Socioeconomic changes predict genome-wide DNA methylation in childhood

Jiaxuan Liu[1,2,†], Janine Cerutti[1,†], Alexandre A. Lussier[1,4,5], Yiwen Zhu[1,2], Brooke J. Smith[1], Andrew D.A.C. Smith[3], Erin C. Dunn[1,4,5,6,*]

[1] Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA
[2] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA
[3] Mathematics and Statistics Research Group, University of the West of England, Bristol, UK
[4] Department of Psychiatry, Harvard Medical School, Boston, MA
[5] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA
[6] Harvard Center on the Developing Child, Harvard University, Cambridge, MA

†Joint First Authors, who contributed equally to the manuscript.

[*]**Corresponding Author**: Erin C. Dunn, ScD, MPH, Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, 185 Cambridge Street, Simches Research Building 6th Floor, Boston, MA 02114; Email: edunn2@mgh.Harvard.edu; Phone: 617-726-9387; Fax: 617-726-0830; Website: www.thedunnlab.com

Abstract

Childhood socioeconomic position (SEP) is a major determinant of health and well-being across the entire life course. To effectively prevent and reduce health risks related to SEP, it is critical to better understand when and under what circumstances socioeconomic adversity shapes biological processes. DNA methylation (DNAm) is one such mechanism for how early life adversity "gets under the skin". In this study, we evaluated the dynamic relationship between SEP and DNAm across childhood using data from 946 mother-child pairs in the Avon Longitudinal Study of Parents and Children (ALSPAC). We assessed six SEP indicators spanning financial, occupational, and residential domains during *very-early childhood* (ages 0-2), *early childhood* (ages 3-5), and *middle childhood* (ages 6-7). Epigenome-wide DNAm were measured at 412,956 CpGs from peripheral blood at age 7. Using an innovative two-stage structured life course modeling approach, we tested three life-course hypotheses for how SEP shapes DNAm profiles — *accumulation*, *sensitive period*, and *mobility*. We showed that *changes* in the socioeconomic environment were associated with the greatest differences in DNAm, and that middle childhood may be a potential sensitive period when socioeconomic instability is especially important in shaping DNAm. Top SEP-related DNAm CpGs were overrepresented in genes involved in pathways important for neural development, immune function, and metabolic processes. Our findings highlight the importance of socioeconomic stability during childhood and if replicated, may emphasize the need for public programs to help children and families experiencing socioeconomic instability and other forms of socioeconomic adversity.

Introduction

Socioeconomic position (SEP) is a fundamental determinant of health and disease across the lifespan (1). As defined by Krieger et al. (1997) (2), SEP is an "aggregate concept" composed of diverse components of economic and social well-being across individual-, household-, and neighborhood-level domains, including both resources (e.g., weekly income) and rank-based characteristics (e.g., occupational prestige). SEP therefore can be measured across time by various indicators, like job stability, ability to afford basic household needs, and neighborhood quality, which are known to play related, yet distinct roles in health and life outcomes (3-5).

Dozens of observational and quasi-experimental studies examining these indicators have shown that children growing-up in low SEP families have increased risk for both short- and long-term cognitive, socioemotional, behavioral, and physical/mental health deficits compared to their high SEP counterparts (6-9). Some of these SEP-related disparities are evident very early in development, starting shortly after birth (10-13). Yet, the biological mechanisms that explain these well-established SEP and health relationships remain relatively unknown, limiting our ability to disentangle specific pathways of pathophysiology and design targeted interventions.

In the past two decades, epigenetic studies have exploded as a means of potentially unraveling the biological pathways through which SEP "gets under the skin". Most epigenetic studies have focused on DNA methylation (DNAm) (14), which occurs when methyl groups are added to cytosines in the DNA sequence, typically within cytosine-guanine (CpG) dinucleotides (15). These DNA modifications do not alter the sequence of the genome, but can influence how genes are expressed in ways that can have important short and long-term health consequences (16).

Recent reviews summarizing the effects of SEP on epigenetic patterns suggest that SEP is linked to DNAm differences in childhood and adulthood (17-19). In fact, over 30 studies have found a relationship between childhood SEP and DNAm. However, less than a quarter of these studies were longitudinal by design (i.e., including repeated measures of SEP exposure across time). Further, less than half were epigenome-wide association studies (EWAS) analyzing SEP-related DNAm variations. In one recent comprehensive review of the SEP-DNAm literature, the number of significant, SEP-associated CpGs reported across prior EWAS studies ranged from 1 to 2,546 (median = 10), yet relatively no consistent patterns in SEP-associated DNAm changes emerged between studies (see Cerutti, Lussier, Zhu, Liu and Dunn (19)). One possible explanation for these mixed results is that studies have conflated both the *type* of SEP indicator measured and the *timing* of SEP measurement (19). Indeed, few studies have investigated the effects of SEP type and/or timing on DNAm, even though it is well known that both features of SEP can influence the extent of its impact (20).

Prior studies that have analyzed the associations between multiple types of SEP indicators and DNAm have found little to no overlap in DNAm changes across SEP measures (21-23), suggesting that different SEP indicators may result in distinct biological signatures and subsequent cascading health risks. Yet, it remains relatively unknown whether exposure to distinct SEP indicators (e.g., low household income vs. neighborhood disadvantage) during childhood impacts later DNAm to a similar extent.

Even fewer studies have investigated the impact of SEP timing on DNAm, likely because it is difficult to collect multiple, repeated measures across time in large, epigenetic datasets. In some notable exceptions, studies comparing the time-dependent effects of childhood SEP (24-26) on DNAm have found timing differences with respect to SEP's impact, consistent with the

idea that there may be *sensitive periods* of elevated plasticity during childhood when adversity-induced biological changes are most likely to occur. However, whether different aspects of the socioeconomic environment across developmental stages differentially influence DNAm remains largely unexplored.

The current study aimed to address this gap by utilizing a large, longitudinal birth cohort with multiple, repeated measures of socioeconomic-related hardships assessed prospectively across childhood before epigenome-wide DNAm collection at age 7. We specifically sought to assess how different indicators of the socioeconomic environment (e.g., neighborhood quality, job loss, low household income) measured repeatedly across the first seven years of life associated with child epigenetic alterations. Given that different socioeconomic domains may impact health via related, but distinct pathways (4, 27), we analyzed exposure to seven distinct socioeconomic-related hardships. Additionally, because socioeconomic adversity could have multiple time-varying effects on DNAm, we tested three commonly examined hypotheses from the life-course epidemiology literature (28) to evaluate the circumstances under which childhood socioeconomic adversity associates with DNAm changes at age 7: 1) *accumulation* hypothesis, where the impact of low SEP increases with the number of time periods exposed, regardless of when it occurs; 2) *sensitive period* hypothesis, where the impact of low SEP is larger in magnitude during a certain developmental period compared to any other; and 3) *mobility* hypothesis, where the impact of SEP on DNAm is driven by an upward or downward change in SEP between adjacent developmental time periods.

Uncovering the dynamic relationships between SEP and DNAm across childhood will not only highlight the biological mechanisms driving the effects of SEP on long-term health, but

also will offer clearer insights to guide targeted interventions aimed at reducing the negative consequences of socioeconomic-related adversity in childhood.

Results

Sample characteristics and prevalence of socioeconomic adversity

We analyzed data from 946 mother-child pairs from a longitudinal birth-cohort in the United Kingdom (UK). Children included in our analytic sample were mostly White (97.1%) and from both sexes (49.9% female) (**Table S1**). Among the six SEP indicators analyzed (i.e., job loss, income reduction, low family income, financial hardship, major financial problems, and neighborhood disadvantage), job loss was the least reported socioeconomic adversity (11.5% ever-exposed), and income reduction was the most common (73.8% ever-exposed) (**Table 1**). The prevalence of all adversities decreased over time (**Table 1, Figure S1**). The six SEP indicators were moderately correlated with each other during all three childhood periods (**Figure S2**), suggesting they captured distinct aspects of the socioeconomic environment.

Childhood socioeconomic adversities were associated with differential DNAm at 62 CpGs

We next examined possible time-dependent associations between each of the SEP indicators and DNAm at individual CpGs using a two-stage structured life-course modeling approach (SLCMA) (29-31), which identified the life-course hypothesis most supported in the observed data and estimated the associations. In this and the following three sections, we summarize 1) the top CpGs associated with socioeconomic adversity, 2) the most selected life-course hypotheses, 3) the robustness of findings evaluated through a variety of sensitivity analyses, and 4) the biological relevance of findings.

