View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Bergen

Environmental Epigenetics, 2019, 1–10

doi: 10.1093/eep/dvz023
Research article

RESEARCH ARTICLE

Epigenome-wide association of father's smoking with offspring DNA methylation: a hypothesis-generating study

G.T. Mørkve Knudsen[1,2,*,†], F.I. Rezwan[3,†], A. Johannessen[2,4], S.M. Skulstad[2], R.J. Bertelsen[1], F.G. Real[1], S. Krauss-Etschmann[5,6], V. Patil[7], D. Jarvis[8], S.H. Arshad[9,10], J.W. Holloway[3,‡] and C. Svanes[2,4,‡]

[1]Department of Clinical Science, University of Bergen, N-5021 Bergen, Norway; [2]Department of Occupational Medicine, Haukeland University Hospital, N-5021 Bergen, Norway; [3]Human Genetics and Genomic Medicine, Human Development and Health, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK; [4]Department of Global Public Health and Primary Care, Centre for International Health, University of Bergen, N-5018 Bergen, Norway; [5]Division of Experimental Asthma Research, Research Center Borstel, 23845 Borstel, Germany; [6]German Center for Lung Research (DZL) and Institute of Experimental Medicine, Christian-Albrechts University of Kiel, 24118 Kiel, Germany; [7]Faculty of Medicine, National Heart & Lung Institute, Imperial College, London SW3 6LY, UK; [8]David Hide Asthma and Allergy Research Centre, St. Mary's Hospital, Isle of Wight PO30 5TG, UK; [9]Clinical and Experimental Sciences, University of Southampton, Southampton General Hospital, Southampton SO16 6YD, UK; [10]NIHR Respiratory Biomedical Research Unit, University Hospital Southampton, Southampton SO16 6YD, UK

*Correspondence address. Haukanesvegen 260, N-5650 Tysse, Norway; Tel: +47 977 98 147; E-mail: Gerd.Knudsen@uib.no and torilmknudsen@gmail.com

†Equal first authors.

‡Equal last authors.

Managing Editor: Moshe Szyf

Abstract

Epidemiological studies suggest that father's smoking might influence their future children's health, but few studies have addressed whether paternal line effects might be related to altered DNA methylation patterns in the offspring. To investigate a potential association between fathers' smoking exposures and offspring DNA methylation using epigenome-wide association studies. We used data from 195 males and females (11–54 years) participating in two population-based cohorts. DNA methylation was quantified in whole blood using Illumina Infinium MethylationEPIC Beadchip. Comb-p was used to analyse differentially methylated regions (DMRs). Robust multivariate linear models, adjusted for personal/maternal smoking and cell-type proportion, were used to analyse offspring differentially associated probes (DMPs) related to paternal smoking. In sensitivity analyses, we adjusted for socio-economic position and clustering by family. Adjustment for inflation was based on estimation of the empirical null distribution in BACON. Enrichment and pathway analyses were performed on genes annotated

Received 18 February 2019; **revised** 22 August 2019; **accepted** 4 November 2019

© The Author(s) 2019. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

to cytosine-phosphate-guanine (CpG) sites using the gometh function in missMethyl. We identified six significant DMRs (Sidak-corrected P values: 0.0006–0.0173), associated with paternal smoking, annotated to genes involved in innate and adaptive immunity, fatty acid synthesis, development and function of neuronal systems and cellular processes. DMP analysis identified 33 CpGs [false discovery rate (FDR) < 0.05]. Following adjustment for genomic control ($\lambda = 1.462$), no DMPs remained epigenome-wide significant (FDR < 0.05). This hypothesis-generating study found that fathers' smoking was associated with differential methylation in their adolescent and adult offspring. Future studies are needed to explore the intriguing hypothesis that fathers' exposures might persistently modify their future offspring's epigenome.

Key words: EWAS; population cohorts; paternal smoking exposure; offspring DNA methylation

Introduction

It has been increasingly acknowledged that environmental conditions during *in utero* development and early life may contribute to later onset health and disease. Evolving evidence suggests that paternal line exposures can also affect offspring health (1–6). In particular, recent epidemiological reports have demonstrated that fathers' smoking is associated with an increased asthma risk and adiposity (7, 8) in their children.

Efforts in identifying biochemical mechanisms underlying such altered phenotypes have suggested epigenetic regulatory systems as a possible mechanistic link between environmental exposures and disease risk (9). Epigenetic processes propagate regulatory information through mitosis essential for normal cell tissue function and development (10). However, the epigenome also displays a high degree of structural adaption, and is determined by the combined response to both environmental and genetic factors (11). The plasticity of these systems is important as they affect gene transcriptional activity and lead to long-lasting phenotypic changes in a disease-related manner that may also persist through meiosis, i.e. between generations.

There is clear evidence for altered epigenetic programming in response to tobacco smoke exposure, and several genome-wide studies have identified associations between personal smoking and changes in DNA methylation at single cytosine-phosphate-guanine (CpG) sites in whole blood or isolated peripheral blood mononuclear cells (12–15). Methylation differences in cord blood of offspring born to smoking mothers have also been reported (16–18), and such differences have been shown to persist until adulthood (19, 20). However, to our knowledge, evidence for a persistent methylation effect in offspring due to paternal tobacco use has yet to be demonstrated.

As DNA methylation can be stably propagated through mitotic and possibly meiotic cell divisions (10, 11), it seems theoretically plausible that offspring DNA methylation might be persistently influenced by paternal smoking exposure. We hereby present a hypothesis-generating analysis of a relatively small number of persons, with the aim to investigate the association between paternal smoking and genomic methylation patterns in offspring, and to explore potential biological impact of methylated regions and annotated genes.

Results

Characteristics of the study populations are presented in Table 1. There was an equal gender distribution in both cohorts, with mean age of 26 and 44 years for RHINESSA and European Community Respiratory Health Survey (ECRHS), respectively. A substantial proportion of the subjects had fathers that smoked during their childhood (66%), for RHINESSA participants this was due to enrichment of samples from persons with smoking fathers for DNA methylation.

