

## ARTICLE

## Open Access

# Childhood adversity and DNA methylation in two population-based cohorts

L. C. Houtepen<sup>1</sup>, R. Hardy<sup>2</sup>, J. Maddock<sup>2</sup>, D. Kuh<sup>2</sup>, E. L. Anderson<sup>1</sup>, C. L. Relton<sup>1</sup>, M. J. Suderman<sup>1</sup> and L. D. Howe<sup>1</sup>

## Abstract

Childhood adversity affects later health, but the underlying molecular mechanisms are unclear. Although there is some evidence from animal models and case-control studies of a role for DNA methylation, evidence from human population-based studies is limited. In two cohorts (mothers from the Avon Longitudinal Study of Parents and Children, ALSPAC,  $n = 780$  and women from the MRC National Survey of Health and Development, NSHD,  $n = 552$ ), we assessed the association of seven adverse childhood experiences (ACEs: parental physical illness, parental mental illness, parental death, parental separation, suboptimal maternal bonding, childhood illness and child maltreatment) as well as their combination (ACE score) with genome-wide DNA methylation levels measured using the Illumina Infinium HumanMethylation450 BeadChip in peripheral blood at mean age 47 years (ALSPAC) and in buccal cells at age 53 years (NSHD). CpG sites with a genome-wide false discovery rate (FDR) below 0.05 and differentially methylated regions (DMRs) with one-step Šidák correction  $p$ -values below 0.05 in each cohort were examined in the other cohort. No individual CpG sites replicated across cohorts. However, nine DMRs replicated across cohorts respectively associated with the ACE score (one region), parental mental illness (two regions), parental physical illness (three regions) and parental death (three regions). These observations indicate that some adverse childhood experiences, notably those related to parental health, may leave imprints on peripheral DNA methylation that persist to mid-life.

## Introduction

Childhood adversity is related to a broad range of negative outcomes across the lifespan including poorer mental and physical health<sup>1,2</sup> as well as lower educational attainment, income and economic participation<sup>3,4</sup>. The molecular mechanisms underlying these associations remain uncertain, although a role for epigenetic marks has been suggested. DNA methylation, the addition of a methyl group to a cytosine base that is followed by a guanine (CpG site)<sup>5</sup>, is the epigenetic modification most widely analysed in population-based studies. Previous studies indicate DNA methylation in adulthood can be affected by environmental factors early in life<sup>6,7</sup>.

Childhood adversity is related to altered DNA methylation in both animal and human studies (see review<sup>8</sup>). In fact, several genome-wide studies of childhood adversity in humans have identified associations throughout the genome<sup>6,9–15</sup>. Moreover, there is some evidence that associations near stress-related genes may be relevant for later life health<sup>16,17</sup>. However, most previous studies either focused on candidate genes or sampled from at-risk populations. One recent genome-wide population-based study on early-life victimisation stress did not identify robust changes in DNA methylation<sup>18</sup>, but did not examine other forms of childhood adversity (parental separation, parental mental or physical illness, child illness).

Therefore, we examine the association between various types of childhood adversity and genome-wide DNA methylation in two large population-based cohorts, the mothers of the Avon Longitudinal Study of Parents and Children (ALSPAC) and women from the MRC National

Correspondence: L. D. Howe ([laura.howe@bristol.ac.uk](mailto:laura.howe@bristol.ac.uk))

<sup>1</sup>MRC Integrative Epidemiology Unit at the University of Bristol, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

<sup>2</sup>MRC Unit for Lifelong Health and Ageing, University College London, London, UK

These authors contributed equally: M.J. Suderman, L. D. Howe

© The Author(s) 2018



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

Survey of Health and Development (NSHD). In these cohorts, previous studies have demonstrated associations between adversity and health<sup>19–23</sup>. Given the known co-occurrence of different types of adversity and the potential cumulative effects on health<sup>24,25</sup>, we analyse a count score of the adverse childhood experiences (ACE) in addition to exploring whether each different type of childhood adverse experience was associated with DNA methylation variation in separate epigenome-wide analyses (EWAS).

## Materials and methods

### Data

The ALSPAC is a prospective, population-based birth cohort study that recruited 14,541 pregnant women resident in Avon, UK, with expected delivery dates between the 1st April 1991 and 31st December 1992<sup>26,27</sup>. The mothers, their partners and the index child have been followed-up through clinics, questionnaires and links to routine data. The ALSPAC mothers form the participants for this study. The study website contains details of all the data that is available through a fully searchable data dictionary (<http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/>). The Accessible Resource for Integrated Epigenomic Studies (ARIES) project<sup>28</sup> includes 1018 mother-child pairs from ALSPAC that had DNA methylation measured using the Infinium HumanMethylation450BeadChip (Illumina, Inc). Our analysis used DNA methylation profiles of peripheral blood collected from mothers at a follow-up clinic (age of participants ranged from 44 to 50 years (mean = 47.1)). DNA methylation profiles derived from peripheral blood collected during the study pregnancy approximately 18 years earlier were used to replicate findings (mean age = 29.1). Ethical approval was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees and informed consent was obtained from all subjects.

The MRC National Survey of Health and Development (NSHD) is based on a nationally representative sample of 5362 births out of all the single births that took place within marriage in one week in March 1946 in England, Scotland and Wales;<sup>29</sup> study members have been followed up 24 times, up to age 69<sup>30</sup>.

For 810 women, DNA methylation at age 53 years was measured using the Infinium HumanMethylation450BeadChip (Illumina, Inc) on buccal epithelial samples and in a subsample of 156 women also on whole blood cells (WBC)<sup>31</sup>. All women gave written informed consent for their samples to be used in genetic studies of health, and the Central Manchester Ethics Committee approved the use of these samples for epigenetic studies of health in 2012.

### Childhood adversity

Seven ACEs were assessed in both ALSPAC and NSHD: (i) suboptimal maternal bonding, (ii) childhood physical illness, (iii) parental mental illness, (iv) parental divorce or separation, (v) parental physical illness or disability, (vi) parental death in childhood and (vii) child maltreatment. In ALSPAC child maltreatment was further refined into five distinct ACEs, three forms of abuse (sexual, emotional and physical) and two forms of neglect (emotional and physical). In ALSPAC, women reported ACEs in retrospective questionnaires at study enrolment (during pregnancy, mean age 28 years) or soon thereafter. In NSHD, ACEs were prospectively reported in interviews and questionnaires by the participants' mothers up to age 16, except for parental bonding and maltreatment, which were self-reported when participants were age 43. All ACEs were defined as binary indicators of presence/absence of the adversity. An ACE score was generated by adding up whether an individual was exposed to the ACEs, creating a score with a maximum possible value of seven in NSHD and eleven in ALSPAC (further details in Supplementary information).

