

Supplemental Tables

Table S1. Overview of cohort specific DNAm pre-processing steps

Cohort	QC*	Pre-processing [†]	QC probes [‡]	Normalization	Extra batch correction	Probe types excl [§]	N autosomal probes
ABOS	GenomeStudio	Minfi	1,4	Enmix -RCP	Combat	1,2,3	Look-up N=6
TEENDIAB	GenomeStudio	Minfi	1,4	Enmix- RCP	Combat	1,2,3	Look-up N=6
NFBC1966	Minfi	Minfi	3	Functional Normalization	-	1	Look-up N=6
EDEN	GenomeStudio	Minfi	1	DASEN	-	1,2,4	439,306
ENVIRONAGE	MethylAid	Minfi	3	Functional Normalization	Combat		839,073
FinnGeDi	GenomeStudio	ChAMP	1,4	Enmix -RCP	Combat	1,2,3	724,671
Gen3G	Minfi	Minfi	2,4	Functional Normalization + Enmix -RCP	Combat	1,2,3	791,324
GUSTO	GenomeStudio	Minfi	1,4	Quantile Normalization	Combat	1,2,3	367,840
Healthy Start	GenomeStudio	Minfi	2,4	Quantile Normalization	Combat	-	484,261
PREDO	Minfi	Minfi	1	Functional Normalization	-	1,2,3	428,619
Generation R	GenomeStudio	Minfi	3	Quantile Normalization	-	1	458,563

*Method by which the quality of individual IDAT files was assed.

[†]Preprocessing software of IDAT files.

[‡]Thresholds used to filter out sub performing probes during preprocessing; 1 = detection P value limit >0.01, 2= detection P value limit >0.05, 3 = detection P value limit > E-16, 4= number of beads required > 2,

[§]Probes excluded from the dataset were: 1. Cross-hybridizing probes as assessed by Chen et al. Am J. Hum. Genet. 2012 (450K) or Pidsley et al. Gen. Biol. 2016 (850K) , 2. non-CpG probes , 3. SNP probes as determined by Chen or Pidsley et al. 4. SNP probes as determined by 1000-genomes Minor allele frequencies (MAF) >10%, 5. SNP probes as determined by 1000-genomes MAF>5%,

Table S2. Bayesian inflation and bias for the EWAS on maternal FG

			Pre-bacon correction†			Post-bacon correction‡		
Study	N	Median SE*	“classic” λ	λ	μ	“classic” λ	λ	μ
ENVIRONAGE	103	0.002	1.61	1.09	0.55	1.06	1.01	0.005
FinnGeDi- controls	236	0.002	1.02	1.00	0.009	1.01	1.00	0.000
FinnGeDi-GDM	266	0.001	1.03	1.00	0.06	1.01	1.00	0.001
Gen3G	451	0.001	1.02	0.91	-0.07	1.26	0.99	-0.012
GUSTO	264	0.002	1.01	0.99	-0.07	1.01	0.99	-0.001
Healthy Start	532	0.002	0.92	0.96	0.001	1.00	0.99	0.000
PREDO	552	0.004	1.12	1.03	0.13	1.04	1.00	0.001
Generation R§	1099	0.0007	0.98	0.98	0.02	1.02	1.00	0.000

*The median of all the standard errors of the regression coefficients for the rlm between DNAm and glucose genome-wide.

†Inflation/deflation (λ) and bias (μ) of the test statistics as determined with R package *bacon* (λ , μ) and genomic control (“classic” λ).

‡Inflation/deflation (λ) and bias (μ) of the test statistics genome-wide after adjustment with *bacon*, as determined with *bacon* (λ , μ) and genomic control (“classic” λ).

§Generation R results are for an EWAS on non-fasting maternal glucose measurements.

Table S3. Bayesian inflation and bias for the EWAS on maternal insulin

Study	N	Median SE*	Pre-bacon correction †			Post-bacon correction‡		
			"classic" λ	λ	μ	"classic" λ	λ	μ
Gen3G	438	0.0006	2.04	1.28	-0.37	1.15	1.01	-0.006
Healthy Start	523	0.0008	0.94	0.97	-0.02	1.01	1.00	0.000
Generation R [§]	1101	0.0004	1.00	0.99	0.03	1.03	1.00	0.000

*The median of all the standard errors of the regression coefficients for the rlm between DNAm and glucose genome-wide.

†Inflation/deflation (λ) and bias (μ) of the test statistics as determined with R package *bacon* (λ , μ) and genomic control ("classic" λ).

‡Inflation/deflation (λ) and bias (μ) of the test statistics genome-wide after adjustment with *bacon*, as determined with *bacon* (λ , μ) and genomic control ("classic" λ).

§Generation R results are for an EWAS on non-fasting measurements.

Table S4. Bayesian inflation and bias for the EWAS on maternal AUC of an OGTT

			Pre-bacon correction †			Post-bacon correction‡		
Study	N	Median SE*	“classic” λ	λ	μ	“classic” λ	λ	μ
EDEN	32	2.03×10^{-5}	1.08	0.59	0.014	3.07	0.92	0.001
FinnGeDi- controls	236	5.23×10^{-6}	0.99	0.98	0.134	1.00	1.00	0.002
FinnGeDi-GDM	266	3.71×10^{-6}	0.97	0.98	0.059	1.01	1.00	0.001
Gen3G	451	4.39×10^{-6}	0.7	0.89	0.051	0.98	0.99	-0.006
Healthy Start	86	1.62×10^{-5}	1.01	0.99	0.002	1.09	1.00	0.001
PREDO	552	1.36×10^{-5}	1.22	1.07	-0.009	1.06	1.00	-0.001
ENVIRONAGE	48	1.68×10^{-5}	0.91	0.89	0.076	1.13	0.99	0.001

*The median of the standard errors of the regression coefficients for the rlm between DNAm and glucose

†Inflation/deflation (λ) and bias (μ) of the test statistics as determined with R package *bacon* (λ , μ) and genomic control (“classic” λ).

‡Inflation/deflation (λ) and bias (μ) of the test statistics genome-wide after adjustment with *bacon*, as determined with *bacon* (λ , μ) and genomic control (“classic” λ).

§Generation R results are for an EWAS on non-fasting measurements.

Table S5. Cross-sectional associations of blood DNA methylation at cg26974062 and metabolic phenotypes in childhood and adulthood

	TEENDIAB participants* (Age 4-19y; German Europeans; 49.6% females)				NFBC1966 participants † (Age 46y; Finnish Europeans; 56% females)			
cg26974062	N	Beta	SE	P	N	Beta	SE	P
Fasting plasma glucose (mmol/l)	366	-0.3	0.3	0.41	680	-0.76	0.15	3.6 x10 ⁻⁷
Fasting plasma insulin (pmol/l)	369	-0.3	0.2	0.15	685	-0.04	0.01	3.2 x10 ⁻³
AUC OGTT	232	-8.52 x10 ⁻⁴	2.05 x10 ⁻³	0.68	589	-2.3 x10 ⁻³	6.0 x10 ⁻⁴	1.72 x10 ⁻⁴
BMI (kg/m ²)	383	-0.1	0.06	0.03	693	-0.06	0.021	5.0 x10 ⁻³
WHR	365	-0.6	2.	0.80	NA	NA	NA	NA
HOMA-IR	366	-0.09	0.3	0.74	NA	NA	NA	NA
HbA1c (mmol/mol)	361	4.27 x10 ⁻³	5.8 x10 ⁻³	0.94	693	-0.085	0.024	5.2 x10 ⁻⁴
Type 2 diabetes	NA	NA	NA	NA	507	-1.54	0.54	4.2 x10 ⁻³

*Outcome of analyses in the TEENDIAB cohort. Columns denote the results from a robust linear model with robust standard errors adjusting for sex, age at blood draw, batch, imputed cell proportions, maternal type 1 diabetes status and parental socio-economic status.

†Outcome of analyses in the NFBC1966 cohort. The results from a robust linear model with robust standard errors adjusting for sex, batch, imputed cell proportions, and socio-economic status in the NFBC-46y cohort.

NA means the variable in question was unavailable for assessment.

Table S6. Cross-sectional associations of adult blood DNA methylation at identified CpGs nominally associated in main analyses ($P < 1.0 \times 10^{-6}$; see Table 2; CpGs other than within *TXNIP*) with glycemic and metabolic traits in the NFBC1966 cohort

	cg26104143 <i>TMEM33</i>			cg09049566 <i>SHROOM1</i>			cg21686486 <i>CYBRD1</i>			cg21139325 <i>HLA-DQB2</i>		
	N	Beta [SE]	P	N	Beta [SE]	P	N	Beta [SE]	P	N	Beta [SE]	P
Fasting plasma glucose (mmol/L)	680	-0.024 [0.063]	0.71	680	0.14 [0.16]	0.38	680	0.059 [0.23]	0.8	680	-0.78 [0.68]	0.26
Fasting plasma insulin (pmol/L)	685	0.052 [0.050]	0.30	685	0.093 [0.09]	0.30	685	-0.04 [0.21]	0.85	685	-0.35 [0.40]	0.38
AUC _{gluc} (mmol*min/L)	589	-6.9x10 ⁻⁵ [2.x10 ⁻⁴]	0.78	589	4.0x10 ⁻⁴ [4.4x10 ⁻⁴]	0.36	589	-5.3x10 ⁻⁴ [9.6x10 ⁻⁴]	0.58	589	-1.1x10 ⁻³ [2.0x10 ⁻³]	0.57
BMI (kg/m ²)	693	-8.6x10 ⁻³ [7.9x10 ⁻³]	0.28	693	0.015 [0.015]	0.31	693	3.4x10 ⁻³ [0.039]	0.93	693	-0.16 [0.070]	0.024
Body fat (bio-impedance)	671	-4.2x10 ⁻³ [5.3x10 ⁻³]	0.42	671	0.013 [8.9x10 ⁻³]	0.14	671	7.5x10 ⁻³ [0.022]	0.73	671	-0.057 [0.040]	0.16
Waist-to-hip-ratio	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
HOMA-IR	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
HbA1c (mmol/L)	693	8.1x10 ⁻³ [8.6x10 ⁻³]	0.35	693	-9.2x10 ⁻³ [0.020]	0.65	693	1.4x10 ⁻³ [0.034]	0.97	693	0.016 [0.077]	0.84
Type 2 diabetes	507	0.20 [0.17]	0.24	507	0.60 [0.47]	0.20	507	0.97 [0.64]	0.13	507	0.67 [1.6]	0.67

Outcome of analyses in the NFBC1966 cohort. The results from a robust linear model with robust standard errors adjusting for sex, batch, imputed cell proportions, and socio-economic status in the NFBC-46y cohort.