Differentially Methylated Region Analysis

Analysis of differentially methylated regions (DMRs) using comb-p identified six significant DMRs (Sidak-corrected P values: 0.0006–0.0173) (Table 2). Among these DMRs, spanning between 3 and 5 DNA methylation sites, five were mapped to known genes. Two of the annotated genes were related to innate immune system pathways (*ATP6V1E1*, *C2*), whereas one

Table 1: characteristics of study participants by cohort, RHINESSA ($n = 95$), and ECRHS2 ($n = 100$)

Descriptive variables	RHINESSA N = 95	ECRHS N = 100	*P*-value[a]
Sex, *n* (%)			
Male	46 (48)	44 (44)	0.63
Female	49 (52)	56 (56)	
Age, mean ± SD	26 ± 7.5	44 ± 6.2	<0.001
Range	11–45	31–54	
Education, *n* (%)			
Primary	5 (5)	10 (10)	0.52
Secondary	33 (35)	37 (37)	
College/university	51 (54)	53 (53)	
Smoke status, *n* (%)			
Never	68 (72)	41 (41)	<0.001
Ex	13 (14)	29 (29)	
Current	14 (15)	30 (30)	
Pack years, median (range)	2 (0–23)	8 (1–37)	<0.001
Childhood smoke exposure, *n* (%)			
Father smoked[b]	66 (69)	63 (63)	0.67
Mother smoked	31 (33)	31 (31)	0.56
Father and mother smoked	31 (33)	24 (24)	0.44
No parent smoked	25 (26)	28 (28)	0.44
Father education, *n* (%)[c]			
Primary	10 (11)	46 (46)	<0.001
Secondary	38 (40)	22 (22)	
College/university	45 (47)	25 (25)	
Mother education, *n* (%)[c]			
Primary	11 (12)	62 (62)	
Secondary	30 (32)	24 (24)	
College/university	43 (45)	7 (7)	
Father age, childbirth, mean ±SD[d] range	31 ± 5.8	32 ± 6.5	0.69
	20–54	20–58	

[a]Chi square test for categorical variables; *t*-test for continuous (norm. distributed); Wald test for continuous (non-norm. distributed).
[b]RHINESSA sample included 23 persons with father smoking starting <age 15 years, 43 with father smoking starting >15 years and smoking for at least 4 years before conception of offspring, and 29 with non-smoking fathers/mothers.
[c]Missing RHINESSA; Educ. 6 (6%); father educ. 2 (2%); mother educ. 6 (6%); ECRHS; father educ./mother educ. 7 (7%).
[d]Father's age in ECRHS obtained from registry data.

Table 2: statistically significant DMRs (Sidak $P < 0.05$) as associated with father's smoking

Location	No. probes	Slk[a] P-value	Sidak[a] P-value	Ref gene name and feature	CpG feature
Chr22:18111277-18111521	4	6.01E-07	0.0019	ATP6V1E1 Intron, 5′UTR, cds	Island
Chr6:31865522-31865866	5	2.49E-06	0.0055	C2 TSS, intron, exon, 5′UTR	Shore
Chr2:80752765-80752967	4	1.69E-06	0.0006	CTNNA2 intron	NA
Chr16:89180587-89180843	3	5.83E-06	0.0173	ACSF3 intron, cds, nc_intron, nc_exon, nc_intron	NA
Chr1:182669050-182669315	3	6.67E-07	0.0019	LINCO1688 intergenic	NA
Chr7:158766826-158767135	3	5.24E-06	0.0129	WDR60 intergenic	Island

[a]Both Slk, uncorrected Stouffer-Liptak-Kechris P values, and Sidak P values corrected for multiple testing are reported.
5′UTR , 5 prime untranslated region; cds, coding sequence; TSS, transcription start site; nc_intron, non-coding intron, nc_exon, non-coding exon.

Table 3: characteristics of DMRs

Genes annotated to DMRs	Putative gene function	Related pathways
ATP6V1E1 (ATPase H + transporting V1 subunit E1)	Encodes component of vacuolar ATPase (V-ATPase) that mediates acidification of intracellular compartments in eukaryotic cells necessary for variety of intracellular processes (32, 66, 67)	Innate immune system Synaptic vesicle cycle
C2 (complement C2)	Serum glycoprotein part of pathway of the complement system responsible for regulating immune responses (33, 68)	Innate immune system Complement pathway
CTNNA2 (catenin alpha 2)	Involved in regulating cell–cell adhesion and differentiation in the nervous system. Essential for proper regulation of cortical neuronal migration and neurite growth (34, 69)	Blood–brain barrier and immune cell transmigration Sertoli–sertoli cell junction dynamics
ACSF3 (acyl-CoA synthetase family member 3)	Catalyzes initial reaction in mitochondrial fatty acid synthesis (70)	Regulation of lipid metabolism by peroxisome proliferator-activated receptor alpha (PPARalpha) Fatty acid biosynthesis
Linc01688 (long intergenic non-protein coding RNA 1688)	Unknown	
WDR60 (Wd repeat domain 60)	Encodes a member of the WD repeat protein family. Involved in variety of cellular processes including cell cycle progression, signal transduction, apoptosis, and gene regulation (71)	Organelle biogenesis and maintenance Intraflagellar transport

DMR was involved in lipid metabolism regulation and fatty acid biosynthesis (*ACSF3*). One DMR overlapped with the catenin alpha 2 gene (*CTNNA2*), which are related to development of the nervous system. One DMR mapped to the WD repeat domain 60 gene (*WDR60*), which regulates a variety of cellular processes including cell cycle progression, signal transduction, and gene regulation (Table 3).

Differentially Mediated Probe Analysis

Epigenome-wide association between father's smoking and offspring DNA methylation at a single probe level identified 33 CpGs that passed epigenome-wide significance at a FDR rate $P < 0.05$ (Fig. 1). However, the EWAS exhibited a genomic inflation factor (lambda) of 1.462 (Supplementary Fig. S1). After applying correction for genomic inflation using the BACON method, epigenome-wide association between father's smoking and offspring DNA methylation identified 37 significantly differentially methylated CpG sites (inflation-adjusted P-value <0.0001) (Supplementary Figs S2 and S3). After subsequent filtering of data and removal of CpG sites having SNPs within the region of ± 50 bp of the CpG, and with minor allele frequency ≥ 0.05, we retained 32 differentially mediated probes (DMPs) with differential methylation between exposure groups for enrichment analysis (Supplementary Table S1). The top 10 DMPs are presented in Table 4. Among these, four were related to innate and adaptive immunity and various immune cell subsets (*BCAS1*, *MFGE8*, *UNC93B1*, and *RALB*) (21–24). Another DMP (*DLGAP1*) was related to neuronal systems and behavioural disorders (25).

Enrichment Analysis

Enrichment analysis of the 32 DMPs (Supplementary Table S1) using Enrichr for transcription factor-binding sites identified by the Encyclopedia of DNA elements (ENCODE) and Epigenomic roadmap project did not identify significant enrichment in regulatory regions (Supplementary Tables S2–S4 and Figs S4–S6).

