DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood

Aryeh Stein, Emory University
EW Tobi, Leiden University Medical Center - LUMC
RC Slieker, Leiden University Medical Center - LUMC
R Luijk, Leiden University Medical Center - LUMC
KF Dekkers, Leiden University Medical Center - LUMC
KM Xu, Leiden University Medical Center - LUMC
PE Slagboom, Leiden University Medical Center - LUMC
EW Van Zwet, Leiden University Medical Center - LUMC
LH Lumey, Leiden University Medical Center - LUMC
BT Heijmans, Leiden University Medical Center - LUMC

Only first 10 authors above; see publication for full author list.

Journal Title: Science Advances
Volume: Volume 4, Number 1
Publisher: (publisher) | 2018-01-01, Pages eaao4364-eaao4364
Type of Work: Article
Publisher DOI: 10.1126/sciadv.aao4364
Permanent URL: https://pid.emory.edu/ark:/25593/s857t

Final published version: http://dx.doi.org/10.1126/sciadv.aao4364

Accessed March 27, 2018 10:48 PM EDT
INTRODUCTION

Since the early 1970s, epidemiological studies have reported associations between an adverse prenatal environment and an increased disease risk in later life (1). Animal studies provided potential mechanisms for these observations (2), including environmentally induced changes to epigenetic marks during development (3). Epigenetic marks such as DNA methylation (DNAm) influence the transcription potential of genomic regions and, once changed, can result in long-term effects (4). Animal experiments show that epigenetic changes that are established during early development contribute to phenotypes later in life (5, 6). In parallel, human studies show changes in DNAm after exposure to a range of adverse prenatal conditions (7–12). These DNAm differences may mediate part of the association between adverse prenatal conditions and childhood phenotypes (13–15). Systematic epigenome-wide studies investigating the associations among these adverse conditions, DNAm changes, and specific phenotypes later in life are still largely lacking (16).

To fill this gap, we examined these associations in the quasi-experimental setting of the Dutch Hunger Winter of 1944–1945 (17), a 6-month famine at the end of World War II. Exposure to famine during gestation is associated with an increased risk of obesity, dyslipidemia, type 2 diabetes, and schizophrenia (18). We focus on our previous studies of the Dutch Hunger Winter, where we documented that prenatal adversity is associated with increases in adverse metabolic phenotypes in adulthood such as body mass index (BMI) and fasting glucose, serum triglycerides (TG), and low-density lipoprotein (LDL) cholesterol concentrations (LDL-C) (19–23). Here, we adopt a genome-wide approach to systematically identify the potential of DNAm to act as a mediator of the relationship between prenatal famine exposure and adult outcomes in our study population of prenatally exposed individuals and time and sibling controls (24). We use mediation analysis as a helpful statistical tool to further explore the nature of the relationship between an independent (famine exposure) and dependent variable (metabolic health) through a hypothesized mediator (DNAm).

To formally establish mediation (25), we first re-examined the relation between famine exposure any time during gestation (“famine exposure”) and adult BMI, glucose, TG, and LDL-C outcomes in individuals with genome-wide DNAm data. Next, we examined genome-wide whether DNAm at specific cytosine-phosphate-guanine (CpG) dinucleotides was associated with both famine exposure and the outcome of interest and subjected the identified candidate CpGs to a formal mediation test (26) to determine the extent to which DNAm mediated the association between prenatal famine and adult outcomes. These analyses were repeated for exposure during early gestation, which is an especially sensitive period of gestation (27, 28). To address the potential functional impact of mediating CpGs, we tested for their
association with gene expression (29, 30) and correlation in methylation levels across postmortem studies (31, 32).

Using this approach, we show that the increase in BMI and TG among individuals exposed in utero is mediated by DNAm, specifically near genes involved in development and metabolic processes including energy metabolism (PIM3), β cell function (TXNIP), glycolysis (PFKFB3), and adipogenesis (METTL8).

**RESULTS**

**Genome-wide mediation analysis**

We first confirmed the previously reported association between famine exposure and BMI, fasted glucose, TG, and LDL-C in the Dutch Hunger Winter Families Study for individuals with complete methylation data (19–23). Individuals with prenatal famine exposure (n = 348) had a 5.6% (+0.36 SD) higher BMI (P = 5.7 × 10−3), a 13.5% (+0.23 SD) higher serum TG (P = 3.8 × 10−3), and a 3.8% (+0.22 SD) higher fasted glucose (P = 0.023) but no difference in LDL-C (P > 0.05), as compared with non-exposed controls (n = 463; Table 1). Therefore, LDL-C was excluded from further analyses. Next, we applied a genome-wide screen (342,596 CpGs) to identify potentially mediating CpGs. We simultaneously tested the association of DNAm with famine exposure and with any of the metabolic outcomes (BMI, fasted glucose, or TG) in a single model. For BMI and TG, this resulted in 8 and 16 associated CpG simultaneously tested the association of DNA methylation (DNAm) with famine exposure and with both genome-wide DNAm (Illumina 450k) and expression [RNA sequencing (RNA-seq)] data (29, 30) showed that methylation of cg09349128 was associated with the expression of PIM3 and also with two additional genes in cis, namely, ZBED4 and CRELD2 (table S1).