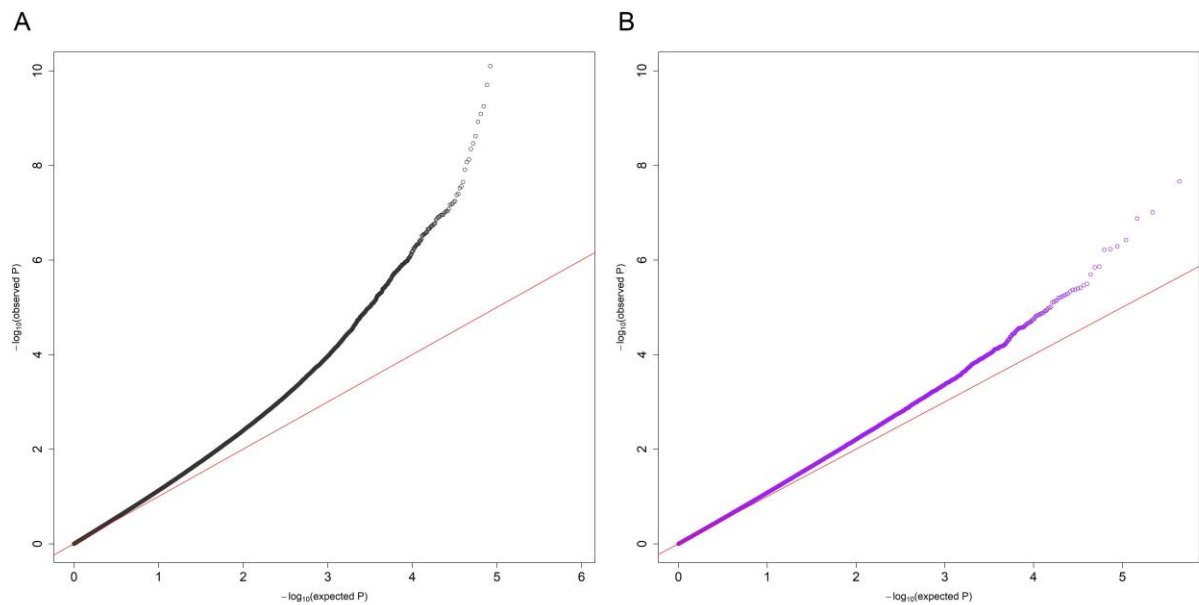
NA means the variable in question was unavailable for assessment.

Table S7. Associations between DNAm and RNA expression in muscle and liver tissues collected in ABOS for all CpG identified in our main analyses (see Table 2)

	Tissue	N	Beta	S.E.	P-value
CG21139325 <i>HLA-DQB2</i>	Muscle	71	1.67×10^{-2}	2.29×10^{-2}	0.47
CG26104143 <i>TMEM33</i>	Muscle	71	-1.03×10^{-2}	1.05×10^{-2}	0.33
CG26974062 <i>TXNIP</i>	Muscle	71	-1.75×10^{-3}	1.49×10^{-3}	0.24
CG09049566 <i>SHROOM1</i>	Muscle	71	8.39×10^{-4}	2.40×10^{-3}	0.73
CG21686486 <i>CYBRD1</i>	Muscle	71	-2.35×10^{-3}	2.16×10^{-3}	0.28
CG02988288 <i>TXNIP</i>	Muscle	71	-4.29×10^{-3}	3.01×10^{-3}	0.15
CG21139325 <i>HLA-DQB2</i>	Liver	319	1.11×10^{-2}	8.77×10^{-3}	0.21
CG26104143 <i>TMEM33</i>	Liver	319	-7.75×10^{-3}	5.32×10^{-3}	0.15
CG26974062 <i>TXNIP</i>	Liver	319	-1.11×10^{-2}	5.16×10^{-3}	0.031
CG09049566 <i>SHROOM1</i>	Liver	319	2.15×10^{-4}	9.84×10^{-4}	0.83
CG21686486 <i>CYBRD1</i>	Liver	319	-3.05×10^{-5}	1.22×10^{-3}	0.98
CG02988288 <i>TXNIP</i>	Liver	319	-4.45×10^{-2}	1.24×10^{-2}	3.2×10^{-4}

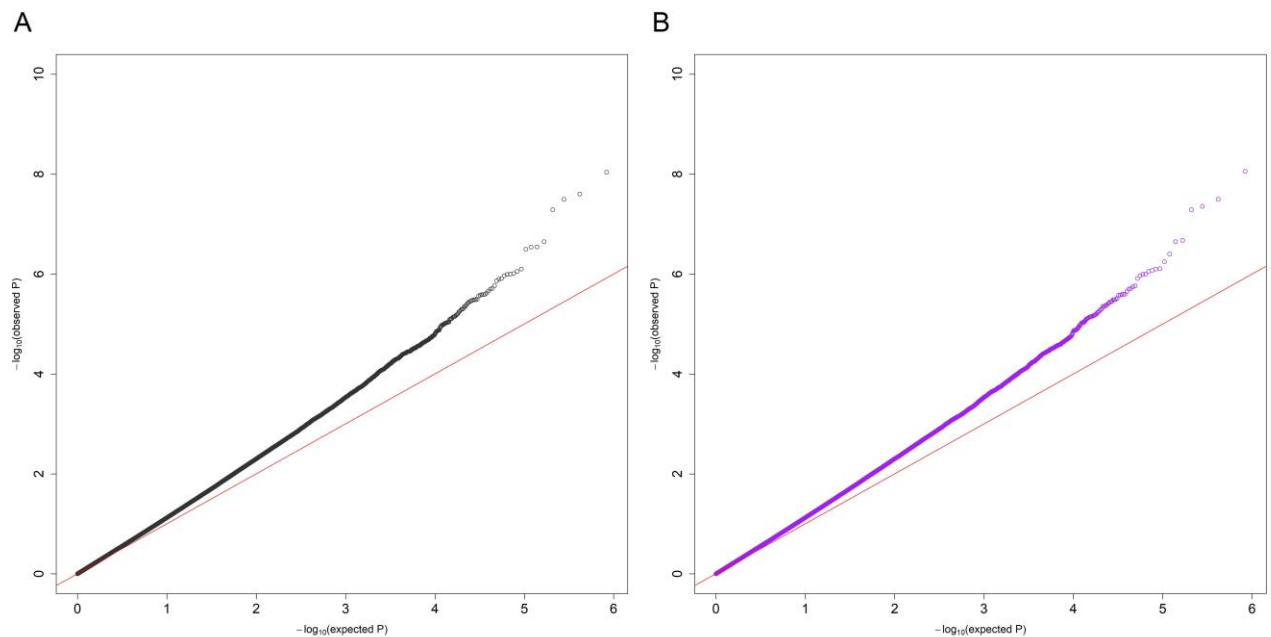
The results from a robust linear model with robust standard errors between DNAm and expression read counts (RPKM) measured in biopsies of women with obesity undergoing gastric bypass surgery; adjusting for sex, age at blood draw, BMI, batch, and log(HOMA2-IR) in the ABOS study.

Figure S1. QQ-plot of the Cochran's Q heterogeneity test for the glucose meta-analysis



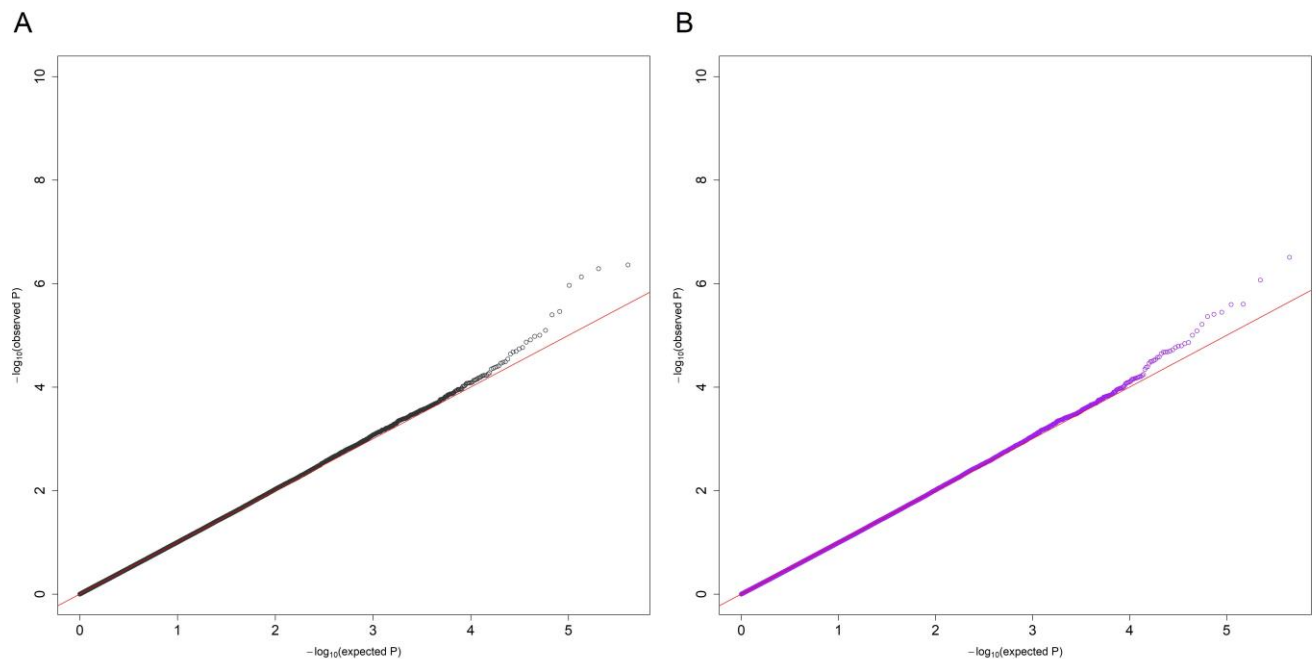
We plotted the P-value from a Cochran's Q test for heterogeneity (y-axis) against the p-value that can be expected by chance (x-axis). Each single dot is the test for heterogeneity for the association between maternal FG and cord blood DNAm at a single CpG dinucleotide. A. The FG meta-EWAS has many more CpGs showing heterogeneity than expected by chance ($\lambda=1.101$). B. Without the EDEN cohort, the genome-wide heterogeneity is reduced to a degree that for no single CpG the heterogeneity would be significant after FDR correction ($\lambda=1.031$).