Analyses using ontologies defined in the KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology) databases retrieved pathways and terms, and although not statistically significant, results from top 10 KEGG pathways showed enrichment of addiction behaviours (nicotine addiction). Summary statistics of top 10 GO and KEGG enrichment results are shown in Tables 5 and 6, respectively.

Sensitivity Analyses

To address the issue of relatedness among some of the participants (siblings in RHINESSA, $n = 44$), we performed linear mixed

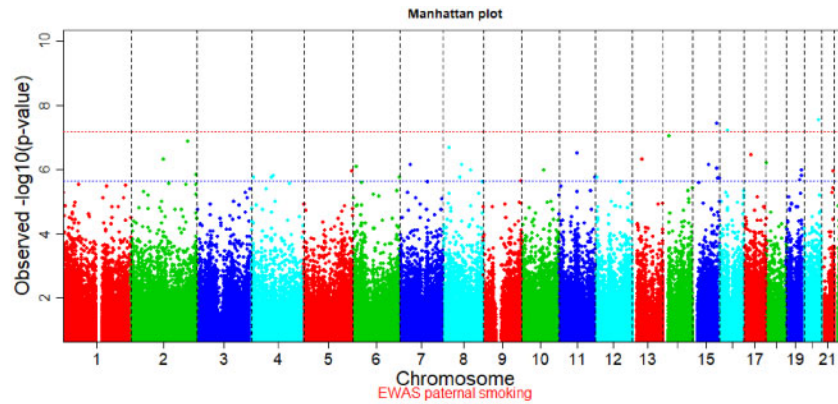

Figure 1: Manhattan plot for paternal smoking EWAS (before adjusted for genomic inflation). In the plot, the vertical axis indicates (−log10 transformed) observed P values, and the horizontal axis indicates chromosome positions with the points indicating individual CpG. Red line: Bonferroni threshold and blue line: Multiple testing correction threshold (FDR < 0.05)

Table 4: differentially methylated probe analysis (corrected P-value <0.00001)

PROBEID	BETA	SE	P-value	Adj P-value	CHR	MAPINFO	Gene
cg05019203	−0.018	0.003	2.83E-08	4.40E-06	20	52612962	BCAS1
cg25727029	0.013	0.002	3.56E-08	5.16E-06	15	89482453	MFGE8
cg00626693	−0.014	0.003	6.27E-08	7.64E-06	16	30622810	ZNF689
cg19754387	0.006	0.001	1.33E-07	1.29E-05	2	208576057	CCNYL1
cg24534854	−0.013	0.003	2.09E-07	1.76E-05	8	22582613	PEBP4
cg20272935	0.024	0.005	3.02E-07	2.27E-05	11	67765720	UNC93B1
cg04164584	−0.010	0.002	3.44E-07	2.49E-05	17	27235821	PHF12
cg06876354	0.017	0.003	4.65E-07	3.07E-05	2	121020189	RALB
cg25012097	−0.012	0.002	4.74E-07	3.11E-05	13	39263863	FREM2
cg07217718	0.025	0.005	6.17E-07	3.73E-05	18	3585484	DLGAP1

PROBEID, probe identifiers; BETA, estimates; SE, standard error; Adj P-value, P-value adjusted by multiple test correction; CHR, chromosome; MAPINFO, position of the CpGs in the chromosome; Gene, UCSC RefGene.

Table 5: top 10 enriched pathways in GO molecular function, biological processes, and cell compartment identified using genes CpGs (threshold: inflation-adjusted P-value <0.0001)

Ontology and term[a]		ID	CpGs in tern	Meth CpGs	P-value
MF	Selenomethionine adenosyltransferase activity	GO:0098601	1	1	<0.001
MF	Methionine adenosyltransferase activity	GO: 0004478	2	1	0.001
MF	Extracellularly glutamate-gated chloride channel activity	GO:0008068	1	1	0.002
BP	Regulation of exocyst assembly	GO:0001928	1	1	0.002
BP	Regulation of exocyst localization	GO:0060178	1	1	0.002
CC	Excitatory synapse	GO:0060076	48	2	0.002
BP	S-adenosylmethionine biosynthetic process	GO:0006556	3	1	0.002
BP	Sequestering of neurotransmitter	GO:0042137	2	1	0.003
BP	Synaptic vesicle lumen acidification	GO:0097401	2	1	0.003

[a]Ontology: BP, biological process; CC, cell compartment; MF, molecular function; ID, GO identifier; CpG in term, number of CpGs in GO term; Meth.CpGs, number of significant CpGs.

models on the 32 significant (inflation-adjusted P-value <0.0001) CpG sites, where family ID was included as random effect. All 32 CpGs were sustained in these analyses (Supplementary Table S5).

To account for potential confounding by social class, we conducted a sensitivity analysis adjusting for paternal socio-economic background by adding education as a proxy for socio-economic status to the regression model. Methylation at all the selected CpG sites (inflation-adjusted P-value <0.0001) was still associated with paternal smoking in this analysis (Supplementary Table S6).

Replication Analysis

Due to the amount of missing CpG sites between the EPIC and the 450 K microarray, we could not pursue replication of the significant DMRs identified in the DMR analysis. We undertook replication of the selected CpG sites (inflation-adjusted P-value <0.0001) in a subsample from Isle of Wight (IoW) with available data from cord blood DNA samples ($N = 159$, study characteristics presented in Supplementary Table S7). However, due to different methylation array platforms, and because some CpGs were discarded by pre-processing, only 13 out of the 32 CpGs

Table 6: top 10 enriched pathways in KEGG using genes CpGs (threshold: inflation-adjusted *P*-value <0.0001)

KEGG	Pathway	ID	CpGs in path	Meth. CpGs	P-value
KEGG	ECM–receptor interaction	path:hsa04512	86	2	0.006
KEGG	Glutamatergic synapse	path:hsa04724	114	2	0.011
KEGG	Nicotine addiction	path:hsa05033	40	1	0.047
KEGG	Cysteine and methionine metabolism	path:hsa00270	48	1	0.049
KEGG	Biosynthesis of amino acids	path:hsa01230	74	1	0.063
KEGG	Synaptic vesicle cycle	path:hsa04721	78	1	0.093
KEGG	Pancreatic cancer	path:hsa05212	75	1	0.095
KEGG	Colorectal cancer	path:hsa05210	86	1	0.104
KEGG	Retrograde endocannabinoid signalling	path:hsa04723	141	1	0.149
KEGG	Cytokine–cytokine receptor interaction	path:hsa04060	289	1	0.167

Pathway, KEGG pathway; ID, pathway identifier; CpG in path, number of CpGs in pathway; Meth.CpGs, number of significant CpGs.

identified in the ECRHS/RHINESSA cohort were available for replication in the IoW cohort (Supplementary Table S7).