We identified 62 CpGs where exposure to socioeconomic adversity explained more than 3% variance in DNAm (R^2>3%, **Table S2**). Most of the 62 CpGs were linked to the two least commonly-reported adversities in ALSPAC: neighborhood disadvantage (17 CpGs) and job loss (15 CpGs, **Table 2**). Only four of the 62 CpGs identified using the R^2 cutoff also passed an FDR<0.05 significance threshold, all of which were associated with neighborhood disadvantage (**Table 2**).

Of note, 61 of these CpGs showed the same direction of effect as that reported in at least two prior EWASs examining SEP and DNAm. Furthermore, 17 out of 62 (27%) CpGs showed at least a nominal (p<0.05) association in at least two prior EWASs. Of these 17 CpGs, two (cg23685969 and cg19260606) exceeded a statistical significance threshold of FDR<0.05 in at least one prior EWAS (**Table S3, Figure S3**).

Mobility and sensitive period hypotheses were most often selected

The SLCMA allowed us to determine which of the following three life-course hypotheses were most supported in the observed data: *accumulation*, *sensitive period*, and *mobility* (**Figure 1**). Of the life-course hypotheses we tested, *mobility* and *sensitive period* effects showed the strongest associations with DNAm (**Figure 2a**).

We first focused on the four socioeconomic adversities for which we tested all three life-course hypotheses (low family income, financial hardship, major financial problem, and neighborhood disadvantage, **Table S4**). Here, 44 CpGs (R^2>3%) were identified, of which four passed an FDR<0.05 threshold. The majority of CpGs reflected *mobility* (20 CpGs) or *sensitive period* (22 CpGs) relationships. The most selected life-course hypothesis varied by socioeconomic adversity. *Sensitive period* hypotheses were selected for all nine CpGs identified from financial hardship, with middle childhood selected for eight of them (**Figure 2a**). By

contrast, *mobility* (worsening SEP) explained more DNAm variability resulting from neighborhood disadvantage (11 of 17 CpGs) and major financial problem (4 of 5 CpGs). The time period when *mobility* had the greatest impact differed across SEP indicators, with very early to early childhood most often selected for neighborhood disadvantage, and early to middle childhood most selected for major financial problem (**Figure 2a**). *Accumulation* was only selected for two CpGs, linked to low family income. Of note, *mobility* hypotheses were selected for all four FDR-significant CpGs, with a worsening hypothesis (meaning *downward mobility*) selected for three of them (**Table S2**). **Figure 2b** shows at these three CpGs, children exposed to worsening SEP had the greatest shift in DNAm as compared to children with other types of SEP trajectories, including those who had persistently low SEP, worsening SEP, improved SEP, or persistently high SEP.

For our instability indicators (job loss and income reduction), which innately capture the effects of socioeconomic mobility, we only tested *accumulation* and *sensitive period* hypotheses (**Table S4**). The strongest evidence was again for *sensitive period* effects, with middle childhood (age 3-5) most selected for job loss (9 of 15 CpGs) and very early childhood (age 0-2) most selected for income reduction (2 of 3 CpGs, **Figure 2a**). *Accumulation* was only selected for one CpG linked to job loss.

Overall, exposure to socioeconomic changes (captured through instability indicators or mobility hypotheses) was associated with, on average, a 3.8% difference in DNAm levels, explaining 3.4% of the variance in DNAm across CpG sites after controlling for covariates (**Table S2**). The same patterns were found at the epigenome-wide level, with most CpGs showing most variability in response to adversity from *mobility* and *sensitive periods*, rather than the *accumulation* of exposure across development (**Figure S4**).

SLCMA results were robust to sensitivity analyses

Additional covariate adjustment had minimal impact on results

To assess residual bias in the identified SEP-DNAm associations and further ensure the robustness of our findings, we additionally controlled for time-invariant SEP indicators, population substructure estimated from epigenetic data, cord blood DNAm, genetic variation, and exposure to the other five time-varying SEP indicators. After additional covariate adjustments, the life-course hypothesis selected by LARS remained the same for all 62 CpGs with R^2>3% (**Table S5, Table S6**). Almost all CpGs remained significant at the nominal $p<0.05$ threshold after adjusting for time-invariant SEP indicators (60 CpGs), population substructure (61 CpGs), cord blood DNAm (61 CpGs), and exposure to the other five SEP indicators (62 CpGs, **Table S5**). The associations between socioeconomic adversities and DNAm were also independent of genetic variation previously linked to significant CpGs **(Table S6**).

Mobility hypotheses improved our ability to identify CpGs related to SEP changes

SEP *mobility* during childhood had never been previously tested on childhood DNAm to our knowledge. Therefore, we assessed the insights gained from adding mobility hypotheses. We re-analyzed the CpGs with an R^2 >3% for low family income, financial hardship, major financial problem, and neighborhood disadvantage using only *accumulation* and *sensitive period* hypotheses. Considering only accumulation and sensitive period hypotheses, we were unable to fully detect shifts in DNAm patterns related to changes in socioeconomic environment. When mobility hypotheses were omitted from the SLCMA analyses, there were minimal changes to the main results showing effects of *sensitive period* on DNAm (n=22 CpGs), as the same hypothesis was selected with similar effect estimates (**Table S7**). However, for CpGs originally linked to *mobility* (n=20), there were substantial attenuations in the estimated SEP-DNAm associations:

sensitive period hypotheses were selected instead, which in turn, showed smaller R^2 (ranging from 0.04-1.6%) and much larger p-values (ranging from 0.001 to 0.84, **Table S7**). These findings suggest that when the underlying association structure is misspecified, important DNAm signatures may not be identified.

EWAS of ever-exposed vs. never-exposed failed to identify time-dependent associations

To evaluate the loss (or gain) of information from the SLCMA compared to more conventional epigenetic approaches, we performed an epigenome-wide association study (EWAS) of any exposure to each type of SEP adversity before age 7 and DNAm, thus ignoring the timing or change of SEP over time. For 59 of the top 62 CpGs (including the 4 FDR-significant CpGs), the effect estimates from the SLCMA were larger in magnitude than those from EWAS (**Figure S5**). In addition, no CpGs with an FDR<0.05 were identified using EWAS of *any exposure*, meaning ever-exposed vs. never-exposed. These findings suggest the SLCMA was better able to identify developmentally sensitive effects of socioeconomic adversity on DNAm profiles, whereas EWAS might fail to detect signals if the true underlying hypothesis was time-dependent (24).

Biological significance of SLCMA findings

DNAm at significant CpGs was weakly correlated across blood and brain

To examine the relevance of SEP-related DNAm pattern identified in peripheral blood tissues to brain health, we examined the correlation of DNAm at the top 62 CpGs in blood and brain samples, using data from the Blood Brain DNA Methylation Comparison Tool (http://epigenetics.essex.ac.uk/bloodbrain) (32). Overall, DNAm was weakly, but positively, correlated between blood and brain regions (**Table S8**) (prefrontal cortex: r_{avg}=0.06; entorhinal cortex: r_{avg}=0.10; superior temporal gyrus: r_{avg}=0.08; cerebellum: r_{avg}=0.09). Some CpGs showed

particularly strong correlations between blood and brain (e.g., cg24938210, r=0.78 to 0.81 across brain regions).

Distinct biological pathways emerged across SEP indicators

The top 62 CpGs showed no significant differences in distributions of genomic features, CpG island locations, or enhancers, as compared to all tested CpGs (Chi-squared tests p>0.05, **Figure S6**).

Gene set enrichment showed that SEP-related DNAm patterns were more likely to occur within or near genes involved in neural system regulation, developmental processes, immune functions, metabolic processes, substance localization, and membrane transport (**Figure S7, Figure S8**). However, there was little overlap observed in the significant gene ontology (GO) terms across SEP indicators (**Figure S7**), except for one GO term (morphogenesis of a branching epithelium), which emerged in the enrichment analysis for both financial hardship and major financial problem. These findings suggest different socioeconomic adversities may lead to shifts in distinct biological pathways.

Discussion

The main finding from this study was that *changes* in the socioeconomic environment may coincide with subsequent changes at a biological level as measured through DNAm signatures. Reports of a change in the socioeconomic environment, particularly worsening neighborhood quality (i.e., mobility) and parental job loss during middle childhood (i.e., sensitive period), were associated, on average, with a 3.8% difference in DNAm levels. These patterns were detected even after accounting for other dimensions of the socioeconomic environment, ancestry, DNAm levels at birth, and genetic variation. To our knowledge, this

224 study is the first to evaluate the role of socioeconomic *changes* in relation to epigenome-wide
225 DNAm within childhood.