Figure S2. QQ-plot of the Cochran's Q heterogeneity test for the glucose meta-analysis with or without Generation R



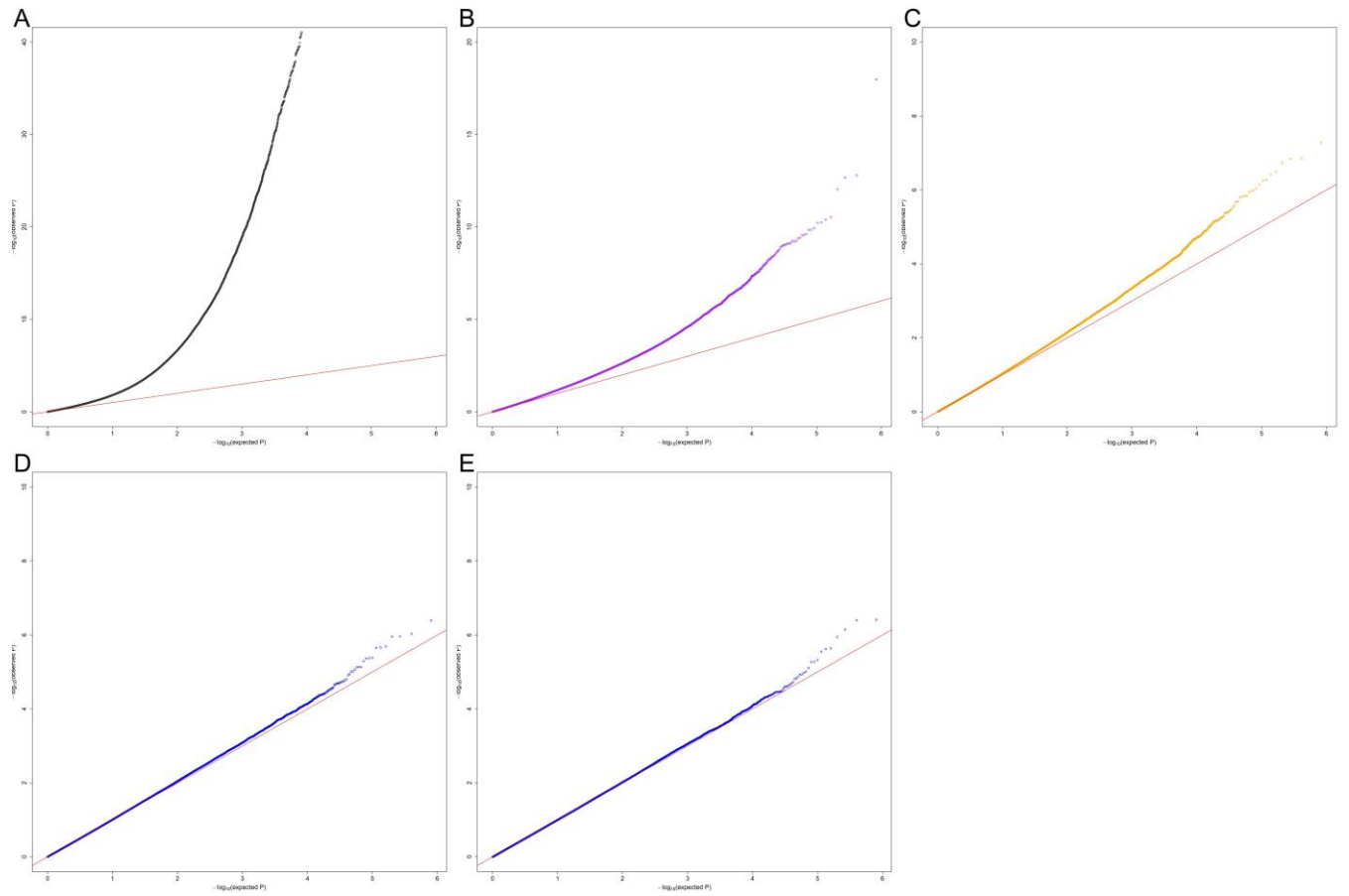
We plotted the P-value from a Cochran's Q test for heterogeneity (y-axis) against the p-value that can be expected by chance (x-axis). Each single dot is the test for heterogeneity for the association between maternal glucose and cord blood DNAm at a single CpG dinucleotide. A. The Cochran's heterogeneity test statistics distribution for the FG meta-analysis ($\lambda=1.11$). B. The meta-analysis of both FG and non-fasting glucose (thus including Generation R, $\lambda=1.11$). The addition of Generation R to the meta-analysis did not induce heterogeneity. For both meta-analyses, no single CpG shows statistical evidence for heterogeneity when both the inflation and number of tests is taken into account ($P_{\text{FDR}} > 0.05$).

Figure S3. QQ-plot of the Cochran's Q heterogeneity test for the insulin meta-analysis with or without Generation R



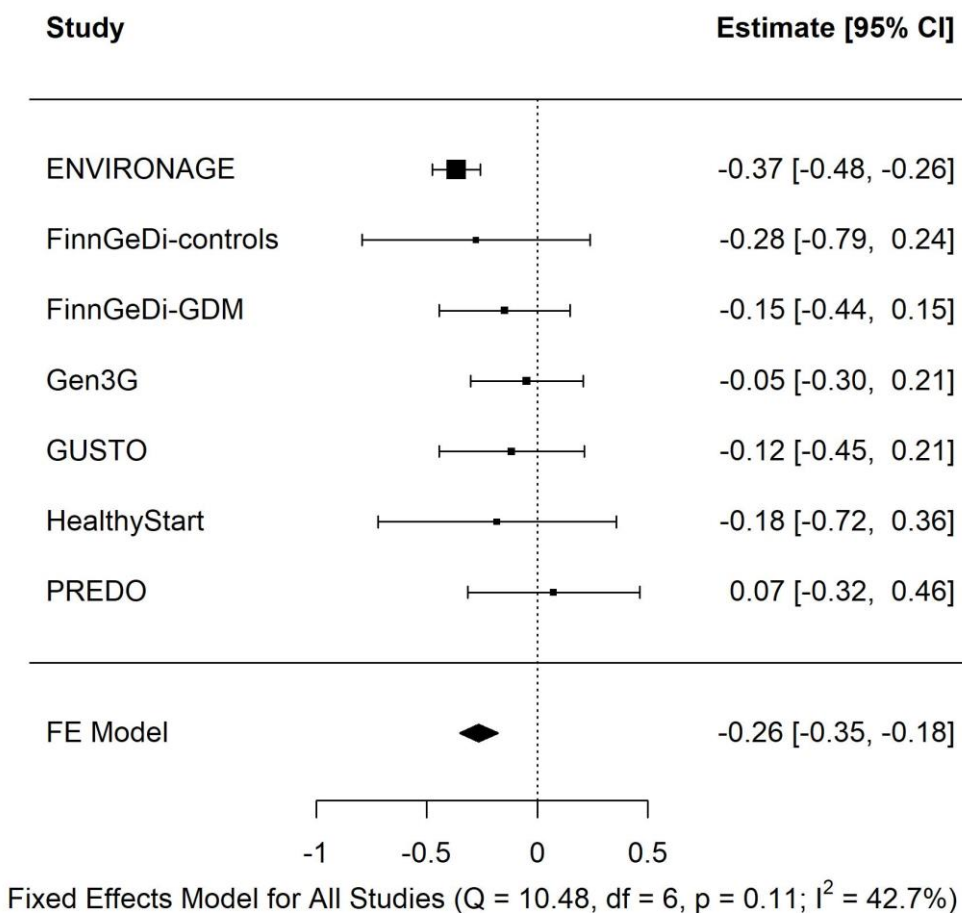
We plotted the P-value from a Cochran's Q test for heterogeneity (y-axis) against the p-value that can be expected by chance (x-axis). Each single dot is the test for heterogeneity for the association between maternal insulin and cord blood DNAm at a single CpG dinucleotide. A. The Cochran's heterogeneity test statistics distribution for the FI meta-analysis ($\lambda=0.999$). B. The meta-analysis of both FI and non-fasting insulin (thus including Generation R, $\lambda=0.988$). The addition of Generation R to the meta-analysis did not induce heterogeneity. For both meta-analyses, no single CpG shows statistical evidence for heterogeneity when both the inflation and number of tests is taken into account ($P_{\text{FDR}} > 0.05$).