Associations between the ACE score and mental health have been demonstrated previously in the NSHD cohort<sup>23</sup>, but not in the ALSPAC mothers. Therefore, to demonstrate that the ACE score behaves as expected in relation to mental health, we examined the association of the ACE score with depression (defined according to a score of 13 or more on the Edinburgh Postnatal Depression Score from questionnaires completed by the mother when the child was 11 years old), both in the full ALSPAC cohort and in the subsample included in our EWAS.

### DNA methylation

In both cohorts following DNA extraction, DNA was bisulfite converted using the Zymo EZ DNA Methylation™ kit (Zymo, Irvine, CA, USA). Genome-wide DNA methylation levels at over 485,000 CpG sites were measured with the Infinium HumanMethylation450 BeadChip. Both datasets were pre-processed in R version 3.2.4 with the meffil package<sup>32</sup>, using functional normalisation<sup>33</sup> to reduce non-biological differences between probes. After pre-processing, 976 samples and 453,965 probes were available for further analysis in ALSPAC, while 766 buccal as well as 153 blood samples and 455,971 probes were available for further analysis in NSHD (further details in Supplementary information). Methylation level at a CpG site is expressed as a 'beta' value ( $\beta$ -value), ranging from 0 (no cytosine methylation) to 1 (complete cytosine methylation).  $\beta$ -Values are reported as percentages. To reduce the influence of outliers in regression models, normalized  $\beta$ -values were 90%-Winsorized.

### Epigenome-wide association studies (EWAS)

Epigenome-wide association studies (EWAS) for the ACE score and each separate ACE were conducted in R version 3.2.4<sup>34</sup>, using linear regression models with untransformed methylation beta values as the outcome. EWAS were conducted separately in each cohort; 8 separate EWAS were performed in NSHD (7 individual ACEs and the ACE score), and 13 EWAS were performed in ALSPAC (11 individual ACEs, child maltreatment, and the ACE score). Multiple testing was accounted for by controlling the false discovery rate (FDR) at 5%, implementing the method by Benjamini and Hochberg<sup>35</sup>.

All models were adjusted for the age at DNA sampling. To adjust for technical batch effects, independent surrogate variables (ISV) were calculated and included in all models<sup>36</sup>. In ALSPAC, Houseman-estimated cell proportions<sup>37</sup> were used to adjust for cellular heterogeneity in blood DNA samples.

As smoking may potentially lie on the causal pathway between childhood adversity exposure and DNA methylation, adjustment would remove part of the pathway of interest and could potentially result in collider bias. Thus, the main analyses were not adjusted for smoking status. However, the influence of smoking for the CpGs and DMRs of interest was examined in a sensitivity analysis by adding a methylation-based smoking score, derived from the effect size estimates of the 185 probes with FDR <0.05 in a recent smoking EWAS<sup>38</sup>, as a covariate to the main model.

When the ACE score was the exposure, the regression coefficients in the EWAS model are interpreted as the difference in mean methylation when the score increases by one. For the individual ACEs, regression coefficients are interpreted as the mean difference in methylation between the adversity exposed and unexposed groups.

### Differentially methylated regions

The EWAS results were used to detect differential methylation across larger regions of the genome with Comb-P<sup>39</sup> (further details in Supplementary information).

### Replication

CpGs with FDR <0.05 or DMRs with Šidák-corrected *P*-value <0.05 in either ALSPAC or NSHD were examined in the other cohort.

### Enrichment for probes previously reported in literature

To examine whether the top 1000 CpGs in our EWAS analyses were enriched for CpGs previously reported in literature, we identified nine publications that reported CpG sites on the Illumina Infinium HumanMethylation450 BeadChip associated with (a form of) childhood adversity (further details in Supplementary information, Supplementary Table 5 and the probe lists in

Supplementary data sheet 1). We also included lists for potentially unreliable probes<sup>40,41</sup> and CpGs related to smoking<sup>38</sup>. For all lists, enrichment of the previously reported CpGs in the top 1000 CpGs in our EWAS analyses was examined using Fisher's exact test.

### Cohort comparison

We compared the associations of adversity in ALSPAC and NSHD by calculating Pearson's correlation of the regression coefficients of the top 1000 CpG sites identified by each cohort. For DMRs, correlation was calculated from the average of the regression coefficients of the CpGs that were part of a DMR. To calculate an informative correlation coefficient, the cohort comparison was only performed if, for a particular ACE, more than three DMRs were identified.

### Cross-tissue analysis

In a subset of 98 women in NSHD with buccal DNA and ACE data, who also had blood DNA methylation measured at the same time point, the tissue specificity of childhood adversity related changes in DNA methylation was examined (cross-tissue subset). Parental mental illness could not be evaluated in this subsample due to a lack of cases (see Supplementary Table 4). For each of the available ACE measures, an EWAS analysis was performed for each tissue type. We then examined the tissue specificity of adversity by calculating the correlation of the regression coefficients for the top 1000 CpGs. For DMRs, we calculated the correlation of the average of the regression coefficients of the CpGs that were part of each DMR. The cross-tissue analysis was only performed if, for a particular ACE, more than three DMRs were identified.

### Genetic variants near replicated DMRs

In a sensitivity analysis, the influence of genetic variation on the association between childhood adversity and methylation at our DMRs of interest was investigated in ALSPAC. Genetic data from genome-wide assays (Illumina 660W-quad BeadChip) was available for 700 of the 780 participants in the ALSPAC EWAS analyses<sup>26</sup>.

For each replicated DMR, we included SNPs as covariates in the EWAS model that were located within 1MB (observed and imputed SNPs pruned for independence (R-squared <0.0001) with a major allele frequency >1% and less than 3% missing) and were associated ( $p < 0.05$ ) with at least one CpG in the DMR.

## Results

### Characteristics of participants

Complete adversity and DNA methylation data were available for 780 women in the ALSPAC cohort and 552 women in NSHD. The mean age at DNA methylation assessment was 47.4 years in ALSPAC, and 53.4 years in

