UCSF
UC San Francisco Previously Published Works

Title

The biological embedding of early-life socioeconomic status and family adversity in children's genome-wide DNA methylation

Permalink

https://escholarship.org/uc/item/95r2c9cz

Journal

Epigenomics, 10(11)

ISSN

1750-1911

Authors

Bush, Nicole R
Edgar, Rachel D
Park, Mina
et al.

Publication Date

2018-11-01

DOI

10.2217/epi-2018-0042

Peer reviewed

Epigenomics

The biological embedding of early-life socioeconomic status and family adversity in children's genome-wide DNA methylation

Nicole R Bush*[1,2], Rachel D Edgar[3], Mina Park[4], Julia L MacIsaac[3], Lisa M McEwen[3], Nancy E Adler[1], Marilyn J Essex[5], Michael S Kobor**[‡,3] & W Thomas Boyce[‡,1,2]

[1]Department of Psychiatry, Center for Health & Community, Weill Neuroscience Institute, University of California, San Francisco, 3333 California Street, Suite 465, San Francisco, CA 94118, USA
[2]Department of Pediatrics, Division of Developmental Medicine, University of California, San Francisco, 550 16th Street, San Francisco, CA 94158, USA
[3]Department of Medical Genetics, BC Children's Hospital, Centre for Molecular Medicine & Therapeutics, University of British Columbia, 950 West 28th Ave., Vancouver, BC V5Z 4H4, Canada
[4]School of Population & Public Health, BC Children's Hospital, University of British Columbia, 4480 Oak St, Vancouver, BC V6H 3N1, Canada
[5]Department of Psychiatry, University of Wisconsin, Madison,16330 Ellendale Road, Dallas, OR 97338, USA
*Author for correspondence: Nicole.bush@ucsf.edu
**Author for correspondence: msk@cmmt.ubc.ca
[‡]Authors contributed equally

Aim: To examine variation in child DNA methylation to assess its potential as a pathway for effects of childhood social adversity on health across the life course. **Materials & methods:** In a diverse, prospective community sample of 178 kindergarten children, associations between three types of social experience and DNA methylation within buccal epithelial cells later in childhood were examined. **Results:** Family income, parental education and family psychosocial adversity each associated with increased or decreased DNA methylation (488, 354 and 102 sites, respectively) within a unique set of genomic CpG sites. Gene ontology analyses pointed to genes serving immune and developmental regulation functions. **Conclusion:** Findings provided support for DNA methylation as a biomarker linking early-life social experiences with later life health in humans.

Individual socioeconomic status (SES) is one of the strongest predictors of health. Research has demonstrated consistent, graded associations between various components of SES and a wide range of health indicators, such that higher income, education and/or occupational level, predict lower rates of morbidity and mortality [1]. These social disparities in physical and mental health begin at birth, continue throughout childhood and contribute to risk for poor mental and physical health across the life course [2–4]. Although there is increasing appreciation for the idea that early-life social experience 'gets under the skin' to affect health outcomes, prospective evidence for potential mechanisms of this 'biological embedding' [5] of early-life stress is limited, particularly in young children. Accumulating theory and evidence point to activation or blunting of neurobiological response systems (i.e., immune, cardiovascular and hypothalamic pituitary adrenal axis system functioning, as well as brain circuitry and function) as correlates and potential mechanisms for these effects [6–18]. Nonetheless, there are likely more fundamental, proximal mechanisms linking early SES and health.

Recent advances in understanding of the apparent responsiveness, plasticity and memory characteristics of DNA methylation [19] place it among the more attractive current models in the search for mechanisms responsible for socioeconomic disparities in health [20]. This paradigm posits that early social environmental signals activate, repress and maintain genomic responses, via adaptive calibration of the DNA methylome, and such effects may contribute in turn to emotional, behavioral and health outcomes. Although it falls short of that mechanistic paradigm, DNA methylation has the potential to operate as a stable biomarker of a child responding to stress and allows for interventions to maintain healthy trajectories. Prospective longitudinal studies are underway, but in

Future Medicine

the meantime, evidence from cross-sectional studies is accumulating that social adversities associate with variation in DNA methylation marks and that childhood SES is a stronger predictor of DNA methylation in adulthood than is adult SES. For example, within a sample of 92 male and female adults, findings suggested that early-life SES was associated with genome-wide DNA methylation in adult peripheral blood mononuclear cells, whereas concurrent adult SES was not reliably related [21]. Within another sample of 40 male adults, socioeconomic position in childhood was a stronger predictor of genome-wide DNA methylation in gene promoter regions than was adult SES [22]. In a sample of 87 adult females, lower family SES at birth was associated with white blood cell methylation of repetitive element Sat2 DNA sequences [23]. In one prospective study of 494 male and females in the Philippines, targeted pathway analyses found that household SES in infancy and childhood was associated with DNA methylation during early adulthood at two loci known to regulate inflammation, which was then associated with participants' concurrent overall levels of plasma inflammatory biomarkers, providing compelling evidence for the potential pathways of SES health effects [24].

Cumulative measures of SES across childhood show even stronger associations with adult health, and this is also true for DNA methylation patterns associated with the SES health gradient. In a sample of 800 male and female adults, SES across the life course predicted variation in adult DNA methylation of candidate genes involved in inflammation, in a dose-dependent manner [25]; in particular, chronic low SES beginning in childhood was a stronger predictor than singular time points of low SES. Despite this accumulating evidence, prospective evidence for these associations is lacking, as is understanding of sensitive periods of exposure. Studies assessing DNA methylation profiles in early childhood are needed to determine how early in life association to DNA methylation might be detected and to clarify target periods for interventions [26].

Although low-SES families are at greater risk of experiencing higher levels of psychosocial adversity [1,27], family psychosocial adversity occurs across the spectrum of SES and both childhood SES and family psychosocial adversity are independently associated with health [16,17,28,29]. For example, prospective reports of parental stress in infancy and early childhood have been found to predict differential buccal epithelial cell DNA methylation in adolescence [30]. Studies of SES effects on DNA methylation that concurrently examine family psychosocial adversity are needed to distinguish effects of these related but distinctive predictors and clarify the pathways by which they potentially impact health.

SES is itself a multifactor variable, and its components, such as income and education, provide different types of resources that may differentially impact children's biological development (see, e.g., [7]). SES is also entwined with race/ethnicity, and examination of SES effects independent of race/ethnicity requires stratified sampling and/or explicit modeling to address this confound. Ancestry is one of the stronger predictors of variation in DNA methylation [31–33], and a recent recommendation suggested accounting for confounding genetic variation within epigenome-wide analyses [34]. Studies run on self-reported single-ethnicity samples may not require those adjustments, but such sampling can lead to limited variation in SES exposure and/or lack of consideration of differences in the correlates of ancestry within a sample. Taken together, these issues call for prospective studies linking childhood SES and psychosocial adversity with genomic DNA methylation within socioeconomically diverse, mixed-sex samples that account for ancestral descent.

Building upon the theoretical centrality of epigenetic processes in health disparities [20], and the considerations described above, we sought to fill important gaps in the evidence base connecting DNA methylation with the health effects of early-life social conditions. We used prospectively collected data on a socioeconomically and ethnically diverse sample of 178 young male and female children to test whether family income, parental education and family adversity during kindergarten were associated with variation in those children's genome-wide DNA methylation in buccal epithelial cells at age 9 years, adjusting for both genetic ancestry and self-reported ethnic minority status. To advance understanding of the complex interplay of SES and psychosocial adversity, we also tested whether the three social environment predictors had overlapping or independent associations with DNA methylation patterning across the genome. The community sample of young children utilized for this study lacked significant disease prevalence (as expected at this age), preventing full examination of whether DNA methylation mediated SES health relations. Therefore, we performed exploratory analyses to examine whether the differential DNA methylation of CpG sites identified were on genes known to relate to health and disease outcomes. Findings suggested that family income-per-dependent in early childhood, as well as parental education and family psychosocial adversity, were each associated with children's DNA methylation marks in middle childhood, across genes with possible health-relevant functions.

