Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight

Leanne K. Kupers, University of Bristol
Claire Monnereau, University Medical Center Rotterdam
Gemma C. Sharp, University of Bristol
Paul Yousefi, University of Bristol
Lucas A. Salas, Dartmouth College
Akram Ghantous, International Agency Research on Cancer
Christian M. Page, Norwegian Institute of Public Health
Sarah E. Reese, National Institutes of Health
Allen J. Wilcox, National Institutes of Health
Darina Czamara, Max-Planck-Institute of Psychiatry

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

Journal Title: Nature Communications
Volume: Volume 10, Number 1
Publisher: Nature Research (part of Springer Nature): Fully open access journals | 2019-04-23, Pages 1893-1893
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1038/s41467-019-09671-3
Permanent URL: https://pid.emory.edu/ark:/25593/tqfxb

Final published version: http://dx.doi.org/10.1038/s41467-019-09671-3

Copyright information:

© 2019, The Author(s).

This is an Open Access work distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

Accessed December 12, 2023 12:35 AM EST
Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight

Leanne K. Küpers et al.#

Birthweight is associated with health outcomes across the life course, DNA methylation may be an underlying mechanism. In this meta-analysis of epigenome-wide association studies of 8,825 neonates from 24 birth cohorts in the Pregnancy And Childhood Epigenetics Consortium, we find that DNA methylation in neonatal blood is associated with birthweight at 914 sites, with a difference in birthweight ranging from −183 to 178 grams per 10% increase in methylation \( (P_{\text{Bonferroni}} < 1.06 \times 10^{-7}) \). In additional analyses in 7,278 participants, <1.3% of birthweight-associated differential methylation is also observed in childhood and adolescence, but not adulthood. Birthweight-related CpGs overlap with some Bonferroni-significant CpGs that were previously reported to be related to maternal smoking (55/914, \( p = 6.12 \times 10^{-74} \)) and BMI in pregnancy (3/914, \( p = 1.13 \times 10^{-3} \)), but not with those related to folate levels in pregnancy. Whether the associations that we observe are causal or explained by confounding or fetal growth influencing DNA methylation (i.e. reverse causality) requires further research.
nte. Exposures, such as maternal smoking, pre-pregnancy body mass index (BMI), hyperglycaemia, hypertension, folate and famine are associated with fetal growth and hence birthweight\(^1\)–\(^6\). Observational studies show that birthweight is also associated with later-life health outcomes, including cardio-metabolic and mental health, some cancers and mortality\(^7\)–\(^11\). In these long-term associations, birthweight may act as a proxy for potential effects of intrauterine exposures\(^12\)–\(^13\). Several mechanisms may explain the associations of intrauterine exposures with birthweight and later-life health as we illustrate in Fig. 1. Our overall conceptual framework in this study was that the intrauterine environment induces epigenetic alterations, which influence fetal growth and hence correlate with birthweight. This is partly supported by previous large-scale epigenome-wide association studies (EWAS) that have reported associations of relevant maternal pregnancy exposures, including smoking, air pollution and BMI, with DNA methylation in offspring neonatal blood\(^14\)–\(^16\). However, whilst four previous EWAS have observed associations of DNA methylation with birthweight\(^17\)–\(^20\), the evidence to date has been limited in scale and power with sample sizes ranging from approximately 200 to 1000.

In this study, we hypothesised that there are associations between DNA methylation and birthweight. We further aimed to explore if these epigenetic alterations are associated with later disease outcomes (Fig. 1). If birthweight is a proxy for a range of adverse prenatal exposures, we might expect neonatal blood DNA methylation to be associated with birthweight. However, we acknowledge that any associations of DNA methylation with birthweight may be explained by confounding\(^21\) or reflect fetal growth influencing DNA methylation.

Here we present a large meta-analysis of multiple EWAS to explore associations between neonatal blood DNA methylation and birthweight. In further analyses, we explore whether any birthweight-associated differential methylation persists at older ages. To aid functional interpretation, we (i) explore the overlap of identified cytosine-phosphate-guanine sites (CpGs) that are differentially methylated in relation to birthweight with those known to be associated with intrauterine exposure to smoking, famine and different levels of BMI and folate; (ii) associate DNA methylation at identified CpGs with gene expression and (iii) explore potential causal links with birthweight and later-life health using Mendelian randomization (MR)\(^22\). We show that DNA methylation in neonatal blood is associated with birthweight and some of the differential methylation is also observed in childhood and adolescence, but not in adulthood.

Also, we show overlap between birthweight-related CpGs and CpGs related to intrauterine exposures. Potential causality of the associations needs to be studied further.

**Results**

**Participants.** We used data from 8825 neonates from 24 studies in the Pregnancy And Childhood Epigenetics (PACE) Consortium, representing mainly European, but also African and Hispanic ethnicities with similar proportions of males and females. Details of participants used in all analyses are presented in Table 1, Supplementary Data 1 and study-specific Supplementary Methods.

**Meta-analysis.** Primary, secondary and follow-up analyses are outlined in the study design in Fig. 2. Methylation at 8170 CpGs, measured in neonatal blood using the Illumina HumanMethylation450 BeadChip assay and adjusted for cell-type heterogeneity\(^23\)–\(^25\), was associated with birthweight (false discovery rate (FDR) <0.05), of which 1029 located in or near 807 genes survived the more stringent Bonferroni correction (\(p < 1.06 \times 10^{-7}\), Supplementary Data 2). We observed both positive (45%) and negative (55%) directions of associations between methylation levels of these 1029 CpGs and birthweight (Fig. 3) and these CpGs were spread throughout the genome (orange track (1) in Fig. 4 and Supplementary Fig. 1). We found evidence of between-study heterogeneity (\(I^2 > 50\%\)) for 115 of the 1029 sites (Supplementary Data 2), thus we prioritised 914 CpGs, located in or near 729 genes, based on \(p < 1.06 \times 10^{-7}\) and \(I^2 \leq 50\%\) for further analyses (Fig. 3 and orange track (1) in Fig. 4). The CpG with the largest positive association was cg06378491 (in the gene body of MAP4K2). For each 10% increase in methylation at this site, birthweight was 178 g higher (95% confidence interval (CI): 138, 218 g). The CpG with the smallest negative association was cg10073091 (in the gene body of DHCR24), which showed a 183 g decrease in birthweight per 10% increase in methylation (95% CI: −225, −142 g). The CpG with the smallest P-value and \(I^2 \leq 50\%\) was cg17714703 (in the gene body of UHRF1), which showed a 130 g increase in birthweight for 10% increase in methylation (95% CI: 109, 151 g).

Findings were consistent with results from our main analyses when restricted to participants of European ethnicity, with a Pearson correlation coefficient for effect estimates of 0.99 for the 914 birthweight-associated CpGs (Supplementary Fig. 2, blue track (2) in Fig. 4 and Supplementary Data 3) and 0.90 for all 450k CpGs. Comparing the main meta-analysis to the four Hispanic cohorts and

![Fig. 1](https://example.com/fig1.png)
the two African cohorts revealed that 94.9% and 74.0% of the 914 CpGs showed consistent direction of association, with Pearson correlation coefficients for point estimates of 0.82 and 0.48, respectively (Supplementary Data 3). In leave-one-out analyses, in which we reran the main meta-analysis repeatedly with one study removed each time, there was no strong evidence that any of the 161 CpGs potentially contained a single-nucleotide polymorphism (SNP) at cytosine or guanine positions differently methylated CpGs among neonates born between 2500 and 4000 g, low birthweight (<2500 g, high birthweight: ≥4000 g, low birthweight: <2500 g). Studies with mixed ethnicities analysed all participants together with adjustment for ethnicities, g., grams, wk, weeks; y., years. Full study names can be found in study-specific Supplementary Methods. For some studies the sample size for defining normal/high BW was too small.