Discussion

In the present study, we have measured epigenome-wide CpG site-specific DNA methylation in adolescent and adult offspring and identified six significant DMRs (Sidak-corrected *P* values 0.0006–0.0173) related to father's smoking. To our knowledge, this is the first study suggesting persisting effects of paternal smoking on offspring DNA methylation. Although previous genome-wide associations of maternal smoking suggest that associations with DNA methylation changes in offspring tend to weaken with increasing age of the offspring (26), and our study subjects will have accumulated a range of exposures influencing DNA methylation, it is remarkable that we were able to detect methylation differences associated with paternal smoking in persons aged 11–54 years.

Of the six statistically significant DMRs identified, one region overlapped with intron 11 within the catenin alpha-2 (*CTNNA2*) gene. *CTNNA2* has previously been shown to be differentially methylated in relation to smoking (18, 27, 28). It is expressed across the central nervous system and suggested involved in behavioural dysfunction and addiction (29). Although it did not harbour a CpG island, which would have provided additional support for a regulatory role for this region, DNA methylation at intronic sequences outside CpG islands may also be of functional important (30). Two DMRs (*ATP6V1E* and *WDR60*), colocalized with CpG islands, and the region within ATP6V1E1 covered parts of the 5 prime untranslated region (5′UTR) and the coding sequence of the gene. One DMR, annotated to the *C2* gene on chromosome 6, was located to a CpG island shore (regions within 2000 bp of a CpG island), and overlapped with the transcription start site (TSS) as well as the 5′UTR and exon 1 of C2. Although this indicates regulatory functions of the DMRs, they consist of CpGs of only nominal significance and differential methylation could reflect irregular spacing of probes and should be interpreted with caution as they may introduce false-positive results.

When exploring the biological impact of annotated genes, there were similar patterns in the DMR and DMP analyses, although the identified DMPs did not remain significant at epigenome-wide levels of significance. Two of the significant DMRs (*ATP6V1E1* and *C2*) and four of the top DMPs (*BCAS1*, *MFGE8*, *UNC93B1*, and *RALB*) were annotated to genes related to innate and adaptive immunity and to different immune cell subsets (21–24, 31, 32). Furthermore, one DMR (*CTNNA2*) and one DMP (*DLGAP1*) mapped to genes involved in function and development of neuronal systems (25, 33), and to behavioural dysfunction (29, 34, 35).

Except *CTNNA2* (18), none of our significant DMRs or top DMPs are previously reported in epigenome-wide studies of the effect of maternal smoking (16, 17, 36–39), or current or lifetime personal smoking exposure (12–15, 40–42). This is also in agreement with Joubert *et al.* who demonstrated that the CpGs differentially methylated in relation to maternal smoking were not associated with paternal smoking (43). Given the differences in gamete development in males and females, it seems biologically plausible that exposure effects through the maternal and paternal line may differ and induce epigenetic modifications at different loci. Further, it seems plausible that effects transmitted across generations may differ from those of personal smoking. To investigate whether the DMP-specific DNA methylation differences were driven by relatedness among participants, we conducted a sensitivity analysis accounting for family. All the top DMPs remained suggesting that our findings were not due to residual confounding by genetic or family-related environmental factors.

There is increasing evidence of shared pathophysiology between nicotine dependence and neuropsychiatric disorders (44), and smoking has been reported to modify genes that predispose to addictive behaviours (27, 45). In previous literature, maternal smoking during pregnancy has been associated with adverse neurodevelopmental outcome (46) and behavioural alterations in offspring (20, 47). Enrichment analysis of the top 32 differentially methylated probes (adj. $P < 0.00001$) identified GO terms and KEGG pathways involved in developmental and regulatory processes of the brain and the central nervous system and nicotine addiction, suggesting that paternal smoking may also induce aberrant methylation in genes related to neurodevelopment. However, as the identified CpGs did not remain significant epigenome-wide after adjustment for inflation, results from the KEGG and GO enrichment analysis should be interpreted with caution and may not be valid.

When we explored the biological and regulatory role of differentially methylated loci by investigating ENCODE and Epigenomic roadmaps annotated regulatory domains, we found no significant enrichment for histone modification signatures and transcription factor sites among our significant CpG sites. Whether the detected methylation differences can introduce functional changes at the gene transcriptional level needs further investigation.

The present study cannot differentiate whether the observed association of father's smoking with offspring DNA methylation may be due to second-hand smoke exposure during the

gestational period and/or childhood (post-conception) or due to altered sperm DNA methylation patterns transmitted to the offspring (pre-conception). A pre-conception effect is suggested by previous studies showing that the strongest effect of father's smoking on offspring phenotype was observed when smoking occurred before conception and particularly at an early age (7, 8, 48). However, further studies with detailed information about exposure onset in large samples will be required to address this.

The identified DMPs associated with father's smoking showed relatively small effect estimates, with top 10 CpG beta values relative to offspring of smoking and non-smoking fathers ranging from −0.02 to 0.03. This is in line with previous findings where DNA methylation differences associated with environmental exposures are characterized by small changes on the scale of 2–10% (30, 49). However, previous studies have demonstrated that even small changes can impact transcriptional activity and be consistent in different populations and across age groups (17, 49). Although associations with *in utero* maternal smoking have shown higher estimates, ranging from −0.28 to 0.18 (16, 18, 26), we would expect DNA methylation changes related to paternal exposures to be subtler when compared to direct effects from placenta–foetus interactions. Further, smaller effect estimates could be expected considering that we analyzed associations of father's smoking with DNA methylation in adolescents and adults. The fact that we found epigenomic regions (DMRs) associated with paternal smoking, adds functional relevance to our discoveries, as it implies differential methylation in regions that may affect regulation of transcription. These regional changes are also more robust as they are less prone to SNP effects and risk of false-positive findings as compared to site-by-site analysis, and they improve the specificity and potentially functional relevance of our findings (50).

A main limitation of our study is the relatively small study population. The present study was underpowered to allow stratification by offspring's sex or age, hence we did not address potential variability of effect estimated by gender or in different age groups. On the other hand, the study participants come from population-based cohorts, which is a strength of the study and to some degree allows for generalization of the results. In thorough analyses, we have accounted appropriately for the study design with two linked cohorts and family members. Further, we had information on personal smoking as well as smoking in both parents and have been able to account for main confounding factors (potentially associated with both the exposure and the outcome) in the analyses. However, rest confounding from included and unknown factors may still be present.