Methods

Human child & adult caregiver participants

Participants were originally recruited in three waves from 29 kindergarten classrooms within six public schools selected to represent a variety of sociodemographic characteristics in the San Francisco Bay area during the fall seasons of 2003, 2004 and 2005. At that time, they were enrolled in the Peers and Wellness Study (PAWS), a larger longitudinal study (n = 338: M age at kindergarten entry: 5.34 years; standard deviation [SD]: 0.31; range: 4.75–6.24) of social dominance, biological responses to adversity, and mental and physical health (full details of the study methods are provided in [59]). Follow-up was funded and conducted in 2010, when children were 9–11 years old (M: 10.92; SD: 0.94), to collect genetic data, resulting in an n of 192 with genetic data. This study and all methods were approved by the University of California Berkeley and San Francisco institutional review board (for the protection of human subjects; IRBs), protocols were followed according to those guidelines, and informed consent was obtained from all participating families. The genotyped subsample was less ethnically diverse and had families with higher incomes and educational attainment (p-values <0.001), compared with nongenotyped children, reflecting the higher rate of loss of contact with lower-SES and minority families after the initial study was completed (Supplementary Table 1). The genetic subsample with complete data used for this study (n = 178) was ethnically diverse (self-reported 11.8% African–American, 9.6% Asian, 52.2% Caucasian, 3.4% Latino, 20.2% multiethnic, 0.6% other) and child sex was well balanced (53% male, n = 94).

Measures of early-life socioeconomic & social adversity

Caregivers (87% mothers) provided socioeconomic information, including household income and the highest level of parental education during their child's kindergarten year. Average household income ranged from less than US$10,000 to more than US$400,000 (M: US$60,000–US$79,000). This value was divided by the number of dependents for that income to create the income-per-dependent variable. The highest level of caregiver educational attainment ranged from less than a high school degree (2.7%) to advanced degrees (54%), with 78% of the sample having at least a college degree. Though higher than national statistics, this level of income and education is representative of populations in the San Francisco Bay area. Children's exposure to family adversity was assessed by six indices of potential sources of family stress. All measures of adversity were based on parent report, and full details of the measures and their reliabilities in the full study sample can be found in [16]. In brief, family adversity was derived from seven measures that tapped six domains of parental experiences of financial stress (four items), parenting overload (five items), marital conflict (ten items) and depression (20 items), as well as their endorsement of levels of negative/anger expressiveness in the family (20 and 18 items), and harsh and restrictive parenting (18 items). As we were interested in capturing children's overall exposure to family adversity, in terms of cumulative risk, the six indices of adversity were standardized and composited into one adversity index as has been done in numerous previous studies and publications [16,30,59,60]. Descriptive statistics and reliabilities for the adversity measures and their composite score in this genetic subsample are shown in Supplementary Table 2.

Genetic data collection

Children's DNA was collected using two methods. DNA for genotyping was collected using Oragene OG-500 DNA all-in-one system for the collection, stabilization, transportation and purification of DNA from saliva. Children were asked to spit into a small vial and samples were immediately mixed with stabilizing solution and stored at room temperature until assayed. DNA for methylation analyses was collected using Isohelix T-swab Buccal Swabs (Isohelix), after participants rinsed their mouths with water twice for 15 s each. Buccal epithelial cells were first mixed with stabilizing solution, then stored at room temperature until DNA was extracted.

Genomic DNA methylation

Genomic DNA was isolated using Simhelix Buccal Swab DNA Isolation Kits (Cell Projects, Harrietsham, Kent, UK). DNA was extracted using DNAeasy DNA extraction kit (Qiagen, CA, USA). Bisulfite conversion of DNA was performed with the EZ-DNA methylation kit (Zymo Research, CA, USA). Bisulfite-converted DNA was interrogated with the Illumina Infinium HumanMethylation450 BeadChip (450K, Illumina Inc., CA, USA) according to manufacturer's instructions.

Genotyping data analysis for genetic ancestry

DNA was interrogated with the PsychChip (Illumina Inc.) according to manufacturer's instructions. SNPs from the PsychChip were filtered based on the GenCall score. SNPs with a low tenth percentile score (<0.355) or with a low average GenCall score (<0.1) were removed. SNPs on the PsychChip not annotated to a chromosome were also removed (32,811 SNPs). A total of 52,151 SNPs were filtered for quality control. Genetic ancestry was called based on the PsychChip SNP alleles using identical by shared state complete-linkage hierarchical clustering in the statistical program PLINK (version 1.07) [61].

In order to define the number of clusters to pull as genetic ancestry clusters, a principal component analysis (PCA) was run on the SNP data, and the first principal component (PC) was used to represent genetic variation in the cohort. Then, as in Solovieff *et al.* [62], k-means clustering was run on the first PC and four clusters were determined to be the smallest number of clusters required to minimize the within-group sum of squares. Four clusters agreed with the population structure observed in the reported ethnicity information, and therefore four clusters from the PLINK identity-by-state clustering output were used as a genetic ancestry covariate in the association analysis.

DNA methylation data preprocessing

Background subtraction and color correction of the DNA methylation data was performed using Illumina GenomeStudio software, at which point data were imported into R for further preprocessing. Technical replicates and expected sex of each participant were checked to ensure consistency. The 65 SNP probes on the 450K that directly measured a SNP instead of a CpG were not needed in this analysis beyond confirming replicate ID and were, therefore, removed. CpGs on the X and Y chromosomes were removed (11,648 CpGs). Probes with evidence of cross-hybridization to regions other than the probe's target in the genome were removed (40,590 probes) [61,63]. Additional filters included 2433 probes not detected above background in at least one sample and 2972 probes with less than three beads contributing to the signal in at least nine samples. This left a final total of 428,407 probes. DNA methylation data were normalized using the Beta-mixture quantile (BMIQ) normalization to correct for probe type [64].

Finally, ComBat was performed to remove the batch effects of chip and row sequentially [65] An additional 1514 CpGs failed in ComBat as samples with less than three beads contributing to the signal at a CpG were confounded with batch. Through the steps of normalization and ComBat, the root-mean-square error (RMSE) of 15 technical replicates in the PAWS generally improved, suggesting that both normalization and ComBat reduced the technical variability between samples (Supplementary Figure 8). Finally, 20,869 probes have a documented SNP in the C or G of the target CpG (polymorphic CpG). Of these SNP CpG probes 17,015 had not been filtered yet, these probes were also removed as the majority of their variability seems to be due to the SNP status, leaving 409,878 CpGs to be tested.

DNA methylation data analysis

We first examined the data for potential covariates for each variable of interest by determining which measures were correlated with each variable. A combination of *a priori* assumptions and exploratory data-driven PCA was used. In this study, the first two PCs are not shown in visualizations, as the first represents shifts in CpG means from one CpG to the next and the second captures the effect of the blood cell contamination already examined. As expected age, child sex and genetic ancestry were major contributors to methylation variation (Supplementary Figure 2). Child sex was well balanced in the study design, thus child sex was not included as a covariate due to concern for overfitting models. Additional testing for confounding of major contributors with main effects of interest led to the final model selection (Supplementary Table 4, see Supplementary Text).

After determining the covariates needed in each model, a differential DNA methylation analysis was performed on all probes on the array, using linear mixed effects modeling [66] including a combination of genetic ancestry, ethnic minority group status and age as covariates, and with twin pair (eight pairs of twins in cohort) as a random effect (Table 1). Statistical significance was assessed using an association p-value corrected for multiple comparisons using Benjamini–Hochberg correction [38]. In addition to the statistical significance, a biological potency cut-off was also used to define significant hits, and a delta beta was calculated as the difference in methylation between the highest and lowest level of a variable. *Post hoc* sensitivity analysis of model covariates was then performed as described in the Supplementary Text.

Table 1. Results of linear mixed effects models at multiple levels of false discovery rates and delta beta cut points.					
Predictor variable	**Covariates**	**Delta beta threshold**	**FDR 0.20**	**FDR 0.1**	**FDR 0.05**
Income-per-dependent	Genetically determined ancestry, self-reported ethnic minority status, child age and twin status	None	27,283	9600	3307
		0.05	488[†]	201	85
		0.1	35	16	8
Family adversity	Genetically determined ancestry and twin status	None	3945	464	10
		0.05	102[†]	15	1
		0.1	8	1	0
Parental education	Genetically determined ancestry, self-reported ethnic minority status and twin status	None	25,203	6431	737
		0.05	354[†]	111	10
		0.1	21	13	1

Statistical thresholds are given at increasing false discovery rates. Increasingly statistically conservative, from left to right, FDR 0.05 is most conservative. Rows down represent increasing delta beta thresholds, indicating increasing potential biological potency, within predictor variable.
[†]Indicates the CpG lists for each variable selected for further interrogation.
FDR: False discovery rate.

