Title

Cesarean delivery and blood DNA methylation at birth and childhood: Meta-analysis in the Pregnancy and Childhood Epigenetics Consortium.

Journal

Science Advances, 10(48)

Authors

Wang, Siwen
Casey, Emma
Sordillo, Joanne
et al.

Publication Date

2024-11-29

DOI

10.1126/sciadv.adr2084

Peer reviewed

PUBLIC HEALTH

Cesarean delivery and blood DNA methylation at birth and childhood: Meta-analysis in the Pregnancy and Childhood Epigenetics Consortium

Siwen Wang[1], Emma Casey[2], Joanne Sordillo[3], Sofía Aguilar-Lacasaña[4,5,6], Fernanda Morales Berstein[7,8], Richard J. Biedrzycki[9], Sonia Brescianini[10], Su Chen[11], Amy Hough[12], Elena Isaevska[13], Woo Jin Kim[14], Marion Lecorguillé[15], Sebastian Shaobo Li[16], Christian M. Page[17,18], Jaehyun Park[19], Stefan Röder[20], Kristina Salontaji[21], Gillian Santorelli[12], Yidan Sun[22,23], Sungho Won[19,24], Eric Zillich[25], Lea Zillich[25], Isabella Annesi-Maesano[26,27], S. Hasan Arshad[28], Mariona Bustamante[4,5,6], Charlotte A. M. Cecil[21,29], Hannah R. Elliott[7,8], Susan Ewart[30], Janine F. Felix[31,32], Luigi Gagliardi[33], Siri E. Håberg[17,34], Gunda Herberth[20], Barbara Heude[15], John W. Holloway[35], Anke Huels[2,36], Wilfried Karmaus[37], Gerard H. Koppelman[22,23], Stephanie J. London[38], Sunni L. Mumford[39,40], Lorenza Nisticò[10], Maja Popovic[13], Franca Rusconi[33], Enrique F. Schisterman[39,40], Dan J. Stein[41], Tabea Send[42], Henning Tiemeier[43,44], Judith M. Vonk[23,45], Martine Vrijheid[4,5,6], Joseph Leo Wiemels[16], Stephanie H. Witt[25], John Wright[12], Edwina H. Yeung[40], Heather J. Zar[46], Ana C. Zenclussen[20], Hongmei Zhang[37], Jorge E. Chavarro[1,44]*†, Marie-France Hivert[3,47]†

Children born via cesarean delivery have a higher risk of metabolic, immunological, and neurodevelopmental disorders compared to those born via vaginal delivery, although mechanisms remain unclear. We conducted a meta-analysis of epigenome-wide association studies to examine the associations between delivery mode and blood DNA methylation at birth and its persistence in early childhood. Participants were from 19 pregnancy cohorts (9833 term newborns) and 6 pediatric cohorts (2429 children aged 6 to 10 years). We identified six CpGs in cord blood associated with cesarean delivery (effect size range: 0.4 to 0.7%, $P < 1.0 \times 10^{-7}$): *MAP2K2* (cg19423175), *LIM2* (cg01500140), *CNP* (cg13917614), *BLM* (cg18247172), *RASA3* (cg22348356), and *RUNX3* (cg20674490), independent of cell proportions and other confounders. In childhood, none of these CpGs were associated with cesarean delivery, and no additional CpGs were identified. Delivery mode was associated with cell proportions at birth but not in childhood. Further research is needed to elucidate cesarean delivery's molecular influence on offspring health.

INTRODUCTION

Cesarean delivery accounts for more than one in five births worldwide, with a projected growth to approximately one in three births by 2030 (*1*). Although medically indicated cesarean delivery reduces maternal and infant morbidity and mortality, emerging evidence suggests that babies born via cesarean delivery have different hormonal, physical, microbial, and medical exposures from those born through vaginal birth. These exposures may subtly alter neonatal physiology, potentially leading to risk of certain health conditions later in life (*2*), such as metabolic risk phenotypes (*3*, *4*), immune diseases (*5*), malignancies (*6*, *7*), and neurodevelopmental disorders (*8*). One of the potential mechanisms through which cesarean delivery may increase the risk of long-term adverse health outcomes is via infant epigenetic alterations, such as DNA methylation (DNAm) (*9*, *10*).

In support of this, some studies have observed associations between cesarean delivery and offspring DNAm patterns at birth. The identified DNAm sites are mostly in genes that are implicated in the development of the immune system (*9*–*11*). However, most of the studies were limited by small sample size (<100 participants), adjusting for few covariates, or failing to account for differences in cell-type proportions, resulting in a lack of replication of results across studies (*12*). In addition, it is unknown whether epigenetic variations related to cesarean delivery are temporary or long-lasting.

There has been a call for more comprehensive epigenetic research among children born via cesarean delivery, including at later postnatal ages (*9*). Therefore, we conducted a meta-analysis of epigenome-wide association studies (EWASs) from an international consortium [Pregnancy and Childhood Epigenetics Consortium (PACE Consortium)] to investigate whether cesarean delivery, as compared to vaginal delivery, was associated with differential methylation at cytosine-phosphate-guanine (CpG) oligodeoxynucleotide sites. Specifically, we investigated whether: (i) cesarean delivery was associated with DNAm in cord blood at birth; (ii) cesarean delivery was associated with DNAm in blood collected at older ages (6 to 10 years of age); and (iii) cesarean delivery was associated with differences in blood cell-type proportions at birth or in childhood.

RESULTS

Consortium characteristics

Offspring in our analyses were predominantly white (93%); some cohorts included South Asian, Asian, Black, and mixed ancestries (Table 1 and table S1). Male offspring percentage ranged from 47 to 62%. Among cohorts with information on maternal prepregnancy body mass index (BMI) [all cohorts except Drakenstein Child Health Study (DCHS), Isle of Wight (IOW) Birth Cohort–F2 (IOW-F2), and Lifestyle and environmental factors and their Influence on Newborns Allergy risk (LiNA)], the mean (SD) value ranged from 22.2 (3.1) kg/m^2 [Mothers and Children's Environmental Health

(MOCEH)] to 26.7 (6.1) kg/m^2 [Born in Bradford (BiB)]. The mean gestational age at delivery was very similar across cohorts, ranging from 39.0 (1.2) to 40.3 (1.2). Among cohorts contributing to childhood blood DNAm analyses, IOW-F1 had the highest offspring age [mean (SD), 10.0 (0) years], with other cohorts ranging from 5.7 (0.1) to 8.1 (0.3) years. Cesarean delivery proportions varied from 7.9% (IOW-F1, the United Kingdom) to 36.3% (MOCEH, South Korea), potentially reflecting regional and temporal differences in clinical practice.

λ Values for each EWAS performed by individual cohorts are presented in table S2. Overall, we did not observe strong evidence of genomic inflation in individual cohort results, with the exception of DCHS (cord blood, cell-type unadjusted, λ = 2.08) and LINA (cord blood, cell-type adjusted, λ = 2.15). The effects of these highly inflated cohorts were investigated through leave-one-out analyses.

Meta-analysis for cord blood DNAm (9833 newborns, 19 cohorts)

In analyses adjusting for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, gestational age (model 1) but not cell-type proportions, 2691 CpGs passed the Bonferroni-adjusted P value cutoff of 1.0×10^{-7} (fig. S1). As expected, we observed some evidence of genomic inflation (λ = 1.45) from failure to account for cell-type heterogeneity. In the main meta-analysis (model 2), which further adjusted for cell-type proportions, genomic inflation was reduced (λ = 1.08, Fig. 1). In these cell-type–adjusted analyses (model 2), we identified six CpG sites at which cord blood DNAm levels were statistically significantly related to delivery mode ($P < 1.0 \times 10^{-7}$; Table 2 and figs. S2 to S7). The identified CpGs were located in/near the following genes: *MAP2K2* (cg19423175, $P = 2.91 \times 10^{-10}$, $I^2 = 58\%$), *CNP* (cg13917614, $P = 1.11 \times 10^{-9}$, $I^2 = 37\%$), *LIM2* (cg01500140, $P = 4.31 \times 10^{-9}$, $I^2 = 39\%$), *BLM* (cg18247172, nearest gene, $P = 9.51 \times 10^{-9}$, $I^2 = 47\%$), *RASA3* (cg22348356, $P = 2.62 \times 10^{-8}$, $I^2 = 64\%$), and *RUNX3* (cg20674490, $P = 9.42 \times 10^{-8}$, $I^2 = 62\%$). Compared to newborns born via vaginal birth, those delivered through cesarean delivery exhibited higher cord blood DNAm levels at these specific sites, ranging from 0.4 to 0.7%. No additional CpGs were identified in analyses restricted to cohorts using Infinium MethylationEPIC (EPIC) BeadChip arrays.

The six identified CpGs were previously associated with gestational age (cg19423175, cg1391761, and cg18247172), adult smoking (cg19423175, cg13917614, cg01500140, cg2234835, and cg20674490), and fasting glucose (cg1942317 and cg18247172) upon lookup in EWAS catalog (table S3) (*13*). However, none of these sites were significantly associated with cis-expression quantitative trait methylation (cis-eQTMs from cord blood DNAm and RNA expression levels in blood) after multiple testing correction (*14*, *15*), except for cg01500140 (table S4). Higher cord blood DNAm at this CpG site (cg01500140) was associated with lower RNA expression of the gene *NKG7* in whole blood ($\log_2$ fold change = −0.88, $P = 1.17 \times 10^{-14}$).