**Mediation: Famine exposure and BMI**

Of the eight mediation candidates for the EWAS of famine exposure and BMI, only cg09349128 was associated with both exposure and outcome (P_{exposure|BMI} = 1.3 × 10−5, P_{outcome} = 1.5 × 10−3, P_{outcome} = 6.5 × 10−8; Table 2). Of interest, DNAm at this CpG has been consistently reported to be associated with BMI (33–36). A formal mediation analysis showed that DNAm at cg09349128 mediates the association between famine exposure and BMI (b_{mediation} = 0.7%/log(BMI); 95% confidence interval (CI), 0.3 to 1.2%; P = 0.001). The mediation path explained 13.4% (95% CI, 5 to 28%); P = 0.001) of the association between famine exposure and BMI (R^2_{LMM} = 0.27). This finding persisted after statistical adjustment for smoking, socioeconomic status (SES), and reported dietary intake at the time of examination (calories from fat, protein, or carbohydrates or from any source).

Genomic annotation of cg09349128 revealed that it mapped to an enhancer region in multiple cell types (37). The enhancer is linked to PIM3 in epigenome reference data, a gene implicated in cell growth and energy metabolism (38), stem cell renewal (39), and glucose-stimulated insulin secretion in β cells (40). Analysis of 2044 whole blood samples with both genome-wide DNAm (Illumina 450k) and expression [RNA sequencing (RNA-seq)] data (29, 30) showed that methylation of cg09349128 was associated with the expression of PIM3 and also with two additional genes in cis, namely, ZBED4 and CRELD2 (table S1).

**Mediation: Famine exposure and TG**

Of the 16 mediation candidates for the association between famine exposure and TG, 6 CpGs were associated with both exposure and outcome (table S2). With the exception of cg09693052, methylation at these CpGs was previously reported to be associated with TG or other metabolic traits in EWASs (30, 41–46). All six CpGs mediated the relationship between famine exposure and TG (P < 0.007; Table 3). The mediation path of each CpG explained between 19.6 and 28.0% of the association between famine exposure and TG (R^2_{LMM} = 0.30). Although not mapping close together and often located on different chromosomes, DNAm levels of the six CpGs were correlated (p = 0.17 to 0.73, P < 0.001). We therefore tested mediation for the mean standardized DNAm level across all six CpGs. This aggregate measure likewise mediated the association between famine exposure and TG [b_{mediation} = 0.075 SD/log(TG); 95% CI, 0.047 to 0.104; P < 0.001]. The mediation path explained 80.0% (95% CI, 38.5 to 100%; P = 0.02) of the association. The role as mediator persisted, for both individual CpGs and their aggregate, after additional adjustment for smoking, SES, and current diet. Adjustment for BMI did somewhat attenuate this association [b_{mediation} = 0.053 SD/log(TG); 95% CI, 0.03 to 0.08; P < 0.001], and as a result, the proportion mediated was affected (P = 0.27). Because an additional analysis showed that the CpGs involved did not mediate the association between famine exposure and BMI (P > 0.076), it is unlikely that an extended mediation path including BMI is involved in terms of DNAm.

The six mediating CpGs mapped to (intronic) enhancers, open chromatin regions, and exons. Analysis of blood data (29, 30) showed that all but cg18120259 were associated with the expression of their nearest genes (table S1). This included cg19693031 with TXNIP and cg07397296 with ABCG1. DNAm at both CpGs has been previously associated with TG (30).

