adversity that could potentially impact adult health. Although current study is only observational in nature, findings from the study can be used as a reference for future studies employing a more causal framework.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2021.09.018.

Data sharing statement

According to current Danish and EU legislations, transfer and sharing of individual-level data require prior approval from the Danish Data Protection Agency. Our present local data protection rules do not allow individual-level data to be shared in public databases. For these reasons, the raw data cannot be deposited in a public database. However, we welcome any enquiries regarding collaboration and individual requests for data sharing. Requests can be directed to Qihua Tan at qtan@health.sdu.dk.

Funding

This study was financially supported by Independent Research Fund Denmark, the Lundbeck Foundation, and the Novo Nordisk Foundation Medical and Natural Sciences Research.

Author contributions

QT, SL, KC, DZ, TK deceived and designed the study; DZ, WW collected samples and clinical data; KC, JM provided Danish twin data; WL, JL, WW analyzed data; SL, WW wrote the paper. KC, JM, SL, TK, QT interpreted results and discussed findings. All authors read and approved the paper.

Declaration of Competing Interest

None declared.

Acknowledgements

This study was financially supported by the Independent Research Fund Denmark (grant number DFF – 6110-00114), the Lundbeck Foundation (grant number R170-2014-1353), and the Novo Nordisk Foundation Medical and Natural Sciences Research Grant (grant number NNF13OC0007493).

References

[1] M. Aryee, A.E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A.P. Feinberg, K. D. Hansen, R.A. Irizarry, Minfi: a fexible and comprehensive bioconductor package for analysis of Infinium DNA methylation microarrays, Bioinformatics 30 (2014) 1363–1369.
[3] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B Methodol. (1995) 289–300.
[4] V. Bol, F. Desjardins, B. Reusens, J.-L. Balligand, C. Remacle, Does early mismatched nutrition predispose to hypertension and atherosclerosis, in male mice? PLoS One 5 (9) (2010), e12656.
[5] M. Borowiec, C.W. Liew, R. Thompson, W. Boonyasrisawat, J. Hu, W.M. Mlynarski, I. El Khattabi, S.H. Kim, L. Marselli, S.S. Rich, A.S. Krolewski, S. Bonner-Weir, A. Sharma, M. Sale, J.C. Mychaleckyj, R.N. Kulkarni, A. Doria, Mutations at the BLK locus linked to maturity onset diabetes of the young and beta-cell dysfunction, Proc. Natl. Acad. Sci. U. S. A. 106 (34) (2009) 14460–14465.
[6] F. Brenet, M. Moh, P. Funk, E. Feierstein, A.J. Viale, N.D. Socci, J.M. Scandura, DNA methylation of the first exon is tightly linked to transcriptional silencing, PLoS One 6 (1) (2011), e14524.
[7] Y. Cai, W. Feng, Famine, social disruption, and involuntary fetal loss: evidence from Chinese survey data, Demography. 42 (2) (2005) 301–322.
[8] K.S. Crider, T.P. Yang, R.J. Berry, L.B. Bailey, Folate and DNA methylation: a review of molecular mechanisms and the evidence for folate's role, Adv. Nutr. 3 (1) (2012) 21–38.
[9] A. de Boer, K. Vermeulen, J.I.M. Egger, J.G.E. Janzing, N. de Leeuw, H.E. Veenstra-Knol, N.S. den Hollander, H. van Bokhoven, W. Staal, T. Kleefstra, EHMT1 mosaicism in apparently unaffected parents is associated with autism spectrum disorder and neurocognitive dysfunction, Mol. Autism. 9 (2018) 5.
[10] S.R. de Rooij, R.C. Painter, F. Holleman, P.M. Bossuyt, T.J. Roseboom, The metabolic syndrome in adults prenatally exposed to the Dutch famine, Am. J. Clin. Nutr. 86 (4) (2007) 1219–1224.

[11] P. Du, X. Zhang, C.C. Huang, N. Jafari, W.A. Kibbe, L. Hou, S.M. Lin, Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis, BMC Bioinform. 11 (2010) 587.

[12] M.M.C. Elwenspoek, A. Kuehn, C.P. Muller, J.D. Turner, The effects of early-life adversity on the immune system, Psychoneuroendocrinology. 82 (2017) 140–154.

[13] N. Giannoukakis, C. Deal, J. Paquette, et al., Parental genomic imprinting of the human IGF2 gene, Nat. Genet. 4 (1993) 98–101.

[14] P.D. Gluckman, M.A. Hanson, M.D. Mitchell, Developmental origins of health and disease: reducing the burden of chronic disease in the next generation, Genome Med. 2 (2) (2010) 14.

[15] F. Hahne, R. Ivanek, Statistical genomics: methods and protocols, in: E. Mathé, S. Davis (Eds.), Chapter Visualizing Genomic Data Using Gviz and Bioconductor, Springer New York, New York, NY, 2016, pp. 335–351. ISBN 978-1-4939-3578-9.