We have not been able to verify our findings in an independent cohort. We pursued replication in a sample from the IoW third-generation study, however, replication of significant DMRs proved not be possible as different methylation platforms were used in the two cohorts (Illumina 450K in IoW and Illumina EPIC Beadchip in RHINESSA/ECRHS) and a large number of sites were missing in the replication analysis. Few other cohorts have reliable and extensive information on father's smoking, while personal or maternal smoking are often well documented. Thus, the novel findings of DMRs related to father's tobacco smoking in our analyses, should be considered hypothesis generating and be interpreted with caution.

Conclusion

In conclusion, this hypothesis generating EWAS study is the first to report associations between paternal smoking and DNA methylation characteristics in adult and adolescent offspring. It is notable that differential methylation was detectable in this age group. Our results are intriguing as they indicate that fathers' exposures might persistently modify their future offspring's epigenome. This emphasizes the necessity to focus on male-line exposures in relation to phenotypic variation in their children, and further research to replicate our findings and explore potential mechanisms.

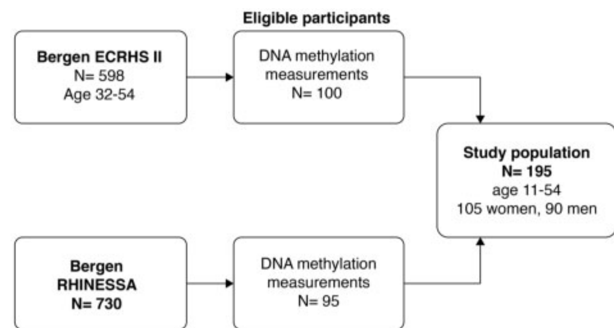

Figure 2: flowchart of study population. Offspring originate from two linked study populations with standardized and harmonized protocols: the ECRHS and the RHINESSA

Methods

Study Population

This study included data from 195 males and females aged 11–54 years participating in two linked population-based cohorts (Fig. 2).

The ECRHS conducted a study of population-based random samples of adult women and men aged 20–44 years in 1990–94 and followed up participants with clinical investigations in 2002–04 and 2012–14. The present analysis included 100 participants from the Bergen study centre with available DNA methylation data from DNA collected in ECRHS II. Information on father's year of birth was obtained from the Norwegian National Registry.

The Respiratory Health in Northern Europe, Spain and Australia study (RHINESSA) (www.rhinessa.net) investigated the offspring of ECRHS study participants in 10 study centres, following standardized protocols harmonized with the ECRHS protocols. The present analysis included 95 participants from the Bergen study centre in which DNA methylation was measured.

For the present analysis, offspring from the two cohorts were merged and analysed together. Information on smoking and other variables were obtained through interviews. Unless otherwise stated, definitions are similar in the two cohorts.

The study was approved by the Regional Committee for Medical and Health Research Ethics in Western Norway (RHINESSA: 2012/2017; ECRHS: 2010/759), and each participant gave written informed consent prior to participation.

Smoking Exposure and Covariates

In the RHINESSA cohort, information on fathers' smoking habits was collected from longitudinal data given by the fathers themselves as participants in the ECRHS II study, responding to the question: (i) *Have you ever smoked for as long as a year?*. In the ECRHS cohort, information on father's smoking was reported by the ECRHS participants and based on the question: *Did your*

father ever smoke regularly during your childhood? Father's smoking was categorized as a binary variable, as having smoked or not during offspring's childhood. In the present analysis paternal smoking was not defined in more detail as information regarding age of smoking onset was only available for RHINESSA participants.

Information on mothers smoking was reported by participants based on the question: *Did your mother ever smoke regularly during your childhood, or while pregnant with you? with the answering* categories 'no' ($n = 128$), 'yes' ($n = 62$), or 'don't know' ($n = 5$) Maternal smoking was dichotomized as either having smoked ('yes') or never smoked ('no') during offspring's childhood, whereas 'don't' know' replies were excluded from further analyses.

Personal smoking was classified as current, ex or never smoking, based on the questions: i. *Have you ever smoked for as long as a year?* (ii) *If yes How old were you when you started smoking?* (iii) *Have you stopped or cut down smoking?* (iv) *How old were you when you stopped or cut down smoking?* Number of pack years was calculated based on the number of years smoked and the average number of daily cigarettes.

Parental educational attainment was categorized in as lower (primary school), intermediate (secondary school) and higher education (college or university). Personal education level was defined the same way in RHINESSA and categorized in three levels based on reported age when education was completed in ECRHS.

Methylation Measurements and Quality Control

DNA was extracted from whole blood using a standard salting out procedure (51). Samples were processed with the Illumina MethylationEPIC Beadchip microarray, which assesses methylation at > 850 000 CpGs. Methylation measurements were performed by the Oxford Genomics Centre (Oxford, UK) using the EZ 96-DNA methylation kit (Zymo Research, CA, USA), following the manufacturer's standard protocol, with multiple identical control samples assigned to each bisulphite conversion batch to assess assay variability. Samples were randomly distributed on microarrays to control against batch effects. The CPACOR pipeline (52) was used to pre-process and normalize the methylation data. We removed probes with CpG loci located on sex chromosomes and probes located at 0 distance to known SNPs. We applied Illumina background correction to all intensity values. Any intensity values having detection P values $>=10^{-16}$ were set as missing data. Samples with call rate <98% were excluded. After pre-processing, 765 082 sites remained for subsequent analysis. A quantile normalization was applied using limma on intensity values separately based on six different probe-type categories (Type-I M red, Type-I U red, Type-I M green, Type-I U green, Type-II red, and Type-II green). Beta values were then calculated from these normalized intensity values. ComBat was used to correct for batch effects (53).

Statistical Analyses

For identification of DMRs, composed of multiple signals across individual CpG positions, we used Comb-p (54) (Python version 2.7). This method identifies regions enriched for low P values based on the probe location and unadjusted P values from the site-specific CpG analysis. For each region the comb-p algorithm adjusts the CpG P values for auto-correction between probes by using the Stouffer-Liptak-Kechris (slk) correction, followed by multiple testing adjustment using a one-step Sidak correction method (54). Regions containing at least two probes and having a Sidak-corrected P-value <0.05 were considered statistically significant.