**Mediation: Early famine exposure**

Early gestation appears to be the most vulnerable period in terms of the later-life phenotype (27) as epigenetic (28) consequences of famine exposure. The association between famine exposure and BMI or TG also

### Table 1. Phenotypic differences and famine exposure.

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>Famine exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>463</td>
<td>348</td>
</tr>
<tr>
<td>Age (years) [SD]</td>
<td>58.0 [5.4]</td>
<td>58.9 [0.5]**</td>
</tr>
<tr>
<td>Male (%)</td>
<td>43.0</td>
<td>46.0</td>
</tr>
<tr>
<td>BMI [SD]</td>
<td>27.0 [4.2]</td>
<td>28.5 [5.0]**</td>
</tr>
<tr>
<td>LDL-C [SD]</td>
<td>3.42 [0.96]</td>
<td>3.45 [0.97]</td>
</tr>
<tr>
<td>Triglycerides [SD]</td>
<td>1.48 [0.86]</td>
<td>1.68 [1.30]**</td>
</tr>
<tr>
<td>Glucose baseline [SD]</td>
<td>5.32 [0.93]</td>
<td>5.52 [1.19]*</td>
</tr>
</tbody>
</table>

Nominal P value either \( P < 0.05, ** P < 0.01 \) or *** \( P < 0.001 \) from a linear mixed-effects model with the denoted variable as the dependent variable and family identifier as random effect. 1Model included an additional random effect for exposure status to control for the difference in variance in age between groups. 2Model-applied correction for age and gender. 3Model-applied correction for age and gender, and statin used, and individuals who were nonfasting at examination were excluded (excluding two controls and five famine-exposed individuals). 4Model-applied correction for age and gender. Individuals who were nonfasting and had prediagnosed diabetes (thus receiving treatment) before the clinical examination were excluded (excluding 19 controls and 32 famine-exposed individuals).
Fig. 1. Manhattan plots: Outcome genome-wide screens for potential mediators. The $-\log(P\text{ value})$ is shown (y axis) for each CpG relative to its genomic locations (x axis) on the 22 autosomal chromosomes tested for (A) an EWAS on both famine exposure and BMI and (B) a nEWAS on both famine exposure and serum TG.

Table 2. Genome-wide screen for potential mediators: famine exposure and BMI.

<table>
<thead>
<tr>
<th>CpG</th>
<th>Location (hg19)</th>
<th>Nearest gene (expression)†</th>
<th>Methylation (SD)‡</th>
<th>Rank</th>
<th>Mediation EWAS</th>
<th>Associations with either famine exposure or BMI*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\beta_{\text{famine}}$</td>
</tr>
<tr>
<td>cg00574958</td>
<td>chr11:68607622</td>
<td>CPT1A</td>
<td>13.6 (2.6)</td>
<td>1</td>
<td>6.7 $\times 10^{-14}$</td>
<td>2.3 $\times 10^{-8}$</td>
</tr>
<tr>
<td>cg06500161</td>
<td>chr21:43656587</td>
<td>ABCG1</td>
<td>65.9 (3.0)</td>
<td>2</td>
<td>2.0 $\times 10^{-12}$</td>
<td>3.3 $\times 10^{-7}$</td>
</tr>
<tr>
<td>cg26950531</td>
<td>chr19:38704515</td>
<td>DPF1</td>
<td>32.5 (6.1)</td>
<td>3</td>
<td>2.2 $\times 10^{-10}$</td>
<td>4.3 $\times 10^{-5}$</td>
</tr>
<tr>
<td>cg11024682</td>
<td>chr17:17730094</td>
<td>SREBF1</td>
<td>53.2 (3.2)</td>
<td>4</td>
<td>2.2 $\times 10^{-8}$</td>
<td>1.9 $\times 10^{-3}$</td>
</tr>
<tr>
<td>cg09349128</td>
<td>chr22:50327986</td>
<td>CRELD2</td>
<td>39.0 (3.8)</td>
<td>5</td>
<td>1.3 $\times 10^{-7}$</td>
<td>8.3 $\times 10^{-3}$</td>
</tr>
<tr>
<td>cg07373197</td>
<td>chr11:68607675</td>
<td>CPT1A</td>
<td>20.4 (4.6)</td>
<td>6</td>
<td>1.5 $\times 10^{-7}$</td>
<td>8.3 $\times 10^{-3}$</td>
</tr>
<tr>
<td>cg23032421</td>
<td>chr3:3152038</td>
<td>ILSRA</td>
<td>73.9 (4.1)</td>
<td>7</td>
<td>2.1 $\times 10^{-7}$</td>
<td>0.01</td>
</tr>
<tr>
<td>cg17058475</td>
<td>chr11:68607737</td>
<td>CPT1A</td>
<td>16.1 (4.1)</td>
<td>8</td>
<td>7.4 $\times 10^{-7}$</td>
<td>0.032</td>
</tr>
<tr>
<td>cg15659713</td>
<td>chr8:38586183</td>
<td>TACCC1</td>
<td>24.6 (3.9)</td>
<td>9</td>
<td>1.7 $\times 10^{-7}$</td>
<td>0.064</td>
</tr>
<tr>
<td>cg21998597</td>
<td>chr12:54764265</td>
<td>ZNF385A</td>
<td>68.9 (5.7)</td>
<td>14</td>
<td>5.2 $\times 10^{-7}$</td>
<td>0.13</td>
</tr>
</tbody>
</table>