Associations were of consistent direction in most cohorts. Leave-one-out analyses did not show evidence of one cohort systematically influencing the meta-analysis results (fig. S8), and there was not a

[1]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA. [2]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA. [3]Division of Chronic Disease Research Across the Lifecourse (CoRAL), Department of Population Medicine, Harvard Medical School, Harvard Pilgrim Health Care Institute, Boston, MA 02215, USA. [4]ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona 08003, Spain. [5]Universitat Pompeu Fabra (UPF), Barcelona 08005, Spain. [6]CIBER Epidemiología y Salud Pública, Madrid 28029, Spain. [7]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol BS8 2BN, UK. [8]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 1UD, UK. [9]Glotech Inc., contractor for Division of Population Health Research, Division of Intramural Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Rockville, MD 20850, USA. [10]Center for Behavioral Sciences and Mental Health, Istituto Superiore di Sanità, Rome 00161 Italy. [11]Department of Biostatistics, College of Public Health, University of Nebraska Medical Center, Omaha, NE 68105, USA. [12]Bradford Institute for Health Research, Bradford BD9 6RJ, UK. [13]Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin, Turin 10126, Italy. [14]Department of Internal Medicine, Kangwon National University, Chuncheon 24289, Korea. [15]Université Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), F-75004 Paris, France. [16]Center for Genetic Epidemiology, Department of Population and Public Health Sciences, University of Southern California, Los Angeles, CA 90033, USA. [17]Centre for Fertility and Health, Norwegian Institute of Public Health, P.O. box 222 Skøyen, 0213 Oslo, Norway. [18]Department of Physical Health and Aging, Division for Mental and Physical Health, Norwegian Institute of Public Health, P.O. box 222 Skøyen, 0213 Oslo, Norway. [19]Interdisciplinary Program of Bioinformatics, College of Natural Sciences, Seoul National University, Seoul 08826, Korea. [20]Department of Environmental Immunology, Helmholtz Centre for Environmental Research - UFZ, Leipzig 04318, Germany. [21]Department of Child and Adolescent Psychiatry / Psychology, Erasmus MC, University Medical Center Rotterdam, Rotterdam 3015 GD, Netherlands. [22]University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Department of Pediatric Pulmonology and Pediatric Allergology, Groningen 9700 RB, Netherlands. [23]Groningen Research Institute for Asthma and COPD, Groningen 9700 RB, Netherlands. [24]Department of Public Health Sciences, Seoul National University, Seoul 08826, South Korea. [25]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim 68159, Germany. [26]Desbrest Institute of Epidemiology and Public Health, Univ Montpellier, INSERM, Montpellier, France. [27]Division of Respiratory Medicine, Allergology, and of Thoracic Oncology, University Hospital of Montpellier, 34093 Montpellier, France. [28]Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK. [29]Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam 3015 GD, Netherlands. [30]College of Veterinary Medicine, Michigan State University, East Lansing, MI 48824, USA. [31]The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam 3015 GD, Netherlands. [32]Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam 3015 GD, Netherlands. [33]Department of Mother and Child Health, Azienda USL Toscana Nord Ovest, Pisa 56121, Italy. [34]Department of Global Public Health and Primary Care, University of Bergen, Kalfarveien 31, N-5018 Bergen, Norway. [35]Human Development and Health, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK. [36]Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA. [37]Division of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, TN 38152, USA. [38]National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA. [39]Department of Biostatistics, Epidemiology and Informatics and Department of Obstetrics and Gynecology, Perelman School of Medicine, Philadelphia, PA 19087, USA. [40]Epidemiology Branch, Division of Population Health Research, Division of Intramural Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD 20817 USA. [41]South African Medical Research Council (SAMRC) Unit on Risk and Resilience in Mental Disorders, Department of Psychiatry and Neuroscience Institute, University of Cape Town, Rondebosch 7700, South Africa. [42]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim 68159, Germany. [43]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA. [44]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA. [45]University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen 9700 RB, Netherlands. [46]South African Medical Research Council (SAMRC) Unit on Child and Adolescent Health, Department of Paediatrics, University of Cape Town, Rondebosch 7700, South Africa. [47]Diabetes Unit, Massachusetts General Hospital, Boston, MA 02114, USA.
*Corresponding author. Email: jchavarr@hsph.harvard.edu
†These authors contributed equally to this work.

Table 1. Characteristics of cohorts participating in the meta-analyses. NA, not applicable.

Cohort	Ancestry	Offspring *n*	Cesarean delivery *n* (%)	Male *n* (%)	Gestational age Mean (SD)	Child age Mean (SD)	Array
Birth							
ALSPAC	European	799	72 (9.0)	47.8%	39.7 (1.3)	NA	450K
BiB	South Asian and European	750	158 (21.1)	53.2%	39.5 (1.2)	NA	EPIC
CCLS 1	European	375	95 (25.3)	58.9%	39.2 (2.0)	NA	450K
CCLS 2	European	298	67 (22.5)	58.1%	39.2 (2.2)	NA	450K
DCHS*	Black and mixed	228	47 (20.6)	55.3%	39.2 (1.3)	NA	450K
EAGeR	European	335	63 (18.8)	49.1%	39.2 (1.2)	NA	EPIC
EDEN	European	155	17 (11.0)	61.0%	39.6 (1.1)	NA	450K
GenR	European	1261	125 (9.9)	51.0%	40.3 (1.2)	NA	450K
INMA	European	362	40 (11.0)	51.4%	39.9 (1.2)	NA	450K
IOW-F2*	European	162	29 (17.9)	58.0%	39.9 (1.6)	NA	450K
LINA	European	457	83 (18.2)	52.8%	39.8 (1.5)	NA	450K
Moba1	European	863	87 (10.1)	54.5%	39.9 (1.2)	NA	450K
Moba2	European	544	64 (11.8)	57.0%	39.9 (1.1)	NA	450K
Moba4	European	825	80 (9.7)	48.4%	39.9 (1.1)	NA	EPIC
Moba8	European	973	99 (10.2)	48.6%	39.9 (1.2)	NA	EPIC
MOCEH	Asian	350	127 (36.3)	50.0%	39.0 (1.2)	NA	EPIC
Piccolipiù	European	354	64 (18.1)	46.6%	39.9 (1.1)	NA	EPIC
POSEIDON	European	258	65 (25.2)	48.1%	39.4 (1.1)	NA	450K
Project Viva	European and Mixed	484	81 (16.7)	52.2%	39.7 (1.6)	NA	450K
Total		9833					
Childhood (6 to 10 years)							
ALSPAC	European	859	78 (9.1)	48.5%	NA	7.5 (0.1)	450K
EDEN	European	149	17 (11.4)	62.0%	NA	5.7 (0.1)	450K
GenR	European	434	38 (8.1)	48.0%	NA	6.0 (0.3)	450K
IOW-F1*	European	330	26 (7.9)	58.0%	NA	10.0 (0.0)	450K
PIAMA	European	202	19 (9.4)	50.0%	NA	8.1 (0.3)	450K
Project Viva	European and Mixed	455	97 (21.3)	52.0%	NA	7.7 (0.6)	450K
Total		2429					

*DNAm levels were measured with either the 450K or EPIC array, and, following quality control, CpG sites overlapping between the arrays (n = 409,033 for DCHS; n = 349,455 for IOW-F1; and n = 365,697 for IOW-F2) were included in the analysis.

single cohort that consistently contributed to heterogeneity across all CpG sites.

Meta-analysis for childhood blood DNAm (2429 children, 6 cohorts)

In meta-analyses of blood collected in childhood, cesarean delivery was not found to be statistically significantly associated with DNAm in either model 1 (fig. S9, adjusted for covariates only, $\lambda = 0.95$) or model 2 (Fig. 2, adjusted for covariates and blood cell types, $\lambda = 1.00$). The estimates {β [95% confidence interval (CI)]} for associations between cesarean delivery and childhood blood DNAm at the six significant sites identified in the cord blood analysis were close to null [Table 3; *MAP2K2* (cg19423175, $P = 0.20$, $I^2 = 0\%$), *CNP* (cg13917614, $P = 0.56$, $I^2 = 0\%$), *LIM2* (cg01500140, $P = 0.17$, $I^2 = 0\%$), *BLM* (cg18247172, nearest gene, $P = 0.23$, $I^2 = 9\%$), *RASA3* (cg22348356, $P = 0.46$, $I^2 = 0\%$), and *RUNX3* (cg20674490, $P = 0.55$, $I^2 = 26\%$)].

Sensitivity analyses

In sensitivity analyses excluding pregnancies affected by maternal pregnancy complications (i.e., gestational diabetes, gestational hypertension, and preeclampsia) or fetal macrosomia, at the six CpG sites previously identified (Table 2), cord blood DNAm showed associations with cesarean delivery exposure of similar magnitudes as in the primary analyses (figs. S10 and S11). However, P values were attenuated possibly because of smaller sample sizes, and thus, lower statistical power. The associations between cesarean delivery and DNAm in blood collected in childhood remained null (fig. S12).

Meta-analyses for blood cell-type proportions at birth and in childhood

In cord blood, cesarean delivery (versus vaginal delivery) was associated with significantly higher proportions of CD4 T cells, CD8 T cells, monocytes, and B cells and significantly lower proportions of granulocytes, natural killer (NK) cells, and nucleated red blood cells

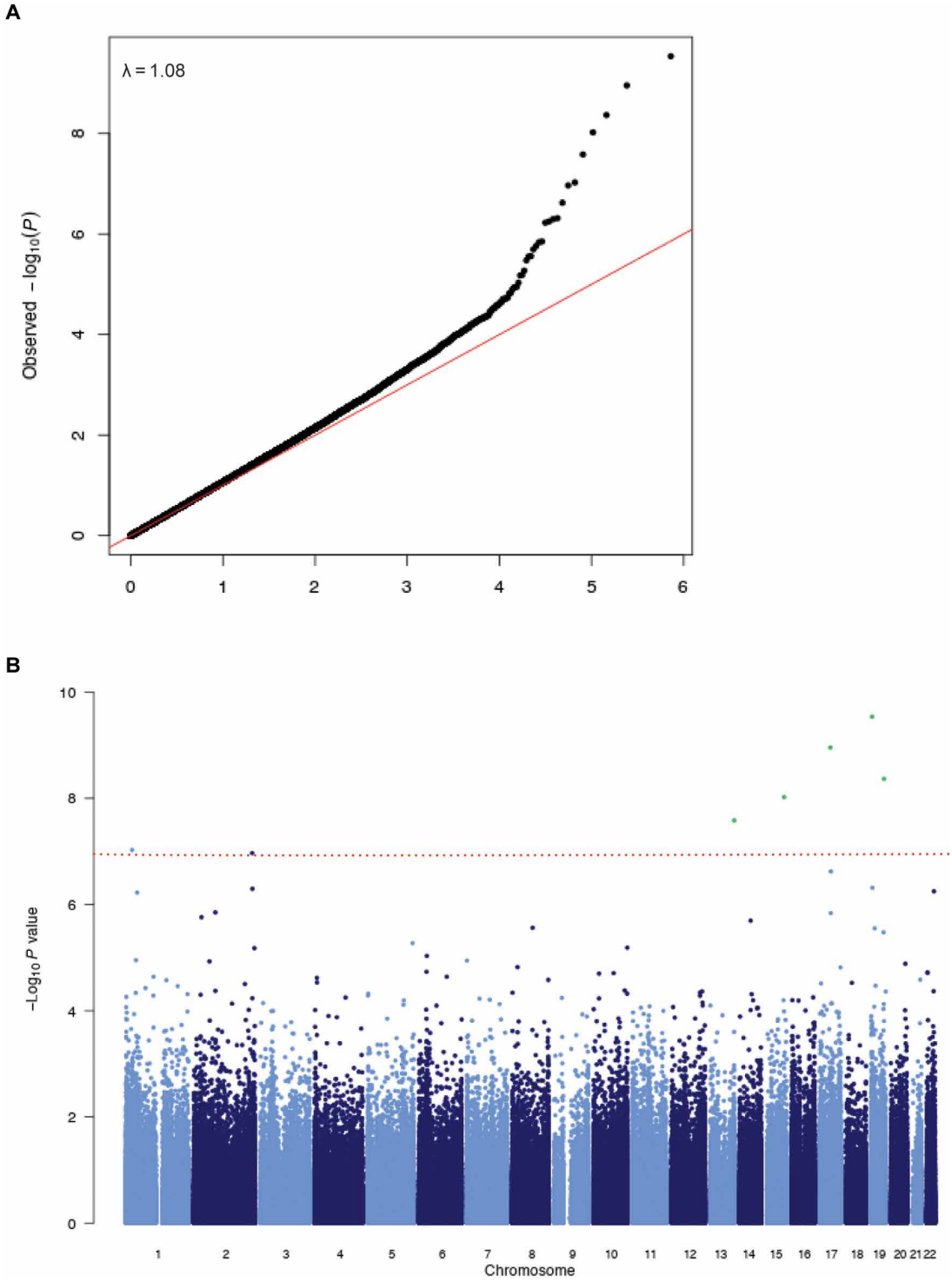

Fig. 1. Fixed-effects meta-analysis of associations between cesarean delivery and DNAm in cord blood collected at birth. $n = 9833$ individuals, 19 cohorts. Model 2, adjusted for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, gestational age, batch, and cell-type proportions. Individual cohorts may additionally have adjusted for ancestry and selection factors if applicable. (**A**) QQ plots for fixed-effects meta-analysis. (**B**) Manhattan plots for fixed-effects meta-analysis. Epigenome-wide significance was considered as $P < 1 \times 10^{-7}$ (dotted line).