Gene ontology enrichment

The lists of significantly differentially methylated genes with each variable gene were used for functional enrichment of gene ontology (GO) terms. GO annotations of the 19,389 genes associated with the 450K probes were used as the background list. Enrichment of GO terms in the list of differentially methylated genes was tested using overrepresentation analysis was done using ErmineJ [67]. Statistical significance is reported as false discovery rates (FDR) computed using the Benjamini–Hochberg method in ErmineJ. Also calculated are the multifunctionality scores of the ontology gene sets [68].

Data availability

The data reported in this paper have been deposited in the Gene Expression Omnibus database accession no. GSE94734.

Results

In order to explore the association of SES factors and family psychosocial adversity on DNA methylation marks, we measured genome-wide DNA methylation on the Illumina Infinium HumanMethylation450 BeadChip (450K, Illumina Inc.). Characteristics of the study participants who provided DNA samples at the follow-up visit and comparison to the original full sample are shown in Supplementary Text (Supplementary Table 1).

Covariate confirmation

Given the diversity of the reported ethnicity in individuals and the known impact of genetic variants on DNA methylation, we built a molecular classification of ancestry, based on SNP variants measured on the PsychChip (Illumina Inc.). Clusters generated by the SNP-based genetic ancestry analysis were moderately consistent with self-reported ethnicity and minority status (p < 0.001, Fisher's Exact Test; Figure 1). However, because these measures were not entirely overlapping and may have separate theoretical relevance [35], and also have recently been shown to each uniquely predict DNA methylation [36], genetic ancestry and self-reported ethnic minority status were both included in analyses when cofounded with the main variables (Table 1; see Supplementary Text).

Having established genetic ancestry and self-reported ethnic-minority status for participants, we next moved forward with our examination of the DNA methylation data. To test for potential confounds due to possible blood contamination during buccal epithelial cell collection, we added age-matched blood and T-cell samples from another dataset and approximated contamination with PCA (see Supplementary Text) [37]. While there was some skewing of samples toward the blood cluster (Supplementary Figure 1), suggesting potential blood contamination, PC2, the PC most related to the skewing of BEC samples to the blood clusters, was not significantly correlated with SES factors or family psychosocial adversity of interest (r_s = -0.09–0.03). This suggested that buccal cell DNA was not appreciatively contaminated with blood or T cells in a manner confounded with our variables of interest. CpGs that did not pass quality controls, were polymorphic, etc., were filtered from the dataset (see Methods section). PC analyses were conducted again, but this time just in the PAWS cohort, on the resulting 409,878 CpGs, to identify
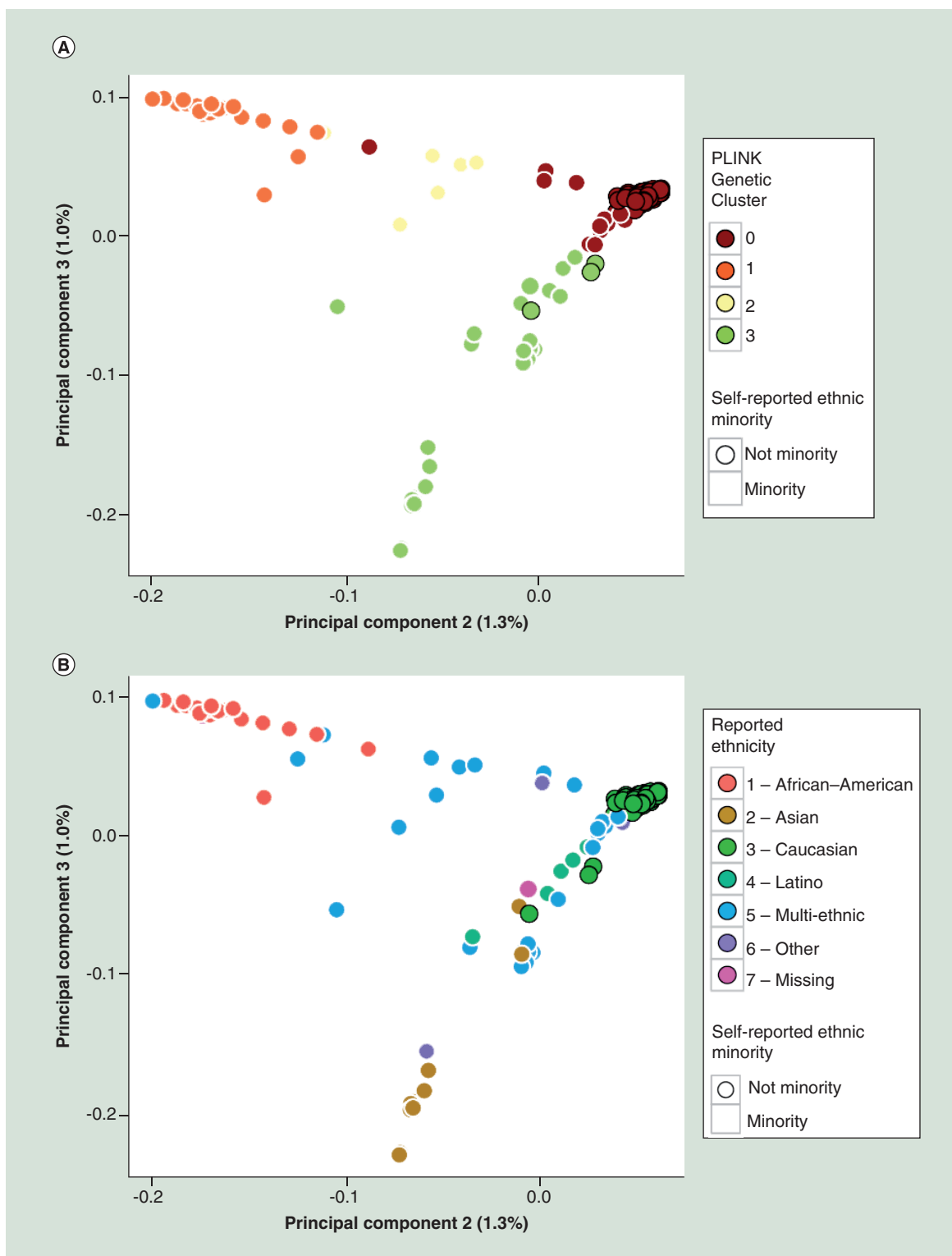
Figure 1.   Principal component analysis of SNP genotyping data showed genetic ancestry to be a major contributor to the variation between individual genotypes. (A) The loadings of second and third principal components representing the variance in the SNP genotyping data. Individual points are colored by four genetic ancestry clusters called by PLINK identity-by-state clustering. **(B)** The same plot as in **(A)** but with points colored by to self-reported ethnicity. In **(A)** and **(B)**, black-outlined points represent individuals self-identified as not minority.

the major contributors to variation in DNA methylation and determine important covariates (Supplementary Figure 2).

Early childhood social environment was associated with DNA methylation

Three early-life social environmental predictors from the PAWS cohort were selected for analysis with respect to patterns of childhood DNA methylation: two components of SES (income-per-dependent and highest household education level) and a composite measure of family psychosocial adversity (financial stress, parenting overload, marital conflict, parental depression, family negative/anger expressiveness, and harsh and restrictive parenting). Supplementary Table 4 presents details on covariates within each model. Bivariate associations showed that the three predictors were weakly to moderately intercorrelated (r = 0.34 income-education, r = -0.13 adversity-education, r = -0.23 income-adversity). However, PC analyses of the DNA methylation patterns showed that while they had overlapping signal in some PCs, they also separated into independent PCs, suggesting each predictor bore some unique DNA methylation signature (Supplementary Figure 2). Self-reported ethnic minority status and genetic ancestry together also accounted for at least 3.4% of the variability in the DNA methylation data (Supplementary Figure 2) mandating their inclusion as covariates.

Having established the necessary covariates for exploration of the variables of interest, we ran linear models for each of the three psychosocial predictors. Table 1 presents the results of linear mixed effects models, with Benjamini–Hochberg correction [38], at multiple levels of FDR by delta beta (difference in DNA methylation between the minimum and maximum levels of the predictor variable) cut points of 0.05 and 0.10, and with covariate adjustments in each model. In order to remove any residual traces of genetic ethnicity confounding in the DNA methylation data, in addition to including genetic ethnicity as a covariate in our regression models, we also removed CpGs associated with genetic ancestry *post hoc* (see Supplementary Text). Given adjustments for genetic and reported ethnicity and our *post hoc* removal of any remaining genetic ancestry-associated CpGs, we report findings with a moderate FDR of 0.20 and delta beta of 0.05. At this level, we identified 488, 354 and 102 significantly differentially methylated CpGs associated with income-per-dependent, highest household education and family adversity, respectively (Table 1; see Supplementary Table 5 for comprehensive table). These differentially methylated CpGs survived despite stringent corrections for genetic ancestry and self-reported ethnicity, which were correlated with both income and education (Supplementary Table 4).