226 Our study extends prior literature on the effects of childhood SEP, providing new insights
227 about the biological embedding of the socioeconomic environment. Only three studies to our
228 knowledge have examined the relationship between socioeconomic mobility and DNAm (22, 33,
229 34). Each of these three studies included just two timepoints of SEP measures, one in childhood
230 and another in adulthood, and only assessed DNAm in adulthood. Our results suggest that acute
231 *changes* in children's socioeconomic environment, compared to exposure to more stable
232 socioeconomic adversity, might play a role in shaping DNAm profiles in childhood as early as
233 age 7. Although our study is the first to measure the impact of exposure to socioeconomic
234 *changes* on DNAm levels in childhood, our results parallel previous findings on SEP-related
235 outcomes in the child development literature. For example, non-epigenetic studies focused on
236 other SEP-related outcomes in childhood have shown that an episode of parental job loss may
237 have a larger impact on child health and behavior than stable employment in low-income jobs
238 (35-37). Indeed, the developmental literature largely suggests that children benefit from stable,
239 predictable environments (38-40) and that changes in the socioeconomic environment can impact
240 cognitive development and other mechanisms implicated in future risk of health and behavioral
241 problems (35-37, 41, 42). Future studies are needed to replicate our findings and investigate how
242 SEP-associated DNAm alterations may influence subsequent health and behavioral outcomes.
243 Insights from such studies will be critical to discern whether SEP-related DNAm changes
244 influence children's vulnerability to disease and other negative health/behavioral outcomes.

245 We found more evidence for the importance of the developmental timing of SEP on
246 DNAm rather than its accumulation. These results parallel previous findings from the ALSPAC

cohort (24) and elsewhere (43), suggesting that sensitive period effects can be detected in the epigenome. Our results also specifically point to the importance of middle childhood as a potential *sensitive period* when the socioeconomic environment might be particularly impactful. SEP plays an important role during school-age years (38, 44), corresponding to our middle childhood time period findings, when children in the cohort began school. Socioeconomic disruptions during school-age years may lead to changes in parent-child interactions, afterschool care center attendance, or extracurricular activities.

Consistent with prior epigenome-wide studies (21, 22), we found little overlap between the top CpGs across SEP domains, suggesting that various aspects of the SEP construct may trigger distinct mechanisms that lead to different alterations in DNAm patterns (19, 45). Across our six SEP indicators, the greatest number of detected CpGs (17 of 62) were related to neighborhood disadvantage, with 4 being the only CpGs to pass an FDR<0.05 significance threshold. These findings point to the important role that neighborhood-level indicators, including more ubiquitous social and physical exposures experienced daily by larger segments of a population, may play in shaping the epigenome during child development. For example, we found that the DNAm alterations linked to neighborhood disadvantage were more likely to occur in genes related to peroxisomes, which are a key component of the biological response to various environmental pollutants (46). By contrast, we found that experiences of financial hardship (e.g., difficulty in affording common household necessities like food, clothing, heat, and rent) and income reduction were linked to biological pathways related to diet quality, such as nutrient transport and metabolic processes. Overall, different clusters of biological pathways emerged across distinct DNAm-associated SEP domains, suggesting that socioeconomic adversities may affect child health through multiple mechanisms.

Many of the genes in which our top CpGs were located on or near have been linked to human health and disease. For example, OAS3, in which our most significant CpG (cg20102336) resides, encodes an enzyme that plays a critical role in innate antiviral response (47) has been linked with the incidence and severity of illness caused by coronavirus disease 2019 (COVID-19) (48, 49). TGFBR3, the nearest gene to another significant CpG (cg08638097), encodes a key receptor in the transforming growth factor-β (TGF-β) superfamily signaling pathways and has been implied in various human cancers including prostate cancer and bladder cancer (50-53). Furthermore, one of the top CpGs showing strong evidence of replication across studies (cg24121967; same direction of effect and $p<0.05$ in 8 and 3 other studies, respectively) was located in a putative oncogene MYEOV whose overexpression has been documented in many cancers such as gastric cancer (54), myeloma (54), and pancreatic cancer (55). These findings suggest that early life socioeconomic adversities are associated with biological disruptions that may ultimately lead to a wide constellation of health risks later in life.

While the current study uncovered many insights into SEP and DNAm associations, a major unanswered question is whether these DNAm changes are adaptive or maladaptive, in both the short- and long-term. Teicher and others have noted that early neurobehavioral changes that occur in response to experiences of childhood adversity often enhance immediate survival at the cost of long-term functioning (56). Thus, are specific epigenomic fluctuations in the face of family socioeconomic adversity reflective of increased risk, resilience, or both? Although we found DNAm differences when comparing children who were exposed vs. unexposed to socioeconomic adversity, we do not know if these SEP-induced shifts represent systemic alterations of biological functions across tissue types, which may cause key impairments that lead to behavioral changes and increase disease risks. With existing publicly available data, we

could only compare the potential implications of our findings to DNAm levels in brain tissue. Additional research comparing DNAm levels between different tissues is warranted to better understand the systemic effects of socioeconomic hardship.

Should these DNAm markers of socioeconomic adversity be replicated and identified as harmful (rather than adaptive) to health, our findings suggest at least two paths forward for prevention and intervention. First, our results suggest that children and families, especially lower-income families who may lack a safety-net to draw from during times of parental job loss or other socioeconomic transitions (57), might benefit from extending policies and social programs aimed at minimizing socioeconomic instability, such as the Supplemental Nutrition Assistance Program (58) and the American Families Plan (59). Second, prevention programs aimed at promoting socioeconomic stability during childhood might benefit from adopting a multisystemic approach that considers the social determinants of health (60) at multiple levels (61). In fact, interventions at the household-level (e.g., parenting-based) and neighborhood-level (e.g., community-based) have revealed measurable biological impacts on children's DNAm profiles (62, 63) and on other biomarkers (64-66).

The current study should be interpreted in light of several limitations. First, like other epigenome-wide studies of this sample size, we identified few specific CpGs passing a stringent correction for multiple testing. However, following the recent movement to move beyond p-value thresholds alone (67, 68), we explored the patterns and implications of SEP-related DNAm profiles among top CpGs passing an effect-size-based threshold. The top CpGs passing this threshold were robust to various sensitivity analyses, and there was consistent evidence for the patterns of CpGs observed, with the majority showing effects in the same direction as previously published findings and two CpGs showing significance in other studies after correcting for

multiple testing. Nevertheless, the results from individual CpGs should be interpreted with caution and validated in larger samples. Second, because this was a population-based sample, extreme cases of socioeconomic disadvantage were likely underrepresented in the ALSPAC cohort. Our results suggest that more severe forms of adversity may have more potent effects, as we identified most top DNAm CpGs (32 out of 62) from the two socioeconomic adversities that showed the lowest prevalence (job loss and neighborhood disadvantage). Future research in populations with more diverse SEP distributions capturing a wider gradient (i.e., extreme poverty) will help fully disentangle the impact of SEP on DNAm patterns. Third, the ALSPAC cohort is mostly White, which limits generalizability of these findings to other individuals and populations of non-European descent. Prior studies (see review (69)) show ancestry-related variation in DNA methylation that may lead to differences in gene regulation across populations. Thus, future replication efforts are needed in more diverse and representative populations. Finally, this study was observational and based on self-report measures of SEP, which could have been influenced by reporter bias, wherein participant responses may have been shaped by factors like social desirability or recall biases, leading to over- or under-estimates of observed associations (70). Although self-reporting bias is common among survey/questionnaire data in observational studies, previous research has shown that individual-level SEP measures like education and income, compared to more objective measures assessed at the census tract-level, can more accurately capture the impact of SEP on a number of health outcomes, such as blood pressure and height (71). Future randomized experiments will help determine the causal effect of socioeconomic adversity on DNAm.

In summary, this study adds to a growing literature showing that early-life socioeconomic adversity can leave biological memories in the form of DNAm differences in childhood.

Uniquely, our findings on socioeconomic mobility and instability suggest changes in the socioeconomic environment during childhood are especially impactful and associated with epigenetic disruptions related to various health outcomes. Ultimately, these findings will enable researchers to build towards better intervention and prevention efforts aimed at reducing socioeconomic disparities and promoting health across the life course.

