Journal of Developmental Origins of Health and Disease

www.cambridge.org/doh

Methylation profiles at birth linked to early childhood obesity

Delphine Lariviere[1], Sarah J.C. Craig[2,3], Ian M. Paul[3,4], Emily E. Hohman[5], Jennifer S. Savage[5,6], Robert O. Wright[7], Francesca Chiaromonte[3,8,9], Kateryna D. Makova[2,3] and Matthew L. Reimherr[3,8]

[1]Department of Biochemistry and Molecular Biology, Penn State University, University Park, PA, USA; [2]Department of Biology, Penn State University, University Park, PA, USA; [3]Center for Medical Genomics, Penn State University, University Park, PA, USA; [4]Department of Pediatrics, Penn State College of Medicine, Hershey, PA, USA; [5]Center for Childhood Obesity Research, Penn State University, University Park, PA, USA; [6]Nutrition Department, Penn State University, University Park, PA, USA; [7]Icahn School of Medicine, Mount Sinai, New York, NY, USA; [8]Department of Statistics, Penn State University, University Park, PA, USA and [9]L'EMbeDS, Sant'Anna School of Advanced Studies, Piazza Martiri della Libertà, Pisa, Italy

Original Article

Cite this article: Lariviere D, Craig SJC, Paul IM, Hohman EE, Savage JS, Wright RO, Chiaromonte F, Makova KD, and Reimherr ML. (2024) Methylation profiles at birth linked to early childhood obesity. *Journal of Developmental Origins of Health and Disease* **15**: e7, 1–10. doi: 10.1017/S2040174424000060

Received: 10 December 2023
Revised: 23 January 2024
Accepted: 1 February 2024

Keywords:
childhood obesity; epigenetics; methylation risk score

Corresponding author:
K. D. Makova; Email: kdm16@psu.edu

Abstract

Childhood obesity represents a significant global health concern and identifying its risk factors is crucial for developing intervention programs. Many "omics" factors associated with the risk of developing obesity have been identified, including genomic, microbiomic, and epigenomic factors. Here, using a sample of 48 infants, we investigated how the methylation profiles in cord blood and placenta at birth were associated with weight outcomes (specifically, conditional weight gain, body mass index, and weight-for-length ratio) at age six months. We characterized genome-wide DNA methylation profiles using the Illumina Infinium MethylationEpic chip, and incorporated information on child and maternal health, and various environmental factors into the analysis. We used regression analysis to identify genes with methylation profiles most predictive of infant weight outcomes, finding a total of 23 relevant genes in cord blood and 10 in placenta. Notably, in cord blood, the methylation profiles of three genes (PLIN4, UBE2F, and PPP1R16B) were associated with all three weight outcomes, which are also associated with weight outcomes in an independent cohort suggesting a strong relationship with weight trajectories in the first six months after birth. Additionally, we developed a Methylation Risk Score (MRS) that could be used to identify children most at risk for developing childhood obesity. While many of the genes identified by our analysis have been associated with weight-related traits (e.g., glucose metabolism, BMI, or hip-to-waist ratio) in previous genome-wide association and variant studies, our analysis implicated several others, whose involvement in the obesity phenotype should be evaluated in future functional investigations.

Introduction

Obesity affects over 40% of Americans,[1] including nearly 20% of children.[2] Childhood obesity is associated with various disorders across the life course, including hypertension, hypercholesterolemia, and insulin resistance.[3–5] To maximize the benefit of preventive interventions,[6–8] early identification of children who are most at risk for developing obesity is paramount.

Weight is a complex trait influenced by many factors, including the environment (e.g., diet, activity level, medications), genetics, epigenetics, the microbiome, and the metabolome of individuals. Previous studies have indicated that 40-80% of variation in BMI can be explained by genetic factors.[9,10] However, the cumulative effect of single nucleotide polymorphisms (SNPs) identified so far does not account for all of the variation attributed to genetics. Specifically, earlier genome-wide association studies (GWASs) have only been able to explain approximately 3% of variation in BMI, and more recent studies[11] considering SNPs significant at the genome-wide level explain up to 6% of such variation. Less stringent studies or meta-analyses raised this percentage to over 20% (reviewed in Bouchard *et al.* 2021),[10] but still failed to explain the observed heritability in obesity, which approaches 50%.[10] Despite this gap, polygenic risk scores (PRSs) are being widely developed to combine variants from GWASs to assess an individual's risk for disease.[12] These scores have been developed for adults[13] and more recently for children.[14]

In addition to genetic factors, epigenetic modifications could provide important insights into an individual's risk for obesity because they can be heritable when located in the germline, and modifiable by environmental factors.[15] Epigenomics, the study of epigenetic modifications on a genome-wide scale, is a field of research that links genes and disease to provide a complex picture accounting for changes due to environmental influences across the lifetime. The most common epigenetic modification of DNA is cytosine methylation at CpG sites. Methylation

plays a role in repressing gene expression when located in regulatory regions[16] and has been linked to active gene transcription when located within the gene body.[17] The proposed molecular mechanisms of gene body methylation range from silencing of repetitive elements[18] to affecting nucleosome positioning[19] and histone modifications.[20] Analogous to constructing PRSs with SNP data, methylation risk scores (MRSs) have been recently developed.[21,22] MRSs are linear combinations of methylation states across multiple CpG sites and may be useful in the clinical setting as these epigenetic marks can be influenced by environmental conditions and thus could be used to monitor changes in disease risk over time.[21]

In the context of childhood obesity, some studies have shown differences in peripheral blood methylation profiles between children with and without overweight.[23,24] Other studies have identified CpG loci whose methylation status in cord blood is linked to adiposity in children between 3 and 7 years of age,[23] as well as up to 18 years of age.[24] Cord blood methylation profiles in children were shown to be influenced by maternal methylation profiles and by environmental factors that impact pregnancy.[25,26] Additionally, MRSs have been found to be associated with BMI in adults,[27] as well as in children.[28] However, the MRSs used in previous children studies were informed by BMI Epigenome-Wide Association Studies in adults,[28,29] so it is still not known whether there are specific gene methylation patterns at birth that are linked to early childhood growth.

In this work, we capitalized on a cohort of second-born siblings to participants in the Intervention Nurses Start Infants Growing on Healthy Trajectories (INSIGHT) Study.[6,8] Specifically, these "SIBSIGHT" study participants were part of an observation-only longitudinal evaluation of second-born siblings. For this study, we investigated whether early childhood growth is associated with methylation in cord blood and placenta samples of 48 children from the SIBSIGHT cohort. Early childhood growth of children from the INSIGHT and SIBSIGHT cohorts has been extensively studied, providing evidence for a successful early-life intervention aimed at preventing childhood obesity for both siblings.[7,8,30] Along with insights into the effects of dietary intake,[31,32] sleep,[33,34] and infant temperament,[35] prior findings by our group have identified associations between early childhood growth and the composition of the oral microbiome,[36] the gut metabolome,[37] the stool microtranscriptome,[38] and the genome.[14] Characterizing an association between gene methylation at birth and weight outcomes in children complements such studies, providing another avenue for identifying risk factors, adapting interventions – and thus preventing early-life obesity and later life comorbidities. Here we used Illumina methylEpic arrays to establish genome-wide methylation profiles for placenta and cord blood tissues, and leveraged a wealth of additional information collected by SIBSIGHT. We tested a hypothesis that gene body methylation profiles at birth could be associated with weight outcomes in the first six months after birth.

Methods

Methylation data collection

We collected 48 matching samples of cord blood and placenta tissue from children enrolled in the SIBSIGHT study.[6,31] This was a convenience sample of second-born siblings from families enrolled in and consented for a clinical trial with their firstborns. We were therefore able to consent them for participation for this study of second-borns as part of a larger NIH-funded effort evaluating differences between first- and second-born children. In short, the families were already familiar with the research protocol and what this meant for their families, making consent and enrollment efficient. A list of the covariates employed in our analysis, with their summary values across the children included in this study can be found in Table 1.