**Table 1 Prevalence of childhood adverse experiences in ALSPAC mothers and NSHD participants**

|   | ALSPAC   |  | NSHD   |  | Comparison <sup>a</sup> |
|---|--|--|--|--|-------------------------|
|   | Participants with complete adversity data (n = 8021) | Participants with DNA methylation data, included in our analysis (n = 780) | Participants with complete adversity data (n = 2438) | Participants with DNA methylation data, included in our analysis (n = 552) | p-value                 |
| Sex = Male (%)                                      | 0  | 0  | 51.4   | 0  | n/a                     |
| Age clinic (mean (sd))                              | 47.90 (4.40)   | 47.14 (4.61)   | 53.48 (0.50)   | 53.45 (0.50)   | <0.001                  |
| Smoking (%)   |  |  |  |  | <0.001                  |
| Never-smoker  | 53.4   | 56.3   | 30.3   | 33.3   |                         |
| Ex-smoker   | 35.4   | 36.2   | 48.5   | 47.3   |                         |
| Current-smoker                                      | 11.1   | 7.4  | 21.2   | 19.4   |                         |
| Childhood adverse experiences before 17 years       |  |  |  |  |                         |
| Parent physically ill (%)                           | 25.8   | 28.6   | 22.7   | 25.9   | 0.308                   |
| Parent mentally ill (%)                             | 4.4  | 5.8  | 2.2  | 1.8  | 0.001                   |
| Parents separated (%)                               | 17.7   | 14.9   | 5  | 4.2  | <0.001                  |
| Parent died (%)                                     | 5.9  | 5.1  | 6.9  | 6.3  | 0.409                   |
| Suboptimal maternal bonding (%)                     | 18.8   | 15.8   | 19.3   | 21.7   | 0.007                   |
| Child illness (%)                                   | 5  | 5.3  | 15.4   | 13.4   | <0.001                  |
| Child maltreatment (%)                              | 24.6   | 22.2   | 6.2  | 6.5  | <0.001                  |
| Sexual abuse (%)                                    | 4.5  | 3.7  | n/a  | n/a  | n/a                     |
| Physical abuse (%)                                  | 3.2  | 3.1  | n/a  | n/a  | n/a                     |
| Emotional abuse (%)                                 | 7.5  | 8.3  | n/a  | n/a  | n/a                     |
| Physical neglect (%)                                | 1.8  | 1.2  | n/a  | n/a  | n/a                     |
| Emotional neglect (%)                               | 20.8   | 19.4   | n/a  | n/a  | n/a                     |
| Adverse Childhood Experiences (ACE) count score (%) |  |  |  |  | <0.001                  |
| 0   | 42.4   | 43.8   | 45.3   | 42   |                         |
| 1   | 29   | 29.4   | 36.8   | 41.1   |                         |
| 2   | 14   | 12.3   | 13.7   | 12.7   |                         |
| 3   | 7.4  | 6.9  | 3.6  | 3.4  |                         |
| 4   | 3.5  | 4  | 0.6  | 0.5  |                         |
| 5   | 1.9  | 1.9  | 0.1  | 0.2  |                         |
| 6   | 1  | 0.9  | 0  | 0  |                         |
| 7   | 0.5  | 0.8  | 0  | 0  |                         |
| 8   | 0.2  | 0  | n/a  | n/a  |                         |
| 9   | 0.1  | 0  | n/a  | n/a  |                         |
| 10  | 0  | 0  | n/a  | n/a  |                         |

<sup>a</sup>The difference between the ALSPAC (n = 780) and NSHD (n = 552) participants that were included in our EWAS analyses was tested using a Pearson's chi-squared test for categorical variables and an ANOVA for numerical variables

NSHD. Table 1 shows the prevalence of each type of ACE, and each category of the ACE score. Parental physical illness and parental death had similar prevalence in both cohorts. Parental mental illness, parental separation and any maltreatment were more prevalent in ALSPAC, and child illness and suboptimal maternal bonding were more common in NSHD. Most ACEs had a similar prevalence in the participants without DNA methylation data, and the ACE score demonstrated similar associations with depression in the full cohort of ALSPAC mothers and in the subsample included in our analysis (see Supplementary Tables 2–4).

The correlation (Cramér's  $V$  for nominal variables) between different adverse experiences varied, with  $\phi_c$  ranging from 0 to 0.5 in ALSPAC and 0 to 0.1 in NSHD (Supplementary Figure 3 and 4). In general, correlations between ACEs were considerably lower in NSHD than ALSPAC.

### Individual CpGs

In both NSHD and ALSPAC, there was no association between the cumulative ACE score and individual CpG sites after multiple testing. In NSHD, DNA methylation levels at three CpG sites were negatively associated ( $FDR < 0.05$ ) with a specific adversity; parental mental illness (cg17164016  $B = -0.07$ ,  $p = 8.8 \times 10^{-08}$ ,  $FDR = 0.043$ ), childhood illness (cg10303653  $B = -0.017$ ,  $p = 2.1 \times 10^{-08}$ ,  $FDR = 0.010$ ) and child maltreatment (cg14296561  $B = -0.040$ ,  $p = 6.5 \times 10^{-09}$ ,  $FDR = 0.003$ ) (see Table 2). Similar associations were found after adjusting for smoking in the EWAS model (cg17164016  $B = -0.07$ ,  $p = 2.5 \times 10^{-07}$ ,  $FDR = 0.12$ ; cg10303653  $B = -0.016$ ,  $p = 2.2 \times 10^{-07}$ ,  $FDR = 0.11$  and cg14296561  $B = -0.043$ ,  $p = 8.0 \times 10^{-10}$ ,  $FDR = 0.0004$ ). None of these associations were replicated in ALSPAC, although cg14296561 was weakly but positively (opposite direction of effect compared with NSHD) associated with child maltreatment ( $B = 0.0091$ , unadjusted  $p = 0.041$ ) (see Table 2).

In ALSPAC, there was no associations of individual CpG sites with any ACE measure after adjustment for multiple tests ( $FDR < 0.05$ ) (see Supplementary Figure 5 and 6 for QQ-plots of the EWAS analyses in NSHD and ALSPAC, respectively).

### Differentially methylated regions

We identified 231 differentially methylated regions (DMRs; one-step Šidák correction  $< 0.05$ ) associated with specific ACE measures in ALSPAC ( $n = 97$ ) and NSHD ( $n = 134$ ). Four of the 134 NSHD DMRs replicated in ALSPAC and six of the 97 ALSPAC DMRs replicated in NSHD (same ACE measure and direction of effect; StoufferLiptak-Kechris corrected  $P$ -value  $< 0.05$  unadjusted for genome-wide tests), even after adjusting for smoking (see Table 3 and Supplementary Table 7).

Of the nine replicated DMRs, the one (chr8:145654565-145654855) DMR negatively associated with the ACE score survived correction for multiple testing (Šidák corrected  $p$ -value  $< 0.05$ ) and was thus discovered independently in both cohorts, two DMRs were positively (chr12:14720726-14721289) or negatively (chr1:3104999-3105327) associated with parental mental illness, three DMRs were either positively (chr15:40364524-40364863 and chr7:24323261-24323940) or negatively (chr2:18766018-18766295) associated with parental death and the last three DMRs were either positively (chr15:81426347-81426670 and chr8:144120335-144120707) or negatively (chr22:27834439-27834630) associated with parental physical illness.

Only two of the eight DMRs associated with a specific ACE were additionally associated with another ACE in both cohorts (chr15:40364524-40364863 with parental mental illness; chr8:144120335-144120707 with parental death). In contrast, the DMR (chr8:145654565-145654855) associated with the ACE score, was also related to three specific ACEs (parental death, suboptimal maternal bond and child maltreatment) in both cohorts (see Table 3).