Materials and Methods

Sample and procedures

Data came from the Accessible Resources for Integrated Epigenomics Studies (ARIES) (72), a subsample of 1,018 mother-child pairs from the Avon Longitudinal Study of Parents and Children (ALSPAC). ALSPAC is a prospective, longitudinal birth-cohort in the UK designed to investigate genetic and environmental determinants of health across the lifespan (73-75). Women living in the county of Avon, UK with estimated delivery dates between April 1991 and December 1992 were invited to participate. Mother-child pairs in the ARIES were randomly selected from ALSPAC based on availability of DNA samples across five waves of data collection (72). We analyzed data from 946 singletons in ARIES with blood-based DNAm profiles generated at age 7. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee. Please note that the ALSPAC study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data). See **Supplemental Methods** for full ALSPAC details.

Measures

<u>Early-life socioeconomic position (SEP)</u>

We analyzed six SEP indicators, spanning financial, occupational, and residential domains: 1) job loss, 2) income reduction, 3) low family income, 4) financial hardship, 5) major financial problems, and 6) neighborhood disadvantage. These were the only available, time-varying SEP indicators that were measured repeatedly via maternal report through mailed questionnaires during three developmental time periods (**Figure 1a**): *very early childhood* (0-2 years), *early childhood* (3-5 years), and *middle childhood* (6-7 years).

For each SEP indicator, children were classified as exposed or unexposed at each period, using criteria described in **Supplemental Methods**. With these repeated, self-reported SEP indicators, we could identify changes occurring *between* time-periods for indicators capturing time-varying status of SEP. For job loss and income reduction, the measures inherently captured change *within* a certain developmental period, because they asked about socioeconomic mobility. To distinguish job loss and income reduction from other indicators, we refer to them throughout the manuscript as "instability indicators".

DNA methylation (DNAm)

DNAm was measured from peripheral blood at age 7 using the Illumina Infinium HumanMethylation450 BeadChip microarray (Illumina, San Diego, CA). DNAm wet laboratory procedures, preprocessing analyses, and quality control are described in **Supplemental Methods**. A total of 412,956 CpGs on autosomal chromosomes passed quality control and were included in this analysis. For each CpG, DNAm level is expressed as a 'beta' value (β-value) ranging from 0 to 1, which represents the proportion of cells methylated at each interrogated CpG.

Covariates

To adjust for baseline demographic differences in ARIES and technical variation in DNAm assessment, we controlled for the following variables measured at birth in all analyses: child age in months at blood draw, child race/ethnicity, child sex, child birthweight, maternal age, number of previous pregnancies, sustained maternal smoking during pregnancy, and cell type proportions estimated using the Houseman method (76). Details can be found in the **Supplemental Methods**.

Data analysis

All our analysis codes are available through our GitHub page: https://github.com/thedunnlab/sep-dnam."

Structured life course modeling approach

We used the two-stage structured life-course modeling approach (SLCMA) (29-31) to evaluate the time-dependent effects of socioeconomic adversity on DNAm. SLCMA is a method that leverages repeated exposure data to simultaneously investigate the relationship between exposure and outcome under multiple a priori-defined life-course hypotheses. In our analyses, we tested three life-course hypotheses, described previously, which were parameterized as follows (**Figure 1b**).

First, to test the *accumulation* hypothesis, we created a sum score (ranging from 0 to 3), which captured the number of time periods across the three developmental stages that children were exposed. Second, to test the *sensitive period* hypothesis, we created three binary variables, one for each of the three developmental periods, to classify children's exposure status (0= unexposed during the period; 1= exposed during that period). Third, to test the *mobility* hypothesis, we created a pair of indicator variables for change in SEP between very early and early childhood, and a pair of indicator variables for change in SEP between early and middle

406 childhood. Each pair consisted of an indicator variable for worsening (1=change from unexposed
407 to exposed, 0=other) and an indicator variable for improvement (1=change from exposed to
408 unexposed, 0=other).

409 We tested all three hypotheses for low family income, financial hardship, major financial
410 problem, and neighborhood disadvantage. Only the *accumulation* and *sensitive period*
411 hypotheses were tested for job loss and income reduction, as these two instability indicators
412 inherently reflect SEP changes (**Table S4**).

413 We performed the SLCMA in two stages: 1) life-course hypothesis *model selection*
414 followed by 2) *post-selection inference* (**Figure 1b**, **Supplemental Methods)**. In the first stage,
415 we tested the variables described above using a Least Angle Regression (LARS) variable
416 selection procedure (77) to identify the life-course hypothesis most supported in the observed
417 data (i.e., explaining the most variation in DNAm). In the second stage, we used selective
418 inference (29, 78) to test the association between the selected variable and DNAm and estimate
419 confidence intervals.

420 Defining CpGs of interest

421 We used two thresholds to identify associations between SEP and CpG CpGs for further
422 investigation. Given recent recommendations discouraging the use of p-values alone for
423 statistical inference (67, 68), we used an effect-size-based threshold of R^2 >3%, meaning that the
424 SEP exposure explained more than 3% of the variance in DNAm. This cutoff was selected based
425 on the effect sizes observed in previous epigenome-wide analyses of childhood adversity in
426 ALSPAC (24, 26) and other well-established environmental exposures, including tobacco
427 smoking (79). We also performed multiple-testing correction using the Benjamini-Hochberg
428 method (80) at a 5% false discovery rate (FDR) to assess the significance of top CpGs.

Sensitivity analyses

We conducted three sensitivity analyses to evaluate the robustness of our SLCMA results. First, we additionally controlled for 1) time-invariant SEP indicators (e.g., maternal education at baseline), 2) population substructure estimated from epigenetic data, 3) cord blood DNAm (to account for differences in DNAm that might have been present at birth), 4) genetic variation (at methylation quantitative trait loci, or mQTL), or 5) exposure to the other five time-varying SEP indicators. Second, we reran the analyses of the CpGs with an R^2 >3% for low family income, financial hardship, major financial problem, and neighborhood disadvantage using only *accumulation* and *sensitive period* hypotheses and compared the results from analysis with and without *mobility* tested. Third, we performed an EWAS of *any exposure* to each type of SEP adversity before age 7 and DNAm and compared the findings with SLCMA results. See **Supplemental Methods** for details.

Secondary analyses

To interpret our findings and place them in the context of prior literature, we conducted two secondary analyses. First, we compared the effect estimates of R^2 >3% CpGs to those reported in previous SEP-related EWAS studies (19) (**Supplemental Methods**). Second, we also evaluated the biological significance of our findings by examining the correlation between DNAm in blood and brain tissue for the R^2 >3% CpGs and testing for the enrichment for genomic features, regulatory elements, and Gene Ontology (GO) terms (**Supplemental Methods**).

Acknowledgments

This work was supported by the award # 96-17-05 from the Russell Sage Foundation and the Ford Foundation (E.C.D., awarded July 2017), who provided core support for this research. Dr. Dunn also received funding support from the National Institute of Mental Health at the National Institutes of Health [grant number R01MH113930]. This publication is the work of the authors, each of whom serve as guarantors for the contents of this paper.

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The UK Medical Research Council and the Wellcome Trust (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. ARIES was funded by the BBSRC (BBI025751/1 and BB/I025263/1). Supplementary funding to generate DNA methylation data which is included in ARIES has been obtained from the MRC, ESRC, NIH and other sources. ARIES is maintained under the auspices of the MRC Integrative Epidemiology Unit at the University of Bristol (MC_UU_12013/2 and MC_UU_12013/8). A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf).

Competing Interest Statement

The authors declare no conflict of interest.

Ethical Standards

All ethical guidelines were followed per research involving use of human subjects. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committee. This study was approved with oversight by the Mass General Brigham Institutional Review Boards (IRB) (Protocol ID 2017P001110).

References

1 Braveman, P. and Gottlieb, L. (2014) The social determinants of health: it's time to consider the causes of the causes. *Public Health Rep.*, **129 Suppl 2**, 19-31.

2 Krieger, N., Williams, D.R. and Moss, N.E. (1997) Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu. Rev. Public Health*, **18**, 341-378.

3 Vable, A.M., Gilsanz, P., Nguyen, T.T., Kawachi, I. and Glymour, M.M. (2017) Validation of a theoretically motivated approach to measuring childhood socioeconomic circumstances in the Health and Retirement Study. *PLoS One*, **12**, e0185898.