To test whether residual noise from potential gender differences and blood contamination – while not confounded with the SES factors or family psychosocial adversity – affected our differential DNA methylation signal, we performed sensitivity analyses with alternative models including more covariates than the models presented so far (Supplementary Text, Supplementary Figure 3). Our sensitivity analysis broadly confirmed our findings, as the delta beta effect sizes from both models were highly correlated (r_s = 0.81–0.94; Supplementary Figure 3A). As expected, we did observe difference in p-values between models and a weak correlation of p-values (r_s = 0.34–0.73; Supplementary Figure 3B). Although this demonstrated p-value is sensitive to the covariates included, delta beta values are generally consistent between models, suggesting the differential DNA methylation signal is not driven by noise from gender and blood contamination. In light of this, and because indiscriminatingly fitting a model with all possible covariates could lead to severe over fitting, especially in a sample of our size, we maintained our previous models.

To test for the effect of individuals with more extreme DNAm values at our hits, we performed a 90% winsorization of DNAm values of our hits and ran the same models on the winsorized DNAm values. Delta betas correlated well between original and winsorized DNAm data (r_s >0.99 for all variables), suggesting trends are not driven by extreme DNAm values. Using the nominal p-value cut-offs corresponding to FDR: 0.2 from each variable and a delta beta of 0.05, however, 69, 63 and 11% of hits from income-per-dependent, parental education, family adversity, respectively, remained significant after winsorization (Supplementary Figure 4). The loss of a substantial portion of family adversity hits after winsorization suggests that nonwinsorized adversity associations may be heavily influenced by individuals with DNAm values near the tails of the distributions. Nevertheless, to capture the entire spectrum of associations, we focused on the full set of CpGs in our subsequent analyses.

There was no general directional trend in the differential DNA methylation for any predictor; low income, low education and high family adversity were each associated with both increased and/or decreased DNA methylation at CpG sites (Figure 2). For higher levels of income-per-dependent, 289 and 199 CpG sites demonstrated increased or decreased DNA methylation, respectively. The corresponding numbers of sites for higher levels of parental education, were 60 and 294, and for higher family adversity, were 65 and 37, respectively.
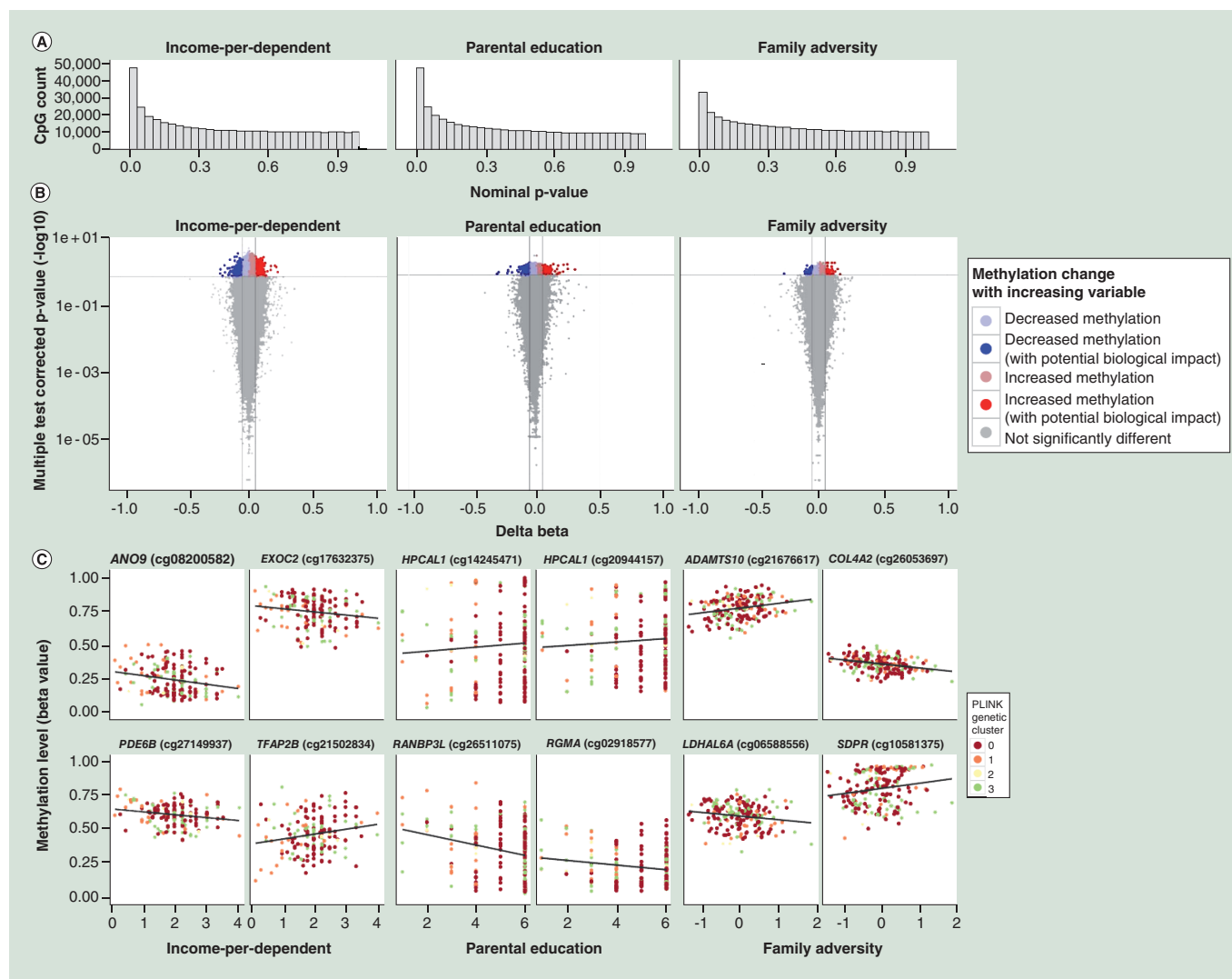
Figure 2.   Variables show different strengths of prediction of DNA methylation at the genome-wide level. (A) Scree plots showing all models had positively skewed nominal p-value distributions. **(B)** Volcano plots show the overall number of CpG hits with each variable. Plots show -log10 multiple test corrected p-values on a log10 scale against delta beta for each CpG (methylation at the highest level of a variable minus methylation at the lowest level of a variable). Horizontal lines show a false discovery rate of 0.2 and vertical lines show a delta beta of 0.05 or -0.05. **(C)** Plots show the relationship between methylation and the variable of interest at representative CpGs. Each plot is labeled with the CpG ID and the associated gene. Lines show a linear model fit through the data. See also Supplementary Figures 1–3 and 5.

Given that nearby CpG loci are often correlated, we explored regional associations with SES factor by examining trends at CpGs within 1 Kb of significantly associated CpGs. While 28–30% of CpGs identified in the original EWAS did not have an adjacent CpG within 1 Kb, when present, the delta betas of adjacent CpGs correlated with the significant EWAS CpGs (Supplementary Figure 5). This suggested that our SES-related factors associated with broader differentially methylated regions beyond the CpGs identified as significant by the epigenome-wide association study (EWAS).