Table 1. Summary of SIBSIGHT covariates used in the analysis

Covariate	Value
Child CWG z-score Average (SD)	−8.868 × 10−4 (1.037)
Child BMI Average (SD)	17.7 (1.5)
Child weight/length Average (SD)	11.8 (1.1)
Mother BMI Average (SD)	24.5 (4.5)
Father BMI Average (SD)	28.6 (4.5)
Child sex N = female (%)	27 (56%)
Gestational Duration (weeks) Average (SD)	38.9 (1.1)
Mode of Delivery N = vaginal (%)	34 (70.8%)
Maternal Age (years) Average (SD)	31.8 (4.3)
Gestational Diabetes N = controlled by diet & exercise (%)	3 (6.25%)
Smoking During Pregnancy N = smoked	1
Maternal Illness during pregnancy (*e.g.*: Thyroid disorders) (N = none)	47
Maternal Medications During Pregnancy N = took medications (%)	33 (68.8%)
Infant Feeding Mode at 4 weeks $N \geq 80\%$ breast milk (%)	32 (66.7%)
Infant Feeding Mode at 16 weeks $N \geq 80\%$ breast milk (%)	25 (52.1%)
Infant Feeding Mode at 28 weeks $N \geq 80\%$ breast milk (%)	19 (39.6%)

SD, standard deviation.

At the time of birth, cord blood samples were collected in K_2EDTA coated vacutainers (Becton, Dickinson, and Company) and stored at 4°C until picked up by the research team. Samples were then stored at −80°C. DNA was isolated using the Qiagen DNeasy Blood and Tissue kit (Qiagen). Purified genomic DNA was then bisulfite-converted using EZ Methylation Kit (Zymo Research).

Placentas were stored at 4°C after delivery before processing. 1 cm^3 pieces of the placenta were dissected from the fetal side, proximal to the area where the umbilical cord attaches. Tissues were formalin-fixed and paraffin-embedded. DNA from these tissues was extracted with the ReliaPrep FFPE gDNA Miniprep System (Promega) and then assessed with the FFPE QC kit (Illumina) for quality. Samples passing quality thresholds were then bisulfite-converted with the EZ Methylation Kit (Zymo Research), and then treated DNA was restored following the Infinium HD FFPE Restoration protocol (Illumina).

Bisulfite-converted DNA from both tissues was then analyzed on the Infinium MethylationEPIC chip (Illumina) in the Genome Sciences Facility at Penn State Hershey College of Medicine.

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

Weight outcomes data collection

For each child enrolled in this study, weight and length (via recumbent length board, Shorr Productions) were collected at birth and six months after birth, and BMI (kg/m^2) and weight-for-length (kg/m) were calculated.[39,40] Additionally, conditional weight gain (CWG) *z-scores* were calculated for each child using anthropometrics at birth and six months, adjusted for sex and age.[41] CWG *z*-scores are the standardized residuals from a linear regression of the weight-for-age *z*-score at six months on the weight-for-age *z*-score at birth (length-for-age *z*-score at birth and six months and exact age at the six-month visit are used as covariates in the regression). CWG *z*-scores are normally distributed and have a mean of 0 and a standard deviation of 1. Positive *z*-scores indicate above average weight gain (i.e., rapid infant weight gain) compared to other infants in the sample, and have been shown to be a risk factor for obesity later in life.[42] We standardized the BMI and weight-for-length ratio data by sex to remove the impact of the differences between sexes on the association with methylation profiles. This standardization is done by separating the two populations by sex and, for each, subtracting the mean and dividing by the standard deviation. Tests for normality were performed in R using the base stats package.

Methylation data preprocessing

Raw signal reads from the chip were converted into Beta signals (β = intensity of the methylation signal/[intensity of the methylation signal + intensity of the unmethylated signal + 100]) using the Minfi package in R.[43] The Minfi package was also used to screen the data for quality. The first screening removed CpG sites where the intensities of the methylated signal and the unmethylated signal were both close to background level. 452,567 CpG sites were removed that way. Further quality control included screening the data for outliers, excluding sex chromosomes, and excluding sites with known SNPs that could have caused false positives or negatives (see Table S1 for a summary of CpGs removed during this screening). After quality control, one placenta sample was removed from further analyses due to the low quality of the methylation data. Next, the array signals were normalized. First, we normalized within the array, which included background correction and normalization of signal intensity. Each chip contains control sites used to normalize between samples. Second, we utilized the Beta Mixture Quantile (BMIQ) normalization (one of the most popular methods found in the literature for MethylEpic analyses)[44] to normalize the signal from the Infinium I (InfI) and Infinium II (InfII) probes utilized on the MethylationEPIC array. BMIQ decomposes density profiles in three states: unmethylated, hemimethylated, and fully methylated. It rescales the InfII distribution to the corresponding InfI distribution. Both normalization steps were performed utilizing tools within the Minfi package (Figure S2).

After preprocessing we have one Beta signal for each CpG site, which corresponds to its methylation level. These values range from 0 (fully unmethylated) to 1 (fully methylated). Using the default density plot function included in the Minfi package, we visualize the distribution of individual CpG sites methylation levels. In a sample containing a single cell type, with identical methylation states between cells, we expect two peaks near 0 and 1. Values in between 0 and 1 indicate a mix of unmethylated and methylated sites in the sample, which can indicate a mix of cell types or cell states.

Regression analyses

To identify factors that could impact the weight outcomes (BMI) other than methylation profiles and sex, we performed a LASSO regression analysis[45,46] (a method that performs predictor selection) on environmental factors, such as feeding mode, and family history, such as parental BMI and pregnancy duration (Table 1). We found no significant associations (Figure S2).

After the normalization performed during preprocessing, we grouped the methylation data by genes. Specifically, we averaged the Beta signals of CpG sites contained within the genomic coordinates of a gene to calculate the gene's methylation level (Figure S3). We then ran LASSO regressions separately for the two tissues and, for each tissue, considering the three different weight outcomes – for a total of six regressions. We used the R package glmnet (LASSO and Elastic-Net Regularized Generalized Linear Models). The tuning parameters used for various LASSO runs can be found in Table S2; they were selected minimizing the cross-validation Mean Squared Error, as shown in the standard result plots produced by glmnet. Some of the LASSO fit analyses gave variable results for correlation with weight-to-length outcome. When repeating the analyses with the same parameters, the shape of the LASSO plot was changing, and, while a few genes were repeatedly selected, some results were not reproducible. To select the most predictive genes when the LASSO gave very variable results, we repeated the analysis until we obtained 10 profiles with the "check mark" shape plot, and selected the best model that included the most commonly selected genes across the replicate 10 analyses. After running each of the LASSO regressions, in order to reduce the bias, this technique creates in the estimation of the regression coefficients, we performed a post-selection fit – i.e. an OLS fit restricted to the set of predictors selected by the LASSO. We also ran marginal regressions for each individual predictor selected by the LASSO; the coefficient estimates from these regression can be considered alongside those produced by the post-selection OLS joint fit, as additional quantifications of the effects of each selected predictor.

Methylation risk scores calculation

Methylation risk scores (MRS) were calculated as described in.[22] Briefly, they are a sum of m gene methylation values c (from section "Methylation data preprocessing") with OLS estimated regression coefficients as weights w (from section "Regression analyses"):

$$MRS = \sum_{i=1}^{m} w_i c_i \quad (1)$$

MRSs were calculated for each weight outcome separately using the 31 genes for which a significant relationship was determined by the regression analysis between the weight outcome and methylation status (Tables S3–S6). The association between MRS and weight outcome was determined by linear regression using the lm function in the basic stats package of R[47] using MRS as the predictor and weight outcome as the response.