4 Elgar, F.J., McKinnon, B., Torsheim, T., Schnohr, C.W., Mazur, J., Cavallo, F. and Currie, C. (2016) Patterns of Socioeconomic Inequality in Adolescent Health Differ According to the Measure of Socioeconomic Position. *Soc. Indic. Res.*, **127**, 1169-1180.

5 d'Errico, A., Ricceri, F., Stringhini, S., Carmeli, C., Kivimaki, M., Bartley, M., McCrory, C., Bochud, M., Vollenweider, P., Tumino, R. *et al.* (2017) Socioeconomic indicators in epidemiologic research: A practical example from the LIFEPATH study. *PLoS One*, **12**, e0178071.

6 Caspi, A., Taylor, A., Moffitt, T.E. and Plomin, R. (2000) Neighborhood deprivation affects children's mental health: Environmental risks identified in a genetic design. *Psychol. Sci.*, **11**, 338-342.

7 Duncan, G.J., Magnuson, K. and Votruba-Drzal, E. (2017) Moving Beyond Correlations in Assessing the Consequences of Poverty. *Annu. Rev. Psychol.*, **68**, 413-434.

8 Bradley, R.H. and Corwyn, R.F. (2002) Socioeconomic status and child development. *Annu. Rev. Psychol.*, **53**, 371-399.

9 Hackman, D.A., Farah, M.J. and Meaney, M.J. (2010) Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nat. Rev. Neurosci.*, **11**, 651-659.

10 Noble, K.G., Houston, S.M., Brito, N.H., Bartsch, H., Kan, E., Kuperman, J.M., Akshoomoff, N., Amaral, D.G., Bloss, C.S., Libiger, O. *et al.* (2015) Family income, parental education and brain structure in children and adolescents. *Nat. Neurosci.*, **18**, 773-778.

11 Stringhini, S., Carmeli, C., Jokela, M., Avendaño, M., Muennig, P., Guida, F., Ricceri, F., d'Errico, A., Barros, H., Bochud, M. *et al.* (2017) Socioeconomic status and the 25 x 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1.7 million men and women. *Lancet*, **389**, 1229-1237.

12 Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A. and Cutler, D. (2016) The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, **315**, 1750-1766.

13 Marmot, M. (2020) Health equity in England: the Marmot review 10 years on. *BMJ*, **368**, m693.

14 Szyf, M., McGowan, P. and Meaney, M.J. (2008) The social environment and the epigenome. *Environ. Mol. Mutagen.*, **49**, 46-60.

15 Jin, B., Li, Y. and Robertson, K.D. (2011) DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer*, **2**, 607-617.

16 Portela, A. and Esteller, M. (2010) Epigenetic modifications and human disease. *Nat. Biotech.*, **28**, 1057-1068.

17 McCrory, C., McLoughlin, S. and O'Halloran, A.M. (2019) Socio-Economic Position Under the Microscope: Getting 'Under the Skin' and into the Cells. *Curr. Epidemiol. Rep.*, **6**, 403-411.
18 Wood, N.M., Trebilco, T. and Cohen-Woods, S. (2020) Scars of childhood socioeconomic stress: A systematic review. *Neurosci. Biobehav. Rev.*, **118**, 397-410.
19 Cerutti, J.K., Lussier, A.A., Zhu, Y., Liu, J. and Dunn, E.C. (2021) Associations between indicators of socioeconomic position and DNA methylation: A scoping review. *Clin. Epigenetics*, **13**.
20 Smith, K.E. and Pollak, S.D. (2021) Rethinking Concepts and Categories for Understanding the Neurodevelopmental Effects of Childhood Adversity. *Perspect. Psychol. Sci.*, **16**, 67-93.
21 Bush, N.R., Edgar, R.D., Park, M., MacIsaac, J.L., McEwen, L.M., Adler, N.E., Essex, M.J., Kobor, M.S. and Boyce, W.T. (2018) The biological embedding of early-life socioeconomic status and family adversity in children's genome-wide DNA methylation. *Epigenomics*, in press.
22 McDade, T.W., Ryan, C.P., Jones, M.J., Hoke, M.K., Borja, J., Miller, G.E., Kuzawa, C.W. and Kobor, M.S. (2019) Genome-wide analysis of DNA methylation in relation to socioeconomic status during development and early adulthood. *Am. J. Phys. Anthropol.*, **169**, 3-11.
23 Alfano, R., Guida, F., Galobardes, B., Chadeau-Hyam, M., Delpierre, C., Ghantous, A., Henderson, J., Herceg, Z., Jain, P., Nawrot, T.S. *et al.* (2019) Socioeconomic position during pregnancy and DNA methylation signatures at three stages across early life: epigenome-wide association studies in the ALSPAC birth cohort. *Int. J. Epidemiol.*, **48**, 30-44.
24 Dunn, E.C., Soare, T.W., Zhu, Y., Simpkin, A.J., Suderman, M.J., Klengel, T., Smith, A., Ressler, K.J. and Relton, C.L. (2019) Sensitive periods for the effect of childhood adversity on DNA methylation: results from a prospective, longitudinal study. *Biol. Psychiatry*, **85**, 838-849.
25 Marini, S., Davis, K.A., Soare, T.W., Zhu, Y., Suderman, M.J., Simpkin, A.J., Smith, A.D.A.C., Wolf, E.J., Relton, C.L. and Dunn, E.C. (2020) Adversity exposure during sensitive periods predicts accelerated epigenetic aging in children. *Psychoneuroendocrino.*, **113**, 104484.
26 Lussier, A.A., Zhu, Y., Smith, B.J., Simpkin, A.J., Smith, A.D.A.C., Suderman, M.J., Walton, E., Ressler, K.J. and Dunn, E.C. (2022) Updates to data versions and analytic methods influence the reproducibility of results from epigenome-wide association studies. *Epigenetics*, in press.
27 Vable, A.M., Duarte, C.D., Cohen, A.K., Glymour, M.M., Ream, R.K. and Yen, I.H. (2020) Does the Type and Timing of Educational Attainment Influence Physical Health? A Novel Application of Sequence Analysis. *Am. J. Epidemiol.*, **189**, 1389-1401.
28 Ben-Shlomo, Y. and Kuh, D. (2002) A life course approach to chronic disease epidemiology: conceptual models, empirical challenges, and interdisciplinary perspectives. *Int. J. Epidemiol.*, **31**, 285-293.
29 Zhu, Y., Simpkin, A.J., Suderman, M.J., Lussier, A.A., Walton, E., Dunn, E.C. and Smith, A.D.A.C. (2021) A Structured Approach to Evaluating Life Course Hypotheses: Moving Beyond Analyses of Exposed Versus Unexposed in the Omics Context. *Am. J. Epidemiol.*, **190**, 1101-1112.