*The estimate and (nominal) P value belonging to the EWAS for famine exposure (β = exposed − unexposed) or BMI (β/log(BMI)). †Nearest gene within 100 kb. ‡The Illumina 450k array β value (ranging from 0 to 1) multiplied by 100 for easy interpretation. This is done throughout the presented work. ¶The $P$ value belonging to an analysis of variance (ANOVA) test ($\chi^2$, df = 2) between a generalized estimating equations (GEE) model with and without both famine exposure and BMI. The two CpGs identified in a previous EWAS on famine exposure (28).
extended to early gestational exposure but not to preconceptional exposure (table S3). Therefore, we investigated whether mediation by DNA methylation (DNAm) is also present for the elevated BMI and TG in individuals with "early" exposure. The subsequent genome-wide screens for possible mediators of these associations did not yield candidates for BMI that were associated with both DNAm and outcome (Fig. 2 and table S4).

Seventeen CpGs were potential mediators for the association between early exposure and TG. Two of which were associated with both early exposure and TG (table S4). Both cg08994060 [\(\beta_{\text{mediation}} = 4.1\%\) (95% CI, 0.3 to 8.3%), \(P = 0.033\)] and cg11269166 [\(\beta_{\text{mediation}} = 3.5\%\) (95% CI, 0.4 to 7.2%), \(P = 0.027\)] mediated the association between early exposure and TG (\(R^2_{\text{LMM}} = 0.32\)). The proportion mediated was 24.5% for cg08994060 (\(P = 0.054\)), whereas cg11269166 mediated 19.4% of the association between exposure and TG (95% CI, 1.5 to 81.3%; \(P = 0.039\)).

The aggregate measure mediated the association between early exposure and TG [\(\beta_{\text{mediation}} = 0.069\) SD/log(TG) (95% CI, 0.026 to 0.116), \(P = 0.02\)], explaining 36.3% (95% CI, 15.7 to 100%; \(P = 0.02\)) of the total exposure-phenotype relationship explained by the indirect (mediated) effect as based on 10K Monte Carlo simulations.

**Table 3. Mediation analysis: DNA methylation and the association between famine exposure and triglycerides.**

<table>
<thead>
<tr>
<th>CpG</th>
<th>Location (hg19)</th>
<th>Nearest gene*</th>
<th>Methylation (SD)†</th>
<th>Rank</th>
<th>EWAS (P_{\text{FDR}})</th>
<th>(P_{\text{famine}})</th>
<th>(P_{\text{TG}})</th>
<th>Previous studies</th>
<th>(\beta_{\text{mediation}})</th>
<th>(P_{\text{mediation}})</th>
<th>Proportion mediated (%)</th>
<th>(P_{\text{proportion}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg19693031</td>
<td>chr1:145441552</td>
<td>TXNIP</td>
<td>77.5 (4.3)</td>
<td>6</td>
<td>2.6 \times 10^{-5}</td>
<td>4.8 \times 10^{-3}</td>
<td>2.3 \times 10^{-11}</td>
<td>(41–45)</td>
<td>2.6 [0.7–4.8]</td>
<td>0.005</td>
<td>28.0 [5.7–100]</td>
<td>0.026</td>
</tr>
<tr>
<td>cg18120259</td>
<td>chr6:43894639</td>
<td>LOC100132354</td>
<td>60.4 (4.7)</td>
<td>10</td>
<td>1.8 \times 10^{-3}</td>
<td>6.6 \times 10^{-4}</td>
<td>6.4 \times 10^{-8}</td>
<td>(44)</td>
<td>2.3 [0.8–4.1]</td>
<td>0.001</td>
<td>24.9 [7.5–100]</td>
<td>0.021</td>
</tr>
<tr>
<td>cg15020801</td>
<td>chr17:46022809</td>
<td>PNPO</td>
<td>36.1 (3.4)</td>
<td>12</td>
<td>3.5 \times 10^{-3}</td>
<td>7.1 \times 10^{-4}</td>
<td>6.0 \times 10^{-8}</td>
<td>(30)</td>
<td>2.3 [0.9–4.2]</td>
<td>0.001</td>
<td>25.7 [7.0–100]</td>
<td>0.022</td>
</tr>
<tr>
<td>cg06983052</td>
<td>chr1:90288099</td>
<td>LRRCD1</td>
<td>64.8 (3.8)</td>
<td>13</td>
<td>4.2 \times 10^{-3}</td>
<td>1.0 \times 10^{-5}</td>
<td>5.3 \times 10^{-6}</td>
<td>(44)</td>
<td>2.6 [1.1–4.5]</td>
<td>&lt;0.001</td>
<td>28.0 [8.8–100]</td>
<td>0.024</td>
</tr>
<tr>
<td>cg07397296</td>
<td>chr21:43655316</td>
<td>ABCG1</td>
<td>26.9 (3.8)</td>
<td>14</td>
<td>0.021</td>
<td>5.1 \times 10^{-3}</td>
<td>1.9 \times 10^{-7}</td>
<td>(49)</td>
<td>1.9 [0.6–3.6]</td>
<td>0.005</td>
<td>20.5 [4.6–97.4]</td>
<td>0.027</td>
</tr>
<tr>
<td>cg20496314</td>
<td>chr22:39759864</td>
<td>SYNGR1</td>
<td>40.2 (4.3)</td>
<td>15</td>
<td>0.032</td>
<td>3.9 \times 10^{-3}</td>
<td>1.5 \times 10^{-7}</td>
<td>(45, 46)</td>
<td>1.8 [0.5–3.5]</td>
<td>0.007</td>
<td>19.6 [3.6–88.3]</td>
<td>0.026</td>
</tr>
</tbody>
</table>