Figure S4. Cochrane's heterogeneity tests for meta-analysis for the AUC_{gluc}



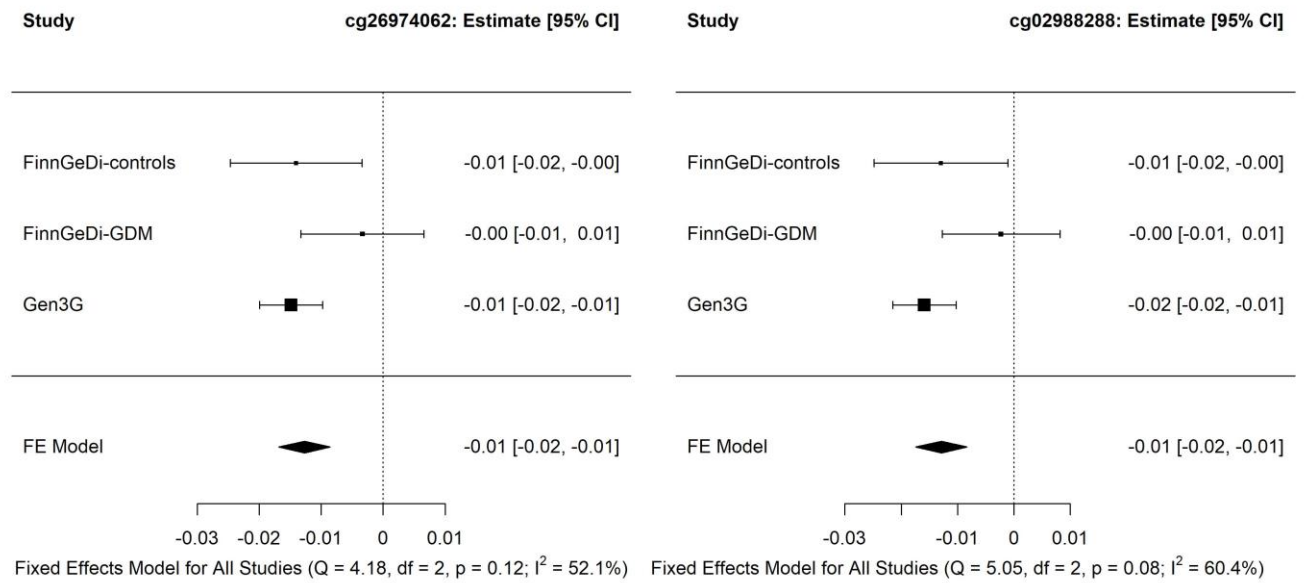
Cochrane's test P-value QQ plot for the meta-analysis with all cohorts with OGTT measurements ($\lambda=1.38$). **B.** Without EDEN ($\lambda=1.08$) **C.** without EDEN and ENVIRONAGE ($\lambda=1.01$). **D.** without EDEN, ENVIRONAGE, and Healthy Start ($\lambda=0.99$) **E.** without EDEN, ENVIRONAGE, Healthy Start, and FinnGeDi-GDM ($\lambda=0.99$).

Figure S5. Forest plot for cg26104143 and the fixed effects meta-analysis for maternal FG



The forest plot gives an overview of the effect estimates of the individual samples and their 95% C.I and the inverse variance fixed effect outcome (FE Model). Clear from the forest plot is the fact that ENVIRONAGE has a very small 95% C.I. for the estimate, giving it much weight in the inverse variance fixed effect outcome (relative weight is depicted by the thickness of the whisker box).

Figure S6. Forest plot for the AUC_{gluc} meta-analysis *TXNIP* CpGs



Additional cohort information

ABOS

The present analysis included 78 skeletal muscle samples and 330 liver samples from participants of the ABOS (Atlas Biologique de l'Obésité Sévère) study ([clinicaltrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01129297), [NCT01129297](https://clinicaltrials.gov/ct2/show/study/NCT01129297)). Tissue samples were collected during bariatric surgery (1). The institutional review board approved the ABOS study protocol, and patients gave written informed consent before enrollment.

EDEN

The EDEN (Etude des Déterminants pré et post natals du développement et de la santé de l'Enfant) study is a prospective Birth Cohort Study (<https://eden.vjf.inserm.fr/>), which has been described in detail elsewhere (2). Pregnant women seen for a prenatal visit at the departments of Obstetrics and Gynecology of the University Hospital of Nancy and Poitiers before their twenty-fourth week of amenorrhea were invited to participate. Enrolment started in February 2003 in Poitiers and September 2003 in Nancy; it lasted 27 months in each centre. Among eligible women, 55% (n=2002) accepted to participate. The study has been approved by the ethical committees Comité Consultatif pour la Protection des Personnes dans la Recherche Biomédicale, Le Kremlin-Bicêtre University hospital, and Commission Nationale de l'Informatique et des Libertés. Collection of blood samples: Immediately after delivery, **cord blood** samples were collected by research midwives from 1367 consenting cohort participants. To prevent any contamination with maternal blood, the cord was doubly clamped immediately after birth (vaginal delivery) or after extraction of the fetus through the uterine incision (elective cesarean section); repeatedly rinsed and venous cord blood serum was sampled between the 2 clamps. **Whole blood** samples were collected at the age of 5 years in 836 participants whose parents gave informed consent. Cord and whole blood samples were centrifuged within 24 hours of collection. The serum plasma was separated and samples were stored at -80°C . Fasted glucose levels and oral-glucose tolerance test: At 24–28 weeks (gestational age, mean \pm SD: 174 ± 19 days; range: 97–223 days), maternal plasma glucose concentrations were measured 1 h after a 50-g glucose challenge. In Nancy, fasted glucose levels were also measured before the 50-g glucose challenge. Women with 1-h glucose concentrations over 130 mg/dL in Nancy and 140 mg/dL in Poitiers were scheduled for a 3-h 100-g oral-glucose tolerance test.

ENVIRONAGE

ENVIRONAGE (ENVIRONmental influence ON early AGEing) (3) is an ongoing population-based prospective birth cohort study that aims at exploring the human ageing and its interaction with the environment. The cohort includes more than 1500 mother-infant pairs have been recruited at the delivery at the East-Limburg Hospital in Genk (Belgium) from February 2010 onward. Inclusion criteria were: delivery without planned caesarean section and ability to fill out a Dutch language questionnaire. Ethical approval was obtained by the Ethical Committee of Hasselt University and the East-Limburg Hospital and written informed consent was given by the parents. Data on lifestyle, diet, air pollution exposures, socio-economic status, phenotypic characteristics as well as clinical medical records and biological specimens have been collected at birth from mothers and children. For this study, DNA methylation was measured from cord blood, collected immediately after delivery. Maternal glucose concentration (mg/dL) was quantified using a Cobas 8000 C702 Analyzer from blood plasma of pregnant women undergoing a 50g oral glucose tolerance test (OGTT). Women were instructed not to eat or drink anything except water during the test and blood was drawn at time 0 and after 60, 120 and 180 minutes. The measure of glucose at time 0 was considered as

fasting glucose. All the measurements were converted to mmol/L for the analysis. When the glucose measurements were taken no women had started treatment for GDM.

FinnGeDi

This subcohort is a part of the Finnish Gestational Diabetes (FinnGeDi) Study including 299 women with gestational diabetes mellitus (GDM) and 238 controls aged 19-45 years (4). Women with GDM were recruited as they came to give birth, and the next consenting woman without GDM was recruited as a control from seven delivery hospitals in Finland (Oulu, Helsinki, Jyväskylä, Pori, Kajaani, Seinäjoki and Lappeenranta) between 2/2009 and 12/2012. Women with prepregnancy diabetes and multiple pregnancy were excluded. GDM was diagnosed by 2-hour 75 g oral glucose tolerance test (OGTT) (according to guidelines at 24-28 weeks of gestation; for high-risk women in addition at 12-16 weeks) where the cut-off concentrations for venous plasma glucose were ≥ 5.3 mmol/l at baseline (fasting glucose), ≥ 10.0 mmol/l at 1 h or ≥ 8.6 mmol/l at 2 h after glucose intake. GDM diagnosis was set if one or more glucose concentrations exceeded the cut-off levels. Controls were confirmed normoglycemic by OGTT after 24 weeks of gestation. Data include cord blood DNA sample; clinical data from hospital and maternal welfare records; register data from national registers; and self-reported lifestyle, medical and family history data from questionnaires. Cohort is located at the Finnish Institute for Health and Welfare (Oulu and Helsinki, Finland).

Gen3G

The Genetics of Glucose regulation in Gestation and Growth (Gen3G) (5) is a prospective pre-birth cohort of mother-child pairs in Sherbrooke, Canada, aimed at increasing our understanding of the determinants of glucose regulation in pregnancy and fetal growth, emphasizing on the interaction between genetic, environmental and lifestyle factors (Guillemette et al. 2016). Pregnant women recruited in the study were representative of women of reproductive age living in the Eastern Townships region in Quebec, Canada. Women recruited were in the first trimester of their pregnancy, and they received prenatal care in the CHUS or at any of the affiliated centers, and planned delivery at the CHUS (Hospital in the Eastern Townships region offering obstetric care). All women attending the CHUS were equally eligible to be included in the study. Exclusion criteria were the presence of pre-existing diabetes and overt hyperglycemia detected at first trimester either by medical or laboratory diagnosis. After overnight fasting, women completed a 75-g oral glucose tolerance testing (OGTT) for gestational diabetes mellitus (GDM) screening at the second visit (24-30 weeks of gestation) (6). Blood samples were collected at each clinical time-point of the OGTT (i.e., fasting, 1-h and 2-h post glucose load) for immediate measure of glucose levels, and to store plasma samples for future analyses. Glucose (in mmol/L) was assessed using the glucose hexokinase method (Roche Diagnostics, Indianapolis, US) at the CHUS biochemistry laboratory rapidly after collection. Insulin (in pg/mL) was measured from previously frozen plasma samples (-80°C) using the multiplexed particle-based flow cytometric assay (Human MILLIPLEX map kit, EMD Millipore, MA, US) at the CR-CHUS endocrinology laboratory (5,6).

Generation R Study

The Generation R Study is a prospective population-based cohort in Rotterdam, the Netherlands. A detailed description can be found elsewhere (7,8). All pregnant women residing in Rotterdam with a

delivery date between April 2002 and January 2006 were invited to participate. The Medical Ethical Committee of Erasmus MC, University Medical Center Rotterdam, approved the study and an informed consent was obtained for all participants. A total of 9,778 mothers were enrolled in the study and follow-up is ongoing. For this analysis, we included mother-child pairs with offspring cord blood DNA methylation data, from those we included only singletons live births, and one child was randomly selected for sibling pairs. Mother-child pairs were excluded if offspring was born before 36 weeks of gestation or if the mother was diagnosed with type 1 or type 2 diabetes, resulting in a total sample size of $n=1,101$. Non-fasting venous blood samples were collected in early pregnancy as previously described (9,10). Although samples were at least 30 minutes post-meal, we had no information on the exact time interval of the post-meal fasting duration and therefore consider all samples random. Glucose concentration (mmol/l) was measured with c702 module on the Cobas 8000 analyzer (Roche, Almere, the Netherlands). Insulin concentration (pmol/l) was measured with electrochemiluminescence immunoassay on the Cobas e411 analyzer (Roche).