Validation datasets

Cord blood methylation data from the PROGRESS cohort[48] was used in validation analyses (dbGaP: phs002754.v1.p1). This cohort consists of 1,001 individuals from Mexico City who were followed from birth through 18 years. Methylation data was downloaded from dbGaP (phs002754.v1.p1) and height and weight data were

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

provided by study authors.[48] CWG *z-scores*, six-month BMI, and six-month weight-for-length were all calculated as described above.

To validate our results on the PROGRESS cohort, we preprocessed the data as we did for the SIBSIGHT cohort (see above), and confirmed the absence of association with the covariates available for this dataset (mother BMI, smoking, and disease during pregnancy). We performed linear regressions on the genes that were selected as predictors in the SIBSIGHT cohort. These regressions have been run both for each individual gene and as joint regression using all of the genes. We used the CWG *z*-scores to perform the linear regression with the genes associated with CWG in the SIBSIGHT cohort, and similarly for the BMI and weight/length ratio.

MRSs were calculated as described above, using the regression coefficients from SIBSIGHT as weights, *w*. As with SIBSIGHT, the association between the MRS and the phenotypes were calculated using the linear regression (lm) function in the basic stats package of R.

Results

We collected placenta and cord blood samples at the time of birth from 48 SIBSIGHT study participants.[6,32] For each sample, we used the Illumina MethylationEPIC array to determine methylation profiles across 575,132 CpG sites genome-wide. After quality control and clustering (see Methods for details), we grouped methylation signals from 293,090 CpGs into 20,108 genes. The number of CpG sites per gene ranged from 1 to 814, with average and median counts of 15 and 7, respectively.

We evaluated three weight outcomes of participating children: the conditional weight gain *z*-score (CWG, a standardized measure of change in weight from birth to six months of age, see Methods for details), BMI (weight/(length)^2) at six months, and the ratio of weight-for-length at six months. All three measures showed regular, Gaussian-like distributions across our participants (Figure S1; the Shapiro-Wilk test did not reject normality; CWG p-value = 0.416, BMI p-value = 0.529, weight-for-length p-value = 0.269). Weight outcomes at six months were chosen as they are the first outcomes we measured after collection of samples at birth.[49,50]

Impact of covariates on weight outcomes

Prior to evaluating the associations between methylation profiles and weight outcomes, we assessed whether non-epigenetic covariates showed significant associations with the latter and should therefore be taken into account in downstream analyses. The non-epigenetic covariates we considered (Table 1) were maternal BMI, age, and health-related variables (presence/absence of gestational diabetes, gestational weight gain, presence/absence of illness, and medication usage during pregnancy), gestational length, sex of the child, and infant feeding mode (i.e. breastfeeding or formula) at the age of 4, 16, and 28 weeks. Only one participating mother reported smoking, so this variable was excluded from the analysis. In order to determine which, if any, of the non-epigenetic covariates had an association with the infant weight outcomes, LASSO regressions were performed[51] using each of the three weight outcomes as the response and the above-listed covariates as predictors. The only significant associations found were those of the sex of the child with weight-for-length and BMI (Figure S2). The initial analyses we performed using BMI and weight-for-length showed associations with only genes differentially expressed in males and females. To account for this, we standardized these two weight outcomes by sex (see Methods for details). The CWG calculation already accounts for sex. We also performed a stratified analysis for each sex, but the low size of the samples led to a low confidence in the results (Supplementary Note 1).

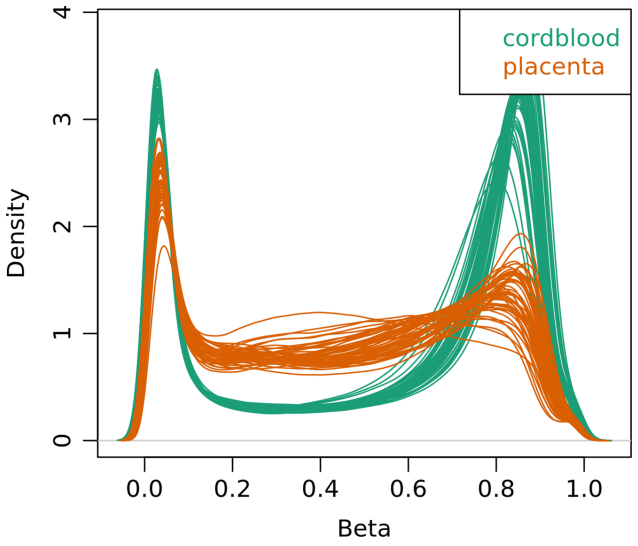

Figure 1. Density plots of Beta values describing the methylation state of CpG sites. Each line corresponds to an individual sample. Smoothing was performed with the function density plot from the minfi package in R. The distributions for the 48 cord blood samples are shown in green, and those for the 48 placenta samples are shown in orange.

Differences in methylation profiles between cord blood and placenta

The methylation state of a CpG site is determined by calculating the ratio between the methylated and unmethylated fluorescent signals from the microarray. This ratio is referred to as the methylation Beta signal.[52] The distribution of the methylation Beta signals across CpG sites differed between the two tissues analyzed (shown for each of the 48 children in Fig. 1). The cord blood samples had the expected bimodal Beta value distribution with a strong peak at $\beta < 0.2$ (hypomethylated CpGs) and a less pronounced peak at $\beta > 0.7$ (hypermethylated CpGs). However, the placenta samples had a poorly defined peak at $\beta > 0.7$, with more CpGs having values between $\beta = 0.2$ and $\beta = 0.7$. This suggests that our placenta samples contained either hemimethylated CpGs or heterogeneous cells with a mix of CpG methylation profiles. This is consistent with previous work showing that up to 37% of the placental genome is partially methylated.[53]

Association study to identify differentially methylated genes

To identify genes with methylation patterns associated with children's weight outcomes, we again used LASSO regression. In total, we performed six regressions, one for each tissue type and weight outcome combination. For each regression, we computed the average methylation states (Beta signals) across the CpGs for each gene, and used these averages as predictors. The results are summarized in Figure 2a. The LASSO regressions for cord blood and placenta identified, respectively, eight and ten genes whose methylation levels were significant predictors of CWG. Additionally, LASSO regressions identified four and 31 genes