*Nearest gene within 100 kb. †The Illumina 450k array \(t\) value (ranging from 0 to 1) multiplied by 100 for easy interpretation. §The percentage of the total exposure-phenotype relationship explained by the indirect (mediated) effect as based on 10K Monte Carlo simulations.

**Cross-tissue comparison**

We measured DNA methylation in whole blood and do not have specimens from other tissues available in our cohort (24). Mean differences in DNA methylation levels vary between tissues (32), but DNA methylation level variation in blood may still reflect variation in other tissues (48). We investigated DNA methylation data across 12 postmortem tissue types from 16 cadavers (30, 31) and found that the DNA methylation patterns in blood showed broad correspondence with those in internal tissues [Spearman’s \(r = 0.42, P < 10^{-5}\); blood, liver, kidney, skin, omentum (visceral fat), subcutaneous fat, and fat from around the kidney], although the correlation between tissues was more modest when CpGs were analyzed individually in this small data set (fig. S3). DNA methylation in blood was correlated with DNA methylation in omentum and subcutaneous fat (both \(r = 0.51, P < 10^{-11}\) and fat taken from around the kidney \(r = 0.42, P = 2.7 \times 10^{-7}\)) but not with DNA methylation in the liver \((r = 0.14, P = 0.14)\).

CpG cg08994060 is located in a deoxyribonuclease I (DNaseI) hypersensitivity cluster in an intron of PFKFB3, a rate-limiting enzyme in glycolysis, and was associated with expression of this gene in whole blood (table S1). Similarly, cg11269166 is located in an enhancer located in an intron of METTL8, a gene involved in adipogenesis (47), and was associated with METTL8 expression in whole blood (table S1).

![Manhattan plot](image-url)  
**Fig. 2.** Manhattan plot: Outcome genome-wide screen for potential mediators. The −log(P value) (y axis) for each CpG relative to its genomic locations (x axis) on the 22 autosomal chromosomes for the EWAS for both early famine exposure and serum TG.
DISCUSSION
A stepwise genome-wide mediation analysis is consistent with the hypothesis that epigenetic mechanisms play a role in mediating the association between prenatal famine exposure and later-life metabolic health. Specifically, our results suggest that DNAm at an enhancer linked to PIM3 expression mediated the association between famine exposure and BMI. PIM3 influences cell growth and energy metabolism, including mitochondrial function (38). DNAm at six CpGs, including at previously TG-associated CpGs at TXNIP and ABCG1 (41–45, 49), mediated the association between famine exposure and TG. DNAm at these CpGs was likewise associated with the expression of genes implicated in cell growth and energy metabolism. Two different CpGs, at PFKFB3 (which plays a key role in glycolysis) and METTL8 (linked to adipogenesis), mediated the association between early exposure and TG. DNAm at the nine identified CpGs correlated with DNAm in various other tissues, including multiple fat deposits. DNAm at the nine CpGs associated with the expression of genes with general roles in growth and metabolism, and although mediators, they are not known to play direct roles in fat and TG metabolism. Therefore, the nine CpGs are unlikely to be directly involved in a mechanistic sense in BMI or TG but may have contributed to adverse morphological or cellular metabolic profiles with an adverse effect on metabolic health in prenatally exposed individuals.