Table 2. Fixed-effects meta-analysis of associations between cesarean deliveries and cord blood DNAm (statistically significant CpGs after Bonferroni correction (threshold $P < 1.0 \times 10^{-7}$) for multiple comparisons; model 2*).

CpG sites	Position	Gene	*n*	β	SE	*P* value	I^2	Effect direction†
cg19423175	Chr19:4094325	*MAP2K2*	9671	0.004	0.0006	2.91×10^{-10}	58%, $P = 0.001$	+++−−+−−+?+++++++++
cg13917614	Chr17:41973641	*CNP*	9833	0.007	0.001	1.11×10^{-9}	37%, $P = 0.06$	−−+−−+−−++++++++++
cg01500140	Chr19:51387331	*LIM2*	9833	0.005	0.0008	4.31×10^{-9}	39%, $P = 0.05$	++-+++++++++++++++
cg18247172	Chr15:90827002	*BLM* (nearest)	9833	0.005	0.0009	9.51×10^{-9}	47%, $P = 0.01$	−−++−+−−++++++++++
cg22348356	Chr13:114125748	*RASA3*	9833	0.005	0.0009	2.62×10^{-8}	64%, $P < 0.001$	+−−+−++−++++++++++
cg20674490	Chr1:2491440	*RUNX3*	9833	0.006	0.001	9.42×10^{-8}	62%, $P < 0.001$	+++−−++−++++++++++

*Model adjusted for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, gestational age, batch, and cell type. Individual cohorts may additionally have adjusted for ancestry and selection factors if applicable. †Effects were presented in the order of ALSPAC, BIB, CCLS 1, CCLS 2, DCHS, EAGeR, EDEN, GenR, INMA, IOW-F2, LiNA, Moba1, Moba2, Moba4, Moba8, MOCEH, Piccolipiù, POSEIDON, and Project Viva. "?" indicates missing probe information in a cohort.

(Fig. 3). In childhood blood, however, cesarean delivery was not associated with differences in estimated cell-type proportions (Fig. 4).

Meta-analyses for the associations between DNAm at birth and cell-type proportions in childhood

As presented in table S5, we investigated whether DNAm levels in cord blood at the six previously identified CpG sites (Table 2) were longitudinally associated with blood cell-type proportions in childhood. Higher cord blood DNAm at cg01500140 was associated with a −4.5% lower NK cells (95% CI: −8.9%, −1.1%; $P = 0.02$) and 14.3% higher granulocytes (95% CI: 2.2%, 25.8%; $P = 0.02$) in childhood, but not other cell types. In addition, higher cord blood DNAm at cg19423175 was associated with 4.1% higher monocytes (95% CI: 0.1%, 7.9%; $P = 0.03$) in childhood, but not other cell-type proportions. We did not observe that cord blood DNAm at other cesarean-related CpGs (cg13917614, cg18247172, cg20674490, and cg22348356) was associated with any blood cell-type proportions in childhood.

DISCUSSION

In this large EWAS meta-analysis in the PACE Consortium across geographically diverse cohorts, we identified differentially methylated DNA sites by delivery mode in offspring blood cells at birth, but not in childhood (6 to 10 years). In cord blood, cesarean delivery was associated with a 0.4 to 0.7% higher DNAm at six CpGs, and the effect sizes remained stable even when excluding offspring born to a pregnancy affected by major complications. Half of these sites showed relatively high heterogeneity across cohorts, likely due to differences in the underlying populations, covariate structures, and, possibly, obstetrical practices. The limited and relatively small difference in DNAm found in our study suggests that DNAm at the CpG sites in blood cells (at birth or later) is unlikely to be strongly affected by delivery mode.

Differences in cord blood cell-type proportions related to delivery mode may explain the lower genomic inflation in cell-type–adjusted models and supported our choice to focus on interpretation of results from model 2 (fully adjusted for covariates and blood cell types). These findings have notable implications for future research involving blood cells DNAm, as they highlight the importance of accounting for cell-type composition when examining the effects of delivery mode on DNAm.

To date, this is the largest effort to investigate offspring DNAm in relation to delivery mode. Several EWAS analyses with smaller sample sizes have tried to characterize the impact of cesarean delivery on cord blood DNAm, yielding conflicting results. A small study conducted among Europeans ($n = 41$, Austria) found significantly higher methylation of *ELA2* and *IRF1* genes in a targeted screening of 96 genes involved in immune response when comparing elective cesarean to vaginal births (*16*). More recently, one study in China ($n = 120$) compared DNAm in cord blood identified five differentially methylated CpG sites in/near five genes (*PIK3CD*, *CXXC5*, *LTBR*, *SPI1*, and *SERPINB9*) that were related to the immune system, specifically the maturation, proliferation, and normal function of B cells (*10*). However, none of these studies adjusted for cell types. In addition, the inconsistencies in these results may be explained by heterogeneity in the study populations, definitions of cesarean delivery (only including elective or not), and adjustments (or lack thereof) for potential confounders (e.g., gestational age, socioeconomic factors, and prenatal exposures). In a prior candidate gene DNAm analysis in a small subset of Project Viva ($n = 96$ uncomplicated full-term pregnancies), we found an association between delivery mode and DNAm at the *IGF2* differentially methylated regions in cord blood (2.8% difference comparing cesarean versus vaginal delivery) (*17*). However, these findings were not replicated in the current large-scale meta-analysis.

In our cord blood analyses, after adjusting for several important sociodemographic, lifestyle, and perinatal confounders, and cell-type proportions, we found associations between cord blood DNAm and cesarean delivery at only six CpG sites. DNAm at some of these CpGs (e.g., cg19423175, cg13917614, and cg18247172) has been previously associated with gestational age (*18*), suggesting that these associations may be partially attributable to residual confounding related to this factor and/or medical indications for cesarean delivery (although we could not rule out the possibility that other studies' findings were confounded by cesarean delivery, as not all studies adjusted for cesarean delivery; table S3). Blood DNAm at cg19423175 (*MAP2K2*) has also been associated with adult smoking in cross-sectional analyses (*19*), suggesting that some of the identified CpG sites may be malleable by various pre- and postnatal stimuli. Of note, some of the (nearest) genes in which the DNAm differences were identified in our study—*BLM*, *RASA3*, and *MAP2K2*—have been associated with various autoimmune disorders, including

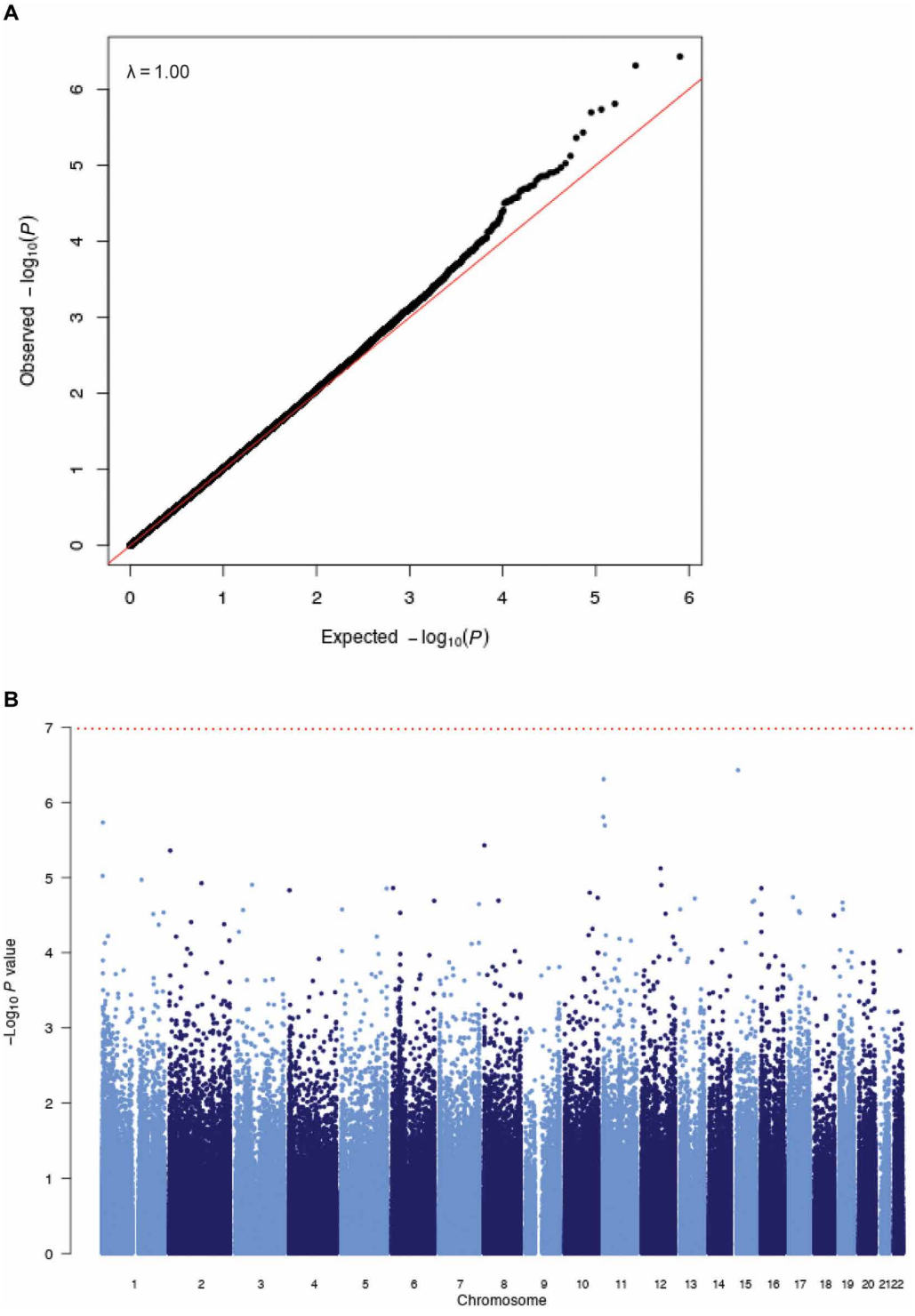

Fig. 2. Fixed-effects meta-analysis of associations between cesarean delivery and DNAm in blood collected in childhood. $n = 2429$ individuals, six cohorts. Model 2, adjusted for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, age at blood draw, batch, and cell-type proportions. Individual cohorts may additionally have adjusted for ancestry and selection factors if applicable. (**A**) QQ plots for fixed effects meta-analysis. (**B**) Manhattan plots for fixed-effects meta-analysis. Epigenome-wide significance was considered as $P < 1 \times 10^{-7}$ (dotted line).