Pyrosequencing verified EWAS findings

To technically verify the data from the 450K, we tested selected CpGs using an independent platform, quantitative bisulfite PCR pyrosequencing (see Supplementary Text). Three CpGs were chosen that had the greatest change in methylation across each one of the predictors (delta beta >0.13). Pyrosequencing results verified DNA methylation levels measured on the 450K array based on a strong correlation, low RMSE and low methylation differences
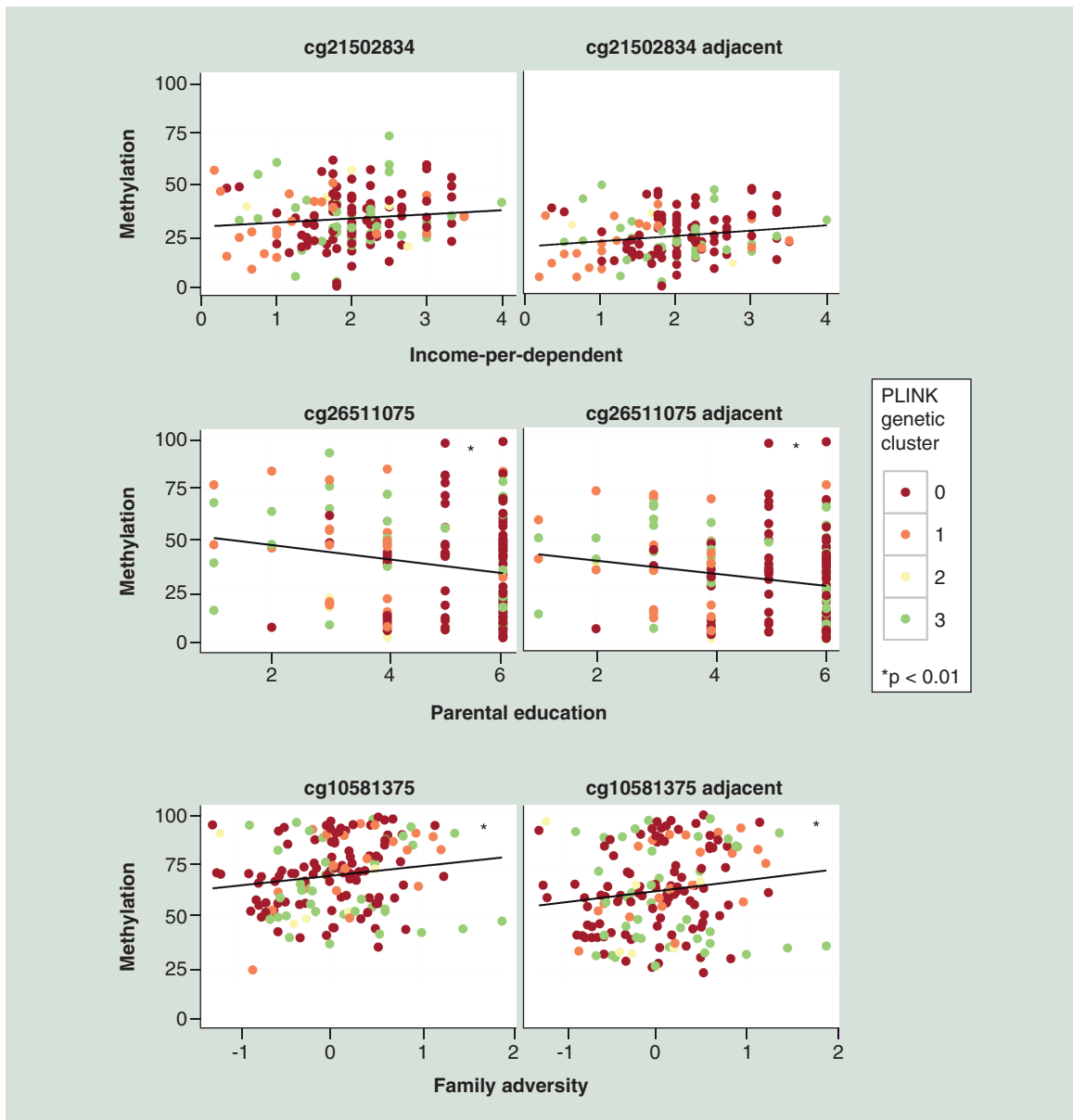
Figure 3.   DNA methylation of adjacent CpGs captured by pyrosequencing. Plots show the relationship between methylation and the variable of interest, at the originally pyrosequenced CpG and at an adjacent CpG captured in each assay. Colors show the genetic ancestry of a sample as used throughout the analysis. Asterisks indicate significant associations in the pyrosequencing data at $p < 0.01$.

between measures from both technologies (r_s = 0.82–0.89; RMSE: 0.11–0.16; Supplementary Figure 6). One additional CpG adjacent to the target CpGs was captured in each of the three pyrosequencing assays. The additional CpGs all correlated with the originally examined CpGs, though associations with income did not reach statistical significance ($p < 0.01$; Figure 3).

Differentially methylated CpGs were depleted in gene promoters & CpG islands

Taking advantage of the design of the 450K array, which often contains more than one CpG per gene, we determined next how many genes were associated with our variables by grouping CpGs belonging to the same gene (see Supplementary Text). Of our differentially methylated CpGs that were significantly associated with income-per-dependent, the highest household education or adversity, 309, 252 and 68 CpGs were associated with 267, 229 and 63 genes, respectively. Genes with multiple CpGs identified with income-per-dependent, the highest
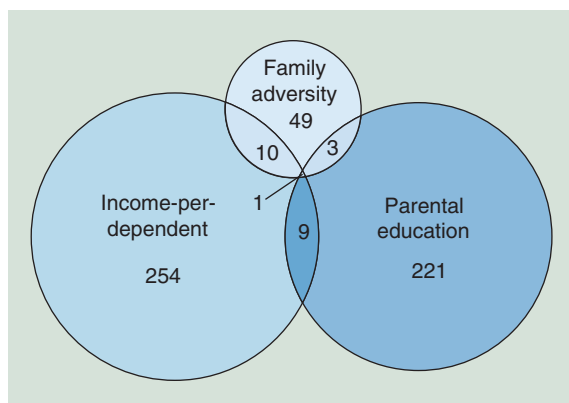
Figure 4.   Venn diagram representing the number of genes associated with significant hit CpGs. False discovery rate of 0.2 and delta beta of 0.05 from Table 1. Only associations that survived very conservative multiple test correction for each of the three variables tested are presented.

household education and adversity had a maximum of six, five and three significant same-gene CpGs associated with those predictors, respectively; 182, 102 and 35 CpGs associated with those predictors, respectively, were not located proximally to any gene and, therefore, were classified as intergenic. The CpGs identified with each predictor were significantly depleted in gene promoters, enriched in intergenic regions and depleted in CpG islands (see Supplementary Text; Supplementary Figure 7; FDR: 0.05). As little is currently known about the function of intergenic CpGs, interpretation can be challenging. Therefore, in line with commonly used approaches in the literature [21,22,25,32], we have focused our interpretation on the CpGs localized to gene features and explored the characteristics of the genes with which they were linked.

Predictors had distinct DNA methylation associations

It was important to ascertain whether the three theoretically related and moderately intercorrelated predictors would vary in which genes they associated with differential DNA methylation. We found that the CpG-associated gene lists for each of the three predictors (income, education, adversity) overlapped at a level consistent with the operation of chance (FDR: 0.4), suggesting that income, education and adversity each had a unique pattern of association with the DNA methylome (Figure 4). Income-per-dependent was the strongest predictor, in terms of number of genes associated with significant differences in CpG methylation. Of note, it was surprising that at this level of analysis, the SES factors were stronger predictors of DNA methylation than was the family adversity composite, which is typically thought to reflect factors more proximal to the child [39].

Health-related biological functions were moderately associated with DNA methylation

To further understand the underlying biology of these associations, we conducted exploratory GO enrichment analysis on the lists of differentially methylated genes using ErmineJ [40], which uses established annotations of functional categories (GO terms) for genes to find shared functional roles among genes. GO enrichment analyses conducted for the genes identified in this study showed few significantly over- or underrepresented GO terms (categories of known gene functions) when adjusted for multiple comparisons (see Supplementary Table 6). Many of the top GO terms enriched for associations with family adversity and income-per-dependent, however, were related to immune functioning (particularly T-cell and immunoglobulin functions), and many of the top GO terms for parental education were related to regulation of developmental processes.

Discussion

There is prior, robust evidence that socioeconomic factors are related to health across the life span and that childhood SES and genomic DNA methylation are associated in adult samples [22,23,25]. Based on these antecedent observations, we hypothesized that SES associations would be detectable in 9-year-old children's genomic DNA methylation patterns. Using a genome-scale approach not focused on candidate genes, this study revealed four main findings. First, we found that two dimensions of a kindergartner's family SES – income-per-dependent and parental education – were each associated with broad and substantial variation in children's genome-wide DNA methylation a few years later in development. Second, exposure to family adversity also showed significant linkages to later DNA methylation, though weaker and potentially driven by individuals with outlier values of DNA methylation. Third, the patterns of associations among these three early-life predictors and DNA methylation were less overlapping

than would be expected given their levels of intercorrelation, suggesting unique pathways for relations between various early-life social environmental factors and transcription regulation biology. Finally, the results also point to the potential prospective association between early childhood SES and adversity and the DNA methylation status of children's genomic DNA within health-relevant gene pathways.