We used a novel approach to identify possible mediators by performing a series of genome-wide analyses that tested for an association with DNAm of both the exposure and a metabolic outcome simultaneously. This resulted in a set of CpGs associated with either famine exposure or a later-life metabolic outcome, or both. The latter were formally tested for mediation. CpGs that associated with (early) famine exposure in our previous EWAS of famine exposure only in the same data set (28) were not identified as potential mediators (that is, were associated with exposure but not with outcome). In concordance with this (28) and earlier work (50–52), all the mediating CpGs overlapped regulatory elements; were linked to the expression of genes involved in growth, differentiation, and metabolism; and were associated with either famine exposure or early famine exposure. It should be noted that the latter analysis was based on a small sample size and had limited power, but the results may point to additional processes unique for those exposed during early gestation.

Several strengths and limitations of this study and mediation analyses in general are important to discuss. We observed that the association between prenatal famine and BMI or TG was mediated by single CpGs (instead of regions). This observation may be related to technical (that is, the sparsity of the Illumina 450k array) and biological factors [differential methylation may be related to very local differences in transcription factor binding, not extending across regions (29)]. We recognize that statistical evidence for mediation does not clarify the nature of the mechanism. However, it does point toward possible pathways that warrant further exploration. It is also relevant to note that DNAm was measured in whole blood, which may not be the most relevant tissue to study for BMI or TG. However, DNAm data from multiple tissues from the same donor (31, 32) show a broad agreement in DNAm patterns of the CpGs in whole blood and various fat deposits. Moreover, two previous studies showed that the associations between DNAm at PIM3 (cg09349128) and BMI (33) and DNAm at ABCG1 (cg07397296) and TG (49), on which we likewise report here, were not only found in blood but could also be validated in DNAm data from adipose tissue. Our findings may therefore point toward more general processes in other tissues, including adipose tissue, and may reflect other epigenetic marks correlated to DNAm.

Mediation analysis is sensitive to unmeasured confounding (53). We address this concern, within the limits of an observational study six decades after the exposure, on several levels with our study design and study execution. First, we analyzed the impact of prenatal famine exposure among same-sex sibling pairs concordant and discordant on famine exposure, thereby mitigating confounding by familial factors. Next, we selected individuals born before and after the famine in the same institutions as controls for those with prenatal famine exposure and rely on the quasi-experimental circumstances of the Dutch Famine, predating confounding by social characteristics related to wealth and nutrition in pregnancy. The inclusion of individuals either conceived and born before the famine or conceived and born after the famine also allows us to investigate and adjust for the influence of age, which is crucial (54). Last, we could effectively exclude confounding effects of several measured lifestyle factors including diet and smoking.

We recognize that the interpretation of mediation analyses requires prudence (53) because mediation analysis is sensitive to differences in measurement accuracy between observed exposure and an empirically measured mediator. However, this risk is reduced by the quasi-experimental setting of our study. Also, mediation analysis is equally sensitive for spurious associations as regular association analysis (53). Of the nine mediating CpGs identified here, all but cg06983052 (LRRC8D) have been related to obesity, TG, or other metabolic phenotypes in previous studies (Tables 2 and 3). In addition, mediation can also be affected by reverse causation (53). Prenatal famine exposure may be considered a causal anchor that, while not as strong as genetic variants in Mendelian randomization (53), reduces the risk of reverse causation for several of the tests performed within the mediation framework with the exception of the association between mediator and outcome (55). Longitudinal data could be used to refute reverse causation (preferably from conception onward), but by design, we did not sample at birth or any other time at a younger age before the potential onset of the metabolic complications.

The nine CpGs on which we report include the well-described CpGs mapping to PIM3 (cg09349128) (33–36, 56) and TXNIP (cg19693031) (30, 41–44) that were also reported in large Mendelian randomization studies (30, 33, 56). No evidence was found for an effect of BMI on DNAm at cg09349128 (33, 56), whereas weak evidence was found for an effect of TG on DNAm at cg19693031 (30). However, DNAm at this latter CpG is also associated with prenatal smoke exposure (45) [and not adult smoke exposure (11)], an exposure likewise linked with elevated serum TG (57). DNAm at this CpG is also associated with future type 2 diabetes risk that is independent of the traditional risk factors, which included dyslipidemia (43), analogs to the association between prenatal famine exposure, and type 2 diabetes (20). The overlap with previous EWAS findings highlights not only the fact that we are looking at epigenetic variation with a modest individual impact on variation in metabolic health but also the idea that the relevance of our results may extend beyond prenatal famine exposure.