Robust multivariate linear regression models were used to analyse the association of offspring differentially associated probes (DMPs) adjusted for paternal and offspring age, offspring gender, as well as personal and maternal smoking status. Educational level was added in sensitivity analyses to account for socio-economic status. Cell proportions (CD8T, CD4T, NK, B Cells, Monocytes, Granulocytes) were estimated using the minfi package (55) (R version 3.4.2), and cell composition coefficients were derived using the Houseman method (56). These were additively included in the model. Multiple hypothesis testing was accounted for by controlling the false discovery rate (FDR), using Benjamini and Hochberg's algorithm (57). CpGs with FDR-corrected P-value <0.1 were considered statistically significant and normalized methylation betas were used as outcome measurements. In order to address possible inflation of our test statistics by systematic biases, a Bayesian method based on estimation of the empirical null distribution was applied using the R/Bioconductor package BACON (58), and P values were estimated.

Some of the study participants originated from the same family. To account for this, we performed linear mixed model analysis on the top CpGs including family IDs as random effect.

For CpG annotation, we used the UCSC Genome browser annotation provided by Illumina in the array manifest together with SNIPPER (version 1.2, http://csg.sph.umich.edu/boehnke/snipper/) to annotate the nearest gene within 10 Mb of each CpG.

To investigate the regulatory context of the top differentially methylated probes (inflation-adjusted P-value <0.00001), we performed enrichment analysis in annotated regulatory elements (TF Chip seq and histone modification signatures) from the ENCODE (59), as well as the Epigenomics roadmap (60) using Enrichr (61).

Pathway analysis was conducted using KEGG (62), and GO databases (63) using gometh function in the missMethyl package (52).

Replication in Isle of Wight Cohort

To pursue replication of findings, we used the IoW third-generation study which since 2010 has enrolled children born to second-generation parents—the original Birth cohort. Extensive descriptions of the IoW multigenerational cohort design and objectives have been published elsewhere (64, 65). Father's smoking information given by the fathers themselves, and DNA methylation measurements using the Illumina Infinium HumanMethylation450 Beadchip array in cord blood DNA available for 159 subjects were included in the present analysis.

Availability of Data and Material

The data that support the findings of this study are available from Bergen study centre of RHINESSA and ECRHS generational population studies, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of RHINESSA and ECRHS.

Ethics Approval and Consent to Participate

The study was approved by the Regional Committee for Medical and Health Research Ethics in Western Norway

(RHINESSA: 2012/2017; ECRHS: 2010/759), and each participant gave written informed consent prior to participation.

Consent for Publication

Not applicable.

Acknowledgements

The authors thank Professor Stephanie London for her contribution to interpretation of early analyses and the reviewer's for their helpful comments on the manuscript.

Funding

The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 633212, and from the Western Norway Regional Health Authorities strategic investment grant no. 912011.

Supplementary Data

Supplementary data are available at *EnvEpig* online.

Conflict of interest statement. None declared.

References

1. Bygren LO, Kaati G, Edvinsson S. Longevity determined by paternal ancestors' nutrition during their slow growth period. *Acta Biotheoret* 2001;**49**:53–9.
2. Kaati G, Bygren LO, Edvinsson S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *Eur J Hum Genet* 2002;**10**:682–8.
3. Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M, Golding J. Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet* 2006;**14**:159–66.
4. Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, Li R, Bock C, Li C, Gu H, Zamore PD, Meissner A, Weng Z, Hofmann HA, Friedman N, Rando OJ. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* 2010;**143**:1084–96.
5. Ng SF, Lin RC, Laybutt DR, Barres R, Owens JA, Morris MJ. Chronic high-fat diet in fathers programs beta-cell dysfunction in female rat offspring. *Nature* 2010;**467**:963–6.
6. Fullston T, Ohlsson Teague EMC, Palmer NO, DeBlasio MJ, Mitchell M, Corbett M, Print CG, Owens JA, Lane M. Paternal obesity initiates metabolic disturbances in two generations of mice with incomplete penetrance to the F2 generation and alters the transcriptional profile of testis and sperm microRNA content. *FASEB J* 2013;**27**:4226–43.
7. Northstone K, Golding J, Davey Smith G, Miller LL, Pembrey M. Prepubertal start of father's smoking and increased body fat in his sons: further characterisation of paternal transgenerational responses. *Eur J Hum Genet* 2014;**22**:1382–6.
8. Accordini SC, Johannessen A, Portas L, Benediktsdóttir B, Bertelsen R. A three-generation study on the association of tobacco smoking with asthma. *Int J Epidemiol* 2018;**47**:1106–17.
9. Bakulski KM, Fallin MD. Epigenetic epidemiology: promises for public health research. *Environ Mol Mutagen* 2014;**55**:171–83.
10. Bird A. Perceptions of epigenetics. *Nature* 2007;**447**:396–8.
11. Feinberg AP. Phenotypic plasticity and the epigenetics of human disease. *Nature* 2007;**447**:433–40.
12. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet* 2011;**88**:450–7.
13. Zeilinger S, Kühnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, Weidinger S, Lattka E, Adamski J, Peters A, Strauch K, Waldenberger M, Illig T. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One* 2013;**8**:e63812.
14. Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, Belvisi MG, Brown R, Vineis P, Flanagan JM. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet* 2013;**22**:843–51.
15. Philibert RA, Beach SR, Brody GH. Demethylation of the aryl hydrocarbon receptor repressor as a biomarker for nascent smokers. *Epigenetics* 2012;**7**:1331–8.
16. Joubert BR, Håberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, Huang Z, Hoyo C, Midttun Ø, Cupul-Uicab LA, Ueland PM, Wu MC, Nystad W, Bell DA, Peddada SD, London SJ. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect* 2012;**120**:1425–31.
17. Murphy SK, Adigun A, Huang Z, Overcash F, Wang F, Jirtle RL, Schildkraut JM, Murtha AP, Iversen ES, Hoyo C. Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke. *Gene* 2012;**494**:36–43.
18. Richmond RC, Simpkin AJ, Woodward G, Gaunt TR, Lyttleton O, McArdle WL, Ring SM, Smith ADAC, Timpson NJ, Tilling K, Davey Smith G, Relton CL. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet* 2015;**24**:2201–17.
19. Tehranifar P, Wu H-C, McDonald JA, Jasmine F, Santella RM, Gurvich I, Flom JD, Terry MB. Maternal cigarette smoking during pregnancy and offspring DNA methylation in midlife. *Epigenetics* 2018;**13**:129–34.
20. Wiklund P, Karhunen V, Richmond R, Rodriguez A, De Silva M, Wielscher M *et al.* DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clin Epigenetics* 2019;**11**:97.
21. Entrez Gene: Breast Carcinoma Amplified Sequence 1. http://www.genecards.org (8 August 2019, date last accessed).
22. Kudo M, Khalifeh Soltani SMA, Sakuma SA, McKleroy W, Lee T-H, Woodruff PG, Lee JW, Huang K, Chen C, Arjomandi M, Huang X, Atabai K. Mfge8 suppresses airway hyperresponsiveness in asthma by regulating smooth muscle contraction. *Proc Natl Acad Sci USA* 2013;**110**:660–5.
23. Entrez Gene: Unc-93 Homolog B1, TLR Signaling Regulator. http://www.genecards.org (8 August 2019, date last accessed).
24. Entrez Gene: RAS Like Proto-oncogene B. http://www.genecards.org (8 August 2019, date last accessed).
25. Entrez Gene: DLG Associated Protein 1. http://www.genecards.org (8 August 2019, date last accessed).
26. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, Reese SE, Markunas CA, Richmond RC, Xu C-J, Küpers LK, Oh SS, Hoyo C, Gruzieva O, Söderhäll C, Salas LA, Baïz N, Zhang H, Lepeule J, Ruiz C, Ligthart S, Wang T, Taylor JA, Duijts L, Sharp GC, Jankipersadsing SA, Nilsen RM, Vaez A, Fallin MD, Hu D, Litonjua AA, Fuemmeler BF, Huen K, Kere J, Kull I, Munthe-Kaas MC, Gehring U, Bustamante M, Saurel-Coubizolles MJ, Quraishi BM, Ren J, Tost J, Gonzalez JR, Peters