Besides the nine replicated DMRs, nineteen DMRs were associated with the abuse and neglect childhood adversity measures that were solely available in ALSPAC (3 DMRs emotional abuse, 5 DMRs emotional neglect, 8 DMRs physical abuse, 2 DMRs physical neglect and 1 DMR sexual abuse) even after adjusting for smoking (see Supplementary Table 8). Although these nineteen DMRs could not be replicated in NSHD, the DMR on chr3:87138203-87138701 was positively associated with emotional neglect in ALSPAC as well as child maltreatment in both ALSPAC and NSHD. Finally, the DMR on chr11:67417958-67418406 was negatively associated with several abuse and neglect measures in ALSPAC (emotional neglect, emotional abuse and child maltreatment) but did not replicate in NSHD where the DMR was positively associated with child maltreatment.

Since there are DNA methylation profiles for many of the same ALSPAC participants measured about 18 years earlier during pregnancy ( $n = 769$ ), we tested the 9 replicated DMRs for replication at that earlier time point. Five of the 9 were replicated, including the single DMR that survived genome-wide adjustment for multiple testing in both cohorts (chr8:145654565-145654855). Details are provided in Supplementary Table 9.

### Enrichment for probes previously reported in literature

Although some of our top 1000 CpGs were enriched for certain previously reported CpGs (1 CpG<sup>18</sup>, 4 CpGs<sup>11</sup>, 2 CpGs<sup>6</sup>, 2 CpGs<sup>15</sup> and 2 CpGs<sup>10</sup>) (Supplementary Figure 7), in most cases the type of adversity in our analysis was very different from the original study. Only three

**Table 2 CpGs surviving FDR correction in an epigenome-wide analysis (EWAS)**

| Discovery                | ProbeID    | Gene | Position        | B      | P        | fdr   | Look up NSHD  | Look up ALSPAC   |
|--------------------------|------------|------|-----------------|--------|----------|-------|---|--|
| NSHD-Parent mentally ill | cg17164016 |      | chr8: 110990365 | -0.067 | 8.86E-08 | 0.043 | ACE score $B = -0.0026, p = 0.20$<br>Parent physically ill $B = -0.0054, p = 0.16$<br><b>Parent mentally ill <math>B = -0.067, p = 8.9e-08^*</math></b><br>Parents separated $B = -0.00068, p = 0.94$<br>Parent died $B = -0.019, p = 0.006^*$<br>Suboptimal maternal bonding $B = 0.0061, p = 0.14$<br>Child illness $B = 0.0012, p = 0.81$<br>Child maltreatment $B = 0.0052, p = 0.46$   | ACE score $B = 0.00093, p = 0.33$<br>Parent physically ill $B = -0.0018, p = 0.56$<br><b>Parent mentally ill <math>B = 0.0039, p = 0.49</math></b><br>Parents separated $B = -0.0014, p = 0.72$<br>Parent died $B = 0.0075, p = 0.22$<br>Suboptimal maternal bonding $B = 0.0016, p = 0.65$<br>Child illness $B = 0.0073, p = 0.23$<br>Child maltreatment $B = 0.0035, p = 0.28$<br>Sexual abuse $B = 0.0044, p = 0.53$<br>Physical abuse $B = 0.0077, p = 0.33$<br>Emotional abuse $B = -0.00036, p = 0.94$<br>Physical neglect $B = 0.0049, p = 0.70$<br>Emotional neglect $B = 0.0035, p = 0.29$  |
|                          |            |      |                 | -0.017 | 2.10E-08 | 0.010 | ACE score $B = -0.0032, p = 0.012^*$<br>Parent physically ill $B = -0.0013, p = 0.59$<br>Parent mentally ill $B = 0.0057, p = 0.47$<br>Parents separated $B = -0.0033, p = 0.54$<br>Parent died $B = 0.0022, p = 0.63$<br>Suboptimal maternal bonding $B = 0.00074, p = 0.78$<br><b>Child illness <math>B = -0.017, p = 2.1e-08^*</math></b><br>Child maltreatment $B = -0.00091, p = 0.84$ | ACE score $B = 0.00034, p = 0.76$<br>Parent physically ill $B = 0.0021, p = 0.54$<br>Parent mentally ill $B = -0.0077, p = 0.24$<br>Parents separated $B = 0.0087, p = 0.054$<br>Parent died $B = -0.0066, p = 0.34$<br>Suboptimal maternal bonding $B = -0.00064, p = 0.88$<br><b>Child illness <math>B = 0.0057, p = 0.41</math></b><br>Child maltreatment $B = 0.001, p = 0.78$<br>Sexual abuse $B = 0.003, p = 0.71$<br>Physical abuse $B = -0.0023, p = 0.8$<br>Emotional abuse $B = 0.00043, p = 0.94$<br>Physical neglect $B = -0.015, p = 0.31$<br>Emotional neglect $B = -0.00034, p = 0.93$<br>ACE score $B = 0.0017, p = 0.20$<br>Parent physically ill $B = 0.0025, p = 0.55$<br>Parent mentally ill $B = -0.0019, p = 0.81$<br>Parents separated $B = 0.012, p = 0.026^*$<br>Parent died $B = -0.0028, p = 0.74$<br>Suboptimal maternal bonding $B = -0.0028, p = 0.59$<br>Child illness $B = 0.0043, p = 0.61$ |
|                          |            |      |                 | -0.040 | 6.51E-09 | 0.003 | ACE score $B = -0.0083, p = 4.4e-05^*$<br>Parent physically ill $B = -0.0062, p = 0.12$<br>Parent mentally ill $B = -0.0058, p = 0.66$<br>Parents separated $B = -0.0033, p = 0.7$<br>0<br>Parent died $B = -0.0072, p = 0.32$<br>Suboptimal maternal bonding $B = -0.0041, p = 0.33$   | ACE score $B = 0.0017, p = 0.20$<br>Parent physically ill $B = 0.0025, p = 0.55$<br>Parent mentally ill $B = -0.0019, p = 0.81$<br>Parents separated $B = 0.012, p = 0.026^*$<br>Parent died $B = -0.0028, p = 0.74$<br>Suboptimal maternal bonding $B = -0.0028, p = 0.59$<br>Child illness $B = 0.0043, p = 0.61$  |



**Table 2** continued

| Discovery | ProbeID | Gene | Position | B | P | fdr | Look up NSHD  | Look up ALSPAC  |
|-----------|---------|------|----------|---|---|-----|---|---|
|           |         |      |          |   |   |     | Child illness $B = -0.0094, p = 0.061$<br><b>Child maltreatment <math>B = -0.04, p = 6.5e-09^*</math></b> | <b>Child maltreatment <math>B = 0.0091, p = 0.041^*</math></b><br>Sexual abuse $B = 0.016, p = 0.10$<br>Physical abuse $B = 0.0017, p = 0.87$<br>Emotional abuse $B = 0.013, p = 0.048^*$<br>Physical neglect $B = -0.02, p = 0.26$<br>Emotional neglect $B = 0.0055, p = 0.25$ |