GUSTO

The Growing Up in Singapore Towards healthy Outcomes (GUSTO) Study is a prospective birth cohort, recruiting pregnant women aged 18 years and above, at Singapore's two major public maternity hospitals (National University Hospital and KK Women's and Children's Hospital), between June 2009 and September 2010 (11). GUSTO participants were Singapore citizens or permanent residents with homogeneous parental ancestries of Chinese, Malay, or Indian origin. Women with significant health conditions such as those on chemotherapy or psychotropic drugs were excluded from the study. Written informed consent was obtained from the participants. The GUSTO study was approved by the Centralized Institutional Review Board of KKH and the Domain-Specific Review Board of NUH. In this meta-analysis, the study cohort was further restricted to live singleton full-term births with no congenital anomalies, with an Apgar score of nine or above and complete information of maternal OGTT results and infant umbilical cord blood DNA methylome ($n=264$). Antenatal (approximately 26 weeks gestation) maternal blood samples were collected from participants post overnight fasting (at least 8 hours), and again 2h post-75g OGTT (routine screening). Plasma samples were measured on a colorimeter [Advia 2400 Chemistry system (Siemen Medical Solutions Diagnostics) in NUH, and Beckman LX20 Pro analyser (Beckman Coulter) in KKH] within 4 hours of sampling.

Healthy Start

Healthy Start is an ongoing, longitudinal pre-birth cohort study that enrolled 1,410 pregnant women in Colorado, USA from outpatient obstetrics clinics in 2009-2015. There were 600 mother-infant pairs in the Healthy Start study with Illumina Infinium HumanMethylation450 array data in cord blood. Of these, 13 were missing data on gestational age at maternal glucose/insulin measurement during pregnancy, 22 were missing information on maternal GDM diagnosis, and 3 were diagnosed with GDM prior to maternal glucose/insulin measurement. No women with pre-existing diabetes were enrolled in the Healthy Start study. For mothers with two children in the study, one of the siblings was selected at random for exclusion ($n=7$). No multiple births were enrolled in the Healthy Start study. Mother-infant pairs were also excluded if mothers were non-fasting at the time of sample collection ($n=7$) or if the infant was born prior to 252 days of gestation ($n=7$), leading to a potentially eligible sample size of 536 mother infant-pairs. Of these, 532 had data on maternal glucose during pregnancy, 523 had data on maternal insulin during pregnancy, and 86 had data on

maternal 100g OGTT results during pregnancy. Only women who failed an initial glucose challenge test during pregnancy or had other GDM risk factors were subjected to the 100g OGTT. Pregnant women were instructed not to eat or drink anything but water for 8-10 hours prior to the study visit. Fasting venous blood samples were drawn and promptly separated via centrifugation. Blood plasma glucose concentration (mg/dL) was quantified using an Olympus AU400e Chemistry Analyzer, and measurements were converted to mmol/L for analysis. Blood plasma insulin concentration (μ U/mL) was quantified using a radioimmunoassay from Millipore Corporation, and measurements were converted to pmol/L and \log_2 transformed for analysis.

PREDO

The Prediction and Prevention of Preeclampsia and Intrauterine Growth Restriction (PREDO) is a prospective birth cohort study of Finnish women who were pregnant between 2005 and 2010 and their children (12). The PREDO study cohort was set up to identify novel risk factors and biomarkers in pregnant women associated with the development of preeclampsia and intrauterine growth restriction (IUGR), to (a) identify effective methods for prediction and prevention of preeclampsia in at-risk women, and (b) determine the association between exposure to preeclampsia, IUGR, or their risk factors and child developmental/health outcomes. Women with a singleton, intrauterine pregnancy who visited antenatal clinics at ten study hospitals in Finland for their first ultrasound screening at 12+0-13+6 weeks+days of gestation were recruited in the PREDO study. Two groups of pregnant women were enrolled: first, pregnant women with a known clinical risk factor status for preeclampsia and IUGR, and second, pregnant women who volunteered to participate regardless of their risk factor status for preeclampsia and IUGR. The sample with a known risk factor status comprises 1,079 pregnant women who gave live birth (969 of these women had at least one and 110 had none of the known risk factors for preeclampsia and IUGR). The community-based sample comprises 3,698 pregnant women who gave live birth. The sample with a known risk factor status visited antenatal clinics up to four times during pregnancy and both samples filled in bi-weekly self-reports. The post-delivery follow-up has taken place at approximately 2 weeks, 6 months, and 3.5 years after the delivery. The most recent follow-up started in 2016 and is ongoing. The study protocol was approved by the Ethics Committee of Obstetrics and Gynaecology, and Women, Children and Psychiatry of the Helsinki and Uusimaa Hospital District and by the participating hospitals. All participants provided written informed consent. Consent of participating children were provided by parent(s)/guardian(s).

TEENDIAB

The TEENDIAB study investigates the period of puberty and adolescence in the natural course of type 1 diabetes development (13). Between 2009 and 2015, a total of 610 children and adolescents aged 6-16 years with at least one first degree relative with type 1 diabetes and who were free of any diabetes-associated antibodies have been enrolled in this prospective observational cohort study. Follow-up visits took place every 6 months (on average) until the age of 18 years by 2016. The primary aims are to explore the roles of anthropometric, metabolic and genetic factors, as well as diet, physical activity, psychosocial factors and pubertal status in the development of islet autoimmunity and type 1 diabetes. At baseline visit, a physical examination was performed by trained staff to measure height, weight, waist and hip circumference. Additionally, an oral glucose tolerance test (OGTT) using 1.75 g of glucose per kg bodyweight, or 75 g at maximum, was carried out with blood samples taken at fasting, 30, 60, 90, and 120 min. Separate whole blood samples

were taken at fasting for Hemoglobin A1c (HbA1c) analysis and DNA isolation. Plasma glucose and HbA1c levels were determined by the hospital laboratories at the two study sites (Hannover and Munich, Germany). Plasma Insulin was determined with an automated immunoassay analyzer (AIA 360; Tosoh, San Francisco, CA) at the Institute of Diabetes Research or the Clinical Chemistry Laboratory, Klinikum rechts der Isar, Technische Universität München. Parental education was collected by standardized questionnaires and categorized into three groups (low, medium and high) using the highest education level of the parents. The study was conducted according to the Declaration of Helsinki and approved by the ethical committee of the Technical University Munich (No. 2149/08) and Medizinische Hochschule Hannover (No. 5644). Written informed consent was obtained from children's parents.

NFBC1966

The Northern Finland Birth Cohort 1966 (NFBC1966) is a prospective birth cohort study that recruited all pregnant women living in the Oulu and Lapland provinces of Finland with expected delivery in 1966. NFBC1966 study population comprised 12,231 children and their parents with follow-ups have conducted at 1, 14, 31 and 46 years of age (University of Oulu. <http://urn.fi/urn:nbn:fi:att:bc1e5408-980e-4a62-b899-43bec3755243>). In 2012 when cohort members were 45-46 years of age, all individuals with known address in Finland (N=10,321) were sent postal questionnaires and an invitation for clinical examination. Clinical examination data were collected for 5,861 participants. Serum glucose, insulin and HbA1c were measured and 75-g OGTT was performed at 46y follow-up after 12h fasting.

Procedures - Cohort specific adjustments to EWAS model

This section gives additional clarification on cohort specific variables added to the robust linear models of the EWAS and other details they wished to make explicit on their analysis. If batch effects were not addressed by adjustment but pre-analysis via Combat, a note is also made.

EDEN

All models were adjusted by the clinical centre (Nancy or Poitiers) where the subjects were recruited.

ENVIRONAGE

Methylation betas values were corrected for batch effect on sample plate using the Combat algorithm (25) and all the analyses were additionally adjusted for the row on the micro-array slide (as numeric variable from 1-6) to account for positional effect.

FinnGeDi

A variable for GDM treatment was included to adjust for GDM (categorical ordinal, 0= no GDM, 1= treated with diet or exercise, 2= treated with insulin).

Gen3G

Beta values were corrected for dye bias/probe type using the Regression on Correlated Probes (RCP) method (20). Latent variation in methylation due to batch effects, regarded here as the sample plate (i.e. 10 unique batches), was removed using ComBat (25). For ComBat, we protected the following covariates: glucose levels during the OGTT (0min-60min-120min) at second trimester visit, maternal age, gestational age at birth, parity, sex of the newborn, and smoking during pregnancy (yes/no/unknown). GDM was diagnosed in 39 women (8.6%) in the second visit (2nd trimester). These participants were included in the analysis. A variable for GDM treatment was included to adjust for GDM (categorical ordinal, 0= no GDM, 1= treated with diet or exercise, 2= treated with insulin).

Generation R Study

For these analyses, mothers with pre-existing diabetes were excluded from the Generation R Study. Additionally, there were 10 sibling-pairs in the dataset and since siblings in each pair had the same data availability, one of each was excluded randomly. Plate number was added as batch covariate. Child sex and gestational age at birth were collected from midwife and obstetric records and parity and maternal age were collected from questionnaires during pregnancy.

GUSTO

GUSTO is a tri-ethnic Asian cohort, so a variable coded as categorical representing genotyped ethnic groups was included. GDM treatment was added as categorical where 0 = non-GDM, 1 = treatment via diet, and 2 = treatment via insulin.

Healthy Start

In addition to the covariates denoted as required in the analysis plan GDM treatment was adjusted as a categorical variable in all models as described in the analysis plan: no GDM (0), GDM with

treatment via diet (coded=1), and GDM with medication treatment (coded=2). All models were additionally adjusted for self-reported race/ethnicity (4 categories: Non-Hispanic white, Hispanic, non-Hispanic African-American, all others).

PREDO

We used ComBat to check and adjust for the batch effects (slide and well). In the regression models, we adjusted for the following: Model A: gestational age and 3 MDS components derived from the genome-wide genotypes; Model B: Model A + sex of the child, and 7 cell type estimates; Model C: Model A + sex of the child, parity, and maternal age; Model D: Model B + parity and maternal age.

TEENDIAB

To remove unwanted variation/batch effects, Combat method was applied as implemented in the R package (25). All childhood/adolescence models were adjusted for sex of the child/adolescent, age at DNA sample, parental socio-economic status, imputed cell types (cd4t, cd8t, nk, bcell, mono, gran) and sentrix position.