**Table 1: Characteristics of the participating studies in the main meta-analysis for the association between neonatal blood DNA methylation and birthweight**

| Study                | Total N | Normal birthweight, N (%) | High birthweight, N (%) | Birthweight (g) | Gestational age (wk) | Ethnicity | Boys, N (%) |
|----------------------|---------|---------------------------|-------------------------|----------------|---------------------|-----------|-------------|
| ALSPAC               | 633     | 547 (86.4)                | 79 (12.5)               | 3512 ± 443     | 39.7 ± 1.3          | European  | 301 (47.6)  |
| CBCa: Hispanic       | 127     | 106 (83.5)                | 19 (15.0)               | 3445 ± 484     | 39.8 ± 1.3          | Hispanic  | 74 (58.3)   |
| CBCa: Caucasian      | 136     | 108 (79.4)                | 26 (19.1)               | 3625 ± 472     | 39.7 ± 1.5          | European  | 79 (58.1)   |
| CHAMACOS             | 283     | 236 (83.4)                | 44 (15.5)               | 3520 ± 446     | 39.3 ± 1.2          | Hispanic  | 142 (50.1)  |
| CHS                  | 199     | 168 (84.4)                | 28 (14.1)               | 3486 ± 476     | 40.2 ± 1.2          | Mixed     | 79 (39.7)   |
| ERLI                 | 131     | 113 (86.3)                | 16 (12.2)               | 3507 ± 480     | 39.3 ± 1.0          | Mixed     | 70 (53.4)   |
| EXPOSOMICS: Rhea, Environica and Picoliupi | 324 | 297 (91.7) | 22 (6.8) | 3368 ± 437 | 39.4 ± 1.2 | European | 169 (52.1) |
| GECKO                | 255     | 206 (80.8)                | 46 (18.0)               | 3543 ± 533     | 39.7 ± 1.3          | European  | 136 (53.3)  |
| INMA                 | 162     | 145 (89.5)                | 15 (9.3)                | 3408 ± 431     | 39.5 ± 1.1          | European  | 74 (45.7)   |
| Generation R         | 717     | 589 (82.1)                | 122 (17.0)              | 3572 ± 465     | 40.2 ± 1.1          | European  | 372 (51.9)  |
| GOYAb                | 947     | 649 (68.5)                | 294 (31.0)              | 3750 ± 501     | 40.4 ± 1.3          | European  | 483 (51.0)  |
| Healthy Start: African American | 77 | -                      | -                       | 3059 ± 358     | 38.9 ± 1.3          | African   | 42 (54.5)   |
| Healthy Start: Hispanic | 115   | -                        | -                       | 3322 ± 395     | 39.1 ± 1.1          | Hispanic  | 55 (47.8)   |
| Healthy Start: Caucasian | 240  | 220 (91.7)                | 14 (5.8)                | 3325 ± 425     | 39.3 ± 1.1          | European  | 125 (52.1)  |
| INMA                 | 166     | -                        | -                       | 3297 ± 400     | 39.9 ± 1.2          | European  | 82 (49.4)   |
| IOW F2               | 118     | 97 (82.2)                 | 17 (14.4)               | 3432 ± 525     | 39.7 ± 1.6          | European  | 59 (50.0)   |
| MoBa1                | 1066    | 795 (74.6)                | 251 (23.5)              | 3644 ± 544     | 39.5 ± 1.6          | European  | 568 (53.3)  |
| MoBa2                | 587     | 435 (74.1)                | 146 (24.9)              | 3701 ± 487     | 40.1 ± 1.2          | European  | 329 (56.0)  |
| MoBa3                | 205     | 153 (74.6)                | 51 (24.9)               | 3706 ± 491     | 39.8 ± 1.2          | European  | 106 (51.7)  |
| NCLa                 | 792     | 752 (74.7)                | 192 (24.2)              | 3671 ± 506     | 40.0 ± 1.3          | European  | 453 (57.2)  |
| NEST: African American | 99    | -                        | -                       | 3197 ± 534     | 39.3 ± 1.2          | African   | 47 (47.5)   |
| NEST: Caucasian      | 111     | 94 (84.7)                 | 13 (11.7)               | 3446 ± 471     | 39.5 ± 1.2          | European  | 50 (45.0)   |
| NHBCS                | 96      | 84 (87.5)                 | 12 (12.5)               | 3509 ± 453     | 39.6 ± 1.1          | European  | 53 (55.2)   |
| PREDO                | 540     | 428 (79.3)                | 99 (18.3)               | 3572 ± 478     | 40.1 ± 1.2          | European  | 264 (48.8)  |
| PRISM                | 138     | -                        | -                       | 3385 ± 441     | 39.5 ± 1.1          | Mixed     | 76 (55.1)   |
| PROGRESS             | 143     | -                        | -                       | 3124 ± 387     | 38.6 ± 1.1          | Hispanic  | 77 (53.8)   |
| RICHS                | 89      | 52 (58.4)                 | 23 (25.8)               | 3335 ± 734     | 38.9 ± 1.2          | European  | 35 (39.3)   |
| Project Viva         | 329     | 263 (79.9)                | 64 (19.5)               | 3623 ± 473     | 40.0 ± 1.0          | European  | 168 (51.1)  |
| Total N              | 8825    | 6377 (7193)               | 1593                    |                |                     |           |             |

Results are presented as mean ± SD or N (%). Normal birthweight: 2500–4000 g, low birthweight: >4000 g, low birthweight: <2500 g. Studies with mixed ethnicities analysed all participants together with adjustment for ethnicities, g., grams, wk, weeks; y., years. Full study names can be found in study-specific Supplementary Methods. For some studies the sample size for defining normal/high BW was too small.