\*In the last two columns indicates a nominal significant association between the CpG and one of the childhood adversities in the respective cohort, NSHD or ALSPAC. Bold in the final two columns represents the ACE to which the discovery analysis relates.

previously reported CpGs replicated ( $p < 0.05$ ) and were amongst the top 1000 CpGs of our analyses: (i) cg19335412 near the *ACTA2* gene that was previously linked to childhood abuse<sup>6</sup> and to sexual abuse in ALSPAC ( $B = 0.039, p = 0.0014$ ), (ii) cg00973947 near the *C3orf58* gene that was previously linked to the childhood trauma questionnaire score<sup>10</sup> and to ACE score in NSHD ( $B = 0.001, p = 0.0014$ ), and (iii) cg27347930 near the *LIG3* and *CDS* genes that was previously linked to sexual abuse in the Dunedin study ( $B = -0.008, p = 2.39 \times 10^{-08}$ )<sup>18</sup> and also to sexual abuse in ALSPAC ( $B = -0.016, p = 0.004$ ).

There was some evidence for enrichment of smoking-related associations<sup>38</sup> and potentially unreliable probes<sup>40,41</sup> in our top CpGs (Supplementary Figure 7).

**Cohort comparison**

For all ACEs measured in both cohorts, the regression coefficients of the top 1000 CpGs were only weakly correlated between cohorts (Spearman  $\rho$  range between  $-0.12$  and  $0.08$ , mean =  $-0.003$ , sd =  $0.05$ ) (see Supplementary Figure 8). Furthermore, the average regression coefficients of the CpGs that are part of DMRs were also only generally weakly and inconsistently correlated between cohorts ( $\rho$  range between  $-0.59$  and  $0.60$ , mean =  $0.009$ , sd =  $0.29$ ) (see Supplementary Figure 9).

**Cross-tissue analysis**

In the NSHD subsample with both blood and buccal cells ( $n = 98$ ), the regression coefficients of the top 1000 CpGs in the buccal EWAS were moderately to highly correlated with the regression coefficients of the same CpGs in blood (Spearman  $\rho$  range =  $0.20-0.59$ , mean =  $0.50$ , sd =  $0.13$ ). Similar moderate to high cross-tissue correlations were present for the top 1000 CpGs in the blood EWAS analyses ( $\rho$  range =  $0.36-0.69$ , mean =  $0.54$ , sd =  $0.10$ ) (see Supplementary Figure 10). The average regression coefficients of the CpGs that are part of DMRs were also highly correlated across tissue type ( $\rho$  range =  $0.66-0.95$ , mean =  $0.85$ , sd =  $0.12$  for blood-based ACE DMRs and  $\rho$  range =  $0.45-0.86$ , mean =  $0.73$ , sd =  $0.15$  for buccal-based ACE DMRs) (see Supplementary Figure 11).

**Influence of genetic variants on replicated DMRs**

In ALSPAC, between eight and 40 genetic variants were within 1MB and associated with at least one CpG for one of the nine replicated DMRs (see Table 4). After adding the appropriate genetic variants as covariates to the original EWAS model, the regression coefficients for most CpGs in the DMRs were in the same direction as the original EWAS model and all DMRs were still statistically significant (one-step Šidák correction  $< 0.05$ ).

**Table 3 DMRs that replicated across cohorts**

| ACE                   | DMR                           | Nr CpGs | Gene                | ALSPAC (Direction <sup>a</sup> ,<br><i>p</i> -value) | NSHD (Direction <sup>a</sup> ,<br><i>p</i> -value) | Replicated ACEs in<br>lookup <sup>b</sup>                                |
|-----------------------|-------------------------------|---------|---------------------|--|--|--|
| ACE count score       | chr8: 145654565-<br>145654855 | 5       | <i>VPS28, TONSL</i> | ----<br><i>p</i> = 7.5e-08*                          | ----<br><i>p</i> = 1.8e-06*                        | ACE count score,<br>Parent died,<br>Maternal bond,<br>Child maltreatment |
| Parent mentally ill   | chr12: 14720726-<br>14721289  | 10      | <i>PLBD1</i>        | -+++++<br><i>p</i> = 0.0058                          | +++++<br><i>p</i> = 9.5e-07*                       | Parent mentally ill  |
| Parent mentally ill   | chr1: 3104999-<br>3105327     | 5       | <i>PRDM16</i>       | ----<br><i>p</i> = 0.001                             | ----<br><i>p</i> = 7.4e-08*                        | Parent mentally ill  |
| Parent physically ill | chr15: 81426347-<br>81426670  | 9       | <i>C15orf26</i>     | +++++<br><i>p</i> = 0.011                            | +++++<br><i>p</i> = 7.4e-07*                       | Parent mentally ill  |
| Parent physically ill | chr22: 27834439-<br>27834630  | 3       | n/a                 | --<br><i>p</i> = 6.9e-06*                            | --<br><i>p</i> = 0.045                             | Parent mentally ill  |
| Parent physically ill | chr8: 144120335-<br>144120707 | 7       | <i>C8orf31</i>      | +++++<br><i>p</i> = 4.4e-06*                         | +++++<br><i>p</i> = 0.00092                        | Parent mentally ill,<br>Parent died                                      |
| Parent died           | chr15: 40364524-<br>40364863  | 3       | n/a                 | +++<br><i>p</i> = 6.3e-07*                           | +++<br><i>p</i> = 0.0044                           | Parent died,<br>Parent mentally ill                                      |
| Parent died           | chr7: 24323261-<br>24323940   | 9       | <i>NPY</i>          | +++++<br><i>p</i> = 9.2e-07*                         | +++++<br><i>p</i> = 0.014                          | Parent died  |
| Parent died           | chr2: 18766018-<br>18766295   | 4       | <i>NT5C1B</i>       | --+<br><i>p</i> = 9.8e-06*                           | ----<br><i>p</i> = 0.014                           | Parent died  |

<sup>a</sup>Direction of effect for each individual CpG that is part of the DMR was derived from the regression coefficient in the epigenome-wide analysis for individual CpGs

<sup>b</sup>If the DMR additionally replicated for another ACE (same direction of effect and StoufferLiptak-Kechris corrected *P*-value <0.05 in both ALSPAC and NSHD), the additional ACE is mentioned in this column

\*DMR passed Šidák correction for multiple testing in comb-p

## Discussion

In this epigenome-wide study of childhood adversity in two population-based cohorts of mid-life women, we discovered nine novel differentially methylated genomic regions (DMRs) but no individual CpGs that replicated across cohorts. One DMR was associated with a measure for cumulative adversity (ACE count score), whereas the other eight regions were associated with specific types of adversity, namely parental mental illness (two regions), parental physical illness (three regions) and parental death (three regions).