Table 3. Fixed-effects meta-analysis of association between cesarean deliveries and childhood blood DNAm (model 2*) at the six CpG sites from the EWAS using cord blood DNAm that reached statistical significance after Bonferroni correction for multiple comparisons.

CpG sites	Position	Gene	*n*	β	SE	*P* value	I^2	Effect direction†
cg19423175	Chr19:4094325	*MAP2K2*	2099	0.002	0.001	0.20	0%, $P = 0.96$	+++?++
cg13917614	Chr17:41973641	*CNP*	2429	0.002	0.003	0.56	0%, $P = 0.92$	+−+−++
cg01500140	Chr19:51387331	*LIM2*	2429	0.002	0.002	0.17	0%, $P = 0.68$	+++−−−
cg18247172	Chr15:90827002	*BLM* (nearest)	2429	0.002	0.002	0.23	9%, $P = 0.36$	++++−+
cg22348356	Chr13:114125748	*RASA3*	2429	−0.001	0.002	0.46	0%, $P = 0.65$	−−+−−+
cg20674490	Chr1:2491440	*RUNX3*	2429	0.003	0.005	0.55	26%, $P = 0.24$	+−−+−+

*Model adjusted for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, age at blood draw, batch, and cell type. Individual cohorts may additionally adjust for ancestry and selection factors if applicable. †Effects were presented in the order of ALSPAC, EDEN, GenR, IOW-F1, PIAMA, and Project Viva. "?" indicates missing probe information in a cohort.

rheumatoid arthritis and inflammatory bowel disease (*20*, *21*). This aligns with prior evidence from epidemiological and genomic analyses that cesarean delivery may influence offspring health through immune dysregulation (*5*, *10*, *11*, *16*).

Our study has several notable strengths. This extensive meta-analysis includes samples from 21 independent cohorts with diverse regional and ethnic backgrounds, improving the generalizability of our results. The large sample size provides sufficient power to detect even small DNAm variations associated with cesarean delivery, both at birth and in childhood. This allows for the possibility of external replication/verification of loci that have been previously reported in smaller EWAS analyses, although we did not confirm any such sites in our analysis.

Our study has several limitations. First, we did not have detailed information regarding the circumstances, leading to cesarean delivery. However, results were similar in analyses excluding pregnancies affected by pregnancy complications that are common indications for cesarean delivery. Nevertheless, we cannot rule out the possibility that certain DNAm patterns might affect placental function or increase the risk of intra-amniotic infection, potentially influencing the medical decision to perform a cesarean delivery. Second, we measured DNAm in blood, which may not be generalizable to other tissues. Third, the probes assayed by the Illumina arrays used covered only a small fraction of the human genome, leaving a large number of CpGs unexplored. Moreover, DNAm is only one of many epigenetic modifications. Research on histone modification and post-transcriptional regulation may be needed to further investigate the role of epigenetic modifications in cesarean deliveries. Fourth, our analysis did not explicitly study the mediation effect of DNAm on the association between delivery mode and offspring health phenotypes. Despite the lack of persistence into childhood, these early DNAm changes may still have long-term effects on organ structure and development, including the immune system. However, we did not observe strong signals for such effects in the exploratory meta-analysis of the associations between the DNAm of the six significant CpGs in cord blood and blood cell-type proportions in childhood. Fifth, although we used standardized, prespecified protocol in all the cohorts, it is possible that the different bioinformatic pipeline across cohorts may have biased the meta-analysis results toward the null, but prior studies have shown minimal impact (*22*). Last, our results may be subject to residual confounding by ethnicity, socioeconomic status, and maternal health factors. In addition, because the study participants were mostly white, future studies in diverse populations are needed to replicate our findings.

In summary, in this extensive consortium-based meta-analysis, we identified only six CpG sites with DNAm differences in cord blood associated with cesarean delivery. However, none of these DNAm differences, nor any others, were observed in blood samples collected during childhood. Further research is warranted to elucidate the molecular influence of cesarean delivery on offspring health outcomes, potentially including more diverse populations and considering the collection of tissue types other than blood cells while acknowledging that this is highly challenging in otherwise healthy newborn populations.

MATERIALS AND METHODS

Study population

We invited cohorts within the PACE Consortium, with DNAm measured in cord blood and/or childhood blood samples (with a mean age of 6 to 10 years at blood collection) and information on delivery mode and relevant covariates to contribute data to this meta-analysis. The participating cohorts comprised 21 independent longitudinal cohorts from 10 countries (Italy, France, Germany, Norway, Spain, South Africa, South Korea, The Netherlands, the United States, and the United Kingdom) (Table 1). These cohorts included the Avon Longitudinal Study of Parents and Children (ALSPAC); the BiB cohort; the California Childhood Cancer Study 1 (CCLS1); CCLS2; the DCHS; the Effects of Aspirin in Gestation and Reproduction (EAGeR) trial; the Etude des Déterminants pré et post natals du développement et de la santé de l'Enfant (EDEN) birth cohort; the Generation R (GenR) Study; Proyecto Infancia y Medio Ambiente (INMA); the IOW-F1; IOW-F2; the LiNA study; the Norwegian Mother, Father and Child Cohort study 1 (Moba1); Moba2; Moba4; Moba8; the MOCEH study; the Prevention and Incidence of Asthma and Mite Allergy (PIAMA) birth cohort; the Piccolipiù study; the Pre-, Peri-, and POstnatal Stress: Epigenetic Impact on DepressiON (POSEIDON) study; and Project Viva. Detailed methods for each cohort are provided in supplementary text. In the cord blood meta-analysis, we included 9833 newborns from 19 cohorts. In the childhood blood meta-analysis, we included 2429 children from six cohorts (Table 1). Four cohorts (ALSPAC, EDEN, GenR, and Project Viva) have both cord blood and childhood blood measures from the same cohort of children. All cohorts acquired

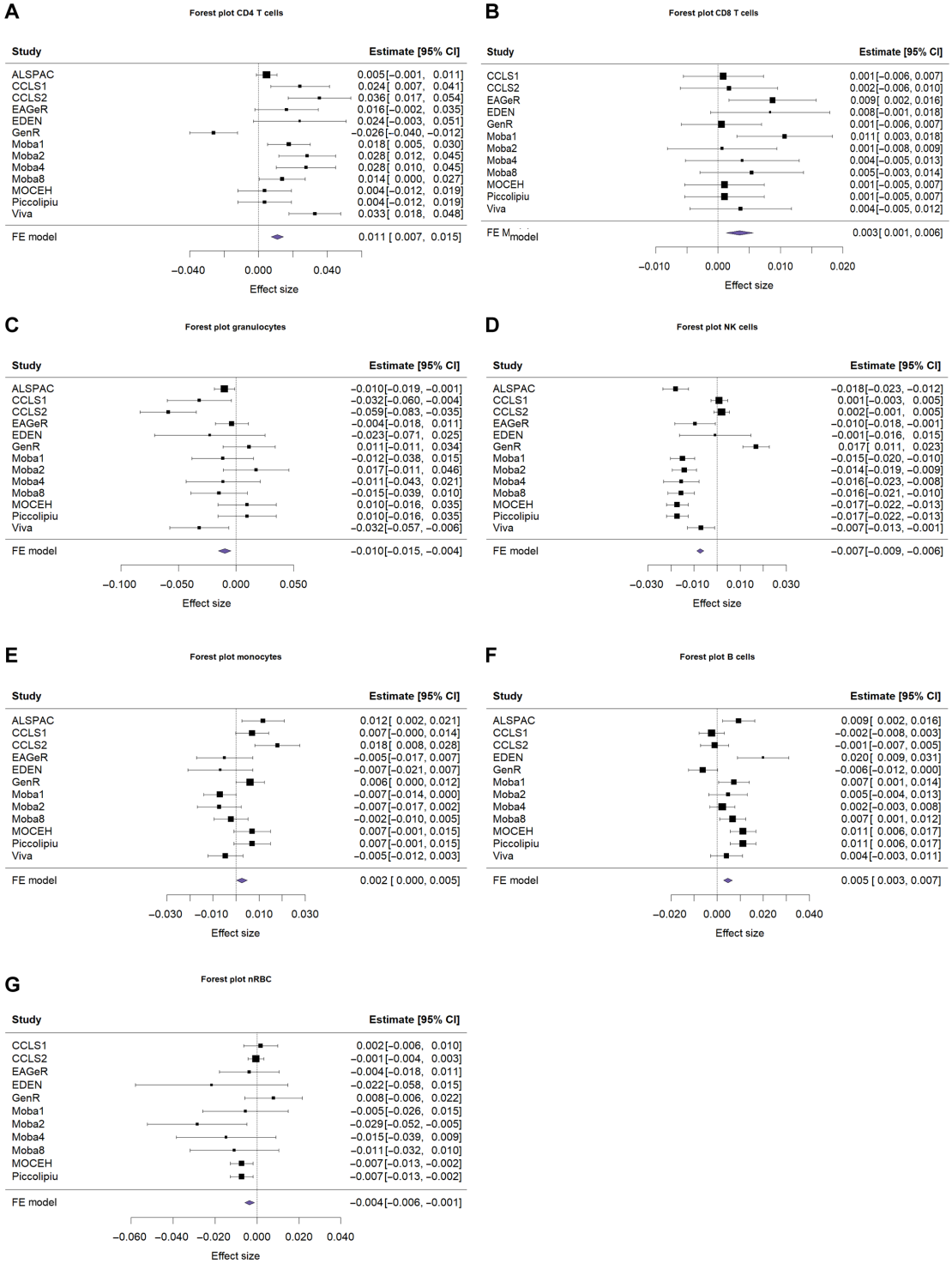

Fig. 3. Fixed-effects meta-analysis of associations between cesarean delivery and cell-type proportions in cord blood collected at birth. $n = 7616$ individuals, 13 cohorts. FE model, fixed-effects model; nRBC, nucleated red blood cells. Models adjusted for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, gestational age, and batch. Individual cohorts may additionally have adjusted for ancestry and selection factors if applicable. (**A**) CD4 T cells. (**B**) CD8 T cells. (**C**) Granulocytes. (**D**) NK cells. (**E**) Monocytes. (**F**) B cells. (**G**) nucleated red blood cells. Note: ALSPAC was excluded from the CD8 T cells analysis because of low estimated proportions. ALSPAC was excluded from the nucleated red blood cells analysis because these data were not available. Moba4 was excluded from the monocytes analysis because these data were not available.