30 Smith, B.J., Smith, A.D.A.C. and Dunn, E.C. (2022) Geyer, M.A., Marsden, C.A., Ellenbroek, B.A., Barnes, T.R.E. and Andersen, S.L. (eds.), In *Curr. Top. Behav. Neurosci.* Springer, Berlin, Heidelberg, in press.
31 Smith, A.D., Heron, J., Mishra, G., Gilthorpe, M.S., Ben-Shlomo, Y. and Tilling, K. (2015) Model Selection of the Effect of Binary Exposures over the Life Course. *Epidemiol.*, **26**, 719-726.
32 Hannon, E., Lunnon, K., Schalkwyk, L. and Mill, J. (2015) Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*, **10**, 1024-1032.
33 Needham, B.L., Smith, J.A., Zhao, W., Wang, X., Mukherjee, B., Kardia, S.L., Shively, C.A., Seeman, T.E., Liu, Y. and Diez Roux, A.V. (2015) Life course socioeconomic status and DNA methylation in genes related to stress reactivity and inflammation: The multi-ethnic study of atherosclerosis. *Epigenetics*, **10**, 958-969.
34 Stringhini, S., Polidoro, S., Sacerdote, C., Kelly, R.S., van Veldhoven, K., Agnoli, C., Grioni, S., Tumino, R., Giurdanella, M.C., Panico, S. *et al.* (2015) Life-course socioeconomic status and DNA methylation of genes regulating inflammation. *Int. J. Epidemiol.*, **44**, 1320-1330.
35 Johnson, R.C., Kalil, A. and Dunifon, R.E. (2012) Employment patterns of less-skilled workers: links to children's behavior and academic progress. *Demography*, **49**, 747-772.
36 Philip Oreopoulos, Marianne Page and Ann Huff Stevens. (2008) The Intergenerational Effects of Worker Displacement. *J. Labor. Econ.*, **26**, 455-000.
37 Brand, J.E. (2015) The Far-Reaching Impact of Job Loss and Unemployment. *Ann. Rev. Soc.*, **41**, 359-375.
38 Sandstrom, H. and Huerta, S. (2013), In *The Negative Effects of Instability on Child Development: A Research Synthesis*. Urban Institute, Washington, D.C., in press.
39 Ellis, B.J., Figueredo, A.J., Brumbach, B.H. and Schlomer, G.L. (2009) Fundamental Dimensions of Environmental Risk : The Impact of Harsh versus Unpredictable Environments on the Evolution and Development of Life History Strategies. *Hum. Nat.*, **20**, 204-268.
40 Belsky, J., Schlomer, G.L. and Ellis, B.J. (2012) Beyond cumulative risk: distinguishing harshness and unpredictability as determinants of parenting and early life history strategy. *Dev. Psychol.*, **48**, 662-673.
41 Levesque, A.R., MacDonald, S., Berg, S.A. and Reka, R. (2021) Assessing the Impact of Changes in Household Socioeconomic Status on the Health of Children and Adolescents: A Systematic Review. *Adol. Res. Rev.*, in press.
42 Shonkoff, J.P. (2010) Building a New Biodevelopmental Framework to Guide the Future of Early Childhood Policy. *Child. Dev.*, **81**, 357-367.
43 Essex, M.J., Boyce, W.T., Hertzman, C., Lam, L.L., Armstrong, J.M., Neumann, S.M. and Kobor, M.S. (2013) Epigenetic vestiges of early developmental adversity: childhood stress exposure and DNA methylation in adolescence. *Child. Dev.*, **84**, 58-75.
44 Hosokawa, R. and Katsura, T. (2018) Effect of socioeconomic status on behavioral problems from preschool to early elementary school - A Japanese longitudinal study. *PLoS One*, **13**, e0197961.
45 Ellwood-Lowe, M.E., Humphreys, K.L., Ordaz, S.J., Camacho, M.C., Sacchet, M.D. and Gotlib, I.H. (2018) Time-varying effects of income on hippocampal volume trajectories in adolescent girls. *Dev. Cogn. Neurosci.*, **30**, 41-50.

46 Cajaraville, M.P., Cancio, I., Ibabe, A. and Orbea, A. (2003) Peroxisome proliferation as a biomarker in environmental pollution assessment. *Microsc. Res. Tech.*, **61**, 191-202.
47 Kristiansen, H., Gad, H.H., Eskildsen-Larsen, S., Despres, P. and Hartmann, R. (2011) The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. *J. Interferon. Cytokine Res.*, **31**, 41-47.
48 Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A.D., Rawlik, K., Pasko, D., Walker, S., Parkinson, N., Fourman, M.H., Russell, C.D. *et al.* (2021) Genetic mechanisms of critical illness in COVID-19. *Nature*, **591**, 92-98.
49 Velavan, T.P., Pallerla, S.R., Rüter, J., Augustin, Y., Kremsner, P.G., Krishna, S. and Meyer, C.G. (2021) Host genetic factors determining COVID-19 susceptibility and severity. *eBioMedicine*, **72**.
50 Gatza, C.E., Oh, S.Y. and Blobe, G.C. (2010) Roles for the type III TGF-beta receptor in human cancer. *Cell Signal*, **22**, 1163-1174.
51 Sharifi, N., Hurt, E.M., Kawasaki, B.T. and Farrar, W.L. (2007) TGFBR3 loss and consequences in prostate cancer. *Prostate*, **67**, 301-311.
52 Liu, X.L., Xiao, K., Xue, B., Yang, D., Lei, Z., Shan, Y. and Zhang, H.T. (2013) Dual role of TGFBR3 in bladder cancer. *Oncol. Rep.*, **30**, 1301-1308.
53 Fang, W.Y., Kuo, Y.Z., Chang, J.Y., Hsiao, J.R., Kao, H.Y., Tsai, S.T. and Wu, L.W. (2020) The Tumor Suppressor TGFBR3 Blocks Lymph Node Metastasis in Head and Neck Cancer. *Cancers (Basel)*, **12**.
54 Leyden, J., Murray, D., Moss, A., Arumuguma, M., Doyle, E., McEntee, G., O'Keane, C., Doran, P. and MacMathuna, P. (2006) Net1 and Myeov: computationally identified mediators of gastric cancer. *Br. J. Cancer*, **94**, 1204-1212.
55 Liang, E., Lu, Y., Shi, Y., Zhou, Q. and Zhi, F. (2020) MYEOV increases HES1 expression and promotes pancreatic cancer progression by enhancing SOX9 transactivity. *Oncogene*, **39**, 6437-6450.
56 Teicher, M.H., Samson, J.A., Anderson, C.M. and Ohashi, K. (2016) The effects of childhood maltreatment on brain structure, function and connectivity. *Nat. Rev. Neurosci.*, **17**, 652-666.
57 Hardy, B., Hill, H.D. and Romich, J. (2019) Strengthening Social Programs to Promote Economic Stability During Childhood. *Soc. Policy Rep.*, **32**, 1-36.
58 Nestle, M. (2019) The Supplemental Nutrition Assistance Program (SNAP): History, Politics, and Public Health Implications. *Am. J. Public Health*, **109**, 1631-1635.
59 Tankersley, J. and Goldstein, D. (2021), In *New York Times*, in press.
60 Nishimi, K., Howe, E. and Dunn, E.C. (2019) *Social Determinants of Psychiatric Illness*. Oxford University Press.
61 Benzies, K. and Mychasiuk, R. (2009) Fostering family resiliency: a review of the key protective factors. *Child Fam. Soc. Work*, **14**, 103-114.
62 Hoye, J.R., Cheishvili, D., Yarger, H.A., Roth, T.L., Szyf, M. and Dozier, M. (2020) Preliminary indications that the Attachment and Biobehavioral Catch-up Intervention alters DNA methylation in maltreated children. *Dev. Psychopathol.*, **32**, 1486-1494.
63 Purewal Boparai, S.K., Au, V., Koita, K., Oh, D.L., Briner, S., Burke Harris, N. and Bucci, M. (2018) Ameliorating the biological impacts of childhood adversity: A review of intervention programs. *Child Abuse Neglect*, **81**, 82-105.

64 Poehlmann-Tynan, J., Engbretson, A., Vigna, A.B., Weymouth, L.A., Burnson, C., Zahn-Waxler, C., Kapoor, A., Gerstein, E.D., Fanning, K.A. and Raison, C.L. (2020) Cognitively-Based Compassion Training for parents reduces cortisol in infants and young children. *Infant. Ment. Health J.*, **41**, 126-144.
65 Slopen, N., McLaughlin, K.A. and Shonkoff, J.P. (2014) Interventions to improve cortisol regulation in children: a systematic review. *Pediatrics*, **133**, 312-326.
66 Raby, K.L., Bernard, K., Gordon, M.K. and Dozier, M. (2020) Enhancing diurnal cortisol regulation among young children adopted internationally: A randomized controlled trial of a parenting-based intervention. *Dev. Psychopathol.*, **32**, 1657-1668.
67 Amrhein, V. and Greenland, S. (2018) Remove, rather than redefine, statistical significance. *Nat. Hum. Behav.*, **2**, 4-4.
68 McShane, B.B., Gal, D., Gelman, A., Robert, C. and Tackett, J.L. (2019) Abandon Statistical Significance. *Am. Stat.*, **73**, 235-245.
69 Gurdasani, D., Barroso, I., Zeggini, E. and Sandhu, M.S. (2019) Genomics of disease risk in globally diverse populations. *Nat. Rev. Genet.*, **20**, 520-535.
70 Althubaiti, A. (2016) Information bias in health research: definition, pitfalls, and adjustment methods. *J. Multidiscip. Healthc.*, **9**, 211-217.
71 Krieger, N. (1992) Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am. J. Public Health*, **82**, 703-710.
72 Relton, C.L., Gaunt, T., McArdle, W., Ho, K., Duggirala, A., Shihab, H., Woodward, G., Lyttleton, O., Evans, D.M., Reik, W. *et al.* (2015) Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *Int. J. Epidemiol.*, **44**, 1181-1190.
73 Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S. and Davey Smith, G. (2013) Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.*, **42**, 111-127.
74 Golding, J., Pembrey, M., Jones, R. and the ALSPAC Study Team. (2001) ALSPAC: The Avon Longitudinal Study of Parents and Children I. Study methodology. *Paediatr. Perinat. Epidemiol.*, **15**, 74-87.
75 Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A. *et al.* (2013) Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int. J. Epidemiol.*, **42**, 97-110.
76 Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K. and Kelsey, K.T. (2012) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform.*, **13**, 86.
77 Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004) Least angle regression. *Ann. Stat.*, **32**, 407-499.
78 Tibshirani, R.J., Taylor, J., Lockhart, R. and Tibshirani, R. (2016) Exact Post-Selection Inference for Sequential Regression Procedures. *J. Am. Stat. Assoc.*, **111**, 600-620.
79 Zeilinger, S., Kühnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., Weidinger, S., Lattka, E., Adamski, J., Peters, A. *et al.* (2013) Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*, **8**, e63812.