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

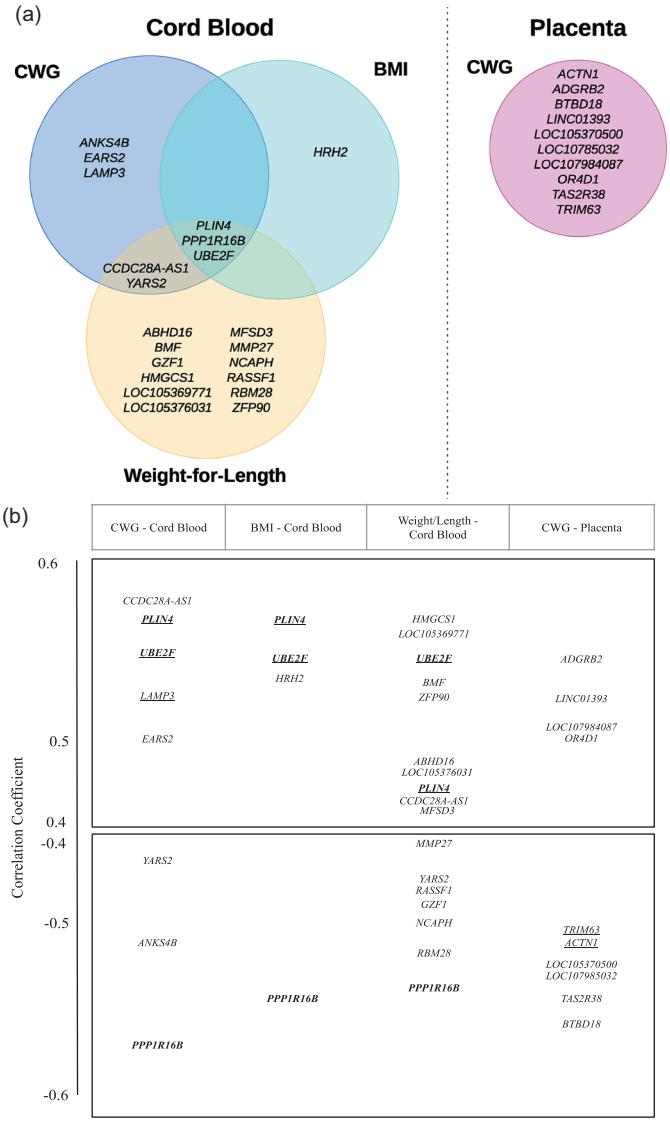

Figure 2. Genes whose methylation levels in cord blood and placenta are predictive of weight outcomes. The outcomes considered are conditional weight gain (CWG), body mass index (BMI), and weight-for-length (weight divided by length). (***a***) a Venn diagram of the relevant genes, as identified by LASSO regressions. (***b***) Gene placement along the vertical axis corresponds to the correlation coefficient between each gene selected by the LASSO fit and the weight outcome. In bold are genes selected across multiple outcomes, and underlined are genes associated with weight outcomes in previous studies (see discussion). Only CWG was associated with differentially methylated genes in the placenta.

whose methylation levels in cord blood were significant predictors of BMI and weight-for-length, respectively (Table S4). In contrast, we did not identify any genes whose methylation in placenta was significantly associated with these two weight outcomes. Notably, in cord blood, there were three genes (*PLIN4*, *PPP1R16B,* and *UBE2F*) whose methylation levels were selected as significant predictors of all three weight outcomes, with similar coefficient estimates in the three regressions. There were no "shared genes" among those selected for cord blood and placenta (Fig. 2b) and there was no correlation between the methylation levels of the selected genes in the two tissues *(PLIN4* – Pearson's correlation = −0.2192, *p*-value = 0.1388*; UBE2F* – Pearson's correlation = −0.0265, *p-value* = 0.8595*; PPP1R16B* – Pearson's correlation = 0.0967, *p-value* = 0.5179*).* We report estimated coefficients from the LASSO regressions in Tables S3-S6; these express effect strength and sign: a positive regression coefficient can be interpreted as a higher methylation level being associated with an increased weight outcome, and a negative regression coefficient as a higher methylation level being associated with a decreased weight outcome.

We took a closer look at the CpG sites in the selected genes and performed a LASSO analysis using these sites. We then mapped the CpG sites that were associated with the three weight outcomes in cord blood, and with CWG in placenta (Supplementary Note 2). We did not notice any patterns in the location of CpGs within the genes, but we could see that the same sites were driving the gene association for genes linked with several weight outcomes. We also noticed that not all genes showed associated CpG sites, suggesting that the overall methylation state of the gene is driving the association instead of specific sites.

We found that several genes selected in SIBSIGHT were also significantly related to child weight outcomes in an independent dataset the PROGRESS[54] cohort. PROGRESS is a freely accessible dataset of children from Mexico City, and comprises both cord blood DNA methylation data and longitudinal growth information for the children. Considering CWG as the weight outcome, and

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

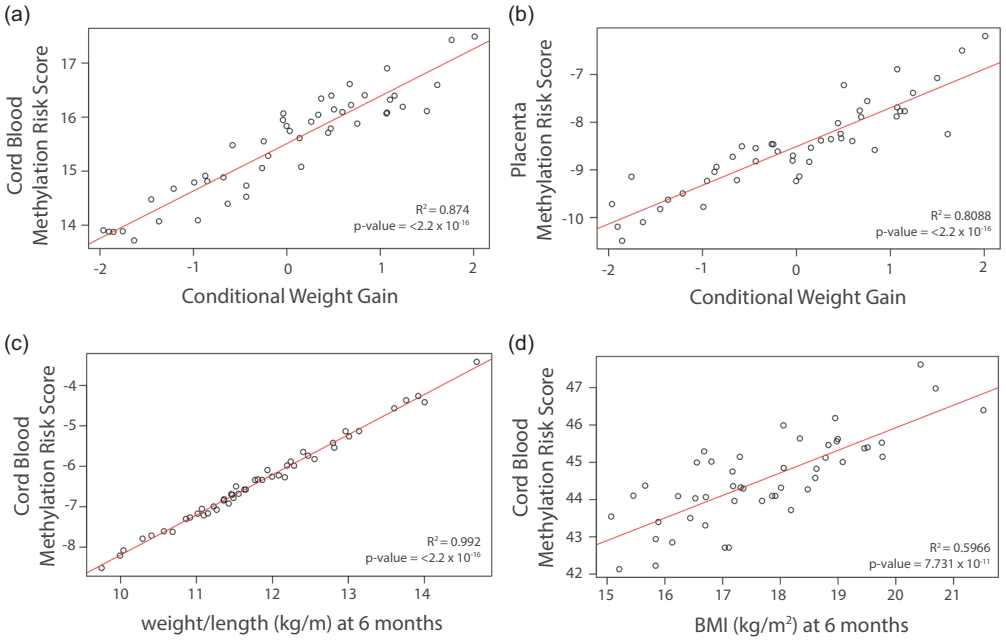

Figure 3. Relationship between MRS and weight outcomes. (***a***) Cord blood MRS vs. conditional weight gain. (***b***) Placenta MRS vs. conditional weight gain. (***c***) Cord blood MRS vs. weight-for-length ratio. (***d***) Cord blood MRS vs. body mass index. Note: placental methylation does not produce a methylation risk score for BMI or weight-for-length as there was no relationship between gene methylation patterns and either of these weight outcomes.

regressing it on one gene at a time, seven out of the eight genes selected in SIBSIGHT had a significant *p*-value also in PROGRESS. When regressing CWG on all eight genes jointly no genes remained significantly associated using 0.05 as a significance threshold; although, *PPP1R16B* had a *p*-value of 0.052 (Table S7). Considering six-month BMI as the weight outcome, two out of four genes (*PLIN4* and *UBE2F*) selected in SIBSIGHT were significant in the joint regression in PROGRESS (Table S8). Finally, considering six-month weight-for-length as the weight outcome, one (*SMIM20*) of the 31 genes selected by SIBSIGHT was significant and one gene (*UBE2F*) had a *p*-value of 0.066 (Table S9). It is notable that the genes that were selected using multiple weight outcomes in SIBSIGHT (*PLIN4, PPP1R16B,* and *UBE2F*) were also significantly associated with phenotypes in the independent PROGRESS cohort.

Methylation risk score

Using results from the above LASSO regressions for the SIBSIGHT cohort, we generated a methylation risk score (MRS) for each growth outcome. These are weighted scores calculated as linear combinations of gene methylation Beta signals weighted by regression coefficient estimates obtained from post-LASSO Ordinary Least Squares fits (see Methods for details). Figure 3 shows the relationship between each MRS and the corresponding growth outcome. The associations were strong and significant in all cases, with high in-sample R-squared (cord blood CWG, adjusted R-squared = 0.874, *p*-value ≤ 2.2 × 10–16, Fig. 3a; placenta CWG, adjusted R-squared = 0.8088, *p*-value ≤ 2.2 × 10–16, Fig. 3b; cord blood BMI, adjusted R-squared = 0.992, *p*-value ≤ 2.2 × 10–16, Fig. 3c; weight-for-length, adjusted R-squared = 0.5966, *p*-value 7.731 × 10–11, Fig. 3d). Furthermore, there was still a significant relationship between these scores (calculated with phenotypes at 6 months) and the corresponding phenotypes at 1 and 2 years (Table S10). Using the independent PROGRESS cohort, however, these MRSs did not have a significant relationship with weight outcomes (Table S11).