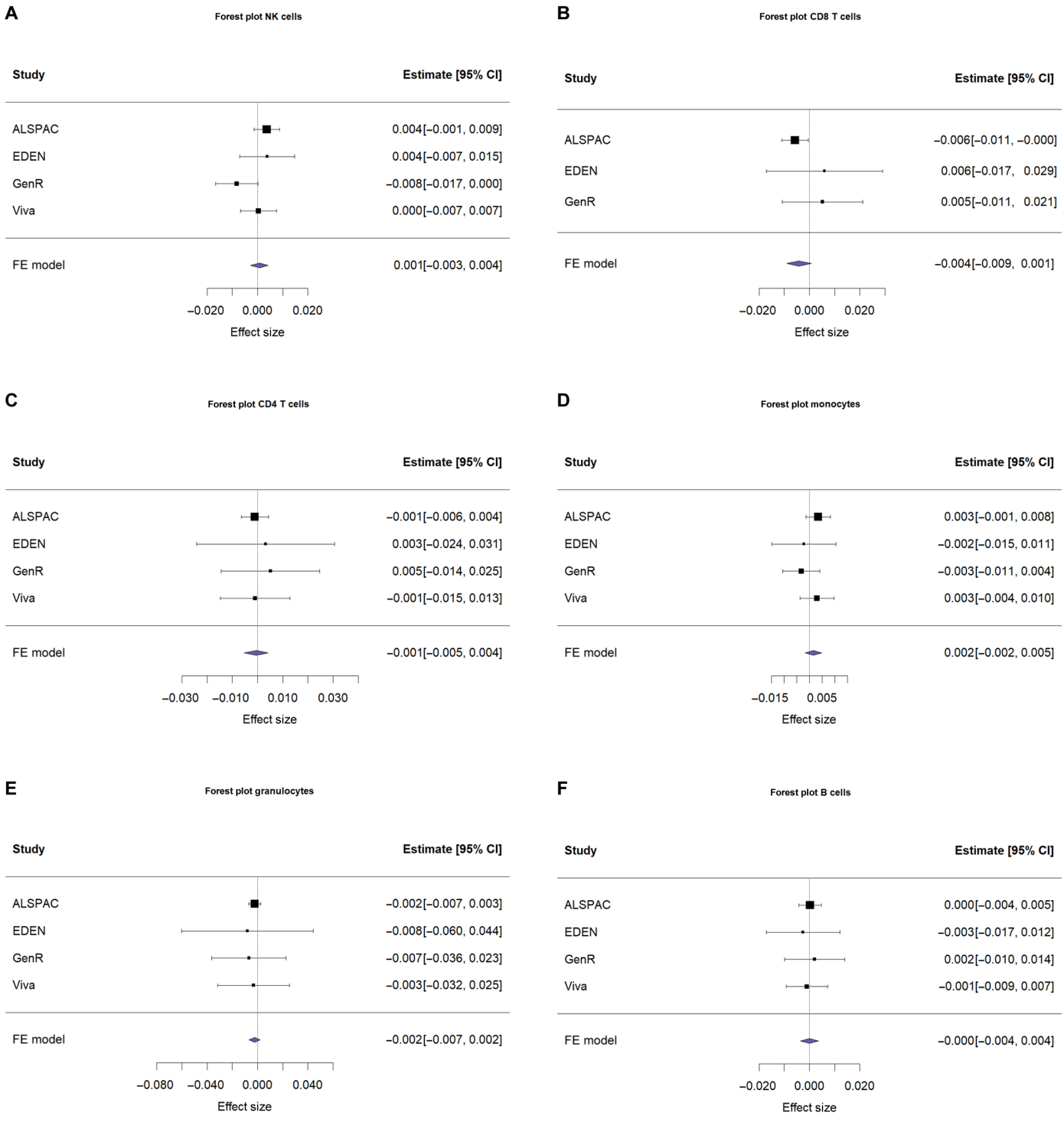

Fig. 4. Fixed-effects meta-analysis of associations between cesarean delivery and cell-type proportions in blood collected in childhood. $n = 1897$ individuals, four cohorts. FE model, fixed effect model. Models adjusted for maternal smoking, socioeconomic status, parity, prepregnancy BMI, offspring sex, birth weight, age at blood draw, and batch. Individual cohorts may additionally have adjusted for ancestry and selection factors if applicable. (**A**) NK cells. (**B**) CD8 T cells. (**C**) CD4 T cells. (**D**) Monocytes. (**E**) Granulocytes. (**F**) B cells. Note: Project Viva was excluded from the CD8 T cells analysis because these data were not available.

ethics approval and participants' informed consent (supplementary text).

Delivery mode assessment

Mode of delivery was ascertained via review of medical records (n = 8464, 68.8%) or self-report on questionnaires (n = 3832, 31.2%) (supplementary text). Cesarean delivery included both planned and unplanned. Vaginal delivery was chosen as the reference category. Offspring from multiple births (twins and triplets+) and preterm births (gestational age <37 weeks at delivery) were excluded. When non-twin siblings were available in the same dataset, only one child of each sibling set was included using methods deemed appropriate by each cohort.

To distinguish between effects of cesarean delivery itself and those from obstetrical/medical indications that commonly lead to cesarean delivery (*23*), in a sensitivity analysis, we excluded all pregnancies that were complicated with gestational diabetes, gestational hypertension, preeclampsia, or fetal macrosomia (birthweight > 4000 g). This analysis was restricted to cohorts with relevant information (table S1) and the number (%) of each of the major complications are shown in table S6. The pooled analytical sample included n = 6371 samples [after excluding n = 2170 (25.4%) pregnancy complications, 15 cohorts] for cord blood analysis and n = 1476 samples [after excluding n = 623 (29.6%) pregnancy complications, 5 cohorts] for childhood blood analysis.

Methylation measurements

Cord blood was collected at delivery, and childhood blood was collected at an average age between 6 and 10 years (Table 1). DNA was isolated according to standard protocols. Cohorts assayed DNAm from blood using the Infinium HumanMethylation450 (450K) or MethylationEPIC (EPIC) BeadChip array (Illumina, San Diego, CA) or both (Table 1). On the basis of PACE consortium standard procedures, each cohort conducted its own quality control (e.g., exclusion based on Illumina's detection P value, failed bisulfite conversion, and sex mismatches) and normalization of DNAm data (supplementary text).

Cohort-specific statistical analyses

Analysts from individual cohorts performed independent EWAS analyses using the same predetermined analytical approach and contributed results files to the primary analyst for meta- analysis. Robust linear regression modeling [rlm() function in R] was used for each CpG site individually. Cesarean delivery was the independent variable, and the DNAm level (β value, for which 0 indicated no methylation and 1 indicated complete methylation) at each CpG was the dependent variable. Model 1 adjusted for offspring biological sex (female and male), maternal smoking (preferred categorization: no smoking, quit early in pregnancy, and sustained smoking across pregnancy; but a binary categorization of any versus no smoking was used in 14 cohorts), maternal prepregnancy BMI (continuous, kilograms per square meter), socioeconomic status (defined by each cohort), parity (nulliparous and parous), gestational age (continuous, weeks, replaced with age at blood draw in years for childhood blood analysis), birth weight (continuous, grams), and batch variables (when applicable). For some cohorts, batch correction (e.g., using COMBAT) (*24*) was applied before analysis (CCLS1, CCLS2, DCHS, EAGeR, INMA, IOW-F1, IOW-F2, and Project Viva), while, for the others, batch variables were included as covariates in the regression analysis. Optionally, cohorts could include race and/or ethnicity and/or selection factors (e.g., factors associated with being selected to a case-control study), if relevant (supplementary text). To account for cell-type heterogeneity, in model 2, all cohorts further adjusted for blood cell types (estimated cell proportions of CD8T, CD4 T, NK cells, B cells, monocytes, granulocytes, and nucleated red blood cells) (*25*–*28*). On the basis of our conceptual model and due to the known variability in DNAm associated with cell-type components (*29*), we chose model 2 as our main models for the interpretation of findings.

Cohort-specific details for covariate assessment are described in supplementary text. EWAS results from each cohort were evaluated with the R package QCEWAS to ensure that results were comparable across participating cohorts (*30*).

Meta-analyses of EWAS

Before meta-analysis, we filtered out all probes that (i) were control and non-CpG probes, (ii) in sex chromosomes, (iii) have hybridizing problems, (iv) are affected by the presence of SNPs, or (v) are known to yield different results for the 450K and EPIC arrays in blood (*31*–*33*). We conducted two separate meta-analyses, for blood samples collected at birth and in childhood, respectively. As our primary modeling approach, we used an inverse variance weighting fixed-effects meta-analysis because of the homogeneous research protocol (e.g., study design, definition of exposure and outcome, and statistical analysis) across participating cohorts (*34*) (using R package EASIER) (*33*). This approach was also recommended for analyses in the genetic field (*35*). The prespecified missingness threshold was 80%, meaning the code executes a meta-analysis only among CpGs present in 80% of the cohorts. Consequently, the main analysis included only probes common to EPIC and 450K arrays, leaving 363,015 probes in cord blood and 400,681 probes in childhood blood. In a secondary analysis, we also explored additional probes among six cohorts using the EPIC array (764,471 probes). All reported P values were two sided, and we corrected for multiple testing using Bonferroni adjustment (for ease of interpretation, a common $P < 1.0 \times 10^{-7}$ threshold was used for all analyses).

Heterogeneity was assessed using the I^2 statistics for each CpG site, and statistical significance for heterogeneity was assessed with the Cochran's Q test. As post hoc analyses, we conducted leave-one-out analyses (i.e., excluding one cohort at a time) for all CpGs that reached genome-wide significance. We compared the effect estimates and SEs with the estimates for our primary model to evaluate whether any meaningful heterogeneity may have influenced our results. The study-level quality control and meta-analyses were performed by two independent study groups. All analyses were performed using R (version 4.2.2).

We annotated the genes for differentially methylated CpGs that reached the Bonferroni-adjusted threshold from Illumina's annotation. If annotation was unavailable, then the University of California, Santa Cruz Genome Browser was referenced to identify the nearest gene within 10 Mbp. In addition, CpGs demonstrating genome-wide significance after Bonferroni correction were cross-referenced in the EWAS catalog (*13*)(www.ewascatalog.org; accessed 3 January 2024) to compare our findings with previously reported associations of delivery mode-related CpGs. Last, we examined the CpGs in relation to 13.6 million blood autosomal cis-eQTMs data in children (mean age, 8 years) from the Human Early Life Exposome project (*14*, *15*). For each CpG DNAm site, we selected the CpG-gene pair with the strongest cis-eQTM association

after correcting for multiple testing at the CpG level, after cell-type adjustment.