Procedures – methylation arrays

ABOS

DNA was extracted from liver and muscle biopsies. Bisulfite conversion of genomic DNA was performed and bisulfite-converted DNA was subjected to a genome-wide DNA methylation analysis performed using the Illumina Infinium MethylationEPIC 850k Bead-Chip array (San Diego, CA). All samples, for liver and muscle separately, were randomized across the chips and analyzed on the same machine by the same technician to reduce batch effects. After single-base extension and staining, the BeadChips were imaged with the Illumina iScan. Raw fluorescence intensities of the scanned images were extracted with the GenomeStudio Methylation module (Illumina). The fluorescence intensity ratio was used to calculate the β -value. All samples had high bisulfite conversion efficiency (signal intensity >4,000) and were included for further analysis based on GenomeStudio quality control. Quality control was performed using R software (version 4.0.0). The DNA methylation IDAT files were imported using the R package *minfi* (14) for pre-processing and quality control. The following probes were excluded from further analysis: probes with a detection P value ≥ 0.01 for at least one sample, cross-hybridizing probes, probes with a bead count less than three in at least 5% of the samples, non-CpG probes, and probes that lie near single nucleotide polymorphisms. Probes on chromosomes X and Y were used for sex estimation and then excluded from downstream analyses. Samples with <99% probes with a detection P value <0.01 were excluded. Probe design biases and batch effects were normalized using R packages *ENmix* and *sva*, respectively. After quality control, cell composition was estimated for both liver and muscle using the R package *RefFreeCellMix*.

EDEN

DNA was extracted using the QIAamp blood kit (Qiagen or equivalent protocols), followed by precipitation-based concentration using GlycoBlue (Ambion). DNA concentration was determined by Nanodrop measurement and Picogreen quantification. 500 ng of DNA was bisulphite-converted using the EZ 96-DNA methylation kit (Zymo Research), following the manufacturer's standard protocol. After verification of the bisulphite conversion step using Sanger Sequencing, genome-wide DNA methylation was measured using the Illumina Infinium HumanMethylation450 BeadChip. After normalization of the concentration, the samples were randomized to avoid batch effects, and all paired samples were hybridized on the same chip. Standard male and female DNA samples were included in this step as control samples. DNA methylation data were pre-processed in R with the Bioconductor package *minfi* (14), using the original IDAT files extracted from the HiScanSQ scanner. Samples that did not provide significant methylation signals in more than 10% of probes (detection $P=0.01$) were excluded from further analysis. Samples were also excluded in cases of low staining efficiency, low single base extension efficiency, low stripping efficiency of DNA from probes after single base extension, poor hybridization performance, poor bisulphite conversion and high negative control probe staining. Further, we used the 65 SNP probes to check for concordances between paired DNA samples from the sample individual and assessed the methylation distribution of the X-chromosome to verify gender. Paired samples with Pearson correlation coefficients <0.9 were regarded as sample mix-ups and were excluded from the study. Probes on sex chromosomes, probes that mapped on multi-loci, the 65 random SNPs assay and probes that contained SNPs at the target CpG sites with a minor allele frequency >10% were excluded during probe filtering (15). The allele frequencies of a list of SNPs were obtained from 1000 Genomes, release 20110521 for the CEU population. Finally, to correct Type I and Type II bias, the "DASEN" method was implemented to

perform signal correction and normalization (16). After quality control, 439,306 autosomal probes remained in EDEN.

ENVIRONAGE

In a subset of 377 children from the ENVIRONAGE cohort cord blood DNA was extracted from buffy coats according to standard protocol. DNA concentration was measured using the Quant-IT assay from Thermo Fisher. Gel-electrophoresis was performed to assess the integrity of DNA samples. Methylation data was produced at GenomeScan in Netherland. Bisulphite conversion using 100-500 ng genomic DNA inputs was performed using the EZ DNA Methylation Gold kit (Zymo Research, Irvine, CA, USA). The converted samples (4µl) were amplified and hybridized on the Illumina HumanMethylation 850K BeadChip arrays and scanned using an Illumina iScan. Data quality was assessed using the R script MethyAid (17) using analysis default settings and 1 sample was removed because of low quality (sample call rate <99%). DNA methylation data were preprocessed using the minfi package in R (14). Briefly, 4 samples were removed because of wrongly predicted sex using shinyMethyl (14), the data were normalized using functional normalization, measurements with detection p-values >10e-16 (18) were set to missing leading to removal of 7961 probes with a call rate <95%. No additional sample had call rate <98%. For each CpG site methylation levels were expressed as beta values calculated as the ratios of intensities arising from methylated probes over those arising from the sum of methylated and unmethylated probes. Batch effect was corrected using the Combat algorithm using the *sva* package. After filtering sex chromosome probes, leaving 839,073 CpGs available measured for 372 samples for the present analysis. No additional filtering was applied for the present analysis.

FinnGeDi

IDAT files were imported using functions in the ChAMP package (19) for R software. Samples with a call rate below 95% were excluded and probes with a call rate below 100%. Call rates were computed using a 1% threshold on the detection p-values. Probes which align to multiple locations, are non-CpG, with bead counts less than 3 were excluded. Type I and type II probes were normalised using the Regression on Correlated Probes (RCP) method (20) from the ENmix package (21). Data were normalised for possible batch effect bias using the Combat function from the *sva* package, with batch/slide as covariate.

Gen3G

We extracted DNA from cord blood buffy coats and performed bisulfite conversion of genomic DNA using the EZ-Methylation Kit (Zymo Research, CA, USA), followed by DNAm quantification utilizing the Infinium MethylationEPIC BeadChip (Illumina, CA, USA). To minimize batch effects, we randomly allocated samples to different plates and chips. We applied quality control (QC) of DNAm data at the sample and probe levels. At the sample level, we removed samples with low quality (n = 3), technical replicates (n = 10) and samples with genotype mismatch (n = 6 samples; based on SNPs on EPIC array from paired cord blood and placenta samples). At the probe level, we excluded probes with detection p > 0.05 in at least 5% of the samples (n = 1,754), probes with null variance (n = 125), probes in sex chromosomes (n = 19,128 sites), non-CpG probes (n = 2,836 sites), SNP-associated probes at the single base extension (n = 5,547 sites) or at the target CpG site with MAF ≥ 5% (n = 5668 sites), and cross-reactive probes (n = 40,454 sites) (22). In total, we retained for analyses

791,324 CpG sites in 451 cord blood samples. We corrected for nonspecific background signals using the *minfi* R package and controlled for additional unwanted technical variation using functional normalization with two principal components derived from control probes. To correct for probe-type bias, we implemented the regression on correlated probes method, which uses the correlation between consecutive probes to adjust the distribution of type II probes. Finally, we used ComBat in the *sva* R package to adjust for sample plate. We estimated the proportion of 7 cord blood cells (CD4+T, CD8+T, monocytes, granulocytes, natural killer cells, B cells & nucleated red blood cells) using the Bakulski algorithm (23).

Generation R Study

DNA extracted (using the salting-out method) from blood samples taken at birth (cord blood) was used for this analysis. 500 ng DNA per sample underwent bisulfite conversion using the EZ-96 DNA Methylation kit (Shallow) (Zymo Research Corporation, Irvine, USA). Samples were plated onto 96-well plates in no specific order. Samples were processed with the Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, USA). Preparation and normalization of the HumanMethylation450 BeadChip array data was performed according to the CPACOR workflow (18) using the software package R. In detail, the idat files were read using the *minfi* package (14). Probes that had a detection p-value above background (based on sum of methylated and unmethylated intensity values) $\geq 1E-16$ were set to missing per array. Next, the intensity values were stratified by autosomal and non-autosomal probes and quantile normalized for each of the six probe type categories separately: type II red/green, type I methylated red/green and type I unmethylated red/green. Beta values were calculated as proportion of methylated intensity value on the sum of methylated+unmethylated+100 intensities. Arrays with observed technical problems such as failed bisulfite conversion, hybridization or extension, as well as arrays with a mismatch between sex of the proband and sex determined by the chr X and Y probe intensities were removed from subsequent analyses. Additionally, only arrays with a call rate > 95% per sample were processed further. Probes on the X and Y chromosomes were excluded from the dataset. The final dataset contained information on 458,563 CpGs.

GUSTO

Data was preprocessed and normalized as described (24). Briefly, probes with less than three beads for either methylation or unmethylated channel or detection p-value > 0.01 were set to missing. Color adjustment and normalization of Type 1 and 2 probes was performed. ComBat was used to remove chip effects (most significant technical variable with respect to raw DNA methylation values in GUSTO). Probes were then filtered. Briefly, allosomic probes, cross-hybridizing probes, and probes with multi-model distributions were removed. This gave a total of 367,840 CpGs available for subsequent analysis.

Healthy Start

DNA methylation in 600 cord blood samples was assessed using the Illumina Infinium HumanMethylation450 array. We removed 587 probes with high detection P value (> 0.05). We removed 664 probes with a beadcount < 3 in at least 5% of samples. We compared the predicted sex to the reported sex and removed samples with inconsistent sex (n=5). The *preprocessQuantile* function in *Minfi* was used to normalize. *ComBat* was used for batch adjustment.

PREDO

Methylation analyses were performed at the Max Planck Institute of Psychiatry in Munich, Germany. DNA was bisulphite-converted using the EZ-96 DNA Methylation kit (Zymo Research, Irvine, CA). Samples were ran on Illumina 450K Methylation arrays and the arrays were scanned using the iScan System (Illumina Inc., San Diego, CA). The quality control pipeline pipeline was set up using the R-package minfi. 3 IDs were excluded as they were outliers in the median intensities. Furthermore, 20 IDs showed discordance between phenotypic sex and estimated sex and were excluded.

Methylation beta-values were normalized using the funnorm function. We excluded any probes on chromosome X or Y, probes containing SNPs and cross-hybridizing probes. Furthermore, any Cgs with a detection p-value > 0.01 in at least 50% of the samples were excluded. The final dataset contains 428,619 CpGs and 834 IDs.