*CBC, CHS and NCL used heel prick blood spot samples instead of cord blood

**Analyses at older ages.** We took the 914 neonatal blood CpGs that were associated with birthweight at Bonferroni-corrected statistical significance and with \( \beta \geq 0.50 \) and examined their associations with birthweight when measured in blood taken in childhood (2–13 years; \( n = 2756 \) from 10 studies), adolescence (16–18 years; \( n = 2906 \) from six studies) and adulthood (30–45 years; \( n = 1616 \) from three studies). Only participants from ALSPAC, CHAMACOS and Generation R had also contributed to the main neonatal blood EWAS. In childhood, adolescence and adulthood, we observed 87, 49 and 42 of the 914 CpGs to be nominally associated with birthweight (\( p < 0.05 \)), still limited. Four CpGs were associated with low versus normal birthweight (Bonferroni-corrected threshold), none of which overlapped with the 914 CpGs from the main analysis (Supplementary Data 5B, purple track (4) in Fig. 4). We identified that 161 of the 914 differentially methylated CpGs potentially contained a single-nucleotide polymorphism (SNP) at cytosine or guanine positions (i.e. polymorphic CpGs; Supplementary Data 6). Polymorphic CpGs may affect probe binding and hence measured DNA methylation levels26,27. We used one of the largest studies (ALSPAC; \( n = 633 \)) to explore this. We found no indication of bimodal distributions for any of the 161 CpGs suggesting SNPs had not markedly affected methylation measurements at these sites (dip test \( p \)-values: 0.299–1.00)28–30.
All these CpGs showed consistent directions of association. Ten CpGs showed differential methylation across all four age periods. However, only a minority survived Bonferroni correction for 914 tests ($p < 5.5 \times 10^{-5}$): 12 (1.3%), 1 (0.1%) and 0 CpGs in childhood, adolescence and adulthood, respectively (Supplementary Data 7; the difference in birthweight in grams per 10% methylation difference, the fixed effects inverse variance weighted meta-analysis (6) in Fig.4). Of the 914 CpGs, 50, 52 and 49% showed consistency in direction of association in childhood, adolescence and adulthood, but correlations of the associations of DNA methylation and birthweight between methylation measured in infancy and that measured in childhood, adolescence and adulthood were weak (Pearson correlation coefficients: 0.15, 0.06 and 0.02, respectively).

**Intrauterine factors.** We observed enrichment of previously published maternal smoking-related CpGs in the birthweight-associated CpGs ($p < 1.06 \times 10^{-7}$). Highlighted in orange are the 914 CpGs with enrichment ($p < 1.06 \times 10^{-7}$) and $I^2 \leq 50\%$ and highlighted in blue are the 115 CpGs with $p < 1.06 \times 10^{-7}$ and $I^2 > 50\%$ for folate. For famine we did not find evidence for overlap with plasma folate. For famine, we were unable to explore overlap with DNA methylation at the Bonferroni-significant level as the previous EWAS of famine only reported results that reached a FDR level of statistical significance. In additional analyses for overlap between all FDR hits from the birthweight EWAS with those FDR hits presented in the smoking, maternal BMI and famine EWAS, we found an overlap of 430/8170 CpGs (5.3%, $p_{enrichment} = 3.34 \times 10^{-4}$) for smoking, 584/8170 CpGs (7.1%, $p_{enrichment} = 3.40 \times 10^{-02}$) for maternal BMI and 72/8170 (0.9%, $p_{enrichment} = 0.51$) for famine.
Metastable epialleles and imprinted genes. We tested the birthweight-associated CpGs for enrichment of metastable epialleles (loci for which the methylation state is established in the periconceptional period\textsuperscript{33,34}). We additionally tested for enrichment of CpGs annotated to imprinted genes (loci that depend on the maintenance of parental-origin-specific methylation marks in the pre-implantation embryo, some of which are known to regulate fetal growth\textsuperscript{35,36}). We did not find evidence of enrichment for metastable epialleles (3/1936 metastable epialleles overlap a birthweight-associated CpG), imprinting control regions (0/741) or imprinted gene transcription start sites (5/1728) (Supplementary Data 9).

Comparison with GWAS for birthweight. To compare these EWAS results to those from genetic studies, we used the 60 recently published fetal SNPs associated with birthweight in a GWAS meta-analysis of 153,781 newborns\textsuperscript{37} and mapped the CpG sites identified in the EWAS to these SNPs to seek evidence of co-localisation of genetic and epigenetic variation (Supplementary Data 10). We repeated this for the 10 recently published maternal SNPs associated with birthweight in a GWAS meta-analysis of 86,577 women\textsuperscript{38} (Supplementary Data 11). We observed that one or more of the 914 birthweight-associated CpGs were within +/−2Mb of 34/60 fetal and all 10 maternal SNPs. Ten fetal and four maternal SNPs were within 100 kb of identified CpGs. In a look-up of the fetal and maternal SNPs from GWAS of birthweight in an online cord blood methylation quantitative trait loci (mQTL) database (mqtldb.org\textsuperscript{39}), 35 fetal and four maternal SNPs affected methylation at some CpG(s), but none at the 914 birthweight-associated CpGs specifically.

Functional analyses. We compared the 914 birthweight-related CpGs with a recently published list of 18,881 expression quantitative trait methylation sites (cis-eQTM, +/−250 kb around
the transcription start site), CpG sites known to correlate with gene expression, from whole blood samples of 2101 Dutch adult individuals. We found that 82 of the 914 birthweight-associated CpGs were associated with gene expression of 98 probes (cis-eQTM)\(^{10}\), \(\chi^2_{\text{enrichment}} < 1.73 \times 10^{-11}\), Supplemental Data 12). Additionally, in 112 Spanish 4-year-olds\(^{41}\), we observed that 19 CpGs were inversely associated with whole blood mRNA gene expression and four CpGs were positively associated with gene expression (FDR<0.05, Supplemental Data 13). Of these 23 CpGs, 13 were also found in the publicly available cis-eQTM list\(^{40}\). In 84 Gambian children (age 2 years)\(^{33}\), we found two CpGs that were inversely associated with whole blood mRNA gene expression, but neither were found in the Spanish results or the publicly available cis-eQTM list. The 914 birthweight-associated CpGs showed no functional enrichment of Gene Ontology (GO) terms or Kyoto Encyclopedia of Genes and Genomes (KEGG) terms (FDR<0.05).

**Mendelian randomization.** We aimed to explore causality using MR analysis, in which genetic variants associated with methylation levels (methylation quantitative trait loci (mQTLs)) are used as instrumental variables to appraise causality. For 788 (86%) of the 914 birthweight-associated CpGs, no mQTLs were identified in a publicly available mQTL database\(^{39}\). For 108 (86%) of the remaining 126 CpGs, only one mQTL was identified and for the remainder none had more than four mQTLs (Supplemental Data 14 provides a complete list of all mQTLs identified for these 126 CpGs). Many of the currently available methods that can be used as sensitivity analyses to explore whether MR results are biased by horizontal pleiotropy (a single mQTL influencing multiple traits) require more than one genetic instrument (here mQTLs) and even with two or three this can be difficult to interpret\(^{42}\). Having determined that it was not possible to undertake MR analyses of 86% of the birthweight-related differentially methylated CpGs (because we did not identify any mQTLs), and for the majority of the remaining CpGs we would not have been reliably able to distinguish causality from horizontal pleiotropy (because only one mQTL could be identified), we decided not to pursue MR analyses further.

**Discussion**

This large-scale meta-analysis shows that birthweight is associated with widespread differences in DNA methylation. We observed some enrichment of birthweight-associated CpGs among sites that have previously been linked to smoking during pregnancy\(^{14}\) and pre-pregnancy BMI\(^{15}\), consistent with the hypothesis that epigenetic pathways may underlie the observational associations of those prenatal exposures with birthweight\(^{21,44,45}\). However, the actual overlap in this analysis was modest, likely explained by the adjustments for maternal smoking and BMI in the EWAS analyses. The overlap that we observed with pregnancy smoking-related CpGs may reflect the possibility that smoking-related CpGs capture smoking better than self-report\(^{46,47}\), in line with expectations of pregnant women underreporting their smoking behaviour. Adjustment for maternal smoking and BMI may have masked a greater level of overlap between our results and EWAS of these two maternal exposures. The fact that we find an association of DNA methylation across the genome with birthweight provides some support for our conceptual framework shown in Fig. 1. However, we acknowledge that the associations that we have observed may also be explained by causal effects of maternal pregnancy exposures on both DNA methylation and fetal growth, as well as subtle inflammatory responses in cell-type proportions associated with maternal smoking that might not have been completely captured with the currently available cell type estimation methods.