Meta-analyses for cell-type proportions

As an exploratory analysis that was prespecified in the research protocol (i.e., cesarean delivery modeled as the independent variable and each cell-type modeled as dependent variable in separate models, with adjustment for all covariates), we conducted another set of fixed-effects meta-analysis for the association between cesarean delivery and estimated blood cell-type proportions (cord blood: 7616 samples, 13 cohorts; childhood blood: 1897 samples, 4 cohorts). *P* values were not corrected for multiple comparisons in this exploratory analysis.

Post hoc longitudinal analysis for the association between DNAm at birth and cell-type proportions in childhood

Given the prior epidemiologic reports that C-section is associated with many immune-related outcomes (e.g. allergy and asthma) and that most of prior studies pointed at loci involved in the immune system (*2*, *5*, *9*–*11*), we conducted a post hoc longitudinal analysis to investigate whether significant CpGs identified in the cord blood EWAS meta-analysis are associated with cell-type proportions in childhood, potentially reflecting long-term health effects. This analysis was feasible for four cohorts—ALSPAC, EDEN, GenR, and Project Viva—where DNAm levels were measured in both cord blood at birth and peripheral blood in childhood for the same cohort of children. We used a fixed-effects meta-analysis to examine associations between DNAm levels at birth and estimated childhood blood cell proportions. We selected blood cell proportions as the outcome due to their potential to serve as proxies for immune-related conditions. *P* values were not adjusted for multiple comparisons, as this analysis was exploratory in nature.

Supplementary Materials

This PDF file includes:

Supplementary text
Figs. S1 to S12
Tables S1 to S6
References

REFERENCES AND NOTES

1. World Health Organization (2021); www.who.int/news/item/16-06-2021-caesarean-section-rates-continue-to-rise-amid-growing-inequalities-in-access.
2. J. Sandall, R. M. Tribe, L. Avery, G. Mola, G. H. Visser, C. S. Homer, D. Gibbons, N. M. Kelly, H. P. Kennedy, H. Kidanto, P. Taylor, M. Temmerman, Short-term and long-term effects of caesarean section on the health of women and children. *Lancet* **392**, 1349–1357 (2018).
3. J. E. Chavarro, N. Martin-Calvo, C. Yuan, M. Arvizu, J. W. Rich-Edwards, K. B. Michels, Q. Sun, Association of birth by cesarean delivery with obesity and type 2 diabetes among adult women. *JAMA Netw. Open* **3**, e202605 (2020).
4. C. Yuan, A. J. Gaskins, A. I. Blaine, C. Zhang, M. W. Gillman, S. A. Missmer, A. E. Field, J. E. Chavarro, Association between cesarean birth and risk of obesity in offspring in childhood, adolescence, and early adulthood. *JAMA Pediatr.* **170**, e162385 (2016).
5. C. E. Cho, M. Norman, Cesarean section and development of the immune system in the offspring. *Am. J. Obstet. Gynecol.* **208**, 249–254 (2013).
6. L. A. Williams, M. Richardson, L. G. Spector, E. L. Marcotte, Cesarean section is associated with an increased risk of acute lymphoblastic leukemia and hepatoblastoma in children from minnesota. *Cancer Epidemiol. Biomarkers Prev.* **30**, 736–742 (2021).
7. Y. Cao, L. H. Nguyen, S. Tica, E. Otegbeye, X. Zong, B. Roelstraete, A. T. Chan, B. B. Warner, O. Stephansson, J. F. Ludvigsson, Evaluation of birth by cesarean delivery and development of early-onset colorectal cancer. *JAMA Netw. Open* **6**, e2310316 (2023).
8. T. Zhang, A. Sidorchuk, L. Sevilla-Cermeno, A. Vilaplana-Perez, Z. Chang, H. Larsson, D. Mataix-Cols, L. Fernandez de la Cruz, Association of cesarean delivery with risk of neurodevelopmental and psychiatric disorders in the offspring: A systematic review and meta-analysis. *JAMA Netw. Open* **2**, e1910236 (2019).
9. M. Almgren, T. Schlinzig, D. Gomez-Cabrero, A. Gunnar, M. Sundin, S. Johansson, M. Norman, T. J. Ekström, Cesarean delivery and hematopoietic stem cell epigenetics in the newborn infant: Implications for future health? *Am. J. Obstet. Gynecol.* **211**, 502.e1–502.e8 (2014).
10. Q. Chen, Y. Ming, Y. Gan, L. Huang, Y. Zhao, X. Wang, Y. Liu, J. Zhang, The impact of cesarean delivery on infant DNA methylation. *BMC Pregnancy Childbirth* **21**, 265 (2021).
11. T. Schlinzig, S. Johansson, A. Gunnar, T. J. Ekstrom, M. Norman, Epigenetic modulation at birth–altered DNA-methylation in white blood cells after Caesarean section. *Acta Paediatr.* **98**, 1096–1099 (2009).
12. S. Virani, D. C. Dolinoy, S. Halubai, T. R. Jones, S. E. Domino, L. S. Rozek, M. S. Nahar, V. Padmanabhan, Delivery type not associated with global methylation at birth. *Clin. Epigenetics* **4**, 8 (2012).
13. T. Battram, P. Yousefi, G. Crawford, C. Prince, M. Sheikhali Babaei, G. Sharp, C. Hatcher, M. J. Vega-Salas, S. Khodabakhsh, O. Whitehurst, R. Langdon, L. Mahoney, H. R. Elliott, G. Mancano, M. A. Lee, S. H. Watkins, A. C. Lay, G. Hemani, T. R. Gaunt, C. L. Relton, J. R. Staley, M. Suderman, The EWAS Catalog: A database of epigenome-wide association studies. *Wellcome Open Res.* **7**, 41 (2022).
14. L. Maitre, J. de Bont, M. Casas, O. Robinson, G. M. Aasvang, L. Agier, S. Andrušaitytė, F. Ballester, X. Basagaña, E. Borràs, C. Brochot, M. Bustamante, A. Carracedo, M. de Castro, A. Dedele, D. Donaire-Gonzalez, X. Estivill, J. Evandt, S. Fossati, L. Giorgis-Allemand, J. R. Gonzalez, B. Granum, R. Grazuleviciene, K. Bjerve Gutzkow, L. Smastuen Haug, C. Hernandez-Ferrer, B. Heude, J. Ibarluzea, J. Julvez, M. Karachaliou, H. C. Keun, N. Hjertager Krog, C. E. Lau, V. Leventakou, S. Lyon-Caen, C. Manzano, D. Mason, R. McEachan, H. M. Meltzer, I. Petraviciene, J. Quentin, T. Roumeliotaki, E. Sabido, P. J. Saulnier, A. P. Siskos, V. Siroux, J. Sunyer, I. Tamayo, J. Urquiza, M. Vafeiadi, D. van Gent, M. Vives-Usano, D. Waiblinger, C. Warembourg, L. Chatzi, M. Coen, P. van den Hazel, M. J. Nieuwenhuijsen, R. Slama, C. Thomsen, J. Wright, M. Vrijheid, Human Early Life Exposome (HELIX) study: A European population-based exposome cohort. *BMJ Open* **8**, e021311 (2018).
15. C. Ruiz-Arenas, C. Hernandez-Ferrer, M. Vives-Usano, S. Mari, I. Quintela, D. Mason, S. Cadiou, M. Casas, S. Andrusaityte, K. B. Gutzkow, M. Vafeiadi, J. Wright, J. Lepeule, R. Grazuleviciene, L. Chatzi, A. Carracedo, X. Estivill, E. Marti, G. Escaramis, M. Vrijheid, J. R. Gonzalez, M. Bustamante, Identification of autosomal cis expression quantitative trait methylation (cis eQTMs) in children's blood. *eLife* **11**, e65310 (2022).
16. M. B. Franz, M. Poterauer, M. Elhenicky, S. Stary, P. Birner, U. Vinatzer, P. Husslein, B. Streubel, H. Husslein, Global and single gene DNA methylation in umbilical cord blood cells after elective caesarean: A pilot study. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **179**, 121–124 (2014).
17. Y. H. Chiu, R. P. Fadadu, A. J. Gaskins, S. L. Rifas-Shiman, H. E. Laue, K. H. Moley, M. F. Hivert, A. Baccarelli, E. Oken, J. E. Chavarro, A. Cardenas, Dietary fat intake during early pregnancy is associated with cord blood DNA methylation at IGF2 and H19 genes in newborns. *Environ. Mol. Mutagen.* **62**, 388–398 (2021).
18. J. Bohlin, S. E. Haberg, P. Magnus, S. E. Reese, H. K. Gjessing, M. C. Magnus, C. L. Parr, C. M. Page, S. J. London, W. Nystad, Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol.* **17**, 207 (2016).
19. S. Sikdar, R. Joehanes, B. R. Joubert, C. J. Xu, M. Vives-Usano, F. I. Rezwan, J. F. Felix, J. M. Ward, W. Guan, R. C. Richmond, J. A. Brody, L. K. Kupers, N. Baiz, S. E. Haberg, J. A. Smith, S. E. Reese, S. Aslibekyan, C. Hoyo, R. Dhingra, C. A. Markunas, T. Xu, L. M. Reynolds, A. C. Just, P. R. Mandaviya, A. Ghantous, B. D. Bennett, T. Wang, T. B. Consortium, K. M. Bakulski, E. Melen, S. Zhao, J. Jin, Z. Herceg, J. V. Meurs, J. A. Taylor, A. A. Baccarelli, S. K. Murphy, Y. Liu, M. C. Munthe-Kaas, I. J. Deary, W. Nystad, M. Waldenberger, I. Annesi-Maesano, K. Conneely, V. W. Jaddoe, D. Arnett, H. Snieder, S. L. Kardia, C. L. Relton, K. K. Ong, S. Ewart, H. Moreno-Macias, I. Romieu, N. Sotoodehnia, M. Fornage, A. Motsinger-Reif, G. H. Koppelman, M. Bustamante, D. Levy, S. J. London, Comparison of smoking-related DNA methylation between newborns from prenatal exposure and adults from personal smoking. *Epigenomics* **11**, 1487–1500 (2019).
20. Y. Liu, M. J. Aryee, L. Padyukov, M. D. Fallin, E. Hesselberg, A. Runarsson, L. Reinius, N. Acevedo, M. Taub, M. Ronninger, K. Shchetynsky, A. Scheynius, J. Kere, L. Alfredsson, L. Klareskog, T. J. Ekstrom, A. P. Feinberg, Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nat. Biotechnol.* **31**, 142–147 (2013).
21. N. T. Ventham, N. A. Kennedy, A. T. Adams, R. Kalla, S. Heath, K. R. O'Leary, H. Drummond, IBD CHARACTER consortium, D. C. Wilson, I. G. Gut, E. R. Nimmo, J. Satsangi, Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat. Commun.* **7**, 13507 (2016).
22. M. C. Wu, B. R. Joubert, P. F. Kuan, S. E. Haberg, W. Nystad, S. D. Peddada, S. J. London, A systematic assessment of normalization approaches for the Infinium 450K methylation platform. *Epigenetics* **9**, 318–329 (2014).
23. C. D. Lynch, J. D. Iams, Diseases resulting from suboptimal immune function in offspring: Is cesarean delivery itself really to blame? *Am. J. Obstet. Gynecol.* **208**, 247–248 (2013).