TEENDIAB

DNA was extracted from whole blood. Bisulfite conversion of genomic DNA was performed and bisulfite-converted DNA was subjected to a genome-wide DNA methylation analysis performed using the Illumina Infinium MethylationEPIC 850k Bead-Chip array (San Diego, CA). All samples were randomized across the chips and analyzed on the same machine by the same technician to reduce batch effects. After single-base extension and staining, the BeadChips were imaged with the Illumina iScan. Raw fluorescence intensities of the scanned images were extracted with the GenomeStudio Methylation module (Illumina). The fluorescence intensity ratio was used to calculate the β -value. All samples had high bisulfite conversion efficiency (signal intensity >4,000) and were included for further analysis based on GenomeStudio quality control. Quality control was performed using R software (version 4.0.3). The DNA methylation IDAT files were imported using the R package *minfi* for pre-processing and quality control. The following probes were excluded from further analysis: probes with a detection P value ≥ 0.01 for at least one sample, cross-hybridizing probes, probes with a bead count less than three in at least 5% of the samples, non-CpG probes, and probes that lie near single nucleotide polymorphisms. Probes on chromosomes X and Y were used for sex estimation and then excluded from downstream analyses. Samples with <99% probes with a detection P value <0.01 were excluded. Probe design biases and batch effects were normalized using R packages ENmix and sva, respectively. After quality control, cell composition was estimated from a reference panel using the R package FlowSorted.Blood.EPIC.

NFBC1966 46y

Genomic DNA was extracted from the blood samples by QIAasympyony DSP DNA Midi Kit (Qiagen, Hilden, Germany) at 46y follow-up. DNA methylation was assayed using Infinium MethylationEPIC BeadChip (Illumina, Inc., San Diego, CA, USA) for 766 randomly selected cohort members. Bisulphite conversion of genomic DNA was performed using the EZ DNA methylation kit according to manufacturer's instructions (Zymo Research, Orange, CA). From these, 8 technical replicates, 40 samples that did not reach a call rate of >95% applying a detection P-value filter of 10-16, 1 outlier and 1 sample with gender inconsistency were excluded. Final sample size for DNA methylation data was 716. Associations between DNA methylation beta values of cg26974062 and cg02988288 and metabolic traits in adulthood (serum glucose, insulin, OGTT AUC, HbA1c and BMI) were analyzed.

Procedures - RNA sequencing in ABOS cohort

ABOS

For liver samples, total RNA-seq was performed using the KAPA RNA HyperPrep kit with RiboErase (HMR; Roche Sequencing), in combination with Illumina sequencing on the NovaSeq6000 system, using a paired-end 2×75 bp protocol. On average, 130M reads per sample were generated and 96% were accurately mapped to the human genome (hg38). For muscle samples, mRNA-seq was performed using KAPA mRNA HyperPrep kit (Roche Sequencing), in combination with Illumina sequencing on the NovaSeq6000 system, using a paired-end 2×75 bp protocol. On average, 72M reads per sample were generated and 95% were accurately mapped to the human genome (hg38). For the liver and muscle samples, the demultiplexing of sequence data was performed using bcl2fastq Conversion Software (Illumina; version v2.20.0.422). Quality control checks were performed via FastQC (version v0.11.9). Removing adaptor sequences and low-quality bases was performed with Trimmomatic (version v0.39). Sequence reads were then mapped to the human genome (hg38) using Star aligner (version v2.7.3a). Gene and isoform abundances were quantified via RSEM (version v1.3.0). RSEM counts were normalized using the method vst from the R package DESeq2.

Additional cohort specific acknowledgements

ABOS

The ABOS study acknowledges all the participants.

TEENDIAB

We thank the TEENDIAB team for taking care of the families, for performing visits and challenge tests, and for entering data in the database. We want to thank all the families participating in the TEENDIAB study and also all the pediatricians, diabetologists, and family doctors in Germany for recruitment and continuous support.

PREDO

The PREDO study would not have been possible without the dedicated contribution of the PREDO study group members: E Hamäläinen, E Kajantie, H Laivuori, PM Villa, A-K Pesonen, A Aitokallio-Tallberg, A-M Henry, VK Hiilesmaa, T Karipohja, R Meri, S Sainio, T Saisto, S Suomalainen-Konig, V-M Ulander, T Vaitilo (Department of Obstetrics and Gynaecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland), L Keski-Nisula, Maija-Riitta Orden (Kuopio University Hospital, Kuopio Finland), E Koistinen, T Walle, R Solja (Northern Karelia Central Hospital, Joensuu, Finland), M Kurkinen (Päijät-Häme Central Hospital, Lahti, Finland), P.Taipale. P Staven (Iisalmi Hospital, Iisalmi, Finland), J Uotila (Tampere University Hospital, Tampere, Finland). We thank all the PREDO children and their parents for their enthusiastic participation. We also thank all the research nurses, research assistants, and laboratory personnel involved in the Predo study.

NFBC 1966 46y

We thank all cohort members and researchers who participated in the 46 yrs study. We also wish acknowledge the work of the NFBC project center.

GUSTO

The GUSTO study group includes Allan Sheppard, Amutha Chinnadurai, Anne Eng Neo Goh, Anne Rifkin-Graboi, Anqi Qiu, Arijit Biswas, Bee Wah Lee, Birit Froukje Philipp Broekman, Boon Long Quah, Chai Kiat Chng, Cheryl Shufen Ngo, Choon Looi Bong, Christiani Jeyakumar Henry, Daniel Yam Thiam Goh, Doris Ngiuk Lan Loh, Fabian Kok Peng Yap, George Seow Heong Yeo, Helen Yu Chen, Hugo P S van Bever, Iliana Magiati, Inez Bik Yun Wong, Ivy Yee-Man Lau, Jeevesh Kapur, Jenny L. Richmond, Jerry Kok Yen Chan, Joanna Dawn Holbrook, Joshua J. Gooley, Keith M. Godfrey, Kenneth Yung Chiang Kwek, Kok Hian Tan, Krishnamoorthy Naiduvaje, Leher Singh, Lin Lin Su, Lourdes Mary Daniel, Lynette Pei-Chi Shek, Marielle V. Fortier, Mark Hanson, Mary Foong-Fong Chong, Mary Rauff, Mei Chien Chua, Michael J. Meaney, Mya Thway Tint, Neerja Karnani, Ngee Lek, Oon Hoe Teoh, P. C. Wong, Peter David Gluckman, Pratibha Keshav Agarwal, Rob Martinus van Dam, Salome A. Rebello, Seang Mei Saw, Shang Chee Chong, Shirong Cai, Shu-E Soh, Sok Bee Lim, Stephen Chin-Ying Hsu, Victor Samuel Rajadurai, Walter Stunkel, Wee Meng Han, Wei Wei Pang, Yap Seng Chong, Yin Bun Cheung, Yiong Huak Chan and Yung Seng Lee.

Generation R Study

The Generation R Study is conducted by Erasmus MC, University Medical Center Rotterdam, in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The study protocol was approved by the Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam. Written informed consent was obtained for all participants. The generation and management of the Illumina 450K methylation array data (EWAS data) for the Generation R Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. We thank Mr. Michael Verbiest, Ms. Mila Jhamai, Ms. Sarah Higgins, Mr. Marijn Verkerk and Dr. Lisette Stolk for their help in creating the EWAS database. We thank Dr. A. Teumer for his work on the quality control and normalization scripts.

Gen3G

We thank participants of the Gen3G cohort who contributed to this study, as well as clinical research nurses and research assistants for recruiting women and obtaining their informed consent. We also thank the CHUS biomedical laboratory for performing some of the assays used in this study.

FinnGeDi

We thank the FinnGeDi Study Board Johan G. Eriksson, Mika Gissler, Risto Kaaja, Eero Kajantie, Hannele Laivuori, Anneli Pouta, Marja Vääräsmäki, and the study hospital physicians Tiina-Liisa

Erkinheimo, Ritva Keravuo, Merja Huttunen, Miira M Klemetti, Jenni Metsälä, Beata Stach-Lempinen, and Minna Tikkanen.

ENVIRONAGE

The authors are extremely grateful to the participating women and neonates, as well as the staff of the maternity ward, midwives and the staff of the clinical laboratory of East-Limburg Hospital in Genk.

EDEN

We are extremely grateful to all the families who took part in this study, the midwives and psychologists for recruiting and following them, and the whole EDEN team, including research scientists, engineers, technicians and managers.

Additional information on cohort funding

ABOS

This work was supported by the French government, managed by the National Research Agency (ANR) under the Future Investment Program (PIA) with the references ANR-16-RHUS-0006 PreciNASH, ANR-10-LABX-0046 (European Genomic Institute for Diabetes [E.G.I.D]), ANR-10-EQPX-07-01 (LIGAN-PM). This work was also supported by the National Center for Precision Diabetic Medicine – PreciDIAB, which is jointly supported by the French National Agency for Research (ANR-18-IBHU-0001), by the European Union (FEDER), by the Hauts-de-France Regional Council and by the European Metropolis of Lille (MEL).

TEENDIAB

The TEENDIAB study was supported by Kompetenznetz Diabetes mellitus (Competence Network for Diabetes mellitus) funded by the Federal Ministry of Education and Research (FKZ 01GI0805), and by the 'Diabetes Hilfs- und Forschungsfond Deutschland (DHFD)'. This study was supported by grants from the European Union's Horizon 2020 Research and Innovation programme (no. 633595) for the DynaHEALTH action, by the H2020 Joint Programming Initiative a Healthy Diet for a Healthy Life (JPI HDHL) under proposal number 655 (PRECisE Project) and from the German Federal Ministry of Education and Research (FKZ 01EA1905).

PREDO

The PREDO Study has been funded by the Academy of Finland, EraNet Neuron, EVO (a special state subsidy for health science research), University of Helsinki Research Funds, the Signe and Ane Gyllenberg foundation, the Emil Aaltonen Foundation, the Finnish Medical Foundation, the Jane and Aatos Erkko Foundation, Juho Vainio foundation, Yrjö Jahnsson foundation, The Finnish Society of Sciences and Letters, Jalmari and Rauha Ahokas foundation, the Novo Nordisk Foundation, the Päivikki and Sakari Sohlberg Foundation, the Sigrid Juselius Foundation granted to members of the Predo study board. Methylation assays were funded by the Academy of Finland.