In summary, using a systematic genome-wide approach, we show that DNAm at specific CpGs mediates a considerable proportion of the associations between prenatal famine exposure and later-life adiposity and serum TG levels. Our data are consistent with the hypothesis that the associations between exposure to an adverse environment during early development and health outcomes in adulthood are mediated by epigenetic factors. The specific causal mechanism awaits elucidation.

MATERIALS AND METHODS
Study setting
The Dutch Hunger Winter was a 6-month famine at the end of World War II that resulted from a combination of punitive measures imposed
by occupying German forces after a national railway strike, winter conditions, and fuel shortages related to war operations. Food rations were distributed centrally and rapidly dropped to below 900 kcal/day between 26 November 1944 and 15 May 1945. After 15 May, rations rapidly returned to pre-famine levels. The percentage of calories from proteins, fat, and carbohydrates in the diet was relatively constant as the food rations diminished (17, 58).

**Study subjects**
The Dutch Hunger Winter Families study is described in detail elsewhere (24). We identified 2417 singleton births between 1 February 1945 and 31 March 1946 at three institutions in famine-exposed cities in the western Netherlands whose mothers were exposed to the famine during or immediately preceding that pregnancy. We selected time controls born in 1943 or in 1947 from the same clinics. We also asked whether a same-sex sibling not exposed to the famine would be willing to participate as a sibling control. Both the famine-exposed and the time controls can therefore have a same-sex sibling as family control. Overall, 1075 interviews and 971 clinical examinations were performed six decades after the exposure. Following the Helsinki guidelines, we obtained ethical approval, both from the Institutional Review Board of Columbia University Medical Center and from the Medical Ethical Committee of the Leiden University Medical Center. The study participants provided verbal consent in a telephone interview, and in case of clinical examinations, a written informed consent was obtained.

**Famine exposure definitions**
We defined famine exposure by the number of weeks during which the mother was exposed to <900 kcal/day after the last menstrual period (LMP) recorded on the birth record. Missing or implausible LMP records (12%) were (re-)estimated from the birth weights and the date of birth (24). We considered somebody exposed to famine in gestational weeks 1 to 10, 11 to 20, 21 to 30, or 31 to delivery if such a gestational time window was entirely contained within this period and had an average exposure of <900 kcal/day during this 10-week period. As a result of the 6-month duration, some individuals were exposed to two adjacent 10-week periods. In short, pregnancies with an LMP between 26 November 1944 and 4 March 1945 were exposed in weeks 1 to 10, those with an LMP between 18 September 1944 and 24 December 1944 were exposed in weeks 11 to 20, those with an LMP between 10 July 1944 and 15 October 1944 were exposed in weeks 21 to 30, and those with an LMP between 2 May 1944 and 24 August 1944 were exposed in weeks 31 to delivery. We defined individuals exposed in any of these periods as “famine-exposed,” and those exposed in gestational weeks 1 to 10 were defined as early exposed. In addition, individuals with an LMP between 1 February and 12 May 1945 were exposed to an average of <900 kcal/day for 10 weeks before conception and were defined as the preconception exposure group.

**Characteristics**
Medical examinations were scheduled early in the morning under fasting conditions. Measurement of height was carried out to the nearest millimeter using a portable stadiometer (Seca), and body weight was measured to the nearest 100 g by a portable scale (Seca). BMI was calculated from these measures. A blood draw was performed at the start of a 75-g oral glucose test, and glucose was quickly assessed in serum by hexokinase reaction on a Modular P800 (Roche). Cholesterol measures were reported earlier (19) and were assessed using standard enzymatic assays. LDL-C was calculated for individuals with a TG concentration lower than 400 mg/dl using the Friedewald formula. Dietary intake in the last 12 months was ascertained from a 140-item food frequency questionnaire developed to assess dietary habits in an elderly Dutch population (59). This questionnaire provided estimates of macronutrient intake (total energy and the percentage of fat, protein, and carbohydrate thereof). As a measure of SES, we classified study participants on a five-level education scale, identifying individuals with primary, lower- or middle-level vocational, secondary, higher vocational, and university education.