24. J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, J. D. Storey, The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
25. L. A. Salas, D. C. Koestler, R. A. Butler, H. M. Hansen, J. K. Wiencke, K. T. Kelsey, B. C. Christensen, An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol.* **19**, 64 (2018).
26. K. Gervin, L. A. Salas, K. M. Bakulski, M. C. van Zelm, D. C. Koestler, J. K. Wiencke, L. Duijts, H. A. Moll, K. T. Kelsey, M. S. Kobor, R. Lyle, B. C. Christensen, J. F. Felix, M. J. Jones, Systematic evaluation and validation of reference and library selection methods for deconvolution of cord blood DNA methylation data. *Clin. Epigenetics* **11**, 125 (2019).
27. E. A. Houseman, W. P. Accomando, D. C. Koestler, B. C. Christensen, C. J. Marsit, H. H. Nelson, J. K. Wiencke, K. T. Kelsey, DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
28. L. E. Reinius, N. Acevedo, M. Joerink, G. Pershagen, S. E. Dahlen, D. Greco, C. Soderhall, A. Scheynius, J. Kere, Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility. *PLOS ONE* **7**, e41361 (2012).
29. A. E. Jaffe, R. A. Irizarry, Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol.* **15**, R31 (2014).
30. P. J. Van der Most, L. K. Kupers, H. Snieder, I. Nolte, QCEWAS: Automated quality control of results of epigenome-wide association studies. *Bioinformatics* **33**, 1243–1245 (2017).
31. W. Zhou, P. W. Laird, H. Shen, Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res.* **45**, e22 (2017).
32. O. Solomon, J. MacIsaac, H. Quach, G. Tindula, M. S. Kobor, K. Huen, M. J. Meaney, B. Eskenazi, L. F. Barcellos, N. Holland, Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. *Epigenetics* **13**, 655–664 (2018).
33. D . Pelegrí, M. Cosín, M. Bustamante, J. R. Gonzalez, isglobal-brge/EASIER:EwAS: quality control, meta-analysIs and EnRichment version 0.1.2.8 from GitHub (2023); https://rdrr.io/github/isglobal-brge/EASIER/, vol. 2023.
34. R. Magi, A. P. Morris, GWAMA: Software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288 (2010).
35. K. Rice, J. P. T. Higgins, T. Lumley, A re-evaluation of fixed effect(s) meta-analysis. *J. R. Stat. Soc. Ser. A Stat. Soc.* **181**, 205–227 (2018).
36. K. Northstone, Y. Ben Shlomo, A. Teyhan, A. Hill, A. Groom, M. Mumme, N. Timpson, J. Golding, The Avon Longitudinal Study of Parents and children ALSPAC G0 Partners: A cohort profile [version 2; peer review: 1 approved]. *Wellcome Open Res.* **8**, 37 (2023).
37. A. Fraser, C. Macdonald-Wallis, K. Tilling, A. Boyd, J. Golding, G. D. Smith, J. Henderson, J. Macleod, L. Molloy, A. Ness, S. Ring, S. M. Nelson, D. A. Lawlor, Cohort profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
38. A. Boyd, J. Golding, J. Macleod, D. A. Lawlor, A. Fraser, J. Henderson, L. Molloy, A. Ness, S. Ring, G. Davey Smith, Cohort profile: The 'children of the 90s'—The index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
39. C. L. Relton, T. Gaunt, W. McArdle, K. Ho, A. Duggirala, H. Shihab, G. Woodward, O. Lyttleton, D. M. Evans, W. Reik, Y. L. Paul, G. Ficz, S. E. Ozanne, A. Wipat, K. Flanagan, A. Lister, B. T. Heijmans, S. M. Ring, G. Davey Smith, Data resource profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *Int. J. Epidemiol.* **44**, 1181–1190 (2015).
40. J. L. Min, G. Hemani, G. Davey Smith, C. Relton, M. Suderman, Meffil: Efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* **34**, 3983–3989 (2018).
41. J. Wright, N. Small, P. Raynor, D. Tuffnell, R. Bhopal, N. Cameron, L. Fairley, D. A. Lawlor, R. Parslow, E. S. Petherick, K. E. Pickett, D. Waiblinger, J. West, Born in Bradford Scientific Collaborators Group, Cohort profile: The Born in Bradford multi-ethnic family cohort study. *Int. J. Epidemiol.* **42**, 978–991 (2013).
42. D. Farrar, L. Fairley, G. Santorelli, D. Tuffnell, T. A. Sheldon, J. Wright, L. van Overveld, D. A. Lawlor, Association between hyperglycaemia and adverse perinatal outcomes in south Asian and white British women: Analysis of data from the Born in Bradford cohort. *Lancet Diabetes Endocrinol.* **3**, 795–804 (2015).
43. J. P. Fortin, A. Labbe, M. Lemire, B. W. Zanke, T. J. Hudson, E. J. Fertig, C. M. Greenwood, K. D. Hansen, Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).
44. D. C. Koestler, M. J. Jones, J. Usset, B. C. Christensen, R. A. Butler, M. S. Kobor, J. K. Wiencke, K. T. Kelsey, Improving cell mixture deconvolution by identifying optimal DNA methylation libraries (IDOL). *BMC Bioinformatics* **17**, 120 (2016).
45. D. J. Stein, N. Koen, K. A. Donald, C. M. Adnams, S. Koopowitz, C. Lund, A. Marais, B. Myers, A. Roos, K. Sorsdahl, M. Stern, M. Tomlinson, C. van der Westhuizen, B. Vythilingum, L. Myer, W. Barnett, K. Brittain, H. J. Zar, Investigating the psychosocial determinants of child health in Africa: The Drakenstein Child Health Study. *J. Neurosci. Methods* **252**, 27–35 (2015).
46. L. Myer, D. J. Stein, A. Grimsrud, S. Seedat, D. R. Williams, Social determinants of psychological distress in a nationally-representative sample of South African adults. *Soc. Sci. Med.* **66**, 1828–1840 (2008).
47. S. Abrishamcar, J. Chen, D. Feil, A. Kilanowski, N. Koen, A. Vanker, C. J. Wedderburn, K. A. Donald, H. J. Zar, D. J. Stein, A. Huls, DNA methylation as a potential mediator of the association between prenatal tobacco and alcohol exposure and child neurodevelopment in a South African birth cohort. *Transl. Psychiatry* **12**, 418 (2022).
48. A. Huls, C. J. Wedderburn, N. A. Groenewold, N. Gladish, M. J. Jones, N. Koen, J. L. MacIsaac, D. T. S. Lin, K. E. Ramadori, M. P. Epstein, K. A. Donald, M. S. Kobor, H. J. Zar, D. J. Stein, Newborn differential DNA methylation and subcortical brain volumes as early signs of severe neurodevelopmental delay in a South African Birth Cohort Study. *World J. Biol. Psychiatry* **23**, 601–612 (2022).
49. M. J. Aryee, A. E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A. P. Feinberg, K. D. Hansen, R. A. Irizarry, Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
50. A. M. Morin, E. Gatev, L. M. McEwen, J. L. MacIsaac, D. T. S. Lin, N. Koen, D. Czamara, K. Raikkonen, H. J. Zar, K. Koenen, D. J. Stein, M. S. Kobor, M. J. Jones, Maternal blood contamination of collected cord blood can be identified using DNA methylation at three CpGs. *Clin. Epigenetics* **9**, 75 (2017).
51. M. E. Price, A. M. Cotton, L. L. Lam, P. Farre, E. Emberly, C. J. Brown, W. P. Robinson, M. S. Kobor, Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin* **6**, 4 (2013).
52. R. Pidsley, E. Zotenko, T. J. Peters, M. G. Lawrence, G. P. Risbridger, P. Molloy, S. Van Djik, B. Muhlhausler, C. Stirzaker, S. J. Clark, Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* **17**, 208 (2016).
53. E. F. Schisterman, R. M. Silver, L. L. Lesher, D. Faraggi, J. Wactawski-Wende, J. M. Townsend, A. M. Lynch, N. J. Perkins, S. L. Mumford, N. Galai, Preconception low-dose aspirin and pregnancy outcomes: Results from the EAGeR randomised trial. *Lancet* **384**, 29–36 (2014).
54. E. H. Yeung, W. Guan, S. L. Mumford, R. M. Silver, C. Zhang, M. Y. Tsai, E. F. Schisterman, Measured maternal prepregnancy anthropometry and newborn DNA methylation. *Epigenomics* **11**, 187–198 (2019).
55. N. Touleimat, J. Tost, Complete pipeline for Infinium Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. *Epigenomics* **4**, 325–341 (2012).
56. B. Heude, A. Forhan, R. Slama, L. Douhaud, S. Bedel, M. J. Saurel-Cubizolles, R. Hankard, O. Thiebaugeorges, M. De Agostini, I. Annesi-Maesano, M. Kaminski, M. A. Charles, EDEN mother-child cohort study group, Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int. J. Epidemiol.* **45**, 353–363 (2016).
57. Y. A. Chen, M. Lemire, S. Choufani, D. T. Butcher, D. Grafodatskaya, B. W. Zanke, S. Gallinger, T. J. Hudson, R. Weksberg, Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203–209 (2013).
58. R. Pidsley, C. C. Y. Wong, M. Volta, K. Lunnon, J. Mill, L. C. Schalkwyk, A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* **14**, 293 (2013).
59. M. N. Kooijman, C. J. Kruithof, C. M. van Duijn, L. Duijts, O. H. Franco, M. H. van IJzendoorn, J. C. de Jongste, C. C. W. Klaver, A. van der Lugt, J. P. Mackenbach, H. A. Moll, R. P. Peeters, H. Raat, E. H. H. M. Rings, F. Rivadeneira, M. P. van der Schroeff, E. A. P. Steegers, H. Tiemeier, A. G. Uitterlinden, F. C. Verhulst, E. Wolvius, J. F. Felix, V. W. V. Jaddoe, The Generation R Study: Design and cohort update 2017. *Eur. J. Epidemiol.* **31**, 1243–1264 (2016).
60. R. H. Mulder, A. Neumann, C. A. M. Cecil, E. Walton, L. C. Houtepen, A. J. Simpkin, J. Rijlaarsdam, B. T. Heijmans, T. R. Gaunt, J. F. Felix, V. W. V. Jaddoe, M. J. Bakermans-Kranenburg, H. Tiemeier, C. L. Relton, M. H. van IJzendoorn, M. Suderman, Epigenome-wide change and variation in DNA methylation in childhood: Trajectories from birth to late adolescence. *Hum. Mol. Genet.* **30**, 119–134 (2021).
61. G. S. Monasso, L. K. Kupers, V. W. V. Jaddoe, S. G. Heil, J. F. Felix, Associations of circulating folate, vitamin B12 and homocysteine concentrations in early pregnancy and cord blood with epigenetic gestational age: The Generation R Study. *Clin. Epigenetics* **13**, 95 (2021).
62. B. Lehne, A. W. Drong, M. Loh, W. Zhang, W. R. Scott, S. T. Tan, U. Afzal, J. Scott, M. R. Jarvelin, P. Elliott, M. I. McCarthy, J. S. Kooner, J. C. Chambers, A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol.* **16**, 37 (2015).
63. M. Guxens, F. Ballester, M. Espada, M. F. Fernandez, J. O. Grimalt, J. Ibarluzea, N. Olea, M. Rebagliato, A. Tardon, M. Torrent, J. Vioque, M. Vrijheid, J. Sunyer, INMA Project, Cohort