Genomic sites demonstrated both increased and decreased DNA methylation in association with increased early-life risk (low income and education and high family adversity), which is consistent with other evidence for bidirectional patterns of association [24,25] and specificity in whether social exposures are related to up- or downregulation of specific gene CpGs. Of possible importance, the number of differentially methylated CpGs associated with family adversity was smaller than the number linked to either income or education and less robust to outlier removal, suggesting the relative importance of aspects of SES unrelated to adversity exposures. Nonetheless, our findings relating family adversity to youth buccal epithelial cell DNA methylation are consistent with a prior study linking family adversity (in the form of parental stressors) to buccal epithelial cell methylation in adolescence [30]. However, that study used less advanced array technology than was available in the current research and was conducted within a more socioeconomically advantaged sample, limiting opportunities for analysis of SES in those models.

Our findings move beyond extant data on childhood family adversity associations with later-life DNA methylation by demonstrating prospective, longitudinal associations early in development (early-to-middle childhood), within a larger and more ethnically and socioeconomically diverse sample, and with consideration of multiple social environment exposures that may confer risk. The patterns of associations found here, if replicated, suggest that macrolevel social factors outside of family stressors may relate to DNA methylation more strongly than a composite of family adversities such as maternal depression, marital conflict and anger expression in the home. Although substantial evidence points to family processes as key mediators of SES effects on child neurodevelopment [41], findings from the current study are commensurate with large-scale population-based studies that have found that early-life SES is a strong predictor of adult cumulative allostatic load (a marker of biological risk) [42] and adult health and wellbeing [29]. It is notable that findings in our study were sustained after simultaneously adjusting for the effects of both genetic ancestry and of self-reported ethnic minority status, which are commonly related to SES [43–47] and recently recognized as unique predictors of DNA methylation [36]. This provides greater confidence in the effects found here and allows for discovery of SES effects within a diverse sample, which is uncommon in previous research in this domain [21–23,25]. Moreover, it is important to note that associations were found in our sample with a broad range of SES (particularly considering the very high cost of living in the region) and adversity but relatively low representation of individuals with great deprivation and extreme family adversity, and associations are likely to be stronger in larger samples with greater sampling of families with very high levels of risk. The substantial reduction of significant findings for the adversity model with windsorization, in particular, suggest the need for this. More substantive associations might also occur with explicit measures of children's direct exposure to severe adversities, such as child abuse, rather than more proximal measures of family adversity. However, social disparities in health occur across the full gradient of socioeconomic and adversity conditions, and documentation of associations at all levels of those spectra is important, as was found here.

Although somewhat difficult to compare, given the differences in DNA methylation measurement technologies and statistical approaches, the number of significant associations after corrections appeared to be stronger than many previously published for SES and methylation [25]. While we saw more DNAm associations than previous comparable studies, we note that our results did mirror, to some extent, the findings that differentially DNAm associated to SES do not exclusively occur at isolated CpGs but might often involve regions of correlated DNAm [25]. Potential reasons for our greater detection of associations could include: our larger sample size compared with most of the prior work, increasing our chances of detecting associations, our more detailed measures of SES and adversity compared with prior studies that have relied on a single item or a dichotomized item, our use of SES and adversity measures collected in childhood to prospectively predict later biological outcomes, the greater ethnic and socioeconomic diversity of our sample, offering more substantial variability in predictors, use of an unbiased genome-wide approach rather than a select group of candidate gene loci and utilization of the 450K array, rather than the less advanced technologies used by the two previous genome-wide publications, provided greater coverage of CpG sites.

The functional significance of the SES-associated differences found here is yet to be determined, but it is tempting to speculate that they may serve as one link between SES and risk for the development of disease. Our exploratory GO enrichment analyses suggested that the majority of the sites identified as differentially methylated by the

three predictors played roles in immune functioning and the positive (i.e., initiating or amplifying) regulation of morphogenesis and neurodevelopmental processes. There was minimal overlap in the genes predicted by the three predictors, yet apparent ontological overlap in some of the sites identified across all three. The GO term findings correspond to experimental evidence demonstrating early-life SES associations with DNA methylation of immune gene loci associated with inflammatory biomarkers [24], susceptibility to infection and illness [3] and low SES associations with proinflammatory genomic transcription profiles [6]. Other studies find that childhood SES relates to clinically relevant, structural differences in brain [7] and changes in glucocorticoid signaling [6]. Collectively, these findings implicate potential biological pathways for SES-linked DNA methylation effects on health and development and suggest the possibility of clinically relevant processes.

Limitations & strengths

Confidence in these findings should be tempered by an awareness of the reported study's shortcomings and the technical and interpretive complexities of this research [34]. Although data were prospectively acquired and analyses guided by specific hypotheses, we are limited in our ability to draw causal conclusions from nonexperimental, correlational data, particularly as DNA was collected only once during middle childhood. Additionally, buccal epithelial cells are a tissue of debatable optimality in assessing DNA methylation. Although some advocate for its use due to its easy accessibility and greater representation of hypomethylated DNA regions, others argue that peripheral blood mononuclear cells are more representative of the DNA methylation patterns related to demographic variation [26,48]. Early adversity has broad effects, with consequences on endocrine, immune and neural functions [49,50], and accompanying epigenetic changes are likely observed in a number of tissues; some of these effects are likely to be shared across multiple tissues, thus detectable in peripheral tissues, such as buccal, though not necessarily identical to those found in other tissues of interest. While debate over surrogate tissue use is ongoing, we believe that buccal samples are an appropriate surrogate to identify biomarkers of a systemic embedding of social stress, particularly in very young children. Focusing on an relatively homogeneous tissue limits the noise attributable to cell heterogeneity in our data [51], but the observed differential methylation is still challenging to interpret on the cellular level.

With the exception of imprinted genes, DNA methylation at individual CpGs is a binary signal in individual cells (CpGs are either methylated or unmethylated). However, in DNA methylation studies, we most often measure a population of cells and can therefore observe DNA methylation in the intermediate range of methylation values (e.g., methylation level of 0.75), which can be interpreted as reflecting cellular mosaicism in the tissue sample with respect to a particular DNA methylation mark. The partial changes in DNA methylation reported here and in the majority of most EWAS studies, thus likely reflect a shift in subpopulations of buccal epithelial cells [52]. While such mosaic shifts associating with SES-related stress are interesting, speculation on mechanisms are complicated by the potential effect of both cellular level changes and shifting subpopulations of buccal epithelial cells.

In a further challenge, almost nothing is known about the mechanisms by which SES and family adversity might affect DNA methylation or shifting cell populations, and our study did not identify the presumed molecular mediators of the noted associations. Our findings may be the consequences of SES-related differences in exposures to physical factors such as nutrition and toxins, differences in children's experiences of trauma or stress or myriad other pathways. Understanding those factors is a priority for research in this domain, as is performing transcriptomic studies to improve interpretability of the causes and consequences of changes in transcription regulators [34]. In addition, next steps in this research include incorporating potential genetic heritability for risk of low SES [53] within models examining the role of social environment on DNA methylation patterning. Furthermore, the genome-scale analyses described here, using the data acquired from the specified methylation array, capture only a fraction of the 28 million CpGs within the human genome. Future work will broaden our understanding as technology and representation of CpG sampling across the genome and issues of addressing multiple testing advance. Finally, an important limitation is our lack of ability to identify a cohort of comparable developmental age and sufficient sample size with breadth of SES and family adversity measurement in which we could attempt to replicate our results, which is critical for increasing confidence in the findings. As the popularity of assessing DNA methylation patterns in young children increases (the authors themselves have such work underway), tests for replication of these findings will clarify the robustness of the associations found, particularly if conducted in larger, epidemiologic cohorts [54].

Conclusion

Despite the above limitations, our study's design conformed to many of the suggestions for improving interpretability of epigenome-wide data [34]. These include *a priori* hypothesis testing, accounting for cell-type heterogeneity, examination of effect sizes of differential methylation and adjustment for population genetic sequence variability due to ancestry. We also found a substantial level of differences in DNA methylation among children of divergent SES and adversity exposure, suggesting DNA methylation may serve as a biomarker of early-life environment.

It is important to emphasize that these findings do not suggest that children's early-life contextual influences are invariant or deterministic, as broad variation in the effects of adversity on development are typical, and many individuals show remarkable resilience to adverse experiences [55]. However, in addition to its potential mediating role, DNA methylation patterns associated with early-life adversity may represent one important moderator of social environmental influences on health and development across the life course.