MJ, Håberg SE, Xu Z, van Meurs JB, Gaunt TR, Kerkhof M, Corpeleijn E, Feinberg AP, Eng C, Baccarelli AA, Benjamin Neelon SE, Bradman A, Merid SK, Bergström A, Herceg Z, Hernandez-Vargas H, Brunekreef B, Pinart M, Heude B, Ewart S, Yao J, Lemonnier N, Franco OH, Wu MC, Hofman A, McArdle W, Van der Vlies P, Falahi F, Gillman MW, Barcellos LF, Kumar A, Wickman M, Guerra S, Charles M-A, Holloway J, Auffray C, Tiemeier HW, Smith GD, Postma D, Hivert M-F, Eskenazi B, Vrijheid M, Arshad H, Antó JM, Dehghan A, Karmaus W, Annesi-Maesano I, Sunyer J, Ghantous A, Pershagen G, Holland N, Murphy SK, DeMeo DL, Burchard EG, Ladd-Acosta C, Snieder H, Nystad W, Koppelman GH, Relton CL, Jaddoe VWV, Wilcox A, Melén E, London SJ. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am J Hum Genet* 2016;**98**:680–96.

27. Liu M, Fan R, Liu X, Cheng F, Wang J. Pathways and networks-based analysis of candidate genes associated with nicotine addiction. *PLoS One* 2015;**10**:e0127438.
28. Uhl GR, Liu Q-R, Drgon T, Johnson C, Walther D, Rose JE, David SP, Niaura R, Lerman C. Molecular genetics of successful smoking cessation: convergent genome-wide association study results. *Arch Gen Psychiatry* 2008;**65**:683–93.
29. Mexal S, Berger R, Pearce L, Barton A, Logel J, Adams CE, Ross RG, Freedman R, Leonard S. Regulation of a novel alphaN-catenin splice variant in schizophrenic smokers. *Am J Med Genet* 2008;**147b**:759–68.
30. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet* 2013;**14**:585.
31. Entrez Gene: ATPase H+Transporting V1 Subunit E1. http://www.genecards.org (8 August 2019, date last accessed).
32. Entrez-Gene: Complement C2. http://www.genecards.org (8 August 2019, date last accessed).
33. Entrez-Gene: Catenin Alpha 2. http://www.genecards.org (8 August 2019, date last accessed).
34. Fan Z, Qian Y, Lu Q, Wang Y, Chang S, Yang L. DLGAP1 and NMDA receptor-associated postsynaptic density protein genes influence executive function in attention deficit hyperactivity disorder. *Brain Behav* 2018;**8**:e00914.
35. Mattheisen M, Samuels JF, Wang Y, Greenberg BD, Fyer AJ, McCracken JT, Geller DA, Murphy DL, Knowles JA, Grados MA, Riddle MA, Rasmussen SA, McLaughlin NC, Nurmi EL, Askland KD, Qin H-D, Cullen BA, Piacentini J, Pauls DL, Bienvenu OJ, Stewart SE, Liang K-Y, Goes FS, Maher B, Pulver AE, Shugart YY, Valle D, Lange C, Nestadt G. Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Mol Psychiatry* 2015;**20**:337–44.
36. de Vocht F, Simpkin AJ, Richmond RC, Relton C, Tilling K. Assessment of offspring DNA methylation across the lifecourse associated with prenatal maternal smoking using Bayesian mixture modelling. *Int J Ent Res Public Health* 2015;**12**:14461–76.
37. Markunas CA, Xu Z, Harlid S, Wade PA, Lie RT, Taylor JA, Wilcox AJ. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ Health Perspect* 2014;**122**:1147–53.
38. Kupers LK, Xu X, Jankipersadsing SA, Vaez A, la Bastide-van Gemert S, Scholtens S *et al*. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *Int J Epidemiol* 2015;**44**:1224–37.
39. Wiklund P, Karhunen V, Richmond RC, Parmar P, Rodriguez A, De Silva M, Wielscher M, Rezwan FI, Richardson TG, Veijola J, Herzig K-H, Holloway JW, Relton CL, Sebert S, Järvelin M-R. DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clin Epigenet* 2019;**11**:97.
40. Wan ES, Qiu W, Baccarelli A, Carey VJ, Bacherman H, Rennard SI, Agusti A, Anderson W, Lomas DA, DeMeo DL. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet* 2012;**21**:3073–82.
41. Zhang Y, Yang R, Burwinkel B, Breitling LP, Brenner H. F2RL3 methylation as a biomarker of current and lifetime smoking exposures. *Environ Health Perspect* 2014;**122**:131–7.
42. Monick MM, Beach SRH, Plume J, Sears R, Gerrard M, Brody GH, Philibert RA. Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from smokers. *Am J Med Genet* 2012;**159b**:141–51.
43. Joubert BR, Haberg SE, Bell DA, Nilsen RM, Vollset SE, Midttun O *et al*. Maternal smoking and DNA methylation in newborns: in utero effect or epigenetic inheritance? *Cancer Epidemiol* 2014;**23**:1007–17.
44. Loukola A, Hallfors J, Korhonen T, Kaprio J. Genetics and smoking. *Curr Addict Rep* 2014;**1**:75–82.
45. Wang J, Li MD. Common and unique biological pathways associated with smoking initiation/progression, nicotine dependence, and smoking cessation. *Neuropsychopharmacology* 2010;**35**:702–19.
46. Kiechl-Kohlendorfer U, Ralser E, Pupp Peglow U, Reiter G, Griesmaier E, Trawöger R. Smoking in pregnancy: a risk factor for adverse neurodevelopmental outcome in preterm infants? *Acta Paediatr* 2010;**99**:1016–9.
47. Toledo-Rodriguez M, Lotfipour S, Leonard G, Perron M, Richer L, Veillette S, Pausova Z, Paus T. Maternal smoking during pregnancy is associated with epigenetic modifications of the brain-derived neurotrophic factor-6 exon in adolescent offspring. *Am J Med Genet* 2010;**153b**:1350–4.
48. Svanes C, Koplin J, Skulstad SM, Johannessen A, Bertelsen RJ, Benediktsdottir B, Bråbäck L, Elie Carsin A, Dharmage S, Dratva J, Forsberg B, Gislason T, Heinrich J, Holm M, Janson C, Jarvis D, Jögi R, Krauss-Etschmann S, Lindberg E, Macsali F, Malinovschi A, Modig L, Norbäck D, Omenaas E, Waatevik Saure E, Sigsgaard T, Skorge TD, Svanes Ø, Torén K, Torres C, Schlünssen V, Gomez Real F. Father's environment before conception and asthma risk in his children: a multi-generation analysis of the Respiratory Health In Northern Europe study. *Int J Epidemiol* 2017;**46**:235–45.
49. Breton CV, Marsit CJ, Faustman E, Nadeau K, Goodrich JM, Dolinoy DC, Herbstman J, Holland N, LaSalle JM, Schmidt R, Yousefi P, Perera F, Joubert BR, Wiemels J, Taylor M, Yang IV, Chen R, Hew KM, Freeland DMH, Miller R, Murphy SK. Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: the Children's Environmental Health and Disease Prevention Research Center's Epigenetics Working Group. *Environ Health Perspect* 2017;**125**:511–26.
50. Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. *Nat Biotechnol* 2012;**30**:1095–106.
51. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**:1215.
52. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* 2016;**32**:286–8.
53. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007;**8**:118–27.

54. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics* 2012;**28**:2986–8.
55. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014;**30**:1363–9.
56. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;**13**:86.
57. Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc Ser B Methodol* 1995;**57**:289–300.
58. van Iterson M, van Zwet EW, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol* 2017;**18**:19.
59. Consortium EP. An integrated Encyclopedia of DNA elements in the human genome. *Nature* 2012;**489**:57–74.
60. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, Farnham PJ, Hirst M, Lander ES, Mikkelsen TS, Thomson JA. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* 2010;**28**:1045–8.
61. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;**44**:W90–7.
62. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017;**45**:D353–d61.
63. Consortium GO. Gene Ontology Consortium: going forward. *Nucleic Acids Research* 2015;**43**:D1049–56.
64. Arshad SH, Karmaus W, Zhang H, Holloway JW. Multigenerational cohorts in patients with asthma and allergy. *J Allergy Clin Immunol* 2017;**139**:415–21.
65. Felix JF, Joubert BR, Baccarelli AA, Sharp GC, Almqvist C, Annesi-Maesano I, Arshad H, Baïz N, Bakermans-Kranenburg MJ, Bakulski KM, Binder EB, Bouchard L, Breton CV, Brunekreef B, Brunst KJ, Burchard EG, Bustamante M, Chatzi L, Cheng Munthe-Kaas M, Corpeleijn E, Czamara D, Dabelea D, Davey Smith G, De Boever P, Duijts L, Dwyer T, Eng C, Eskenazi B, Everson TM, Falahi F, Fallin MD, Farchi S, Fernandez MF, Gao L, Gaunt TR, Ghantous A, Gillman MW, Gonseth S, Grote V, Gruzieva O, Håberg SE, Herceg Z, Hivert M-F, Holland N, Holloway JW, Hoyo C, Hu D, Huang R-C, Huen K, Järvelin M-R, Jima DD, Just AC, Karagas MR, Karlsson R, Karmaus W, Kechris KJ, Kere J, Kogevinas M, Koletzko B, Koppelman GH, Küpers LK, Ladd-Acosta C, Lahti J, Lambrechts N, Langie SAS, Lie RT, Liu AH, Magnus MC, Magnus P, Maguire RL, Marsit CJ, McArdle W, Melén E, Melton P, Murphy SK, Nawrot TS, Nisticò L, Nohr EA, Nordlund B, Nystad W, Oh SS, Oken E, Page CM, Perron P, Pershagen G, Pizzi C, Plusquin M, Raikkonen K, Reese SE, Reischl E, Richiardi L, Ring S, Roy RP, Rzehak P, Schoeters G, Schwartz DA, Sebert S, Snieder H, Sørensen TI, Starling AP, Sunyer J, Taylor JA, Tiemeier H, Ullemar V, Vafeiadi M, Van Ijzendoorn MH, Vonk JM, Vriens A, Vrijheid M, Wang P, Wiemels JL, Wilcox AJ, Wright RJ, Xu C-J, Xu Z, Yang IV, Yousefi P, Zhang H, Zhang W, Zhao S, Agha G, Relton CL, Jaddoe VWV, London SJ. Cohort profile: Pregnancy and Childhood Epigenetics (PACE) Consortium. *Int J Epidemiol* 2018;**47**:22–3u.
66. Baud V, Mears AJ, Lamour V, Scamps C, Duncan AM, McDermid HE, Lipinski M. The E subunit of vacuolar H(+)-ATPase localizes close to the centromere on human chromosome 22. *Human Molecular Genetics* 1994;**3**:335–9.
67. van Hille B, Vanek M, Richener H, Green JR, Bilbe G. Cloning and tissue distribution of subunits C, D, and E of the human vacuolar H(+)−ATPase. *Biochem Biophys Res Commun* 1993;**197**:15–21.
68. Ishii Y, Zhu ZB, Macon KJ, Volanakis JE. Structure of the human C2 gene. *J Immunol* (Baltimore, Md : 1950). 1993;**151**:170–4.
69. Schaffer AE, Breuss MW, Caglayan AO, Al-Sanaa N, Al-Abdulwahed HY, Kaymakcalan H *et al*. Biallelic loss of human CTNNA2, encoding alphaN-catenin, leads to ARP2/3 complex overactivity and disordered cortical neuronal migration. *Nature Genetics* 2018;**50**:1093–101.
70. www.genecards.org, "Entrez-Gene: Acyl-CoA synthesase Family member 3" Retrieved 2019-08-08.
71. www.genecards.org, "Gene-Entrez: Wd Repeat Domain 60" Retrieved 2019-08-08.