80 Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.*, **57**, 289 - 300.

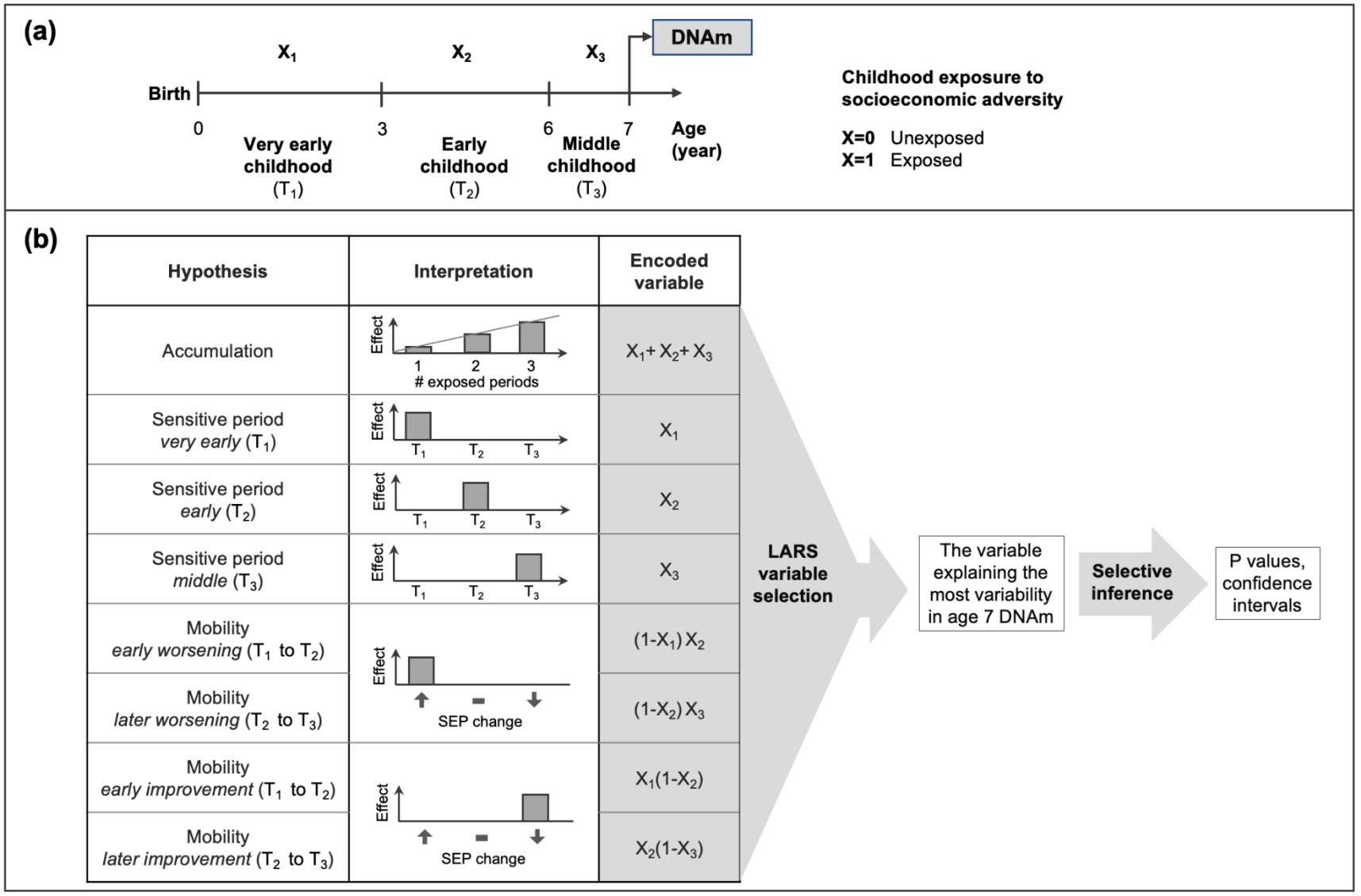
(a)
X1
X2
X3
DNAm
Birth
0
3
6
7
Age (year)
Very early childhood (T1)
Early childhood (T2)
Middle childhood (T3)
Childhood exposure to socioeconomic adversity
X=0 Unexposed
X=1 Exposed
(b)
Hypothesis
Interpretation
Encoded variable
Accumulation
X1+ X2+ X3
Sensitive period very early (T1)
X1
Sensitive period early (T2)
X2
Sensitive period middle (T3)
X3
Mobility early worsening (T1 to T2)
(1-X1) X2
Mobility later worsening (T2 to T3)
(1-X2) X3
Mobility early improvement (T1 to T2)
X1(1-X2)
Mobility later improvement (T2 to T3)
X2(1-X3)
Effect
exposed periods
SEP change
LARS variable selection
The variable explaining the most variability in age 7 DNAm
Selective inference
P values, confidence intervals

Figure 1. Study design and the conceptual life-course models used in the structured life course modeling approach (SLCMA). (a) Measurement of childhood socioeconomic adversity (X) and DNA methylation (DNAm) over time (T). Exposure to socioeconomic adversities, or indicators of low socioeconomic position (SEP), were measured repeatedly across three childhood periods: very early (0-2 years, T_1), early (3-5 years, T_2), and middle childhood (6-7 years, T_3). DNAm was measured around age 7. (b) Illustration of the life-course hypotheses tested in the SLCMA, the least angle regression (LARS) variable selection procedure, and selective inference test. *Accumulation*, *sensitive period*, and *mobility* hypotheses were examined in this study. *Accumulation* assumes that the effect of low SEP increases with the number of exposed periods. *Sensitive period* assumes that low SEP is particularly impactful during one of the three time periods. *Mobility* assumes that changes in SEP across specific periods is particularly impactful. *Early worsening* and *early improvement* refer to adversity getting worse (↓SEP) or better (↑SEP) from very early to early childhood, respectively; *later worsening* and *later improvement* refer to adversity getting worse or better from early to middle childhood, respectively. For each socioeconomic adversity, hypotheses were encoded into variables and then entered into the LARS variable selection procedure to identify the one explaining the most variability in DNAm at age 7 at each CpG site. We then performed post-selection inference to test the association between the selected variable and DNAm as well as estimate confidence intervals. See **Supplemental Methods** for more details about SLCMA.