The differential methylation associated with birthweight in neonates persisted only minimally across childhood and into adulthood. Larger (preferably longitudinal) studies are needed to explore persistent differential methylation in more detail and with better power at older ages. It is possible that inclusion of the Gambia study in the childhood EWAS (which was the only non-European study in these analyses and was not included in the main meta-analyses with neonatal blood) might have impacted these results, although this study made up just 7% of the total child follow-up sample. A rapid attenuation of differential methylation in relation to birthweight in the first years after birth has previously been reported\(^ {19}\), but our sample size for these analyses may have been too small to detect persistence. This rapid decrease, if real, may indicate a reduction in the dose of the child’s exposure to maternal factors such as smoking once the offspring is delivered, with that reduction continuing as the child ages. Persistence of birthweight-related differential DNA methylation may not necessarily be a prerequisite for long-term effects, as transient differential methylation in early life may cause lasting functional alterations in organ structure and function that predispose to later adverse health effects.

In observational studies, birthweight has repeatedly been associated with a range of later-life diseases. Change in DNA methylation has been hypothesized as a potential mechanism linking early exposures, birthweight and later health (Fig. 1). We originally aimed to explore this using MR analysis. For the vast majority of the birthweight-associated CpGs, no genetic instrumental variables were available. For the remaining 126 CpGs, only one mQTL was available, which would make it impossible to disentangle causality from horizontal pleiotropy. To ensure a strong basis for future MR analyses on this topic, there is a clear need for a more extensive mQTL resource.

Strengths of this study are its large sample size and the extensive analyses that we have undertaken. In a post hoc power calculation based on the sample size of 8825 with a weighted mean birthweight of 3560 g (weighted mean standard deviation (SD): 483 g) and with an alpha set at the Bonferroni-corrected level of \(P<1.06 \times 10^{-7}\) we had 80% power, with a two-sided test, to detect a minimum difference of 0.13 SD (63 g) in birthweight for each SD increase in methylation. The difference in methylation corresponding to a 1 SD increase differs per CpG, as it depends on the distribution of the methylation values. We acknowledge that smaller differences which might be clinically or biologically relevant may not have been identified in the current analysis. Nonetheless, to our knowledge this analysis has brought together all studies currently available with relevant data and is the largest published study of this association. DNA methylation patterns in neonatal blood, whilst easily accessible in large numbers, may not reflect the key tissue of importance in relation to birthweight. DNA methylation and gene expression in placental tissue may be important targets for future studies. DNA methylation varies between leucocyte subtypes\(^{49}\) and we used an adult whole blood reference to correct for this in the main analyses\(^ {23,24}\), as the study-specific analyses were completed before the widespread availability of specific cord blood reference datasets\(^{30,51}\). However, we observed very similar findings in two studies.
(Generation R and GECKO) when we compared the results with those using one of the currently available cord blood references. Although we adjusted for potential major confounders that may affect both methylation and fetal growth, we acknowledge that the main results cannot ascertain causality. That is, whilst we have hypothesised that variation in fetal DNA methylation influences fetal growth and hence birthweight, and undertaken the analyses accordingly, we cannot exclude the possibility that differences in neonatal blood DNA methylation are caused by variation in fetal growth itself, or that the association is confounded by factors, including maternal smoking and BMI, that independently influence both fetal growth and DNA methylation (as suggested in Fig. 1). The 450k array that was used to measure genome-wide DNA methylation only covers 1.7% of the total number of CpGs present in the genome and specifically targets CpGs in promoter regions and gene bodies. We removed the CpGs that were flagged as potentially cross-reactive, as the measured methylation levels may represent methylation at either of the potential loci. Also, although we did not find evidence for polymorphic effects for the 161 potentially polymorphic CpGs in ALSPAC, we cannot completely exclude these potential polymorphic effects in the main analysis results. The majority of participants were of European ethnicity and when analyses were restricted to those of European ethnicity the results were essentially identical to those with all studies included. Direct comparisons of the main analysis with analyses in those of Hispanic or of African ethnicity for the 914 hits suggested strong correlations with Hispanic but weaker with African ethnicity. However, these results need to be treated with caution. First, we had very few studies of Hispanic and African populations. Second, we only compared the initial hits from the main meta-analysis with all ethnicities included. A detailed exploration of ethnic differences would require similar large samples for each ethnic group and within ethnic EWAS, which is beyond the scope of the data currently available.

Neonatal blood DNA methylation at sites across the genome is associated with birthweight. Further research is required to determine if these are causal and if so whether they mediate any long-term effect of intrauterine exposures on future health.

Methods

Participants. In the main EWAS meta-analysis we explored associations of neonatal blood DNA methylation with birthweight using data from 8825 neonates from 24 studies in the PACE Consortium (Table 1). We removed multiple births from all analyses and excluded preterm births (<37 weeks) and offspring of mothers with pre-eclampsia or diabetes (three major pathological causes of differences in fetal growth). In follow-up analyses we explored whether any sites found in the main analysis were discernible in relation to birthweight when examined in DNA from blood drawn during childhood (2–13 years: 2756 children from 10 studies), adolescence (16–18 years; 2906 adolescents from six studies) or adulthood (30–45 years; 1616 adults from three studies), see Supplementary Data 1B. Informal consent was obtained from all participants, and all studies received approval from local ethics committees. Study-specific methods and ethical approval statements are provided in Supplementary Methods.

Birthweight, DNA methylation and covariates. Our primary outcome was birthweight on a continuous scale (grams), adjusted for gestational age, and measured immediately after birth or retrospectively reported by mothers in questionnaires. In secondary analyses, we categorised and compared associations with high (≥4000 g, n = 1593) versus normal (2500–4000 g, n = 6377) birthweight. We also explored all associations with (continuous and categorical) birthweight in analyses that did not exclude women with pre-eclampsia, diabetes or preterm delivery, which also resulted in enough cases to explore low (<2500 g, n = 178) versus normal (2500–4000 g, n = 4197) birthweight (Supplementary Data 1C shows the characteristics of participants). Primary, secondary and follow-up analyses are outlined in the study design in Fig. 2. DNA methylation was measured in neonatal blood samples using the Illumina Infinium HumanMethylation450 BeadChip assay. All participants had cord blood samples except for three studies with heel stick blood spots (n = 1254 [14.2%]). After study-specific laboratory analyses, quality control, normalisation, and removal of controls probes (n = 65) and probes that mapped to the X (n = 12,232) and Y (n = 370) chromosomes, we identified 473,864 CpGs. DNA methylation is expressed as the proportion of cells in which the DNA was methylated at a specific site and hence takes values from zero to one. We converted this to a percentage and present differences in mean birthweight per 10% higher DNA methylation level at each CpG. All analyses were adjusted for gestational age at delivery, child sex, maternal age at delivery, parity the 10 Smoke (smoking during pregnancy) were measured in neonatal blood samples (no smoking throughout pregnancy), pre-pregnancy BMI, socio-economic position, technical variation, and estimated white blood cell proportions (B-cells, CD8+ T-cells, CD4+ T-cells, granulocytes, NK-cells and monocytes). In studies with participants from multiple ethnic groups, each group was analysed separately and results were added as fixed-effects meta-analyses as separate studies. Further details are provided in the study-specific Supplementary Methods.