Discussion

In this study, we analyzed the methylation profiles of placenta and cord blood samples collected at birth. Using three outcomes characterizing early childhood growth, we identified genes whose methylation levels in these tissues are associated with weight gain during the first six months after birth. Comparing results of LASSO regression runs as well as the Ordinary Least Squares (OLS) regression of selected predictors across weight outcomes and tissues (Figure S5), we found that CWG and BMI provide more reliable results in cord blood when reproducing the analysis, with CWG having a higher adjusted R-squared than BMI (0.87 and 0.64 respectively). The weight-for-length ratio in cord blood had the highest R-squared (0.97) but the results were more variable. Indeed, attempts to replicate the correlation analysis with the weight-for-length ratio led to frequent low quality LASSO plots (with no associated genes or no clear minimum mean squared error), and a highly variable list of correlated genes, even after filtering for *p*-value (see Methods). For the placenta methylation, only CWG exhibited correlation with a set of gene methylation states, whereas BMI and weight-for-length did not. More generally, we found that, compared with the placenta methylation data, the cord blood methylation data presented a lower number of mixed methylation profiles, more genes associated with weight outcomes, and higher R-squared of the OLS regression on identified predictors (0.68 for predictors associated with CWG in the placenta).

Genes whose methylation levels in cord blood are predictive of weight outcomes

In cord blood, we found three genes whose methylation levels were significantly associated with all three weight outcomes in SIBSIGHT and with outcomes in an independent cohort (PROGRESS). These are discussed below, followed by a discussion of genes identified as significant predictors for only one of the outcomes.

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

PLIN4

One of the genes significantly associated with all three weight outcomes, and always with a positive sign (higher methylation inducing higher weight outcomes), was *PLIN4*. The protein encoded by this gene (Perilipin 4) is a member of the PAT family of lipid storage droplet proteins.[55] It is an important regulator of lipid storage. Low levels of expression of this protein have been associated with an increase in weight status[56] of adults. Changes in *PLIN4* methylation have been observed after weight loss, with hypermethylation in the promoter region before vs. after gastric bypass surgery in adults.[57] *PLIN4* has also been classified as a putative obesogen in children, and was shown to be differentially methylated between obese and non-obese children in another study.[58]

PPP1R16B

Another gene significantly associated with all three weight outcomes, and always with a negative sign (higher methylation inducing lower weight outcomes), was *PPP1R16B*. The protein encoded by *PPP1R16B* is phosphatase 1 (*PP1*) regulatory inhibitory subunit 16B,[59] which is also referred to as *TIMAP* or *ANKRD4*.[60] PP1 is involved in many essential cellular mechanisms and is part of a large interactome with over 200 interactors identified in vertebrates.[61] Studies of *PPP1R16B* showed its high levels of expression in endothelial cells and suggested that PP1 is involved in endothelium stability and permeability.[60] The activity of *PPP1R16B* has been shown to play a role in several diseases, including obesity and diabetes mellitus.[60]

UBE2F

Finally, the third gene significantly associated with all three weight outcomes, and always with a positive sign (higher gene methylation inducing higher weight outcomes), was *UBE2F*. The protein encoded by *UBE2F* (Ubiquitin Conjugating Enzyme E2F) is a ubiquitin-protein ligase involved in post-translational modifications of proteins through the addition of ubiquitin-like protein NEDD8.[62] Previous studies have shown an association between the expression of this gene and BMI in children.[63] In animal models, *UBE2F* has been shown to be expressed at higher levels in the adipose tissue of obese rats compared to lean rats.[64]

Other genes

We also identified several genes whose methylation level was significantly associated with only one weight outcome. Some such genes were also associated with obesity or an obesity-related trait in previous studies. One category of genes we identified were genes linked to nutrient metabolism, e.g. *ANKS4B*, *LAMP3*, as well as *PPP1R16B* (discussed above). The protein encoded by *ANKS4B* (Ankyrin Repeat And Sterile Alpha Motif Domain Containing 4B) plays a role in the epithelial brush border differentiation, controlling the microvilli organization and length.[65] It is involved in pancreas development and function,[66,67] affecting the secretion of insulin. This function could explain its link to weight gain. In our study, we found a negative association between CWG and cord blood methylation levels of *ANKS4B*. The protein encoded by *LAMP3* (Lysosomal Associated Membrane Protein 3) is involved in hepatic lipid metabolism and is overexpressed in patients with non-alcoholic fatty liver disease as well as in obese mice.[68] Our analysis indicated that *LAMP3* methylation is positively associated with CWG. The 33 additional genes implicated by our study but not already documented in the literature as being linked to obesity or metabolism (see Tables S3-S5) should be further analyzed in functional studies aimed at determining how they may influence weight gain in early childhood.

Genes whose methylation levels in placenta are predictive of weight outcomes

In placenta, we found ten genes whose methylation levels were significantly associated with the CWG outcome (see Fig. 3, Table S6). Four were identified as being involved in body weight and weight gain in prior studies, two have not been previously associated with obesity or obesity-related traits in adults, and four are putative and of unknown function. Among previously studied genes, *TRIM63,* encoding for E3 ubiquitin ligase MURF1, has been linked to skeletal muscle atrophy and is overexpressed in obese rats compared to lean rats.[69,70] Methylation levels of *TRIM63* had a negative association with CWG in our study. *ADGRB2* is part of the adhesion G-protein-coupled receptor genes family, which is linked to insulin secretion in humans[71] and modulation of adipogenesis and adipocyte function.[72] We found that methylation levels of *ADGRB2* had a positive association with CWG. *ACTN1* has been shown to be involved in adipogenesis[73] and weight regain after weight loss.[74,75] In rats, it is up-regulated in the brain of animals with a high-fat diet.[76] *ACTN1* had a negative association with CWG in our study. Finally, *TAS2R38* has been shown to be involved in the perception of bitter taste,[77] and unrelated studies documented a link between the perception of bitterness and obesity in adults[78,79] and male children.[80] Methylation levels of *TAS2R38* had a negative association with CWG in our study. These links suggest that these genes should be investigated further.

Methylation of genes previously associated with BMI

To evaluate whether the genetic factors of weight gain overlap with the epigenetic factors, we looked at the methylation states of a set of genes we previously identified as containing variations associated with BMI[14] (Supplemental Note 3). The absence of correlation between the methylation state of these genes and the weight outcomes in our cohort suggests that different genes are involved in the genetic and epigenetic regulation of obesity.

Methylation risk scores as predictors of weight outcomes

A growing trend in genetics is to generate PRSs for complex diseases because these types of disorders are often influenced by a large number of genetic variants, each with a small effect size.[81] PRSs, while not deterministic, can indicate which patients have a higher risk of developing certain conditions, which can aid in the establishment of intervention and/or treatment plans. MRSs have a similar advantage, capturing the cumulative effect of many CpG sites, or in this case the methylation signal of several genes, with small effect sizes. We developed MRSs for three phenotypes at six months after birth with cord blood methylation data from the SIBSIGHT cohort. Importantly, these MRSs remained significantly correlated with weight outcomes up to two years later.

Conclusions and future directions

In this study we identified genes whose methylation levels in cord blood and placenta are significantly associated with three different weight outcomes; conditional weight gain *z*-score, BMI, and weight-for-length. Notably, we identified three genes whose

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

methylation in cord blood is predictive of all three weight outcomes. Two of these genes, *PLIN4* and *UBE2F*, have been associated with weight in prior studies. Also notably, and somewhat in contrast, only one outcome (CWG) was associated with gene methylation in the placenta. This can be explained by a higher number of cell types in the placenta tissue, making it more difficult to identify specific methylation patterns across a large number of methylation profiles. Alternatively, methylation states in the placenta might only be associated with CWG as birthweight is considered in the calculation of this outcome. It is possible that the conditions in the placenta might be more likely to influence birth weight than postnatal growth.[82,83] One limitation of this study is the small sample size (48) compared to traditional Epigenome-Wide Association Studies. In order to increase the power of our analysis, CpG sites were grouped by gene to reduce the dimensionality of the data, with the drawback that this allows us to capture only large-scale associations (i.e. over the whole gene and not individual CpGs). To confirm our findings, our analysis should be replicated using a larger sample. Additionally, there are other variables that can influence methylation levels of the placenta (e.g. maternal diet)[84,85] that were not collected as part of our study but could be informative to include in future studies.

