Data Supplement for Marzi 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 (doi: 10.1176/appi.ajp.2017.17060693)

This Supplement contains additional details about the sample, measurement of victimization experiences in childhood and adolescence, and genome-wide quantification of DNA methylation. In addition, it contains additional figures and tables to accompany statistical analyses reported in the Main Article. The analysis plan for this paper was posted in advance (http://www.moffittcaspi.com).

Environmental Risk (E-Risk) Longitudinal Twin Study

Participants were members of E-Risk, which tracks the development of a 1994-95 birth cohort of 2,232 British children (T. E. Moffitt, 2002). Briefly, the E-Risk sample was constructed in 1999-2000, when 1,116 families (93% of those eligible) with same-sex 5-year-old twins participated in home-visit assessments. This sample comprised 56% monozygotic (MZ) and 44% dizygotic (DZ) twin pairs; sex was evenly distributed within zygosity (49% male). The study sample represents the full range of socioeconomic conditions in Great Britain, as reflected in the families' distribution on a neighborhood-level socioeconomic index (ACORN [A Classification of Residential Neighbourhoods], developed by CACI Inc. for commercial use): 25.6% of E-Risk families live in "wealthy achiever" neighborhoods compared to 25.3% nationwide; 5.3% vs. 11.6% live in "urban prosperity" neighborhoods; 29.6% vs. 26.9% in "comfortably off" neighborhoods; 13.4% vs. 13.9% in "moderate means" neighborhoods; and 26.1% vs. 20.7% in "hard-pressed" neighborhoods. E-Risk underrepresents "urban prosperity" neighborhoods because such households are often childless.

Home visits were conducted when participants were aged 5, 7, 10, 12 and most recently, 18 years (93% participation). The Joint South London and Maudsley and the Institute of Psychiatry Research Ethics Committee approved each phase of the study. Parents gave informed consent and twins gave assent between 5-12 years and then informed consent at age 18.

At age 18, 2,066 participants were assessed, each twin by a different interviewer. The average age at the time of assessment was 18.4 years (SD = 0.36); all interviews were conducted after the 18th birthday. There were no differences between those who did and did not take part at age 18 in terms of socioeconomic status (SES) assessed when the cohort was initially defined (X^2=0.86, p=0.65), age-5 IQ scores (t=0.98, p=0.33), age-5 internalizing or externalizing behavior problems (t=0.40, p=0.69 and t=0.41, p=0.68, respectively), or childhood polyvictimization (z=0.51, p=0.61).

Our epigenetic study used DNA from a single tissue: blood. At age 18, whole blood was collected from 82% (N=1700) of the participants in 10mL K_2EDTA tubes. DNA was extracted from the buffy coat using a Flexigene DNA extraction kit (Qiagen, Hilden, Germany) following manufacturer's instructions. There were no differences between those who did and did not provide blood in terms of SES (X^2=2.14, p=0.34), age 5-IQ scores (t=1.48, p=0.14), age-5 internalizing or externalizing behaviour problems (t=1.21, p=0.23 and t=0.47, p=0.64), childhood polyvictimization (z=0.81, p=0.42), or adolescent polyvictimization (z=1.25, p=0.21). (Study

members who did not provide blood provided buccal swabs, but these were not included in our methylation analysis to avoid tissue-source confounds.)

Genome-wide quantification of DNA methylation

We assayed 1669 blood samples (out of 1700); 31 samples were not useable (e.g., due to low DNA concentration**).** ~500ng of DNA from each sample was treated with sodium bisulfite using the EZ-96 DNA Methylation kit (Zymo Research, CA, USA). DNA methylation was quantified using the Illumina Infinium HumanMethylation450 BeadChip ("Illumina 450K array") run on an Illumina iScan System (Illumina, CA, USA). Twin pairs were randomly assigned to bisulfite-conversion plates and Illumina 450K arrays, with siblings processed in adjacent positions to minimize batch effects. Fully methylated control samples (CpG Methylated HeLa Genomic DNA; New England BioLabs, MA, USA) were included in a random position on each plate; the distinct DNA methylation profile of this sample enabled us to confirm the experiment was successful and to ensure there were no plate mix-ups or rotations.

Data were imported using the *methylumIDAT* function in *methylumi* (Davis, Bilke, Tim Triche, & Bootwalla, 2015), and subjected to quality control analyses. First, we excluded all samples with median methylated ('M') and unmethylated ('U') intensities <2500. Second, using the ten control probes included on the 450K array, we examined the efficiency of the sodium bisulfite conversion reaction; samples were excluded if their "conversion score" was <80. Third, multidimensional scaling was performed for DNA methylation probes on each of the sex chromosomes and compared to the reported gender. Fourth, to confirm genetic identity of the DNA samples, we assessed genotype concordance between SNP probes on the 450K array and data generated using Illumina OmniExpress24v1.2 genotyping BeadChips.