Statistical methods. Robust linear (birthweight as a continuous outcome) or logistic (binary birthweight outcomes) regression EWAS were undertaken within each study according to a pre-specified analysis plan. Quality control, normalisation and regression analyses were conducted independently by each study. After confirming the requirements of study-specific QCs, we performed fixed effects inverse variance weighted meta-analysis. The meta-analysis was done independently by two study groups and the results were compared in order to minimise the likelihood of human error. We show (two-sided) results after correcting for multiple testing using both the FDR<0.05 and the Bonferroni correction (< 1.06 x 10−7). We completed follow-up analyses for differentially methylated CpGs that reached the Bonferroni-adjusted threshold and did not show large between-study heterogeneity. We annotated the nearest gene for each CpG using the UCSC Genome Browser build hg19. We explored whether between-study heterogeneity might be explained by differences in ethnicity between studies, by repeating the meta-analysis including only participants of European ethnicity, which was by far the largest ethnicity subgroup (n = 6023 from 17 studies) (Fig. 2). Ethnicity was defined using maternal or self-report, unless specified otherwise in study-specific Supplementary Methods. We also did meta-analysis only including the Hispanic studies and only including the African American studies and present those results for illustrative purposes only, given the much smaller sample size. All analyses were performed using R, except for the meta-analysis which was performed using METAL. We removed CpGs that co-hybridised to alternate sequences (i.e. cross-reactive sites), because we cannot distinguish whether the differential methylation is at the locus that we have reported or at the one that the probe cross-reacts with. We compared the birthweight-related CpGs to lists of CpGs that are potentially influenced by a SNP (polymorphic sites). For these CpGs, we determined if DNA methylation levels were influenced by nearby SNPs, by assessing whether their distributions deviated from unimodality using Hartigans’ dip test and visual inspection of density plots in n = 742 cord blood samples in the ALSPAC study.

Analyses at older ages. Analyses of the associations with DNA methylation in blood collected in childhood, adolescence and adulthood followed the same covariate adjustment and methods as for the main analyses (p < 5.5 x 10−4 for 914 tests). All participants and studies in these analyses at older ages had not been included in the main meta-analysis in neonatal blood, except for ALSPAC (n = 633 in neonatal analyses, n = 605 in childhood and n = 526 in adolescence), CHA-MACOS (n = 283 in neonatal analyses and n = 191 in childhood) and Generation R (n = 717 in neonatal analyses and n = 372 in childhood). Characteristics are shown in study-specific Supplementary Methods and Supplementary Data 1B.

Intrauterine factors. We used a hypergeometric test to explore the extent to which any of the birthweight-related CpGs overlapped with those that have been associated with intrauterine exposure to smoking, folate and plasma folate and BMI. For this comparison we used the same (Bonferroni-corrected) cut-off for statistical significance. No CpGs reached the Bonferroni-corrected cut-off for smoking or folate. We additionally appraised this overlap using the FDR<0.05 cut-off for all traits (n = 870 birthweight-related CpGs, n = 6703 smoking-related CpGs, n = 16,067 BMI-related CpGs, n = 443 folate-related CpGs, n = 77 folate-related CpGs). These FDR results were available from the publications for smoking, folate and BMI, and we obtained them from the corresponding author for BMI.

Metastable epialleles and imprinted genes. We tested the birthweight-associated CpGs for enrichment of metastable epialleles and CpGs associated with imprinting. The metastable CpGs were derived from one published study that identified 1936 putative metastable epialleles. For imprinted genes, we first identified a set of CpGs falling within a curated set of imprinting control regions; differentially methylated regions controlling the parental-specific expression of one or more imprinted genes. We extracted the set of imprinting control region controlled genes from the above source and identified all 450 CpGs within +/-10kb of the gene transcription start site, including all known alternative TSS identified in grch37.ensembl.org using biomart. We compared with GWAS for birthweight. We compared the birthweight-associated CpGs with the 60 SNPs from the most recent GWAS meta-analyses of fetal genotype associations with birthweight in >150,000 newborns and with a list of CpGs from the most recent GWAS meta-analysis of maternal genotype associations with birthweight in >86,000 women. With this comparison we checked if the EWAS top hits were located within a 4 Mb window (+/− 2 Mb)
surrounding these SNPs. We additionally checked whether SNPs and CpGs were located in the same gene.

Functional analyses. To explore the association of methylation with gene expression, we compared birthweight-related CpGs with a recently published list of 18,881 cis-eQMTs from whole blood samples of 2101 Dutch adult individuals.40 With a hypergeometric test, we calculated enrichment of cis-eQMTs in the list of birthweight-associated CpGs. We further explored methylation of birthweight-associated CpGs in relation to whole blood mRNA gene expression (transcript levels) within a 500 kb region of the CpGs (+/− 250 kb, FDR<0.05) in 122 Spanish 4-year-olds41 and 84 Gambian 2-year-olds42 (Supplementary Methods). To better understand the potential mechanisms linking DNA methylation and birthweight, we explored the potential functions of the birthweight-associated CpGs using GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. We used the missMethyl R package,63 which enabled us to correct for the number of probes per gene on the 450k array, based on the November 2018 version of the GO and KEGG source databases. To filter out the large, general pathways we set the number of genes for each gene set between 15 and 1000, respectively. We calculated FDR at 5% corrected P-values for enrichment.

Mendelian randomization. MR uses genetic variants as instrumental variables to study the causal effect of exposures on outcomes.44,65 We aimed to use two-sample MR to explore (a) evidence of a causal association of methylation levels at the identified CpGs with birthweight and (b) evidence of a causal association of these CpGs with later-life health outcomes (i.e. to explore our hypothesised causal mechanisms shown in Figure 1). We did this by first searching a publicly available mQTL database39 to identify cis-mQTLs within 1 Mb of each of the Bonferroni-corrected, with FDR ≤ 50%, birthweight-related differentially methylated CpGs. These mQTLs could then be used as genetic instrumental variables for methylation levels of the birthweight-related CpGs. We then aimed to determine the association of these mQTLs with birthweight and later-life health outcomes from publicly available summary GWAS results.66

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability
All relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding authors upon reasonable request. All summary statistics from this EWAS meta-analysis are available via doi: 10.5281/zenodo.2222267. A reporting summary for this Article is available as a Supplementary Information file.

Code availability
The code used for this EWAS meta-analysis is available from the authors upon request.