We used the PROGRESS/ELEMENT DNA Methylation Study Dataset to test our selected genes and MRSs in an independent cohort. However, while the SIBSIGHT cohort is largely white and non-Hispanic/Latino,[32] the PROGRESS cohort is composed of individuals located in the Latin American city of Mexico City, Mexico.[54] It has been shown that in adults there is population-to-population variation in DNA methylation related to several diseases and phenotypes (e.g. cancer and diabetes)[86] and that individuals who have similar demographics, life style, etc. have more similar methylation patterns.[87] Interestingly, we found evidence of between-populations differences in the association between weight outcomes and methylation patterns emerging as early as six months after birth. We found that the strongest "gene signals" from SIBSIGHT (*PLIN4, UBE2F,* and *PPP1R16B*) could also be detected in several of the regressions run on PROGRESS data. However, our MRSs were not predictive of weight outcomes in the PROGRESS cohort. We hypothesize that the underlying genetic, demographic, etc. differences between the two populations could be the reason why results from SIBSIGHT could not be more consistently validated in PROGRESS. This is notable because differences between the two cohorts were expected, however such a distinct contrast at such an early age was not. This suggests that external factors influencing the patterning of CpG methylation in early life should be carefully studied in order to determine factors potentially affecting future weight outcomes (e.g. maternal pre-pregnancy BMI[24] or environmental exposures).[54] It will be beneficial to identify if there are shared patterns because these could be used to generate a MRS that could be used universally to identify the children most at risk for developing obesity and therefore benefit the most from targeted obesity prevention programs.

In this study we characterized methylation patterns within the gene body and not within the promoter regions.[88] The relationship between gene body methylation and gene expression has been shown to be U-shaped in some studies, with both high and low expression corresponding to high levels of methylation,[89] but in other studies methylation and transcription have been found to be positively correlated.[20] Additional studies are needed to fully investigate the expression levels of the gene bodies in both placenta and cord blood. Such studies could validate our findings and provide a better understanding of the mechanisms eventually affecting weight outcomes. To our knowledge, there are no gene body methylation studies investigating the large-effect obesity genes, e.g., *LEPTIN* and *FTO*, in infants. Notably, methylation of these two genes was not found to be associated with weight outcomes in our study.

In a prior study by our group,[14] we found that there may be different genetic components influencing infant weight gain vs. adult weight gain. Regulatory mechanisms, including methylation patterns, could therefore differ between adults and infants as well. This represents an interesting direction for future research; overall, methylation levels decrease throughout childhood and adolescence[90] and it would be of great interest to investigate how the signatures we found here would persist as an individual ages.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S2040174424000060.

Acknowledgments. We thank Bonnie Higgins and Kate Anthony for their assistance in sample processing and sample preparation, Dr Ana Kenney and Dr Debmalya Nandy for their statistical help, and the Penn State College of Medicine Genome Sciences Facility for generation of the methylation array data. We are also grateful for and thank the INSIGHT/SIBSIGHT families for their participation and the study nurses for their assistance in sample and data collection. We thank Dr Allan Just, who coordinated the sample collection and participated in the study design of the PROGRESS cohort. He performed the randomization scheme for the sample analysis to reduce batch effects, he performed QA/QC, cleaned the data, organized its deposition into dbGAP and worked on all the IRB proposals related to the study.

Author contribution. Delphine Lariviere and Sarah J. C. Craig have contributed equally.

Funding statement. This study was supported by grants R01DK099354 and R01DK88244 from the National Institute of Diabetes and Digestive and Kidney Disease. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Funding was also provided by Penn State Institute for Computational and Data Sciences, Penn State Eberly College of Sciences, and the Huck Institutes of Life Sciences at Penn State.

PROGRESS dataset: This work is supported by grants from the National Institute of Environmental Health Sciences (R01ES021357S1, R01ES021357, R01ES020268, R01ES013744). We acknowledge Instituto Nacional de Salud Pública, Instituto Nacional de Perinatologia, and the American British Cowdray Hospital in Mexico City for their assistance with sample collection and use of their facilities. The PROGRESS/ELEMENT Mexico City cohort Principal Investigator is Robert Wright, Icahn School of Medicine at Mount Sinai.

Competing interests. None.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national guidelines on human experimentation from the U.S. Department of Health and Human Services, Office of Human Research Protections and with the Helsinki Declaration of 1975, as revised in 2008, and has been approved by the Penn State Institutional Review Board.

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. *Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018*, NCHS Data Brief, no. 360, 2020. National Center for Health Statistics, Hyattsville, MD.
2. Fryar CD, Carroll MD, Afful J, Division of Health and Nutrition Examination Surveys. *Prevalence of Overweight, Obesity, and Severe Obesity Among Children and Adolescents Aged 2-19 Years: United States, 1963-1965 Through 2017-2018*. 2021. https://www.cdc.gov/nchs/data/hestat/obesity-child-17-18/overweight-obesity-child-H.pdf.