Future perspective

Our findings also do not imply that potential effects are immalleable. Rather, given the potential 'reversibility' of epigenetic modifications shown in rodent models [56], it is tempting to speculate that, even beyond the critical developmental periods of childhood, social environmental calibration of gene expression may be possible. Yet extant data reveal that associations between childhood SES and DNA methylation, as we found here, can be lasting and predict adult DNA methylation, suggesting potential sensitive periods in early development. Even in the context of immutable changes in DNA methylation status for some genetic loci, interventions to reduce adversity or improve adaptation may to lead to relevant biological and social benefits that could override early-life effects. Longitudinal sampling and DNA methylation measurements will be necessary to rigorously test these ideas. Moreover, careful consideration of the intersection of DNA methylation with allelic variation to determine what fraction of epigenetic variation is genetic, environmental, G × E or G + E [57] is needed to advance understanding in these realms. In addition, measuring epigenetic marks beyond DNA methylation is also likely to be increasingly informative, as is the integration of different epigenetic marks with other 'omics measures (e.g., immune markers, mRNA). Deeper examinations, such as measuring DNA methylation in single cells to decipher the fundamental structure and relationship to variables beyond examination of cells in bulk, are also needed. In the meantime, many social scientists agree that developing policies that minimize adversities early in life and reduce social disparities in health and related biological outcomes is the best plan for optimizing the health and development of societies [58]. The documentation and identification of endophenotypic biomarkers of effects of early adversity, such as global DNA methylation during childhood, may advance efforts in this regard.

Summary points

- Individual socioeconomic status is one of the strongest predictors of health.
- Prospective evidence for potential mechanisms of the 'biological embedding' of early-life stress is limited, particularly in young children.
- This study examined variation in child buccal epithelial cell DNA methylation to assess its potential as a pathway for effects of childhood social adversity on health across the life course.
- In a diverse, prospective community sample of 178 kindergarten children, associations between three types of social experience and DNA methylation later in childhood were examined.
- Family income, parental education and family psychosocial adversity each associated with increased or decreased DNA methylation within a unique set of genomic CpG sites, after accounting for genetic ancestry, self-reported ancestry and other key covariates.
- Gene ontology analyses pointed to genes serving immune and developmental regulation functions.
- Findings provided support for DNA methylation as a biomarker linking early-life social experiences with later-life health in humans.

Author's contributions

All authors have written and approved the manuscript. WT Boyce and NE Adler designed the original cohort study, and WT Boyce supervised the collection of the questionnaire data for the study. NR Bush conceived and designed manuscript study question, received the funding for the genetic component of the research, led the collection of the DNA samples in a follow-up study to the original cohort study, and led the development of the manuscript. JL MacIsaac and LM McEwen were involved in generation of experimental data, and M Park assisted with the data cleaning. RD Edgar and M Park performed data analyses. JL McEwen and LM McEwen performed all pyrosequencing analyses. NR Bush wrote the introduction and discussion, wrote a portion of the methods and results, created a table and edited the manuscript. RD Edgar wrote a portion of the methods and results, created the tables and figures, and edited the manuscript. MJ Essex and NE Adler contributed to theory development and interpretation of results and edited the manuscript. MS Kobor and WT Boyce supervised the work and commented on the manuscript at all stages.

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest

1.   Adler NE, Stewart J. Health disparities across the lifespan: meaning, methods, and mechanisms. *Ann. NY Acad. Sci.* 1186, 5–23 (2010).

●●   **Reviews the robust evidence for social disparities in health and the variety of measurement factors to consider when assessing disparities.**

2 .   Braveman P, Egerter S. Overcoming obstacles to health: a report to the Robert Wood Johnson Foundation Commission to Build a Healthier America. Robert Wood Johnson Foundation, Princeton, NJ, USA, 1–80 (2008).

3 .   Cohen S, Doyle WJ, Turner RB, Alper CM, Skoner DP. Childhood socioeconomic status and host resistance to infectious illness in adulthood. *Psychosom. Med.* 66(4), 553–558 (2004).

4.   Miller G, Chen E. Unfavorable socioeconomic conditions in early life presage expression of proinflammatory phenotype in adolescence. *Psychosom. Med.* 69(5), 402–409 (2007).

5.   Hertzman C. The biological embedding of early experience and its effects on health in adulthood. *Ann. NY Acad. Sci.* 896, 85–95 (1999).

6.   Miller GE, Chen E, Fok AK *et al.* Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling. *Proc. Natl Acad. Sci. USA* 106(34), 14716–14721 (2009).

7.   Noble KG, Houston SM, Brito NH *et al.* Family income, parental education and brain structure in children and adolescents. *Nat. Neurosci.* 18(5), 773–778 (2015).

8.   Bush NR, Boyce WT. The contributions of early experience to biological development and sensitivity to context. In: *Handbook of Developmental Psychopathology*. Lewis M, Rudolph KD (Eds). Springer, NY, USA, 287–309 (2014).

9.   Hertzman C, Boyce WT. How experience gets under the skin to create gradients in developmental health. *Ann. Rev. Public Health* 31, 329–347 (2010).

•   **Reviews the evidence for associations between early-life social experiences and biomarkers of stress physiology, such as the hypothalamic pituitary adrenal axis, as well as the potential role of epigenetic changes.**

10. Miller GE, Chen E, Parker KJ. Psychological stress in childhood and susceptibility to the chronic diseases of aging: moving toward a model of behavioral and biological mechanisms. *Psychol. Bull.* 137(6), 959–997 (2011).

•   **Provides a robust review of evidence for early-life stress and risk for later problems with health, pointing to key biological mechanisms for the associations.**

11. Farah MJ, Shera DM, Savage JH *et al.* Childhood poverty: specific associations with neurocognitive development. *Brain Res.* 1110(1), 166–174 (2006).

12. Sheridan MA, Sarsour K, Jutte D, D'esposito M, Boyce WT. The impact of social disparity on prefrontal function in childhood. *PLoS ONE* 7(4), e35744 (2012).

13. Chen E, Cohen S, Miller GE. How low socioeconomic status affects 2-year hormonal trajectories in children. *Psychol. Sci.* 21(1), 31–37 (2010).

14. Shonkoff JP, Garner AS, The Committee on Psychosocial Aspects of Child and Family Health *et al.* The lifelong effects of early childhood adversity and toxic stress. *Pediatrics* 129(1), e232–e246 (2012).

15. Alkon A, Boyce WT, Tran L, Harley KG, Neuhaus J, Eskenazi B. Prenatal adversities and Latino children's autonomic nervous system reactivity trajectories from 6 months to 5 years of age. *PLoS ONE* 9(1), e86283 (2014).

16. Bush NR, Obradovic J, Adler N, Boyce WT. Kindergarten stressors and cumulative adrenocortical activation: the "first straws" of allostatic load? *Dev. Psychopathol.* 23(4), 1089–1106 (2011).

17. Blair C, Raver CC, Granger D, Mills-Koonce R, Hibel L. Allostasis and allostatic load in the context of poverty in early childhood. *Dev. Psychopathol.* 23(Special Issue 03), 845–857 (2011).

18 . Ly M, Haynes MR, Barter JW, Weinberger DR, Zink CF. Subjective socioeconomic status predicts human ventral striatal responses to social status information. *Curr. Biol.* 21(9), 794–797 (2011).

19. Meaney MJ. Epigenetics and the biological definition of gene X environment interactions. *Child Dev.* 81(1), 41–79 (2010).

20. Boyce WT, Kobor MS. Development and the epigenome: the 'synapse' of gene–environment interplay. *Dev. Sci.* 18(1), 1–23 (2015).

••   **Summarizes, for an audience of developmental researchers and clinicians, how epigenetic processes may offer a strong, parsimonious account for the convergence of genetic and contextual variation in the genesis of adaptive and maladaptive development.**

21. Lam LL, Emberly E, Fraser HB *et al.* Factors underlying variable DNA methylation in a human community cohort. *Proc. Natl Acad. Sci. USA* 109(Suppl. 2), 17253–17260 (2012).

22. Borghol N, Suderman M, McArdle W *et al.* Associations with early-life socio-economic position in adult DNA methylation. *Int. J. Epidemiol.* 41(1), 62–74 (2012).