[16] C.N. Hales, D.P. Barker, The thrifty phenotype hypothesis: type 2 diabetes, Br. Med. Bull. 60 (1) (2001) 5–20.

[17] Y. He, L.D. de Witte, L.C. Houtepen, D.M. Nispeling, Z. Xu, Q. Yu, Y. Yu, E.M. Hol, R.S. Kahn, M.P. Boks, DNA methylation changes related to nutritional deprivation: a genome-wide analysis of population and in vitro data, Clin. Epigenetics 11 (1) (2019) 80.

[18] B.T. Heijmans, E.W. Tobi, A.D. Stein, H. Putter, G.J. Blauw, E.S. Susser, P. E. Slagboom, L.H. Lumey, Persistent epigenetic differences associated with prenatal exposure to famine in humans, Proc. Natl. Acad. Sci. U. S. A. 105 (44) (2008) 17046–17049.

[19] L. Hjort, D. Martino, L.G. Grunnet, H. Naeem, J. Maksimovic, A.H. Olsson, C. Zhang, C. Ling, S.F. Olsen, R. Saffery, A.A. Vaag, Gestational diabetes and maternal obesity are associated with epigenome-wide methylation changes in children, JCI Insight. 3 (17) (2018), e122572.

[20] J.M.P. Holly, K. Biernacka, C.M. Perks, The neglected insulin: IGF-II, a metabolic regulator with implications for diabetes, obesity, and Cancer, Cells. 8 (10) (2019) 1207.

[21] E.A. Houseman, W.P. Accomando, D.C. Koestler, B.C. Christensen, C.J. Marsit, H. H. Nelson, J.K. Wiencke, Kelsey KTl, DNA methylation arrays as surrogate measures of cell mixture distribution, BMC Bioinform. 13 (2012) 86.

[22] A.E. Jaffe, P. Murakami, H. Lee, J.T. Leek, D.M. Fallin, A.P. Feinberg, R.A. Irizarry, Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies, Int. J. Epidemiol. 41 (1) (2012) 200–209.

[23] T. Kleefstra, H.G. Brunner, J. Amiel, A.R. Oudakker, W.M. Nillesen, A. Magee, D. Geneviève, V. Cormier-Daire, H. van Esch, J.P. Fryns, B.C. Hamel, E. A. Sistermans, B.B. de Vries, H. van Bokhoven, Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome, Am. J. Hum. Genet. 79 (2) (2006) 370–377.

[24] R.G. Krishna, B. Vishnu Bhat, Z. Bobby, et al., Identification of differentially methylated candidate genes and their biological significance in IUGR neonates by methylation EPIC array, J. Matern. Fetal Neonatal Med. (2020), https://doi.org/10.1080/14767058.2020.1727881.

[25] S.C. Langley, A.A. Jackson, Increased systolic blood pressure in adult rats induced by fetal exposure to maternal low protein diets, Clin. Sci. (Lond.) 86 (2) (1994) 217–222.

[26] S.C. Langley, R.F. Browne, A.A. Jackson, Altered glucose tolerance in rats exposed to maternal low protein diets in utero, Comp. Biochem. Physiol. A Physiol. 109 (2) (1994) 223–229.

[27] S. Lee, The association of genetically controlled CpG methylation (cg158269415) of protein tyrosine phosphatase, receptor type N2 (PTPRN2) with childhood obesity, Sci. Rep. 9 (2019) 4855.

[28] C.R. Lewis, M.F. Olive, Early-life stress interactions with the epigenome: potential mechanisms driving vulnerability toward psychiatric illness, Behav. Pharmacol. 25 (5–6) (2014) 341–351.

[29] C. Li, L.H. Lumey, Exposure to the Chinese famine of 1959-61 in early-life and long-term health conditions: a systematic review and meta-analysis, Int. J. Epidemiol. 46 (4) (2017) 1157–1170.

[30] Y. Li, V.W. Jaddoe, L. Qi, et al., Exposure to the chinese famine in early-life and the risk of metabolic syndrome in adulthood, Diabetes Care 34 (4) (2011) 1014–1018.

[31] Y. Li, Y. He, L. Qi, V.W. Jaddoe, E.J.M. Feskens, X. Yang, et al., Exposure to the Chinese famine in early-life and the risk of hyperglycaemia and type 2 diabetes in adulthood, Diabetes 59 (2010) 2400–2406.

[32] W. Li, L. Christiansen, J. Hjelmborg, J. Baumbach, Q. Tan, On the power of epigenome-wide association studies using a disease-discordant twin design, Bioinformatics. 34 (23) (2018) 4073–4078.