References
1. Tyrrell, J. et al. Genetic evidence for causal relationships between maternal obesity-related traits and birth weight. JAMA 315, 1129 (2016).
2. Tyrrell, J. et al. Genetic variation in the 15q25 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) interacts with maternal self-reported smoking status during pregnancy to influence birth weight. Hum. Mol. Genet. 21, 3344–3358 (2012).
3. Bakker, R., Steegers, E. A. P., Hofman, A. & Jaddoe, V. W. V. Blood pressure and kidney function: longitudinal association from birth to adulthood. Circulation 136, 1954–1963 (2017).
4. Lawlor, D. A. et al. Association of existing diabetes, gestational diabetes and glycosuria in pregnancy with macrosomia and offspring body mass index, waist and fat mass in later childhood: Findings from a prospective pregnancy cohort. Diabetologia 53, 89–97 (2010).
5. van Iters, E. M. & Steegers-Theunissen, R. P. M. Influence of maternal fetal status on human fetal growth parameters. Mol. Nutr. Food. Res. 57, 582–595 (2013).
6. Painter, R. C., Roseboom, T. J. & Bleker, O. P. Prenatal exposure to the Dutch famine and disease in later life: an overview. Reprod. Toxicol. 20, 345–352 (2005).
7. Whincup, P. H. et al. Birth weight and risk of type 2 diabetes: a systematic review. JAMA 300, 2886–2897 (2008).
8. Lawlor, D. A., Ronalds, G., Clark, H., Davey Smith, G. & Leon, D. A. Birth weight is inversely associated with incident coronary heart disease and stroke among individuals born in the 1950s: findings from the Aberdeen children of the 1950s prospective cohort study. Circulation 112, 1414–1418 (2005).
9. O’Donnell, K. J. & Meaney, M. J. Fetal origins of mental health: the developmental origins of health and disease hypothesis. Am. J. Psychiatry 174, 319–328 (2017).
10. McCormack, V. A., Silva, I. D. S., Koupil, I., Leon, D. A. & Lethbridge, H. O. Birth characteristics and adult cancer incidence: Swedish cohort of over 11,000 men and women. Int. J. Cancer 115, 611–617 (2005).
11. Risnes, K. R. et al. Birthweight and mortality in adulthood: a systematic review and meta-analysis. Int. J. Epidemiol. 40, 647–661 (2011).
12. Freathy, R. M. Can genetic evidence help us to understand the fetal origins of type 2 diabetes? Diabetologia 59, 1850–1854 (2016).
13. Hansson, M. Birth weight and the fetal origins of adult disease. Pediatr. Res. 52, 471–475 (2002).
14. Joubert, B. R. et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. Am. J. Hum. Genet. 98, 680–696 (2016).
15. Sharp, C. C. et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation findings from the pregnancy and childhood epigenetics (PACE) consortium. Hum. Mol. Genet. 26, 4007–4015 (2017).
16. Gruzieva, O. et al. Epigenome-wide meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. Environ. Health Perspect. 125, 104–110 (2017).
17. Adkins, R. M., Tylavsky, F. A. & Krushkal, J. Newborn umbilical cord blood DNA methylation and gene expression levels exhibit limited association with birth weight. Chem. Biodivers. 9, 888–899 (2012).
18. Engel, S. M. et al. Neonatal genome-wide methylation patterns in relation to birth weight in the Norwegian Mother and Child Cohort. Am. J. Epidemiol. 179, 834–842 (2014).
19. Simpkin, A. J. et al. Longitudinal analysis of DNA methylation associated with birth weight and gestational age. Hum. Mol. Genet. 24, 3752–3763 (2015).
20. Agha, G. et al. Birth weight-for-gestational age is associated with DNA methylation at birth and in childhood. Clin. Epigenetics 8, 1–12 (2016).
21. Lawlor, D. A., Retlon, C., Sattar, N. & Nelson, S. M. Maternal adiposity—a determinant of perinatal and offspring outcomes? Nat. Rev. Endocrinol. 8, 479–488 (2012).
22. Lawlor, D. A. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int. J. Epidemiol. 45, 908–915 (2016).
23. Houseman, E. A. et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 13, 1–16 (2012).
24. Reinius, L. E. et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS ONE 7, e13161 (2012).
25. Aryee, M. J. et al. Minfi: A flexible and comprehensive Biocductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30, 1363–1369 (2014).
26. Chen, Y. A. et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina HumanMethylation450 microarray. Epigenetics. 8, 203–209 (2013).
27. Naeem, H. et al. Reducing the risk of false discovery enabling identification of biologically significant genome-wide methylation status using the HumanMethylation450 array. BMC Genomics 15, 51 (2014).
28. Hartigan, J. & Hartigan, P. The dip test of unimodality. Ann. Stat. 13, 70–84 (1985).
29. Maechler, M. Hartigan’s dip test statistic for unimodality - corrected. R package (2015).
30. Retlon, C. L. et al. Data resource profile: accessible resource for integrated epigenomic studies (ARIES). Int. J. Epidemiol. 44, 1181–1190 (2015).
31. Joubert, B. R. et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. Nat. Commun. 7, 10577 (2016).
32. Tobi, E. W. et al. Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylene. Int. J. Epidemiol. 44, 1211–1223 (2015).
33. Rakyan, V. K., Blewitt, M. E., Druker, B., Preis, J. I. & Whitelaw, E. Metastable epialleles in mammals. Trends Genet. 18, 348–351 (2002).
34. Van Baak, T. E. et al. Epigenetic supersimilarity of monoyzygotic twin pairs. Genome. Biol. 19, 2 (2018).
35. Moore, G. E. et al. The role and interaction of imprinted genes in human fetal growth. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 370, 20140074 (2015).
36. Monk, D. et al. Recommendations for a nomenclature system for reporting methylation aberrations in imprinted domains. Epigenetics 1–5 https://doi.org/10.1080/15592294.2016.1264561 (2016).
37. Horikoshi, M. et al. Genome-wide associations for birth weight and related health outcomes in the UK. Hum. Mol. Genet. 27, 742–756 (2018).
39. Gaunt, T. R. et al. Systematic identification of genetic influences on methylation across the human life course. Genome Biol. 17, 1–14 (2016).

40. Bonder, M. J. et al. Disease variants alter transcription factor levels and methylation of their binding sites. Nat. Genet. 49, 131–138 (2016).

41. Guexens, M. et al. Cohort profile: the INMA—Infancia y Medio Ambiente—(environment and childhood) project. Int. J. Epidemiol. 41, 930–940 (2012).

42. Moore, S. E. et al. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infantile incontinence and growth: the Gambia: the ENID trial: early nutrition and immune development. BMC Pregnancy Childbirth 12, 1–8 (2012).

43. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum. Mol. Genet. 27, R195–R208 (2018).

44. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013.

45. Küpers, L. K. et al. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. Int. J. Epidemiol. 44, 1224–1237 (2015).

46. Valer, L. et al. Misclassified exposure in epigenetic mediation analyses. Does DNA methylation mediate effects of smoking on birthweight? Epigenomics 9, 253–265 (2017).

47. Reese, S. E. et al. DNA methylation score as a biomarker in newborns for sustained maternal smoking during pregnancy. Environ. Health Perspect. https://doi.org/10.1289/EHP333 (2016).

48. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 33(Suppl), 245–254 (2003).

49. Jaffe, A. E. & Irizarry, R. A. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 15, R31 (2014).

50. Rakilj, K. et al. DNA methylation of cord blood cell types: applications for mixed cell birth studies. Epigenetics 11, 354–362 (2016).