23. Tehranifar P, Wu HC, Fan X *et al.* Early life socioeconomic factors and genomic DNA methylation in mid-life. *Epigenetics* 8(1), 23–27 (2013).

24. Mcdade TW, Ryan C, Jones MJ *et al.* Social and physical environments early in development predict DNA methylation of inflammatory genes in young adulthood. *Proc. Natl Acad. Sci. USA* 114(29), 7611–7616 (2017).

••   **Presented an approach to social epigenetic studies, such as our own, that may guide researchers through providing candidates and hypothesis for validation.**

25. Stringhini S, Polidoro S, Sacerdote C *et al.* Life-course socioeconomic status and DNA methylation of genes regulating inflammation. *Int. J. Epidemiol.* 44(4), 1320–1330 (2015).

26. Smith AK, Kilaru V, Klengel T *et al.* DNA extracted from saliva for methylation studies of psychiatric traits: evidence tissue specificity and relatedness to brain. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 168B(1), 36–44 (2015).

27. Adler N, Bush NR, Pantell MS. Rigor, vigor, and the study of health disparities. *Proc. Natl Acad. Sci. USA* 109(Suppl. 2), 17154–17159 (2012).

28. Lengua LJ, Bush NR, Long AC, Kovacs EA, Trancik AM. Effortful control as a moderator of the relation between contextual risk factors and growth in adjustment problems. *Dev. Psychopathol.* 20(2), 509–528 (2008).

29 . Sheikh MA, Abelsen B, Olsen JA. Clarifying associations between childhood adversity, social support, behavioral factors, and mental health, health, and well-being in adulthood: a population-based study. *Front. Psychol.* 7, 727 (2016).

30. Essex MJ, Boyce WT, Hertzman C *et al.* Epigenetic vestiges of early developmental adversity: childhood stress exposure and DNA methylation in adolescence. *Child Dev.* 84(1), 58–75 (2013).

31. Fraser HB, Lam LL, Neumann SM, Kobor MS. Population-specificity of human DNA methylation. *Genome Biol.* 13(2), R8 (2012).

32. Heyn H, Moran S, Hernando-Herraez I *et al.* DNA methylation contributes to natural human variation. *Genome Res.* 23(9), 1363–1372 (2013).

33. Lam LL, Emberly E, Fraser HB *et al.* Factors underlying variable DNA methylation in a human community cohort. *Proc. Natl Acad. Sci. USA* 109(Suppl. 2), 17253–17260 (2012).

34. Birney E, Smith GD, Greally JM. Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genet.* 12(6), e1006105 (2016).
   •• Expertly highlights, for the field of epigenetic research, the limitations faced by the field but often inadequately considered or addressed.

35. Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum. Genomics* 9(1), 1 (2015).

36. Galanter JM, Gignoux CR, Oh SS *et al.* Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *eLife* 6, e20532 (2017).
   • Thoughtfully demonstrated the importance of accounting for both ancestry and ethnicity independently, as a molecular characteristic and as a social construct, respectively.

37. Jones MJ, Farré P, McEwen LM *et al.* Distinct DNA methylation patterns of cognitive impairment and trisomy 21 in down syndrome. *BMC Med. Genomics* 6, 58 (2013).

38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B (Methodol.)* 57(1), 289–300 (1995).

39. Martin P, Martin M. Proximal and distal influences on development: the model of developmental adaptation. *Dev. Rev.* 22(1), 78–96 (2002).

40. The Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucl. Acids Res.* 43(database issue), D1049–D1056 (2015).

41. Hackman DA, Gallop R, Evans GW, Farah MJ. Socioeconomic status and executive function: developmental trajectories and mediation. *Dev. Sci.* 18(5), 686–702 (2015).

42. Friedman EM, Karlamangla A, Gruenewald T, Koretz B, Seeman T. Early life adversity and adult biological risk profiles. *Psychosom. Med.* 77(2), 176–185 (2015).

43. Williams DR, Mohammed SA, Leavell J, Collins C. Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. *Ann. NY Acad. Sci.* 1186(1), 69–101 (2010).

44. Kuzawa CW, Sweet E. Epigenetics and the embodiment of race: developmental origins of US racial disparities in cardiovascular health. *Am. J. Hum. Biol.* 21(1), 2–15 (2009).

45 . Florez JC, Price AL, Campbell D *et al.* Strong association of socioeconomic status and genetic ancestry in Latinos: implications for admixture studies of type 2 diabetes. In: *Racial Identities, Genetic Ancestry, and Health in South America: Argentina, Brazil, Colombia, and Uruguay.* Gibbon S, Santos RV, Sans M (Eds). Palgrave Macmillan, NY, USA, 137–153 (2011).

46. Peralta CA, Ziv E, Katz R *et al.* African ancestry, socioeconomic status, and kidney function in elderly African Americans: a genetic admixture analysis. *J. Am. Soc. Nephrol.* 17(12), 3491–3496 (2006).

47. Freedland KE, Miller GE, Sheps DS. The great debate, revisited. *Psychosom. Med.* 68, 179–184 (2006).

48. Lowe R, Gemma C, Beyan H *et al.* Buccals are likely to be a more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics* 8(4), 445–454 (2013).

49. Bush NR, Lane RD, McLaughlin KA. Mechanisms underlying the association between early-life adversity and physical health: charting a course for the future. *Psychosom. Med.* 78(9), 1114–1119 (2016).
   • Reviews a range of studies examining potential biological mechanisms for effects of early-life adversity on later-health and provides suggestions for future studies in that domain to advance the field.

50 . Teicher MH, Samson JA. Annual research review: enduring neurobiological effects of childhood abuse and neglect. *J. Child Psychol. Psychiatry* 57(3), 241–266 (2016).

51. Theda C, Hwang SH, Czajko A, Loke YJ, Leong P, Craig JM. Quantitation of the cellular content of saliva and buccal swab samples. *Sci. Rep.* 8(1), 6944 (2018).

52. Berko ER, Suzuki M, Beren F *et al.* Mosaic epigenetic dysregulation of ectodermal cells in autism spectrum disorder. *PLOS Genet.* 10(5), e1004402 (2014).

53. Hill WD, Hagenaars SP, Marioni RE *et al.* Molecular genetic contributions to social deprivation and household income in UK biobank. *Curr. Biol.* 26(22), 3083–3089 (2016).

54. Breton CV, Marsit CJ, Faustman E *et al.* 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.* 125(4), 511–526 (2017).

55. Bush NR, Boyce WT. Differential sensitivity to context: implications for developmental psychopathology. In: *Developmental Psychopathology*. Cicchetti D (Ed.). John Wiley and Sons, Inc., Hoboken, NJ, USA (2016).

56. Meaney MJ, Szyf M. Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues Clin. Neurosci.* 7(2), 103–123 (2005).

57. Teh AL, Pan H, Chen L *et al.* The effect of genotype and *in utero* environment on inter-individual variation in neonate DNA methylomes. *Genome Res.* 24, 1064–1074 (2014).

58. Shonkoff JP. Leveraging the biology of adversity to address the roots of disparities in health and development. *Proc. Natl Acad. Sci. USA* 109(Suppl. 2), 17302–17307 (2012).

59. Obradović J, Bush NR, Stamperdahl J, Adler NE, Boyce WT. Biological sensitivity to context: the interactive effects of stress reactivity and family adversity on socioemotional behavior and school readiness. *Child Dev.* 81(1), 270–289 (2010).

60. Essex MJ, Klein MH, Cho E, Kalin NH. Maternal stress beginning in infancy may sensitize children to later stress exposure: effects on cortisol and behavior. *Biol. Psychiatry* 52(8), 776–784 (2002).

61. Purcell S, Neale B, Todd-Brown K *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81(3), 559–575 (2007).

62. Solovieff N, Hartley SW, Baldwin CT, Perls TT, Steinberg MH, Sebastiani P. Clustering by genetic ancestry using genome-wide SNP data. *BMC Genet.* 11, 108 (2010).

63. Price ME, Cotton AM, Lam LL *et al.* Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin* 6(1), 4 (2013).

64. Teschendorff AE, Marabita F, Lechner M *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 29(2), 189–196 (2013).

65. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1), 118–127 (2007).

66. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67(1), 48 (2015).

67. Gillis J, Mistry M, Pavlidis P. Gene function analysis in complex data sets using ErmineJ. *Nat. Protoc.* 5(6), 1148–1159 (2010).

68. Gillis J, Pavlidis P. The impact of multifunctional genes on "guilt by association" analysis. *PLoS ONE* 6(2), e17258 (2011).