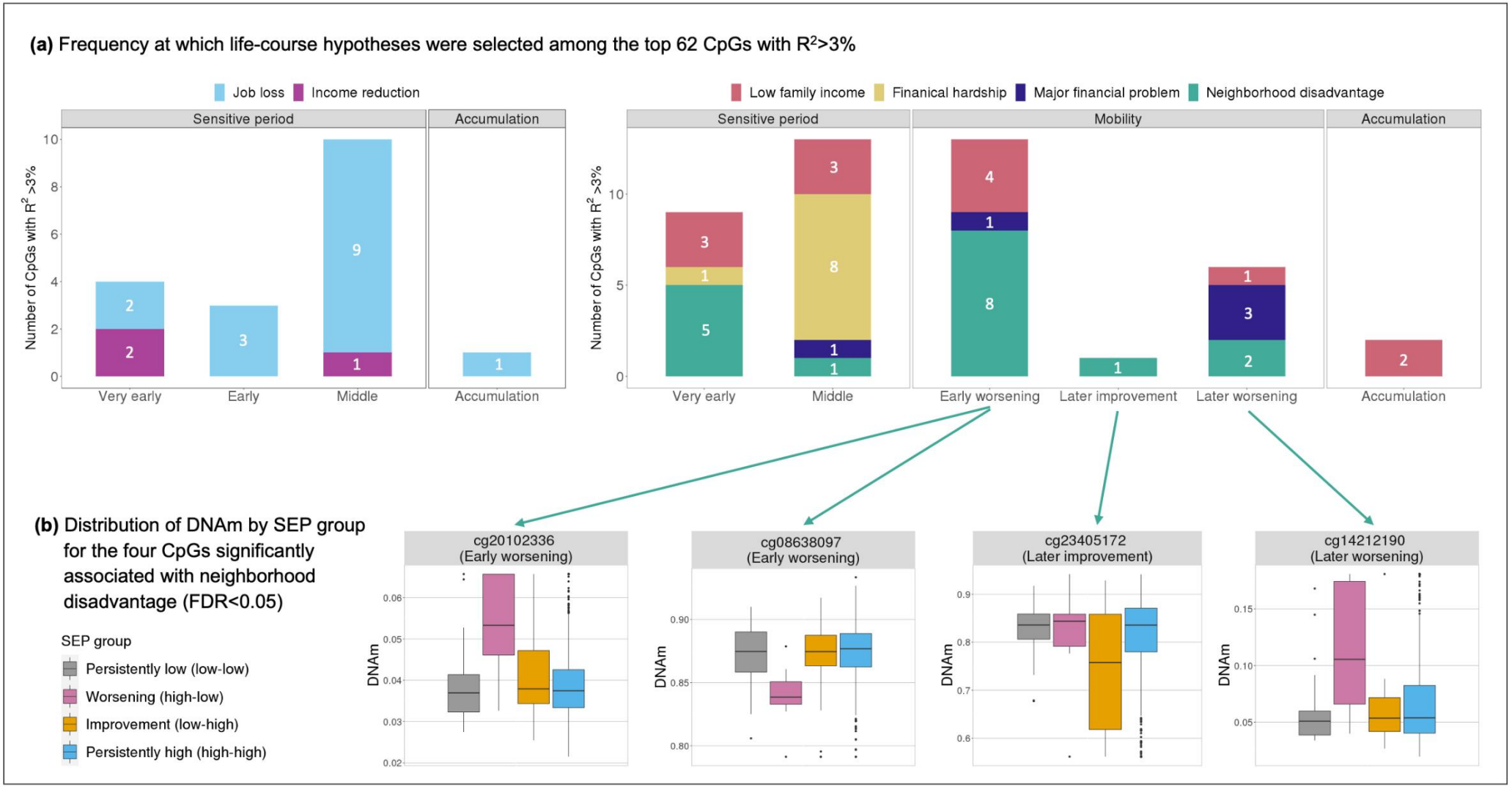
(a) Frequency at which life-course hypotheses were selected among the top 62 CpGs with R²>3%
Job loss
Income reduction
Sensitive period
Accumulation
Number of CpGs with R² >3%
Very early
Early
Middle
Accumulation
Low family income
Finanical hardship
Major financial problem
Neighborhood disadvantage
Sensitive period
Mobility
Accumulation
Very early
Middle
Early worsening
Later improvement
Later worsening
Accumulation
(b) Distribution of DNAm by SEP group for the four CpGs significantly associated with neighborhood disadvantage (FDR<0.05)
SEP group
Persistently low (low-low)
Worsening (high-low)
Improvement (low-high)
Persistently high (high-high)
cg20102336 (Early worsening)
cg08638097 (Early worsening)
cg23405172 (Later improvement)
cg14212190 (Later worsening)
DNAm

Figure 2. Mobility and sensitive period hypotheses were most often selected among the top 62 CpGs linked with socioeconomic adversity (or socioeconomic position, SEP) that explained > 3% variance in DNA methylation (DNAm).
(a) Frequency at which each life-course hypothesis was selected among the 62 CpGs. For job loss and income reduction, we tested *accumulation* and *sensitive period* hypotheses, and middle childhood was the most selected hypothesis. For the other four socioeconomic adversities, we tested *accumulation*, *sensitive period*, and *mobility* hypotheses. *Mobility* hypotheses, specifically worsening SEP, were most selected. *Very early*, *Early*, and *Middle* refer to sensitive period hypotheses related to the three childhood periods: very early (0-2 years), early (3-5 years), and middle childhood (6-7 years). *Early worsening/improvement* refer to mobility hypotheses for changes between very early and early childhood, and *later worsening/improvement* refer to mobility hypotheses for changes between early and middle childhood.
(b) For the four CpGs associated with neighborhood disadvantage at an FDR<0.05, SEP mobility group implied by the selected mobility hypothesis showed the greatest shift in DNAm. The distribution of DNAm by SEP mobility group is shown in boxplots, where the center line indicates the median, box limits indicate the 25th and 75th percentiles, whiskers extend up to 1.5 inter-quartile range (IQR) from the box limits, and individually plotted data points were values further than 1.5 IQR from the box limits. SEP mobility group was defined based on the exposure status at two consecutive childhood periods (*very early* and *early*, or *early* and *middle*) involved in the mobility hypothesis chosen for each CpG; *persistently low SEP* was defined as being exposed during both periods; *worsening SEP* was defined as being unexposed during the former period but exposed during the later period; *improving SEP* was defined as being exposed during the former period but unexposed during the later period; *persistently high SEP* was defined as being unexposed during both periods

Table 1. Prevalence of exposure to socioeconomic adversity by developmental period in the ARIES analytic sample.

	Job loss (N=667)	Income reduction (N=711)	Low family income (N=619)	Financial hardship (N=697)	Major financial problem (N=710)	Neighborhood disadvantage (N=687)
Very early childhood (0-2 years)	42 (6.3%)	458 (64.4%)	95 (15.4%)	127 (18.2%)	138 (19.4%)	83 (12.1%)
Early childhood (3-5 years)	32 (4.8%)	220 (30.9%)	79 (12.8%)	46 (6.6%)	69 (9.7%)	36 (5.2%)
Middle childhood (6-7 years)	18 (2.7%)	134 (18.9%)	55 (8.9%)	29 (4.2%)	60 (8.5%)	29 (4.2%)
Ever-exposed [a]	77 (11.5%)	525 (73.8%)	130 (21.0%)	147 (21.1%)	184 (25.9%)	98 (14.3%)
Average correlation over time [b]	0.49	0.34	0.87	0.70	0.50	0.80

The first four rows present the number (%) of children who were exposed to the specific type of socioeconomic adversity at each developmental period or ever exposed throughout the three periods.

[a] Children who were exposed during at least one period were defined as ever-exposed for the specific type of socioeconomic adversity.

[b] Polychoric correlations were presented, characterizing the average correlation over time within the given type of exposure. The average within-SEP correlations were moderate, suggesting these measures were variable across time, which allowed for detecting differences across periods.

Table 2. Summary of the SLCMA results for the 62 CpGs with R^2>3%.

Adversity	Number of R^2>3% CpGs	Range of R^2	Range of p-values	Number of FDR<0.05 CpGs
Neighborhood disadvantage	17	3.0-4.2%	1.3×10^{-7} – 7.1×10^{-6}	4[a]
Job loss	15	3.1-3.7%	5.8×10^{-7} – 8.8×10^{-6}	-
Low family income	13	3.0-3.8%	1.7×10^{-6} – 2.5×10^{-5}	-
Financial hardship	9	3.0-3.7%	5.9×10^{-7} – 8.5×10^{-6}	-
Major financial problem	5	3.0-3.8%	2.6×10^{-7} – 4.7×10^{-6}	-
Income reduction	3	3.0-3.3%	1.5×10^{-6} – 4.5×10^{-6}	-

SLCMA = structured life course modeling approach; FDR = false discovery rate.
The R^2 values reflect the increase in the variance of DNA methylation explained by the first hypothesis chosen after accounting for covariates. P-values were calculated using selective inference, which assesses the significance of the increase in R^2 explained. See **Table S2** for the full list of the 62 CpGs.
[a] Four CpGs for neighborhood disadvantage passed an FDR<0.05 significance threshold: cg20102336, cg08638097, cg23405172, and cg14212190.

Abbreviations

ALSPAC, Avon Longitudinal Study of Parents and Children
ARIES, accessible Resources for Integrated Epigenomics Studies
CpG, cytosine-guanine
DNAm, DNA methylation
EWAS, Epigenome-wide association study
FDR, false discovery rate
GO, gene ontology
LARS, Least Angle Regression
mQTL, methylation quantitative trait loci
SEP, socioeconomic position
SLCMA, structured life-course modeling approach
UK, United Kingdom