51. Gervin, K. et al. Cell type specific DNA methylation in cord blood: a 450K-reference data set and cell count-based validation of estimated cell type composition. Epigenetics. 11, 690–698 (2016).

52. Deederwaerder, S. et al. Evaluation of the infinium methylation 450K technology. Epigenomics https://doi.org/10.2217/epi.11.105 (2011).

53. Felix, J. F. et al. Cohort profile: pregnancy and childhood epigenetics (PACE) consortium. Int. J. Epidemiol. 16, 10–14 (2017).

54. van der Most, P. J., Kupers, L. K., Snieder, H. & Nolte, I. QC EWAS: automated quality control of results of epigenome-wide association studies. Bioinformatics 33, 1243–1245 (2017).

55. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–2191 (2010).

56. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57, 289–300 (1995).

57. Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. Stat. Med. 21, 1539–1558 (2002).

58. Triche, T., Fdb.InfiniumMethylation.hg19: annotation package for Illumina Infinium DNA methylation probes. R package version 2.2.0 (2014).

59. Carlson, M. & Maintainer, B. TxDb.Hsapiens.UCSC.hg19.knownGene: annotation package for TxDb object(s). R package version 3.2.2 (2015).

60. R Core Team. R: A language and environment for statistical computing. (Austria, 2013).

61. Durinck, S. et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics 21, 3439–3440 (2005).

62. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. 4, 1184–1191 (2009).

63. Phipson, B., Maksimovic, J. & Oshlack, A. missMethyl: an R package for analyzing data from Illumina’s HumanMethylation450 platform. Bioinformatics 32, 286–288 (2016).

64. Davey Smith, G. & Ebrahim, S. ‘Mendelian randomization’: Can genetic epidemiology contribute to understanding environmental determinants of disease? Int. J. Epidemiol. 32, 1–22 (2003).

65. Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N. & Davey Smith, G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat. Med. 27, 1133–1163 (2008).

66. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human genome. eLife 7, e34408 (2018).

Acknowledgements
For all studies, acknowledgments can be found in Supplementary Information: Supplementary Acknowledgements. For all studies, funding statements can be found in Supplementary Information: Supplementary Funding.

Author contributions
L.K.K., D.A.L., C.L.R., H.S. and J.F.F. conceived and designed the study. Study-specific analyses were completed by G.C.S. (ALSPAC and GOYA), S.K.M. (BAMSE, R.R. (CBC), P.Y. (CHAMACOS), C.V.B. (CHS), K.M.B. (EARL), A.G. and A.N. (EP4063MICS, The Gambia and MoBa3), S.A.S. (FLEHS1), L.K.K. (GECKO), C.A. (Gen3G), C.M. (Generation R), J.L. (Glakau), A.P.S. (Healthy Start), L.A.S. (INMA), F.I.R. (IOF FIV), J.W.H. (IOW F2), D.A.Y.D. P. (Lifelines), C.M.P. (MoBa1), S.E.R. (MoBa2), A.J.W. (NCL), D.D.J. (NEST), M.W. (NFBC86 and NFBC06), T.M.E. (NHBCS and RICH5), J.V.D. (NTR), C.L.J. (PIJAMA), D.C. (PRISOM), A.C.J. (PROGRESS), S.L. (Project Viva), R.C.H. (Rainc), V.U. (STOPPA). L.K.K. and C.M. meta-analysed the results. L.K.K., C.M., G.C.S., P.Y., L.A.S., A.G., AN. and M.J.S. performed follow-up analyses. L.K.K., D.A.L., C.L.R., H.S. and J.F.F. interpreted the results. L.K.K., with input from D.A.L., C.L.R., H.S. and J.F.F., wrote the first draft of the manuscript. All authors (L.K.K., C.M., G.C.S., P.Y., L.A.S., A.G., C.M.P., S.E.R., A.J.W., D.C., A.P.S., A.N., S.L., R.R., C.H., C.V.B., C.A., A.C.I., K.M.B., J.W.H., T.M.E., C.-J.X., R.-C.H., D.A.w.d.P., M.W., S.K.M., V.U., F.I.R., J.L., J.D.R., S.A.S.L., T.G.R., M.C.M., E.A.N., Z.X., L.D., S.W., M.D., D.L.D., O.S., J.H.H., D.D.I., L.G., M.B., P.F., R.O.W., H.P., H.Z., M.R.K., U.G., C.J.M., L.J.B., J.M.V., M.-R.I., A.B., A.K.O., S.E., P.M.V., S.E.M., G.W., A.R.I.S., S.E.H., T.L.S., I.A.T., K.R., I.V.V., K.K., T.S.N., M.J.S., Y.Y.G., L.R., M.K., A.A.L., B.E., K.H., H.M., R.L.M., T.D., M.V., L.B., A.A.B., L.A.C., W.K., D.A., M.D.V., S.S., J.K., R.K., S.A.H., E.H., M.N.R., D.I.B., A.P.F., C.J.N., E.G., M.M., M.D.F., E.M., A.M.P., E.K., C.A., E.O., D.D., H.M.B., P.E.M., R.J.W., G.H.K., L.T., M.-F.H., I.S., M.G.M.-K., S.K.M., E.C., I.W., N.H., Z.H., E.B.B., G.D.S., V.W.V.J., R.T.L., W.N., S.J.L., D.A.L., C.L.R., H.S. and J.F.F. read and critically revised subsequent drafts.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-09671-3.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: Nature Communications thanks the anonymous reviewers(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party materials in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019
Leanne K. Küpers\textsuperscript{1,2,3,4}, Claire Monnereau\textsuperscript{5,6,7}, Gemma C. Sharp\textsuperscript{1,8}, Paul Youseti\textsuperscript{1,9}, Lucas A. Salas\textsuperscript{10,11}, Akram Ghantous\textsuperscript{12}, Christian M. Page\textsuperscript{13,14}, Sarah E. Reese\textsuperscript{15}, Allen J. Wilcox\textsuperscript{15}, Darina Czamara\textsuperscript{16}, Anne P. Starling\textsuperscript{17}, Alexei Novoloaca\textsuperscript{12}, Samantha Lent\textsuperscript{18}, Ritu Roy\textsuperscript{19,20}, Cathrine Hoyo\textsuperscript{21,22}, Carrie V. Breton\textsuperscript{23}, Catherine Allard\textsuperscript{24}, Allan C. Just\textsuperscript{25}, Kelly M. Bakulski\textsuperscript{26}, John W. Holloway\textsuperscript{27,28}, Todd M. Everson\textsuperscript{29}, Cheng-Jian Xu\textsuperscript{30,31}, Rae-Chi Huang\textsuperscript{32}, Diana A. van der Plaat\textsuperscript{33}, Matthias Wiel scher\textsuperscript{34}, Simon Kebede Merid\textsuperscript{35}, Vilhelmina Ullema r\textsuperscript{36}, Faisal I. Re zwan\textsuperscript{28}, Jari Lahti\textsuperscript{37,38}, Jenny van Dongen\textsuperscript{39}, Sabine A.S. Langie\textsuperscript{40,41,42}, Tom G. Richardson\textsuperscript{1,2}, Maria C. Magnus\textsuperscript{1,2,13}, Ellen A. Nohr\textsuperscript{43}, Zongli Xu\textsuperscript{44}, Liesbeth Duij ts\textsuperscript{4,44,45}, Shanshan Zhao\textsuperscript{46}, Weiming Zhang\textsuperscript{47}, Michelle Plusquin\textsuperscript{48,49}, Dawn L. DeMee\textsuperscript{50}, Olivia Solomon\textsuperscript{51}, Joosje H. Heimovaara\textsuperscript{3}, Dere je D. Jima\textsuperscript{22,51}, Lu Gao\textsuperscript{23}, Mariona Bustamante\textsuperscript{11,52,53,54}, Patrice Perron\textsuperscript{24,55}, Robert O. Wright\textsuperscript{25}, Iva Hertz-Picciotto\textsuperscript{56}, Hongmei Zhang\textsuperscript{57}, Margaret R. Karagas\textsuperscript{10,58}, Ulrike Gehring\textsuperscript{59}, Carmen J. Marsit\textsuperscript{60}, Lawrence J. Beilin\textsuperscript{60}, Judith M. Vonk\textsuperscript{33}, Marjo-Riitta Jarvelin\textsuperscript{34,61,62,63}, Anna Bergström\textsuperscript{35,64}, Anne K. Örtqvist\textsuperscript{36}, Susan Ewart\textsuperscript{65}, Pia M. Villa\textsuperscript{66}, Sophie E. Moore\textsuperscript{67,68}, Gonneke Willemsen\textsuperscript{39}, Arnout R.L. Standaert\textsuperscript{40}, Siri E. Håberg\textsuperscript{13}, Thorkild I.A. Sørensen\textsuperscript{16,69,70}, Jack A. Taylor\textsuperscript{15}, Katri Räikkönen\textsuperscript{38}, Ivana V. Yang\textsuperscript{71}, Katerina Kechr is\textsuperscript{45}, Tim S. Nawrot\textsuperscript{48,72}, Matt J. Silver\textsuperscript{67}, Yun Yun Gong\textsuperscript{73}, Lorenzo Richardi\textsuperscript{74,75}, Manolis Kogevinas\textsuperscript{11,53,54,76}, Augusto A. Litonjua\textsuperscript{50}, Brenda Esk enazi\textsuperscript{9,77}, Karen Huen\textsuperscript{9}, Hamdi Mbarek\textsuperscript{78}, Rachel L. Maguire\textsuperscript{21,79}, Terence Dwyer\textsuperscript{80}, Martine Vrijheid\textsuperscript{11,53,54}, Luigi Bouchard\textsuperscript{81,82}, Andrea A. Baccarelli\textsuperscript{83,84}, Lisa A. Croen\textsuperscript{85,86}, Wilfried Karmas\textsuperscript{57}, Denise Anderson\textsuperscript{32}, Maaike de Vries\textsuperscript{33}, Sylvain Sebert\textsuperscript{61,62,86}, Juha Ke re\textsuperscript{87,88,89}, Robert Karlsson\textsuperscript{36}, Syed Hasan Arshad\textsuperscript{27,90}, Esa Hämäläinen\textsuperscript{91}, Michael N. Routledge\textsuperscript{92}, Dorret I. Boomsma\textsuperscript{39,93}, Andrew P. Feinberg\textsuperscript{94}, Craig J. Newschaffer\textsuperscript{95}, Eva Govarts\textsuperscript{40}, Matthieu Moisse\textsuperscript{96,97}, M. Daniele Fallin\textsuperscript{98}, Erik Melen\textsuperscript{35,99}, Andrew M. Prentice\textsuperscript{67}, Eero Kajantie\textsuperscript{100,101,102}, Catarina Almqvist\textsuperscript{103,104}, Emily Oken\textsuperscript{104}, Dana Dabelea\textsuperscript{105}, H. Marike Boe zen\textsuperscript{33}, Phillip E. Melton\textsuperscript{106,107}, Rosalind J. Wright\textsuperscript{25}, Gerard H. Koppelman\textsuperscript{108}, Letizia Trevisi\textsuperscript{109}, Marie-France Hivert\textsuperscript{110,104,109}, Jordi Sunyer\textsuperscript{11,53,54,76}, Monica C. Munthe-Kaas\textsuperscript{110,111}, Susan K. Murphy\textsuperscript{112}, Eva Corpeleijn\textsuperscript{13}, Joseph Wiemels\textsuperscript{113}, Nina Holland\textsuperscript{10}, Zdenko Herceg\textsuperscript{12}, Elisabeth B. Binder\textsuperscript{16,114}, George Davey Smith\textsuperscript{12}, Vincent W.V. Jaddoe\textsuperscript{5,6,7}, Rolv T. Lie\textsuperscript{13,115}, Wenche Nystad\textsuperscript{116}, Stephanie J. London\textsuperscript{15}, Debbie A. Lawlor\textsuperscript{12}, Caroline L. Relton\textsuperscript{1,2}, Harold Snieder\textsuperscript{3} & Janine F. Felix\textsuperscript{5,6,7}