[33] J. Maksimovic, L. Gordon, A. Oshlack, SWAN: subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips, Genome Biol. 13 (2012) R44.

[34] R.L. Margolis, C.A. Ross, Neuronal signaling pathways: genetic insights into the pathophysiology of major mental illness, Neuropsychopharmacology. 35 (1) (2010) 350–351.

[35] J.M. McClellan, E. Susser, M.C. King, Maternal famine, de novo mutations, and schizophrenia, JAMA. 296 (5) (2006) 582–584.

[36] C.Y. McLean, D. Bristor, M. Hiller, S.L. Clarke, B.T. Schaar, C.B. Lowe, A. M. Wenger, G. Bejerano, GREAT improves functional interpretation of cis-regulatory regions, Nat. Biotechnol. 28 (2010) 495–501.

[37] Q. Tan, M. Thomassen, M. Burton, K.F. Mose, K.E. Andersen, J. Hjelmborg, T. Kruse, Generalized correlation coefficient for non-parametric analysis of microarray time-course data, J Integr Bioinform. 14 (2) (2017), 20170011.

[38] R. Mu, X. Zhang, Why does the great Chinese famine affect the male and female survivors differently? Mortality selection versus son preference, Econ. Hum. Biol. 9 (1) (2011) 92–105.

[39] T. Nagase, K. Ishikawa, M. Suyama, R. Kikuno, N. Miyajima, A. Tanaka, et al., Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro, DNA Res. 5 (1998) 277–286.

[40] F. Ning, Z.C. Pang, Y.H. Dong, W.G. Gao, H.R. Nan, S.J. Wang, et al., Risk factors associated with the dramatic increase in the prevalence of diabetes in the adult Chinese population in Qingdao, China, Diabet. Med. 26 (2009) 855–863.

[41] B. Novakovic, R. Saffery, The importance of the intrauterine environment in shaping the human neonatal epigenome, Epig4enomics. 5 (1) (2013) 1–4.

[42] R.C. Painter, T.J. Roseboom, O.P. Bleker, Prenatal exposure to the Dutch famine and disease in later life: an overview, Reprod. Toxicol. 20 (3) (2005) 345–352.

[43] R.C. Painter, S.R. de Rooij, P.M. Bossuyt, D.I. Phillips, C. Osmond, D.J. Barker, O. P. Bleker, T.J. Roseboom, Blood pressure response to psychological stressors in adults after prenatal exposure to the Dutch famine, J. Hypertens. 24 (9) (2006) 1771–1778.

[44] T. Roseboom, J. Van der Meulen, A. Ravelli, C. Osmond, D. Barker, O. Bleker, Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview, Twin Res. 4 (5) (2001) 293–298.

[45] T. Roseboom, S. de Rooij, R. Painter, The Dutch famine and its long-term consequences for adult health, Early Hum. Dev. 82 (8) (2006) 485–491.

[46] B.P.F. Rutten, E. Vermetten, C.H. Vinkers, G. Ursini, N.P. Daskalakis, E. Pishva, L. de Nijs, L.C. Houtepen, L. Eijssen, A.E. Jaffe, G. Kenis, W. Viechtbauer, D. van den Hove, K.G. Schraut, K.P. Lesch, J.E. Kleinman, T.M. Hyde, D.R. Weinberger, L. Schalkwyk, K. Lunnon, J. Mill, H. Cohen, R. Yehuda, D.G. Baker, A.X. Maihofer, C.M. Nievergelt, E. Geuze, M.P.M. Boks, Longitudinal analyses of the DNA methylome in deployed military servicemen identify susceptibility loci for post-traumatic stress disorder, Mol. Psychiatry 23 (5) (2018) 1145–1156.

[47] C. Schuurmans, D.M. Kurrasch, Neurodevelopmental consequences of maternal distress: what do we really know? Clin. Genet. 83 (2) (2013) 108–117.

[48] C.E. Schwarze, A. Mobascher, B. Pallasch, et al., Prenatal adversity: a risk factor in borderline personality disorder? Psychol. Med. 43 (6) (2013) 1279–1291.

[49] L. Shen, C. Li, Z. Wang, R. Zhang, Y. Shen, T. Miles, J. Wei, Z. Zou, Early-life exposure to severe famine is associated with higher methylation level in the IGF2 gene and higher total cholesterol in late adulthood: the genomic research of the Chinese famine (GRECF) study, Clin. Epigenetics 11 (1) (2019) 88.

[50] V. Smil, China's great famine: 40 years later, BMJ. 319 (7225) (1999) 1619–1621.

[51] D.R. Smith, C.M. Stanley, T. Foss, R.G. Boles, K. McKernan, Rare genetic variants in the endocannabinoid system genes CNR1 and DAGLA are associated with neurological phenotypes in humans, PLoS One 12 (11) (2017), e0187926.