The most robust observation in this study was for a DMR negatively associated with the ACE score, which was also associated with parental death, suboptimal maternal bonding and child maltreatment in ALSPAC and NSHD. The DMR is located near a CpG island of the promotor of the tonsoku-like DNA repair protein (*TONSL*) gene (see Fig. 1), which is a histone chaperone that may counteract chromatin compaction by preventing the association between the histone H4 tail with the H2A–H2B acidic patch on neighbouring nucleosomes<sup>42</sup>. The DMRs related to a specific ACE tended to be related

to only one adversity, which is in line with a previous study where different DNA methylation CpG sites were affected by specific types of adversity compared to several types of childhood adversity<sup>11</sup>. This appears to suggest that certain parts of the methylome are associated with distinct adverse childhood experiences, while other locations reflect more general exposure to adversity during childhood.

The eight replicated DMRs related to a specific ACE were all associated with parental health measures (parental mental illness, parental physical illness and parental death). This could indicate that parental health has a bigger impact on the child's epigenome than other types of adversity. Alternatively, it may reflect intergenerational transmission of health from parent to child with the woman's own health in turn influencing her epigenome. Parental physical illness and parental death are linked to six of the eight replicated DMRs. These ACEs have the most similar prevalence in ALSPAC and NSHD suggesting that similarity of measurements and hence effective harmonisation across cohorts aided replication. Our lack of associations for other measures of adversity is in line



**Table 4 Influence of genetic variants near the replicated DMRs in ALSPAC**

| ACE                   | DMR                       | Nr CpGs | Gene                           | Original analysis<br>(Direction <sup>a</sup> , p-value) | Nr SNPs (within 1MB,<br>associated one CpG (p<br>< 0.05)) | Sample size | Sensitivity analysis,<br>including SNPs<br>(Direction <sup>a</sup> , p-value) <sup>b</sup> |
|-----------------------|---------------------------|---------|--------------------------------|---|---|-------------|--|
| ACE count score       | chr8: 145654565-145654855 | 5       | <i>VPS28</i> ,<br><i>TONSL</i> | ----<br><i>p</i> = 7.5e−08*                             | 8   | 618         | ----<br><i>p</i> = 0.0020  |
| Parent mentally ill   | chr12: 14720726-14721289  | 10      | <i>PLBD1</i>                   | -+++++<br><i>p</i> = 0.0058                             | 28  | 532         | -+±+++++<br><i>p</i> = 0.0496  |
| Parent mentally ill   | chr1: 3104999-3105327     | 5       | <i>PRDM16</i>                  | ----<br><i>p</i> = 0.001                                | 16  | 672         | ----<br><i>p</i> = 0.001   |
| Parent physically ill | chr15: 81426347-81426670  | 9       | <i>C15orf26</i>                | +++++<br><i>p</i> = 0.011                               | 20  | 609         | +++++<br><i>p</i> = 0.0010   |
| Parent physically ill | chr22: 27834439-27834630  | 3       | n/a                            | ---<br><i>p</i> = 6.9e−06*                              | 35  | 686         | ---<br><i>p</i> = 1.207e−05*   |
| Parent physically ill | chr8: 144120335-144120707 | 7       | <i>C8orf31</i>                 | +++++<br><i>p</i> = 4.4e−06*                            | 13  | 673         | +++++<br><i>p</i> = 7.998e−05*   |
| Parent died           | chr15: 40364524-40364863  | 3       | n/a                            | +++<br><i>p</i> = 6.3e−07*                              | 19  | 605         | +++<br><i>p</i> = 8.915e−06*   |
| Parent died           | chr7: 24323261-24323940   | 9       | <i>NPY</i>                     | +++++<br><i>p</i> = 9.2e−07*                            | 40  | 584         | +++++<br><i>p</i> = 0.0005   |
| Parent died           | chr2: 18766018-18766295   | 4       | <i>NT5C1B</i>                  | ---+<br><i>p</i> = 9.8e−06*                             | 27  | 520         | -±-<br><i>p</i> = 0.0289   |

<sup>a</sup>Direction of effect for each individual CpG that is part of the DMR was derived from the regression coefficients for individual CpGs

<sup>b</sup>The CpGs with a different direction of effect in the sensitivity analysis are underlined and in red.

\*DMR passed Šidák correction for multiple testing in comb-p

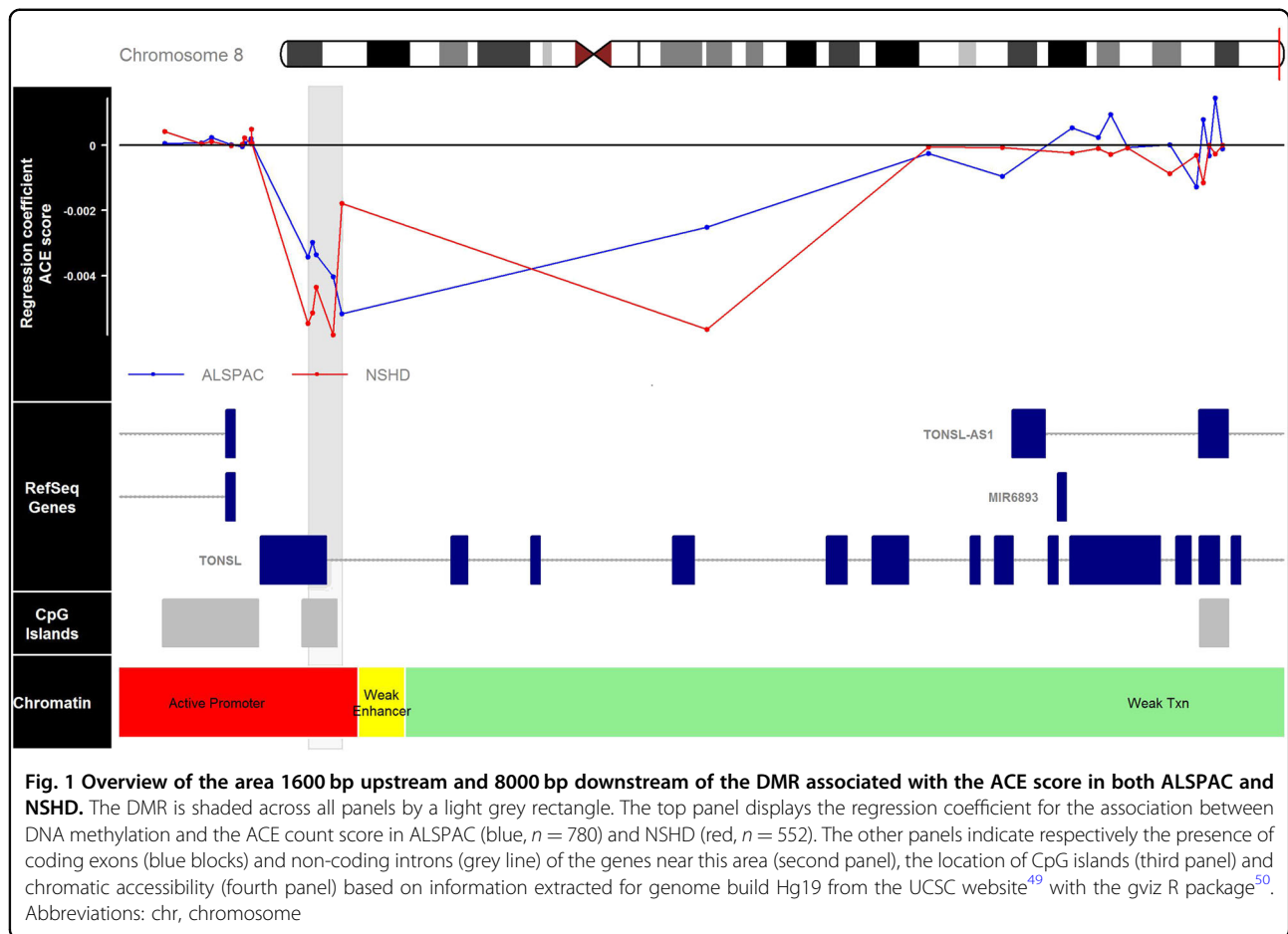
with a recent epigenome-wide study reporting limited evidence for an association between DNA methylation and several forms of early-life victimisation stress<sup>18</sup>.