3. Sahoo K, Sahoo B, Choudhury AK, Sofi NY, Kumar R, Bhadoria AS. Childhood obesity: causes and consequences. *J Family Med Prim Care*. 2015; 4(2), 187–192.
4. Martínez JA, Milagro FI, Claycombe KJ, Schalinske KL. Epigenetics in adipose tissue, obesity, weight loss, and diabetes. *Adv Nutr*. 2014; 5(1), 71–81.
5. Marques-Rocha JL, Samblas M, Milagro FI, Bressan J, Martínez JA, Marti A. Noncoding RNAs, cytokines, and inflammation-related diseases. *FASEB J*. 2015; 29(9), 3595–3611.
6. Paul IM, Williams JS, Anzman-Frasca S, *et al.* The intervention nurses start infants growing on healthy trajectories (INSIGHT) study. *BMC Pediatr*. 2014; 14(1), 184.
7. Savage JS, Birch LL, Marini M, Anzman-Frasca S, Paul IM. Effect of the INSIGHT responsive parenting intervention on rapid infant weight gain and overweight status at age 1 Year. *JAMA Pediatr*. 2016; 170(8), 742. DOI: 10.1001/jamapediatrics.2016.0445.
8. Paul IM, Savage JS, Anzman-Frasca S, *et al.* Effect of a responsive parenting educational intervention on childhood weight outcomes at 3 Years of age: the INSIGHT randomized clinical trial. *JAMA*. 2018; 320(5), 461–468.
9. Chesi A, Grant SFA. The genetics of pediatric obesity. *Trends Endocrinol Metab*. 2015; 26(12), 711–721.
10. Bouchard C. Genetics of obesity: what we have learned over decades of research. *Obesity*. 2021; 29(5), 802–820.
11. Yengo L, Sidorenko J, Kemper KE, the GIANT Consortium, *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of european ancestry. *Hum Mol Genet*. 2018; 27(20), 3641–3649. DOI: 10.1093/hmg/ddy271.
12. Lewis CM, Vassos E. Polygenic risk scores: from research tools to clinical instruments. *Genome Med*. 2020; 12(1), 44.
13. Khera AV, Chaffin M, Wade KH, *et al.* Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell*. 2019; 177(3), 587–596.e9.
14. Craig SJC, Kenney AM, Lin J, *et al.* Constructing a polygenic risk score for childhood obesity using functional data analysis. *Econometr Stat*. 2021; 25, 66–86. DOI: 10.1016/j.ecosta.2021.10.014.
15. Skinner MK, Manikkam M, Guerrero-Bosagna C. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Metab*. 2010; 21(4), 214–222.
16. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. *Lancet*. 2018; 392(10149), 777–786.
17. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012; 13(7), 484–492.
18. Yoder JA, Walsh CP, Bestor TH. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet*. 1997; 13(8), 335–340.
19. Collings CK, Anderson JN. Links between DNA methylation and nucleosome occupancy in the human genome. *Epigenetics Chromatin*. 2017; 10(1), 18.
20. Huang X, Zhang X, Zong L, *et al.* Gene body methylation safeguards ribosomal DNA transcription by preventing PHF6-mediated enrichment of repressive histone mark H4K20me3. *J Biol Chem*. 2021; 297(4), 101195.
21. Thompson M, Hill BL, Rakocz N, *et al.* Methylation risk scores are associated with a collection of phenotypes within electronic health record systems. *BMJ*. 2022; 7(1), 50. DOI: 10.1101/2022.02.07.22270047.
22. Hüls A, Czamara D. Methodological challenges in constructing DNA methylation risk scores. *Ciba F Symp*. 2020; 15, 1–11.
23. Kresovich JK, Zheng Y, Cardenas A, *et al.* Cord blood DNA methylation and adiposity measures in early and mid-childhood. *Clin. Epigenetics*. 2017; 9(1), 86.
24. Si J, Meir AY, Hong X, *et al.* Maternal pre-pregnancy BMI, offspring epigenome-wide DNA methylation, and childhood obesity: findings from the Boston birth cohort. *BMC Med*. 2023; 21(1), 317.
25. Montrose L, Padmanabhan V, Goodrich JM, *et al.* Maternal levels of endocrine disrupting chemicals in the first trimester of pregnancy are associated with infant cord blood DNA methylation. *Ciba F Symp*. 2018; 13(3), 301–309.
26. Abraham E, Rousseaux S, Agier L, *et al.* Pregnancy exposure to atmospheric pollution and meteorological conditions and placental DNA methylation. *Environ Int*. 2018; 118, 334–347.
27. Hamilton OKL, Zhang Q, McRae AF, *et al.* An epigenetic score for BMI based on DNA methylation correlates with poor physical health and major disease in the lothian birth cohort. *Int J Obes*. 2019; 43(9), 1795–1802.
28. Reed ZE, Suderman MJ, Relton CL, Davis OSP, Hemani G. The association of DNA methylation with body mass index: distinguishing between predictors and biomarkers. *Clin. Epigenetics*. 2020; 12(1), 50.
29. Mendelson MM, Marioni RE, Joehanes R, *et al.* Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a mendelian randomization approach. *PLoS Med*. 2017; 14(1), e1002215.
30. Hohman EE, Savage JS, Marini ME, *et al.* Effect of the INSIGHT firstborn parenting intervention on secondborn sleep. *Pediatrics*. 2022; 150(1), e2021055244.
31. Hohman EE, Savage JS, Birch LL, Paul IM. The intervention nurses start infants growing on healthy trajectories (INSIGHT) responsive parenting intervention for firstborns affects dietary intake of secondborn infants. *J Nutr*. 2020; 150(8), 2139–2146.
32. Ruggiero CF, Hohman EE, Birch LL, Paul IM, Savage JS. The intervention nurses start infants growing on healthy trajectories (INSIGHT) responsive parenting intervention for firstborns impacts feeding of secondborns. *Am J Clin Nutr*. 2020; 111(1), 21–27.
33. Paul IM, Savage JS, Anzman-Frasca S, Marini ME, Mindell JA, Birch LL. INSIGHT responsive parenting intervention and infant sleep. *Pediatrics*. 2016; 138(1), e20160762. DOI: 10.1542/peds.2016-0762.
34. Paul IM, Hohman EE, Loken E, *et al.* Mother-infant room-sharing and sleep outcomes in the INSIGHT study. *Pediatrics*. 2017; 140(1), E145–E146. DOI: 10.1542/peds.2017-0122.
35. Anzman-Frasca S, Paul IM, Moding KJ, Savage JS, Hohman EE, Birch LL. Effects of the INSIGHT obesity preventive intervention on reported and observed infant temperament. *J Dev Behav Pediatr*. 2018; 39(9), 736–743. DOI: 10.1097/dbp.0000000000000597.
36. Craig SJC, Blankenberg D, Parodi ACL, *et al.* Child weight gain trajectories linked to oral microbiota composition. *Sci. Rep*. 2018; 8(1), 14030.
37. Nandy D, Craig SJC, Cai J, *et al.* Metabolomic profiling of stool of two-year old children from the INSIGHT study reveals links between butyrate and child weight outcomes. *Pediatr Obes*. 2022; 17(1), e12833.
38. Carney MC, Zhan X, Rangnekar A, *et al.* Associations between stool microtranscriptome, gut microbiota, and infant growth. *J Dev Orig Hlth Dis*. 2021; 12(6), 876–882. DOI: 10.1017/s2040174420001324.
39. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013; 14(10), R115.
40. Fitzgerald KN, Hodges R, Hanes D, *et al.* Potential reversal of epigenetic age using a diet and lifestyle intervention: a pilot randomized clinical trial. *Aging*. 2021; 13(7), 9419–9432.
41. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin. Epigenetics*. 2019; 11(1), 123.
42. Pidsley R, Zotenko E, Peters TJ, *et al.* Critical evaluation of the illumina methylationEPIC beadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016; 17(1), 208.
43. Schroeder DI, Blair JD, Lott P, *et al.* The human placenta methylome. *Proc Natl Acad Sci*. 2013; 110(15), 6037–6042.
44. Heiss JA, Téllez-Rojo MM, Estrada-Gutiérrez G, *et al.* Prenatal lead exposure and cord blood DNA methylation in PROGRESS: an epigenome-wide association study. *Environ Epigenet*. 2020; 6(1), dvaa014.
45. Smith CE, Ordovás JM. Update on perilipin polymorphisms and obesity. *Nutr Rev*. 2012; 70(10), 611–621.
46. Richardson K, Louie-Gao Q, Arnett DK, *et al.* The PLIN4 variant rs8887 modulates obesity related phenotypes in humans through creation of a novel miR-522 seed site. *PLoS One*. 2011; 6(4), e17944.
47. Benton MC, Johnstone A, Eccles D, *et al.* An analysis of DNA methylation in human adipose tissue reveals differential modification of obesity genes before and after gastric bypass and weight loss. *Genome Biol*. 2015; 16(1), 8.
48. Mansego ML, Garcia-Lacarte M, Milagro FI, Marti A, Martinez JA, GENOI members. DNA methylation of miRNA coding sequences putatively associated with childhood obesity. *Pediatr Obes*. 2017; 12(1), 19–27.
49. Heinzel K, Bleul CC. The Foxn1-dependent transcripts PCOLCE2 and mPPP1R16B are not required for normal thymopoiesis. *Eur J Immunol*. 2007; 37(9), 2562–2571.