[52] S. Song, Mortality consequences of the 1959-1961 great leap forward famine in China: debilitation, selection, and mortality crossovers, Soc. Sci. Med. 71 (3) (2010) 551–558.

[53] S. Song, W. Wang, P. Hu, Famine, death, and madness: schizophrenia in early adulthood after prenatal exposure to the Chinese great leap forward famine, Soc. Sci. Med. 68 (7) (2009) 1315–1321.

[54] D. St Clair, M. Xu, P. Wang, Y. Yu, Y. Fang, F. Zhang, X. Zheng, N. Gu, G. Feng, P. Sham, L. He, Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959-1961, JAMA. 294 (5) (2005) 557–562.

[55] A.D. Stein, P.A. Zybert, K. van der Pal-de Bruin, L.H. Lumey, Exposure to famine during gestation, size at birth, and blood pressure at age 59 y: evidence from the Dutch famine, Eur. J. Epidemiol. 21 (10) (2006) 759–765.

[56] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. 102 (43) (2005) 15545–15550.

[57] N. Szejko, J.P. Fichna, K. Safranow, T. Dziuba, C. Żekanowski, P. Janik, Association of a variant of CNR1 gene encoding cannabinoid receptor 1 with Gilles de la Tourette syndrome, Front. Genet. 11 (2020) 125.

[58] Q. Tan, M. Frost, B.T. Heijmans, Hjelmborg J. von Bornemann, E.W. Tobi, K. Christensen, L. Christiansen, Epigenetic signature of birth weight discordance in adult twins, BMC Genomics 15 (1) (2014) 1062.

[59] Q. Tan, L. Christiansen, J.B. Hjelmborg, K. Christensen, Twin methodology in epigenetic studies, J. Exp. Biol. 218 (2015) 134–139.

[60] Q. Tan, The epigenome of twins as a perfect laboratory for studying behavioural traits, Neurosci. Biobehav. Rev. 107 (2019) 192–195.

[61] E.W. Tobi, J.J. Goeman, R. Monajemi, H. Gu, H. Putter, Y. Zhang, R.C. Slieker, A. P. Stok, P.E. Thijssen, F. Müller, E.W. van Zwet, C. Bock, A. Meissner, L.H. Lumey, P. Eline Slagboom, B.T. Heijmans, DNA methylation signatures link prenatal famine exposure to growth and metabolism, Nat. Commun. 5 (2014) 5592.

[62] P.C. Tsai, J.T. Bell, Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation, Int. J. Epidemiol. 44 (4) (2015) 1429–1441.

[63] A.M. Vaiserman, A.K. Koliada, Early-life adversity and long-term neurobehavioral outcomes: epigenome as a bridge? Hum. Genom. 11 (1) (2017) 34.

[64] A.F. van Abeelen, M.V. Veenendaal, R.C. Painter, S.R. de Rooij, M.G. Dijkgraaf, P. M. Bossuyt, S.G. Elias, D.E. Grobbee, C.S. Uiterwaal, T.J. Roseboom, Survival effects of prenatal famine exposure, Am. J. Clin. Nutr. 95 (1) (2012) 179–183.

[65] P.X. Wang, J.J. Wang, Y.X. Lei, et al., Impact of fetal and infant exposure to the Chinese great famine on the risk of hypertension in adulthood, PLoS One 7 (11) (2012), e49720.

[66] X. Wang, P.W. Laird, T. Hinoue, S. Groshen, K.D. Siegmund, Non-specific filtering of beta-distributed data, BMC Bioinform. 15 (2014) 199.

[67] Y. Wang, X. Wang, Y. Kong, J.H. Zhang, Q. Zeng, The great Chinese famine leads to shorter and overweight females in Chongqing Chinese population after 50 years, Obesity (Silver Spring) 18 (3) (2010) 588–592.

[68] Z. Wang, J. Song, Y. Li, B. Dong, Z. Zou, J. Ma, Early-life exposure to the Chinese famine is associated with higher methylation level in the INSR gene in later adulthood, Sci. Rep. 9 (1) (2019) 3354.
[69] L.S. Weinstein, J. Liu, A. Sakamoto, T. Xie, M. Chen, Minireview: GNAS: normal and abnormal functions, Endocrinology. 145 (12) (2004) 5459–5464.
[70] M. Xu, S. Sun, B. Liu, G. Feng, L. Yu, L. Yang, G. He, P. Sham, Z. Susser, D.S. Clair, L. He, Prenatal malnutrition and adult schizophrenia: further evidence from the 1959-1961 Chinese famine, Schizophr. Bull. 35 (3) (2009) 568–576.
[71] X. Yang, H. Han, D.D. De Carvalho, F.D. Lay, P.A. Jones, G. Liang, Gene body methylation can alter gene expression and is a therapeutic target in cancer, Cancer Cell 26 (4) (2014) 577–590.