The replication of DMRs is noteworthy given the considerable differences between ALSPAC and NSHD in tissue type (blood or buccal), source of reporting (retrospective self-report or mother during childhood), and years of birth (1950–1976 and 1946). The cohort differences may explain differences in adversity prevalence which are known to be affected by source of reporting<sup>43–45</sup>. Our own cross-cohort analyses also suggest cohort differences may play a role, as the regression coefficients for the same ACE exposure did not correlate across cohorts but were correlated between tissue types within the same cohort despite large, genome-wide DNA methylation differences between tissues<sup>46</sup>. Furthermore, our genetic analysis does not indicate the methylation differences at the DMRs are driven solely by genetic variation, as the direction of effect for the individual CpGs is similar and the DMRs remain significant after adjusting for genetic variants within 1MB of the DMR.

Our tissue comparison showed medium to high correlation across buccal and blood tissue for CpGs and DMRs associated with the same ACE, whereas the cross-tissue

correlation was lower for different ACEs. This suggests similar childhood adversity related DNA methylation changes could be identified in blood and buccal tissue. Some recommend buccal samples for population epigenetic studies, as they contain more hypomethylated DNA regions, which tend to cluster around disease-associated SNPs;<sup>47</sup> others, however, argue that demographic factors may be better reflected in blood DNA methylation patterns<sup>46</sup>. Identifying DMRs that are associated in both tissue types supports the robustness of our findings.

Although the non-replication of the three CpGs that survived FDR correction in NSHD could be related to cohort differences, the distribution of the p-values also suggests that there were no associations or we were underpowered to identify individual CpGs in each cohort (Supplementary Figure 5 and 6) even though smaller case-control studies previously reported evidence for associations<sup>11,13,15</sup>. Although selection bias—with the DNA methylation sample being slightly more affluent than the full cohort—may bias results towards the null<sup>48</sup>, in our study most of the ACEs had similar prevalence estimates in the DNA methylation subsample, suggesting selection bias is not likely to have affected our findings. One potential explanation for the contrast between findings in



case-control studies and the lack of individual CpGs in this and other population-based cohort studies is that childhood adversity related DNA methylation differences are more subtle in population-based cohorts compared to high risk case-control designs due to differences in severity of the adversity exposures and the instruments used to assess childhood adversity. These differences may also explain the lack of enrichment of most previously identified CpGs in our population-based cohorts. In this context, it is striking that the most robust enrichment was for the Dunedin study, another population-based cohort study<sup>18</sup>. Here, one out of two CpGs associated with sexual abuse replicated in ALSPAC. Note that enrichment in NSHD could not be determined due to a lack of sexual abuse data. The Dunedin cohort had the most similarities to our ALSPAC sample: retrospectively reported sexual abuse exposure between 0–16 years in a relatively large (~800) population-based birth cohort sample, with DNA methylation assessed in mid-life.

Overall, we find evidence that adverse childhood experiences, particularly those related to parental health, can have long-term effects on peripheral DNA

methylation in mid-life. We find some evidence to replicate a previously identified CpG site associated with sexual abuse.

#### Code availability

Code to generate these results is available from the authors on request.

#### Acknowledgements

We are extremely grateful to all the families who took part in this 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. We thank the study participants for their continuing participation in the MRC National Survey of Health and Development (NSHD). We also thank members of the NSHD scientific and data collection teams who have been involved in the NSHD data collections. This work was supported by grants from the UK Economic and Social Research Council (ES/M010317/1 and ES/N000382/1). Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award No. R01AG048835. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. L.D.H. is supported by a Career Development Award fellowship from the UK Medical Research Council (MR/M020894/1). L.C.H., M.S., C.R., E.L.A. and L.D.H. work in a unit that receives funding from the University of Bristol and the UK Medical Research Council (MC\_UU\_12013/2). The UK Medical Research Council and the

Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This work is supported by the UK Medical Research Council which provides core funding for the MRC National Survey of Health and Development and R.H. and D.K. [MC\_UU\_12019/1].

#### Conflict of interest

The authors declare that they have no conflict of interest.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41398-018-0307-3>).