NFBC 1966 46y

NFBC1966 received financial support from University of Oulu Grant no. 24000692, Oulu University Hospital Grant no. 24301140, ERDF European Regional Development Fund Grant no. 539/2010 A31592.

Healthy Start

The Healthy Start study was supported by the following grants from the National Institutes of Health: National Institute of Diabetes and Digestive and Kidney Diseases R01DK076648 (PI Dabelea), Office of the Director UH3OD023248 (PI Dabelea), National Institute of Environmental Health Sciences R01ES022934 (MPI Hamman, Adgate, Dabelea). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

GUSTO

This work was supported by the Translational Clinical Research (TCR) Flagship Program on Developmental Pathways to Metabolic Disease funded by the National Research Foundation (NRF) and administered by the National Medical Research Council (NMRC), Singapore - NMRC/TCR/004-NUS/2008; NMRC/TCR/012-NUHS/2014. Additional funding was provided by the Industry Alignment Fund – Pre-positioning Programme (IAF-PP H17/01/a0/005, available to YSC, PDG and NK) from Agency for Science Technology and Research (A*STAR), Singapore, and NMRC Open Fund Large Collaborative grant (OFLCG19May-003, available to YSC and NK).

Generation R Study

The general design of the Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development and the Ministry of Health, Welfare and Sport. The EWAS data was funded by a grant to VWJ from the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA; project nr. 050-060-810), by funds from the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by a grant from the National Institute of Child and Human Development (R01HD068437). VWJ received an additional Consolidator Grant from the European Research Council (ERC-2014-CoG-648916). This project received funding from the European Union's Horizon 2020 research and innovation programme (733206, LifeCycle; 874739, LongITools; 824989, EUCAN-Connect) and from the European Joint Programming Initiative "A Healthy Diet for a Healthy Life" (JPI HDHL, NutriPROGRAM project, ZonMw the Netherlands no.529051022 and JPI HDHL, PREcisE project, ZonMw, the Netherlands no.529051023). RG received grants from the Dutch Heart Foundation (2017T013); the Dutch Diabetes Foundation (2017.81.002); and the Netherlands Organisation for Health Research and Development (NWO, ZonMw; 543003109).

Gen3G

Work presented in this study was supported by an American Diabetes Association Pathways Award #1-15-ACE-26 (MFH); Gen3G has also been supported by Fonds de recherche du Québec en santé #20697 (MFH); Canadian Institute of Health Research #MOP 115071 (MFH), and Diabète Québec grants (PP and LB). LB is a research scholar from the Fonds de la recherche du Québec en santé (FRQS) and a member of the CR-CHUS, a FRQS-funded Research Center.

FinnGeDi

FinnGeDi has been supported by Academy of Finland, Novo Nordisk Foundation, Juho Vainio Foundation, Yrjö Jahnsson Foundation, Northern Ostrobothnia Hospital District Special State Research Subsidiary (vtr), University of Oulu Graduate School, Finnish Diabetes Research Foundation, Finnish Foundation for Pediatric Research, Signe and Ane Gyllenberg Foundation, Sigrid Juselius Foundation, Diabetes UK. This work was supported by the French government, managed by the National Research Agency (ANR) under the Future Investment Program (PIA) with the references: ANR-10-LABX-0046 (European Genomic Institute for Diabetes [E.G.I.D]), ANR-10-EQPX-07-01 (LIGAN-PM) and ANR-16-CE17-0017-01. This work was also supported by the Fondation Francophone pour la Recherche sur le Diabète (FFRD) that is sponsored by the Fédération Française des Diabétiques (FFD), Abbott, AstraZeneca, Eli Lilly, Merck Sharp & Dohme (MSD) and Novo Nordisk, and by the National Center for Precision Diabetic Medicine – PreciDIAB, which is jointly supported by the French National Agency for Research (ANR-18-IBHU-0001), by the European Union (FEDER), by the Hauts-de-France Regional Council and by the European Metropolis of Lille (MEL).

ENVIRONAGE

The ENVIRONAGE birth cohort is supported by the European Research Council [ERC-2012-StG.310898], and by funds of the Flemish Scientific Research council [FWO, G.0.733.15.N].

EDEN

The EDEN cohort has been funded by Foundation for Medical Research (FRM), National Agency for Research (ANR), National Institute for Research in Public Health (IRESP: TGIR cohorte santé 2008 program), French Ministry of Health (DGS), French Ministry of Research, Inserm Bone and Joint Diseases National Research (PRO-A) and Human Nutrition National Research Programs, Paris–Sud University, Nestlé, French National Institute for Population Health Surveillance (InVS), French National Institute for Health Education (INPES), the European Union FP7 programmes (FP7/2007-2013, HELIX, ESCAPE, ENRIECO projects, FP7-ENV HEALS), Diabetes National Research Program (through a collaboration with the French Association of Diabetic Patients (AFD)), French Agency for Environmental Health Safety (now ANSES), Mutuelle Générale de l'Éducation Nationale (MGEN), French National Agency for Food Security, and the French-speaking association for the study of diabetes and metabolism (ALFEDIAM). The cord and whole blood methylome has been assessed in the frame of the MeDALL project, which was supported by the European Union under the Health Cooperation Work Programme of the 7th Framework programme (grant no. 261357).

References to the additional cohort information

1. Caiazzo R, Lassailly G, Leteurtre E, Baud G, Verkindt H, Raverdy V, et al. Roux-en-Y gastric bypass versus adjustable gastric banding to reduce nonalcoholic fatty liver disease: A 5-Year Controlled Longitudinal Study. In: *Annals of Surgery*. Lippincott Williams and Wilkins; 2014. p. 893–9.
2. Heude B, Forhan A, Slama R, Douhaud L, Bedel S, Saurel-Cubizolles MJ, et al. Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int J Epidemiol*. 2016 Apr 11;45(2):353–63.
3. Janssen BG, Madhloum N, Gyselaers W, Bijmens E, Clemente DB, Cox B, et al. Cohort Profile: The ENVIRonmental influence ON early AGEing (ENVIRONAGE): a birth cohort study. *Int J Epidemiol*. 2017 Oct 1;46(5).
4. Keikkala E, Mustaniemi S, Koivunen S, Kinnunen J, Viljakainen M, Männistö T, et al. Cohort Profile: The Finnish Gestational Diabetes (FinnGeDi) Study. *Int J Epidemiol*. 2020 Jun 1;49(3):762–763g.
5. Guillemette L, Allard C, Lacroix M, Patenaude J, Battista M-CC, Doyon M, et al. Genetics of Glucose regulation in Gestation and Growth (Gen3G): a prospective prebirth cohort of mother-child pairs in Sherbrooke, Canada. *BMJ Open*. 2016 Feb 1;6(2):e010031. Available from: <http://dx.doi.org/10.1136/bmjopen-2015-010031>
6. Hivert MF, Cardenas A, Allard C, Doyon M, Powe CE, Catalano PM, et al. Interplay of placental DNA methylation and maternal insulin sensitivity in pregnancy. *Diabetes*. 2020 Mar 1;69(3):484–92.
7. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IJendoorn MH, et al. The Generation R Study: design and cohort update 2017. *Eur J Epidemiol*. 2016 Dec 1;31(12):1243–64.
8. Kruithof CJ, Kooijman MN, van Duijn CM, Franco OH, de Jongste JC, Klaver CCW, et al. The Generation R Study: Biobank update 2015. *Eur J Epidemiol*. 2014 Dec 28;29(12):911–27.
9. Wahab RJ, Jaddoe VWV, Roest AAW, Toemen L, Gaillard R. Associations of maternal glycemia in the first half of pregnancy with alterations in cardiac structure and function in childhood. *Diabetes Care*. 2020 Sep 1;43(9):2272–80.
10. Geurtsen ML, van Soest EEL, Voerman E, Steegers EAP, Jaddoe VWV, Gaillard R. High maternal early-pregnancy blood glucose levels are associated with altered fetal growth and increased risk of adverse birth outcomes. *Diabetologia*. 2019 Oct 1;62(10):1880–90.
11. Soh S-E, Tint MT, Gluckman PD, Godfrey KM, Rifkin-Graboi A, Chan YH, et al. Cohort profile: Growing Up in Singapore Towards healthy Outcomes (GUSTO) birth cohort study. *Int J Epidemiol*. 2014 Oct 20;43(5):1401–9.
12. Girchenko P, Lahti M, Tuovinen S, Savolainen K, Lahti J, Binder EB, et al. Cohort Profile: Prediction and prevention of preeclampsia and intrauterine growth restriction (PREDO) study. *Int J Epidemiol*. 2017 Oct 1;46(5).
13. Ziegler AG, Meier-Stiegen F, Winkler C, Bonifacio E. Prospective evaluation of risk factors for the development of islet autoimmunity and type 1 diabetes during puberty - TEENDIAB: Study design. *Pediatr Diabetes*. 2012;13(5):419–24.
14. Fortin J-P, Triche T, Hansen K. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array. *Bioinformatics*. 2016;1–3.

15. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013 Feb;8(1559-2308):203–9.
16. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*. 2013 May 1 14(1).
17. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics*. 2014 Dec 1;30(23):3435–7.
18. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol*. 2015 Feb 15;16(1).
19. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: Updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017 Dec 15;33(24):3982–4.
20. Niu L, Xu Z, Taylor JA. RCP: A novel probe design bias correction method for Illumina Methylation BeadChip. *Bioinformatics*. 2016. p. 2659–63.
21. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucleic Acids Res*. 2015 Sep 17;gkv907. Available from: <http://nar.oxfordjournals.org/content/early/2015/09/17/nar.gkv907.long>
22. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016;17(1):208.
23. Bakulski KM, Feinberg JL, Andrews S V., Yang J, Brown S, L. McKenney S, et al. DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics*. 2016 May 3;11(5):354–62.
24. Wu Y, Lin X, Lim IY, Chen L, Teh AL, MacIsaac JL, et al. Analysis of two birth tissues provides new insights into the epigenetic landscape of neonates born preterm. *Clin Epigenetics*. 2019 Feb 11;11(1).
25. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007 Jan;8(1):118–27.