**DNAm data**
DNAm was measured using the Illumina Infinium Human Methylation 450k BeadChip, and preprocessing was previously described by us in detail (28). Briefly, samples were randomly distributed, ensuring similar distributions of exposure periods, sex ratios, and mean ages per 96-well plate and 450k array, keeping sibling pairs together, but were randomly assigned to either the left or right column of the 450k array. We assessed data quality using both sample-dependent and sample-independent quality metrics using the R package MethylAid (60). Bisulfite conversion efficiency was assessed using the dedicated 450k probes and sequencing the IGF2 DMR0 of a random set of samples. We re-measured a subset of the genotypes measured on the 450k array with MassARRAY and checked the gender of samples using all X-chromosomal CpGs to exclude sample swaps. We used noob and Functional Normalization as implemented in the minfi package (61) using six principal components to normalize for batch effects, dye color intensity differences, and background signal. Individual measurements with a detection P value of >0.01 or zero-intensity value in one of the used color channels were set as missing. The measurement success rate per sample was >99%. Next, we removed a-specific/polymeric and non-autosomal probes, probes with <95% success rate, and those probes that were completely methylated or unmethylated in all major cell types in whole blood. Methylation percentages in text, figures, and tables reflect microarray β-value estimates (which range from close to zero to close to one or 0 to 100%, as denoted throughout). For the individuals for whom we have genome-wide DNAm data (Illumina 450k array), 348 individuals met our definition of famine exposure, of whom 73 also meet the definition of exposure during weeks 1 through 10 of gestation (early exposure). In total, 94 individuals met our preconception exposure definition. In addition, we have 463 individuals with no famine exposure directly before conception and during gestation. DNAm data are available upon request.

In addition, we used DNAm data from 16 obductions (31, 32) from which samples were collected within 12 hours of death (mean age, 62.8 years). In concordance with the ethical guidelines in the Code for Proper Secondary Use of Human Tissue in the Netherlands (Dutch Federation of Medical Scientific Societies), these samples were anonymized, and raw data have been deposited in the National Center for Biotechnology Information’s Gene Expression Omnibus (GSE78743). To explore the relation with gene expression and validate associations with BMI or TG, we referred to the DNAm data of 3296 individuals (30) measured within the Genome of the Netherlands reference project [GoNL; deposited at the European Genome-phenome Archive (EGA) under accession number EGAC00001000277]. Preprocessing and normalization were done as described above for both data sets. R Code for the quality control and normalization is provided at https://git.lumc.nl/molepi/DNAmArray.

**Gene expression data**
We used RNA-seq data created from total RNA extracted from whole blood using the TruSeq v2 library protocol and 2 × 50-base pair paired-end
sequencing on an Illumina HiSeq2000 for 2044 of the 3296 individuals for which we have DNAm data from the GoNL reference project (likewise under EGA accession number EGAC0001000277). Data processing is detailed elsewhere (29) and consisted in removal of bar-coded adapters, low-quality reads, alignment against genome build NCBI37, and gene-annotation Ensembl v.71, normalization, and GC-bias correction.

**Statistics**

All analyses were performed in the R programming environment (R.3.2.2). Reanalysis of the associations between phenotypes and famine exposure was performed with linear mixed-effects models (LMMs) with one of the phenotypes (BMI, TG, glucose, and LDL-C) as the dependent variable using the R lmerTest package (62). The sibling-pair identifier was included as a random effect, and where appropriate, we corrected for gender and age at examination. We used the restricted maximum likelihood method and variance components for the covariance matrix structure of both DNAm and gene expression data were corrected using the R package bacon (71). Corrected estimates, standard errors, and test statistics for only those genes within 100 kb were then assessed and corrected by FDR for multiple testing. All P values reported were two-sided.

**SUPPLEMENTARY MATERIALS**

Supplemental material for this article is available at http://advances.sciencemag.org/cgi/content/full/4/eaao4364/DC1

---

**REFERENCES AND NOTES**


Members of the BIOS Consortium: Peter A. ’t Hoen, René Pool, Marleen M. van Greevenbroek, Coen D. Stehouwer, Carla J. van der Kallen, Casper G. Schalkwijk, Cisca Wijmenga, Sasha Zhernakova, Ettje F. Tigchelaar, Marian Beekman, Joris Deelen, Diana van Heemst, Jan H. Veldink, Leonard H. van den Berg, Cornelia M. van Duijn, Albert Hofman, André G. Uitterlinden, P. Mila Jhamai, Michael Verbiest, Marijn Verkerk, Ruud van der Breggen, Jeroen van Rooij, Nico Lakenberg, Hailiang Mei, Jan Bot, Dasha V. Zhernakova, Peter Van ’t Hof, Patrick Deelen, Irene Nooren, Matthijs Hoed, Marijn Vermaat, Marc Jan Bonder, Feerek van Dijk, Michiel van Galen, Wibowo Arindrarto, Szymon M. Kielbasa, Morris A. Swertz, Aaron Isaacs, and Lude Franke.

Submitted 21 July 2017
Accepted 3 January 2018
Published 31 January 2018
10.1126/sciadv.aao4364