\textsuperscript{1}MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. \textsuperscript{2}Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. \textsuperscript{3}University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, The Netherlands. \textsuperscript{4}Division of Human Nutrition and Health, Wageningen University, Wageningen, The Netherlands. \textsuperscript{5}The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. \textsuperscript{6}Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. \textsuperscript{7}Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. \textsuperscript{8}School of Oral and Dental Sciences, University of Bristol, Bristol, UK. \textsuperscript{9}Children’s Environmental Health Laboratory, Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA, USA. \textsuperscript{10}Department of Epidemiology, Geisel School of Medicine at Dartmouth College, Hanover, NH, USA. \textsuperscript{11}ISGlobal, Bacelona Institute for Global Health, Barcelona, Spain. \textsuperscript{12}Epigenetics Group, International Agency for Research on Cancer, Lyon, France. \textsuperscript{13}Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo, Norway. \textsuperscript{14}Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway. \textsuperscript{15}Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, Durham, NC, USA. \textsuperscript{16}Department of Translational Research in Psychiatry, Max-Planck-Institute of Psychiatry, Munich, Germany. \textsuperscript{17}Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. \textsuperscript{18}Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. \textsuperscript{19}HDF Comprehensive Cancer Center, University of California, San Francisco, CA, USA. \textsuperscript{20}Computational Biology and Informatics, UCSF, San Francisco, CA, USA. \textsuperscript{21}Department of Biologics, North Carolina State University, Raleigh, NC, USA. \textsuperscript{22}Center for Human Health and the Environment, North Carolina State University, Raleigh, NC, USA. \textsuperscript{23}Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90089, USA. \textsuperscript{24}Centre de recherche du Centre hospitalier universitaire de Sherbrooke, Sherbrooke, QC, Canada. \textsuperscript{25}Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. \textsuperscript{26}Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA. \textsuperscript{27}Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. \textsuperscript{28}Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. \textsuperscript{29}Department of Environmental Health, Rollins School of Public Health at Emory University, Atlanta, GA, USA. \textsuperscript{30}University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children’s Hospital, Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands. \textsuperscript{31}University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands. \textsuperscript{32}Telethon Kids Institute, University of Western Australia, Perth, Australia. \textsuperscript{33}University of Groningen, University Medical Center Groningen, Department of Epidemiology and Groningen Research Institute for Asthma and