Profile: The INMA—INfancia y Medio Ambiente—(Environment and Childhood) Project. *Int. J. Epidemiol.* **41**, 930–940 (2012).
64. M. van Iterson, E. W. Tobi, R. C. Slieker, W. den Hollander, R. Luijk, P. E. Slagboom, B. T. Heijmans, MethylAid: Visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics* **30**, 3435–3437 (2014).
65. J. P. Fortin, E. Fertig, K. Hansen, shinyMethyl: Interactive quality control of Illumina 450k DNA methylation arrays in R. *F1000Res.* **3**, 175 (2014).
66. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
67. S. A. Miller, D. D. Dykes, H. F. Polesky, A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **16**, 1215 (1988).
68. M. Bibikova, J. B. Fan, GoldenGate assay for DNA methylation profiling. *Methods Mol. Biol.* **507**, 149–163 (2009).
69. P. Magnus, C. Birke, K. Vejrup, A. Haugan, E. Alsaker, A. K. Daltveit, M. Handal, M. Haugen, G. Hoiseth, G. P. Knudsen, L. Paltiel, P. Schreuder, K. Tambs, L. Vold, C. Stoltenberg, Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). *Int. J. Epidemiol.* **45**, 382–388 (2016).
70. K. S. Ronningen, L. Paltiel, H. M. Meltzer, R. Nordhagen, K. K. Lie, R. Hovengen, M. Haugen, W. Nystad, P. Magnus, J. A. Hoppin, The biobank of the Norwegian Mother and Child Cohort Study: A resource for the next 100 years. *Eur. J. Epidemiol.* **21**, 619–625 (2006).
71. B. R. Joubert, S. E. Haberg, R. M. Nilsen, X. Wang, S. E. Vollset, S. K. Murphy, Z. Huang, C. Hoyo, O. Midttun, L. A. Cupul-Uicab, P. M. Ueland, M. C. Wu, W. Nystad, D. A. Bell, S. D. Peddada, S. J. London, 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* **120**, 1425–1431 (2012).
72. B. R. Joubert, H. T. den Dekker, J. F. Felix, J. Bohlin, S. Ligthart, E. Beckett, H. Tiemeier, J. B. van Meurs, A. G. Uitterlinden, A. Hofman, S. E. Haberg, S. E. Reese, M. J. Peters, B. K. Andreassen, E. A. Steegers, R. M. Nilsen, S. E. Vollset, O. Midttun, P. M. Ueland, O. H. Franco, A. Dehghan, J. C. de Jongste, M. C. Wu, T. Wang, S. D. Peddada, V. W. Jaddoe, W. Nystad, L. Duijts, S. J. London, Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat. Commun.* **7**, 10577 (2016).
73. S. E. Haberg, C. M. Page, Y. Lee, H. E. Nustad, M. C. Magnus, K. L. Haftorn, E. O. Carlsen, W. R. P. Denault, J. Bohlin, A. Jugessur, P. Magnus, H. K. Gjessing, R. Lyle, DNA methylation in newborns conceived by assisted reproductive technology. *Nat. Commun.* **13**, 1896 (2022).
74. B.-M. Kim, M. Ha, H.-S. Park, B. E. Lee, Y. J. Kim, Y.-C. Hong, Y. Kim, N. Chang, Y.-M. Roh, B.-N. Kim, S.-y. Oh, E.-H. Ha, MOCEH Study Group, The Mothers and Children's Environmental Health (MOCEH) study. *Eur. J. Epidemiol.* **24**, 573–583 (2009).
75. D. J. Stekhoven, P. Buhlmann, MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2012).
76. J. Park, J. Kim, E. Kim, S. Won, W. J. Kim, Association between prenatal cadmium exposure and cord blood DNA methylation. *Environ. Res.* **212**, 113268 (2022).
77. A. H. Wijga, M. Kerkhof, U. Gehring, J. C. de Jongste, D. S. Postma, R. C. Aalberse, A. P. Wolse, G. H. Koppelman, L. van Rossem, M. Oldenwening, B. Brunekreef, H. A. Smit, Cohort profile: The Prevention and Incidence of Asthma and Mite Allergy (PIAMA) birth cohort. *Int. J. Epidemiol.* **43**, 527–535 (2014).
78. C.-J. Xu, C. Söderhäll, M. Bustamante, N. Baïz, O. Gruzieva, U. Gehring, D. Mason, L. Chatzi, M. Basterrechea, S. Llop, M. Torrent, F. Forastiere, M. P. Fantini, K. C. L. Carlsen, T. Haahtela, A. Morin, M. Kerkhof, S. K. Merid, B. van Rijkom, S. A. Jankipersadsing, M. J. Bonder, S. Ballereau, C. J. Vermeulen, R. Aguirre-Gamboa, J. C. de Jongste, H. A. Smit, A. Kumar, G. Pershagen, S. Guerra, J. Garcia-Aymerich, D. Greco, L. Reinius, R. R. C. Mc Eachan, R. Azad, V. Hovland, P. Mowinckel, H. Alenius, N. Fyhrquist, N. Lemonnier, J. Pellet, C. Auffray, BIOS Consortium, P. van der Vlies, C. C. van Diemen, Y. Li, C. Wijmenga, M. G. Netea, M. F. Moffatt, W. O. C. M. Cookson, J. M. Anto, J. Bousquet, T. Laatikainen, C. Laprise, K.-H. Carlsen, D. Gori, D. Porta, C. Iñiguez, J. R. Bilbao, M. Kogevinas, J. Wright, B. Brunekreef, J. Kere, M. C. Nawijn, I. Annesi-Maesano, J. Sunyer, E. Melén, G. H. Koppelman, DNA methylation in childhood asthma: An epigenome-wide meta-analysis. *Lancet Respir. Med.* **6**, 379–388 (2018).
79. S. H. Witt, J. Frank, M. Gilles, M. Lang, J. Treutlein, F. Streit, I. A. C. Wolf, V. Peus, B. Scharnholz, T. S. Send, S. Heilmann-Heimbach, S. Sivalingam, H. Dukal, J. Strohmaier, M. Sutterlin, J. Arloth, M. Laucht, M. M. Nothen, M. Deuschle, M. Rietschel, Impact on birth weight of maternal smoking throughout pregnancy mediated by DNA methylation. *BMC Genomics* **19**, 290 (2018).
80. E. Oken, A. A. Baccarelli, D. R. Gold, K. P. Kleinman, A. A. Litonjua, D. De Meo, J. W. Rich-Edwards, S. L. Rifas-Shiman, S. Sagiv, E. M. Taveras, S. T. Weiss, M. B. Belfort, H. H. Burris, C. A. Camargo Jr., S. Y. Huh, C. Mantzoros, M. G. Parker, M. W. Gillman, Cohort profile: Project Viva. *Int. J. Epidemiol.* **44**, 37–48 (2015).

Acknowledgments: We thank all the participants for this study. We also acknowledge all funding agencies and coinvestigators of cohorts involved in this study, as mentioned in the supplementary text. **Funding:** The study was supported by National Institutes of Health grant R01HD093761 (J.E.C.). For each participating cohort, funding statements can be found in the supplementary text. **Author contributions:** Conceptualization: S.Wa., J.S., J.E.C., and M.-F.H. Methodology: S.Wa., J.E.C., and M.-F.H. Investigation: S.Wa., E.C., J.S., S.A.-L., F.M.B., R.J.B., S.B., S.C., A.Ho., E.I., W.J.K., M.L., S.S.L., C.M.P., J.P., S.R., K.S., G.S., Y.S., S.Wo., E.Z., and L.Z. Visualization: S.Wa. Supervision: I.A.-M., M.B., C.A.M.C., H.R.E., S.E., J.F.F., L.G., S.E.H., G.H., B.H., J.W.H., A.Hu., W.K., G.H.K., S.J.L., S.L.M., L.N., M.P., F.R., E.F.S., T.S., D.J.S., H.T., J.M.V., M.V., J.L.W., S.H.W., J.W., E.H.Y., A.C.Z., H.J.Z., H.Z., J.E.C. and M.-F.H. Writing—original draft: S.Wa., J.E.C., and M.-F.H. Writing—review and editing: S.Wa., E.C., J.S., S.A.-L., F.M.B., R.J.B., S.B., S.C., A.Ho., E.I., W.J.K., M.L., S.S.L., C.M.P., J.P., S.R., K.S., G.S., Y.S., E.Z., L.Z., I.A.-M., S.H.A., M.B., C.A.M.C., H.R.E., S.E., J.F.F., L.G., S.E.H., G.H., B.H., J.W.H., A.Hu., W.K., G.H.K., S.J.L., S.L.M., L.N., M.P., F.R., E.F.S., T.S., D.J.S., H.T., J.M.V., M.V., J.L.W., S.H.W., J.W., E.H.Y., A.C.Z., H.J.Z., H.Z., J.E.C., and M.-F.H. **Competing interests:** G.H.K. discloses grant support from Netherlands Lung Foundation, European Union (H2020 grant), Zon MW (Vici grant), TEVA The Netherlands, Vertex, and GSK, all outside the submitted work (money to institution); his institution received funding for invited presentations and participations in advisory boards from Astra Zeneca, PURE IMS, Sanofi, and Boehringer Ingelhei, outside the submitted work. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. For access to individual-level data from cohorts, please contact the respective cohorts. Such access may be subject to local, national, and international regulations.

Submitted 22 June 2024
Accepted 25 October 2024
Published 27 November 2024
10.1126/sciadv.adr2084