Received: 15 June 2018 Revised: 5 September 2018 Accepted: 5 October 2018

Published online: 03 December 2018

#### References

- Kessler, R. C. et al. Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *Br. J. Psychiatry* **197**, 378–385 (2010).
- Felitti, V. J. et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. *Am. J. Prev. Med.* **14**, 245–258 (1998).
- Gilbert, R. et al. Burden and consequences of child maltreatment in high-income countries. *Lancet* **373**, 68–81 (2009).
- Pinto Pereira, S. M., Li, L. & Power, C. Child maltreatment and adult living standards at 50 years. *Pediatrics* **139**, e20161595–e20161595 (2017).
- Bird, A. P. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
- Suderman, M. et al. Lymphoblastoid cell lines reveal associations of adult DNA methylation with childhood and current adversity that are distinct from whole blood associations. *Int. J. Epidemiol.* **44**, 1331–1340 (2015).
- Demetriou, C. A. et al. Biological embedding of early-life exposures and disease risk in humans: a role for DNA methylation. *Eur. J. Clin. Invest.* **45**, 303–332 (2015).
- Vinkers, C. H. et al. Traumatic stress and human DNA methylation: a critical review. *Epigenomics* **7**, 593–608 (2015).
- Kumsta, R. et al. Severe psychosocial deprivation in early childhood is associated with increased DNA methylation across a region spanning the transcription start site of CYP2E1. *Transl. Psychiatry* **6**, e830 (2016).
- Prados, J. et al. Borderline personality disorder and childhood maltreatment: a genome-wide methylation analysis. *Genes, Brain Behav.* **14**, 177–188 (2015).
- Cecil, C. A. M. et al. Epigenetic signatures of childhood abuse and neglect: Implications for psychiatric vulnerability. *J. Psychiatr. Res.* **83**, 184–194 (2016).
- Mehta, D. et al. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proc. Natl Acad. Sci. USA* **110**, 8302–8307 (2013).
- Yang, B.-Z. et al. Child abuse and epigenetic mechanisms of disease risk. *Am. J. Prev. Med.* **44**, 101–107 (2013).
- Weder, N. et al. Child abuse, depression, and methylation in genes involved with stress, neural plasticity, and brain circuitry. *J. Am. Acad. Child Adolesc. Psychiatry* 2014; **53**. <https://doi.org/10.1016/j.jaac.2013.12.025>.
- Marinova, Z. et al. DNA methylation profiles of elderly individuals subjected to indented childhood labor and trauma. *Bmc. Med. Genet.* **18**, 21 (2017).
- Klengel, T. et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *NatNeurosci* **16**, 33–41 (2013).
- Tyrka, A. R. et al. Methylation of the leukocyte glucocorticoid receptor gene promoter in adults: associations with early adversity and depressive, anxiety and substance-use disorders. *Transl. Psychiatry* **6**, e848 (2016).
- Marzi, S. J. et al. Analysis of DNA methylation in young people: limited evidence for an association between victimization stress and epigenetic variation in blood. *Am. J. Psychiatry* **75**, 517–529 (2018). [appi.ajp.2017.17060693](https://doi.org/10.1176/appi.ajp.2017.17060693).
- Anderson, E. L. et al. Associations of adversity in childhood and risk factors for cardiovascular disease in mid-adulthood. *Child Abus Negl.* **76**, 138–148 (2018).
- Anderson, E. L. et al. Adversity in childhood and measures of aging in midlife: findings from a cohort of british women. *Psychol. Aging* **32**, 521–530 (2017).
- Anderson, E. L. et al. Prospective associations of psychosocial adversity in childhood with risk factors for cardiovascular disease in adulthood: The MRC National Survey of Health and Development. *Int. J. Equity Health* **16**, 1–7 (2017).
- Magnus, M. C. et al. Childhood psychosocial adversity and female reproductive timing: a cohort study of the ALSPAC mothers. *J. Epidemiol. Community Health* **72**, 34–40 (2018).
- Caleyachetty, R. et al. Modeling exposure to multiple childhood social risk factors and physical capability and common affective symptoms in later life. *J. Aging Health* **30**, 386–407 (2018).
- Dong, M. et al. The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Child Abus Negl.* **28**, 771–784 (2004).
- Van Niel, C., Pachter, L. M., Wade, R., Felitti, V. J. & Stein, M. T. Adverse events in children: predictors of adult physical and mental conditions. *J. Dev. Behav. Pediatr.* **35**, 549–551 (2014).
- Fraser, A. et al. Cohort profile: The avon longitudinal study of parents and children: ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
- Boyd, A., et al. Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* 2013; **42**. <https://doi.org/10.1093/ije/dys064>.
- Relton C. L., et al. Data resource profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *Int. J. Epidemiol.* 2015; **44**. <https://doi.org/10.1093/ije/dyv072>.
- Kuh D., et al. Cohort Profile: Updating the cohort profile for the MRC National Survey of Health and Development: A new clinic-based data collection for ageing research. *Int. J. Epidemiol.* 2011; **40**. <https://doi.org/10.1093/ije/dyq231>.
- Kuh, D. et al. The MRC National Survey of Health and Development reaches age 70: maintaining participation at older ages in a birth cohort study. *Eur. J. Epidemiol.* **31**, 1135–1147 (2016).
- Teschendorff, A. E. et al. Correlation of Smoking-Associated DNA Methylation Changes in Buccal Cells With DNA Methylation Changes in Epithelial Cancer. *JAMA Oncol.* **1**, 476–485 (2015).
- Min J, Hemani G., Davey Smith G., Relton C. L., Suderman M. Meffil: efficient normalisation and analysis of very large DNA methylation samples. *bioRxiv* 2017. <http://biorxiv.org/content/early/2017/04/27/125963.abstract>.
- Fortin J.-P., et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *bioRxiv* 2014. <http://biorxiv.org/content/early/2014/02/23/002956.abstract>.
- R-Core-Team. R: A language and environment for statistical computing. 2014. <http://www.r-project.org/>.
- Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
- Teschendorff, A. E., Zhuang, J. & Widschwendter, M. Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. *Bioinformatics* **27**, 1496–1505 (2011).
- Houseman, E. A. et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
- Joehanes, R. et al. Epigenetic Signatures of Cigarette Smoking. *Circ. Cardiovasc Genet* **9**, 436–447 (2016).
- Pedersen, B. S., Schwartz, D. A., Yang, I. V. & Kechris, K. J. Comb-p: Software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics* **28**, 2986–2988 (2012).
- Naeem H., et al Reducing the risk of false discovery enabling identification of biologically significant genome-wide methylation status using the Human-Methylation450 array. *BMC Genomics* 2014; **15**. <https://doi.org/10.1186/1471-2164-15-51>.
- Zhou, W., Laird, P. W. & Shen, H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res.* **45**, e22 (2017).
- Saredi, G. et al. H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex. *Nature* **534**, 714–718 (2016).
- Naicker S. N., Norris S. A., Mabaso M., Richter L. M. An analysis of retrospective and repeat prospective reports of adverse childhood experiences from the South African Birth to Twenty Plus cohort. *PLoS One* 2017; **12**. <https://doi.org/10.1371/journal.pone.0181522>.
- Reuben, A. et al. Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *J. Child Psychol. Psychiatry Allied Discip.* **57**, 1103–1112 (2016).

45. Newbury, J. B. et al. Measuring childhood maltreatment to predict early-adult psychopathology: comparison of prospective informant-reports and retrospective self-reports. *J. Psychiatr. Res.* **96**, 57–64 (2018).
46. Jiang, R. et al. Discordance of DNA Methylation Variance Between two Accessible Human Tissues. *Sci. Rep.* **5**, 8257 (2015).
47. Lowe, R. et al. Buccals are likely to be a more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics* **8**, 445–454 (2013).
48. Howe, L. D., Tilling, K., Galobardes, B. & Lawlor, D. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. *Epidemiology* **24**, 1–9 (2013).
49. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
50. Hahne F. et al. Gviz: Plotting data and annotation information along genomic coordinates. 2015. Accessed on 27 November 2018. <https://rdrr.io/bioc/Gviz/>.