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press

50. Boratkó A, Csortos C. TIMAP, the versatile protein phosphatase 1 regulator in endothelial cells. *IUBMB Life*. 2017; 69(12), 918–928. DOI: 10.1002/iub.1695.
51. Heroes E, Lesage B, Görnemann J, Beullens M, Van Meervelt L, Bollen M. The PP1 binding code: a molecular-lego strategy that governs specificity. *Febs J*. 2013; 280(2), 584–595. DOI: 10.1111/j.1742-4658.2012.08547.x.
52. Monda JK, Scott DC, Miller DJ, *et al.* Structural conservation of distinctive N-terminal acetylation-dependent interactions across a family of mammalian NEDD8 ligation enzymes. *Structure*. 2013; 21(1), 42–53.
53. Rajakumar K, Yan Q, Khalid AT, *et al.* Gene expression and cardiometabolic phenotypes of vitamin D-deficient overweight and obese black children. *Nutrients*. 2019; 11(9), 2016.
54. Sakamuri SSVP, Putcha UK, Veettil GN, Ayyalasomayajula V. Transcriptome profiling of visceral adipose tissue in a novel obese rat model, WNIN/Ob & its comparison with other animal models. *Indian J Med Res*. 2016; 144(3), 409–423.
55. Crawley SW, Weck ML, Grega-Larson NE, Shifrin DAJr, Tyska MJ. ANKS4B is essential for intermicrovillar adhesion complex formation. *Dev Cell*. 2016; 36(2), 190–200.
56. Low BSJ, Lim CS, Ding SSL, *et al.* Decreased GLUT2 and glucose uptake contribute to insulin secretion defects in MODY3/HNF1A hiPSC-derived mutant β cells. *Nat Commun*. 2021; 12(1), 3133.
57. Sato Y, Hatta M, Karim MF, *et al.* Anks4b, a novel target of HNF4α protein, interacts with GRP78 protein and regulates endoplasmic reticulum stress-induced apoptosis in pancreatic β-cells. *J Biol Chem*. 2012; 287(27), 23236–23245.
58. Liao X, Song L, Zhang L, *et al.* LAMP3 regulates hepatic lipid metabolism through activating PI3K/Akt pathway. *Mol Cell Endocrinol*. 2018; 470, 160–167.
59. Gunder LC, Harvey I, Redd JAR, *et al.* Obesity augments glucocorticoid-dependent muscle atrophy in male C57BL/6J mice. *Biomedicines*. 2020; 8(10), 420.
60. Peris-Moreno D, Taillandier D, Polge C. MuRF1/TRIM63, master regulator of muscle mass. *Int J Mol Sci*. 2020; 21(18), 6663.
61. Olaniru OE, Persaud SJ. Adhesion G-protein coupled receptors: implications for metabolic function. *Pharmacol Ther*. 2019; 198, 123–134.
62. Suchý Táš, Zieschang C, Popkova Y, *et al.* The repertoire of adhesion G protein-coupled receptors in adipocytes and their functional relevance. *Int J Obes*. 2020; 44(10), 2124–2136.
63. Yi X, Yang Y, Wu P, Xu X, Li W. Alternative splicing events during adipogenesis from hMSCs. *J Cell Physiol*. 2020; 235(1), 304–316.
64. Roumans NJT, Wang P, Vink RG, van Baak MA, Mariman ECM. Combined analysis of stress- and ECM-related genes in their effect on weight regain. *Obesity*. 2018; 26(3), 492–498.
65. Roumans NJ, Vink RG, Fazelzadeh P, van Baak MA, Mariman EC. A role for leukocyte integrins and extracellular matrix remodeling of adipose tissue in the risk of weight regain after weight loss. *Am J Clin Nutr*. 2017; 105(5), 1054–1062.
66. Smine S, Obry A, Kadri S, *et al.* Brain proteomic modifications associated to protective effect of grape extract in a murine model of obesity. *Biochim Biophys Acta Proteins Proteomics*. 2017; 1865(5), 578–588.
67. Tepper BJ, Koelliker Y, Zhao L, *et al.* Variation in the bitter-taste receptor gene TAS2R38, and adiposity in a genetically isolated population in southern Italy. *Obesity*. 2008; 16(10), 2289–2295.
68. Pilic L, Graham CA-M, Hares N, *et al.* Bitter taste sensitivity is determined by TAS2R38 haplotype, associated with saturated fat intake and is lower in overweight and obese compared to normal weight UK adults. *Curr Dev Nutr*. 2020; 4, 1271–1271.
69. Karmous I, Plesník Jří, Khan AS, *et al.* Orosensory detection of bitter in fat-taster healthy and obese participants: genetic polymorphism of CD36 and TAS2R38. *Clin Nutr*. 2018; 37(1), 313–320.
70. Feeney EL, O'Brien SA, Scannell AGM, Markey A, Gibney ER. Suprathreshold measures of taste perception in children - association with dietary quality and body weight. *Appetite*. 2017; 113, 116–123.
71. Sugrue LP, Desikan RS. What are polygenic scores and why are they important? *JAMA*. 2019; 321(18), 1820–1821.
72. Koukoura O, Sifakis S, Spandidos DA. DNA methylation in the human placenta and fetal growth (review). *Mol Med Rep*. 2012; 5(4), 883–889.
73. Tekola-Ayele F, Zeng X, Ouidir M, *et al.* DNA methylation loci in placenta associated with birthweight and expression of genes relevant for early development and adult diseases. *Clin Epigenetics*. 2020; 12(1), 78.
74. Daniels TE, Sadovnikoff AI, Ridout KK, Lesseur C, Marsit CJ, Tyrka AR. Associations of maternal diet and placenta leptin methylation. *Mol. Cell. Endocrinol*. 2020; 505, 110739.
75. Thakali KM, Zhong Y, Cleves M, Andres A, Shankar K. Associations between maternal body mass index and diet composition with placental DNA methylation at term. *Placenta*. 2020; 93, 74–82.
76. Kader F, Ghai M. DNA methylation-based variation between human populations. *Mol Genet Genomics*. 2017; 292(1), 5–35.
77. Giri AK, Bharadwaj S, Banerjee P, *et al.* DNA methylation profiling reveals the presence of population-specific signatures correlating with phenotypic characteristics. *Mol. Genet. Genomics*. 2017; 292(3), 655–662.
78. Solomon O, Huen K, Yousefi P, *et al.* Meta-analysis of epigenome-wide association studies in newborns and children show widespread sex differences in blood DNA methylation. *Mutat Res Rev Mut Res*. 2022; 789, 108415.
79. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. *Oncotarget*. 2012; 3(4), 462–474.
80. Robinson N, Brown H, Antoun E, Godfrey KM, Hanson MA, Lillycrop KA, Crozier SR, Murray R, Pearce MS, Relton CL, Albani V, McKay JA. Childhood DNA methylation as a marker of early life rapid weight gain and subsequent overweight. *Clin Epigenetics*. 2021; 13(1), 8.
81. Clinical Growth Charts, 2023. CDC. https://www.cdc.gov/growthcharts/clinical_charts.htm (2022).
82. Term Infant Growth Tools, 2023. American Academy of Pediatrics (AAP). https://www.aap.org/en/patient-care/newborn-and-infant-nutrition/newborn-and-infant-nutrition-assessment-tools/term-infant-growth-tools/.
83. Cole TJ. Conditional reference charts to assess weight gain in british infants. *Arch Dis Child*. 1995; 73(1), 8–16.
84. Griffiths LJ, Smeeth L, Hawkins SS, Cole TJ, Dezateux C. Effects of infant feeding practice on weight gain from birth to 3 years. *Arch Dis Child*. 2009; 94(8), 577–582.
85. Aryee MJ, Jaffe AE, Corrada-Bravo H, *et al.* Minfi: a flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. *Bioinformatics*. 2014; 30(10), 1363–1369.
86. Marabita F, Almgren M, Lindholm ME, *et al.* An evaluation of analysis pipelines for DNA methylation profiling using the illumina HumanMethylation450 beadChip platform. *Ciba F Symp*. 2013; 8, 333–346.
87. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc*. 1996; 58, 267–288.
88. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *J R Stat Soc Series B Stat Methodol*. 2011; 73, 273–282.
89. R Core Team. *R: A language and environment for statistical computing*, 2013. R Foundation for Statistical Computing, Vienna, Austria.
90. Burris HH, Baccarelli AA, Byun H-M, *et al.* Offspring DNA methylation of the aryl-hydrocarbon receptor repressor gene is associated with maternal BMI, gestational age, and birth weight. *Ciba F Symp*. 2015; 10(10), 913–921.

https://doi.org/10.1017/S2040174424000060 Published online by Cambridge University Press