Samples from 1658 participants passed our QC pipeline, including 1468 participants who were members of complete twin pairs (430 MZ pairs and 304 DZ pairs) and 190 participants whose co-twin did not have complete data (e.g., did not provide blood, did not pass QC). Data were processed with the *pfilter* function from the *wateRmelon* package (Pidsley et al., 2013) excluding 0 samples with >1% of sites with a detection p value >0.05, 567 sites with beadcount <3 in 5% of samples and 1448 probes with >1% of samples with detection p value >0.05. The data were normalized with the *dasen* function from the *wateRmelon* package (Pidsley et al., 2013).

Prior to any analyses, probes with common (> 5% MAF) SNPs within 10 bp of the single base extension and probes with sequences previously identified as potentially hybridizing to

multiple genomic loci were excluded (Chen et al., 2013; Price et al., 2013), resulting in a final dataset of 430,802 probes.

To permit control for technical variation, we used methylation-array control-probe principal components (Lehne et al., 2015). 28 principal components were needed to explain 90% of the variance. These principal components were used as covariates in all individual-level analyses. To control for cell type composition, we used as covariates cell-type proportions estimated from the methylation data (Houseman et al., 2012).

Assessment of victimization

Assessment of victimization in childhood.

We have previously reported evidence on the reliability and validity of our measurement of childhood victimization (Danese et al., 2017). Here we summarize the method.

Exposure to several types of victimization was assessed repeatedly when the children were 5, 7, 10, and 12 years of age and dossiers were compiled for each child with cumulative information about exposure to domestic violence between the mother and her partner; frequent bullying by peers; physical maltreatment by an adult; sexual abuse; emotional abuse and neglect; and physical neglect. All the component measures are outlined briefly below.

Physical intimate-partner violence. Mothers reported about perpetration by and victimization of 12 forms of physical violence (e.g., slapping, hitting, kicking, strangling) from the Conflict Tactics Scale (Terrie E. Moffitt, Caspi, Krueger, Magdol, & et al., 1997; Straus, Straus, & Gelles, 1990), on three assessment occasions during the child's first decade of life (when the children were 5, 7, and 10 years of age). Reports of either perpetration or victimization constituted evidence of physical domestic violence. Families in which no physical violence took place were coded as 0 (55.2%); families in which physical violence took place on one occasion were coded as 1 (28.0%); and families in which physical violence took place on multiple occasions were coded as 2 (16.8%).

Bullying by peers. Experiences of victimization by bullies were assessed using both mothers' and children's reports. During the interview, the following standard definition of bullying was read out: "Someone is being bullied when another child (a) says mean and hurtful things, makes fun, or calls a person mean and hurtful names; (b) completely ignores or excludes someone from their group of friends or leaves them out on purpose; (c) hits, kicks, or shoves a person, or locks them in a room; (d) tells lies or spreads rumors about them; and (e) other

hurtful things like these. We call it bullying when these things happen often, and when it is difficult to make it stop. We do not call it bullying when it is done in a friendly or playful way." Mothers were interviewed when children were 7, 10, and 12 years old and asked whether either twin had been bullied by another child, responding never, yes, or frequently. We combined mothers' reports at child age 7 and 10 to derive a measure of victimization during primary school. Mothers' reports when the children were 12 years old indexed victimization during secondary school. During private interviews with the children when they were 12 years old, the children indicated whether they had been bullied by another child during primary or secondary school. When a mother or a child reported victimization, the interviewer asked them to describe what happened. Notes taken by the interviewers were later checked by an independent rater to verify that the events reported could be classified as instances of bullying operationally defined as evidence of (a) repeated harmful actions, (b) between children, and (c) where there is a power differential between the bully and the victim. Although inter-rater reliability between mothers and children was only modest (kappa = 0.20–0.29), reports of victimization from both informants were similarly associated with children's emotional and behavioral problems, suggesting that each informant provides a unique but meaningful perspective on bullying involvement (Shakoor et al., 2011). We thus combined mother and child reports of victimization to capture all instances of bullying victimization for primary and secondary school separately: reported as not victimized by both mother and child; reported by either mother or child as being occasionally victimized; and reported as being occasionally victimized by both informants or as frequently victimized by either mother or child or both (Bowes et al., 2013). We then combined these primary and secondary school ratings to create a bullying victimization variable for the entire childhood period (5–12 years). Children who were never bullied in primary or secondary school or occasionally bullied during one of these time periods were coded as 0 (55.5%); children who were occasionally bullied during primary and secondary school, or frequently bullied during one of these time periods were coded as 1 (35.6%); and children who were frequently bullied at both primary and secondary school were coded as 2 (8.9%).

Physical and sexual harm by an adult. When the twins were aged 5, 7, 10 and 12, their mothers were interviewed about each twins' experience of intentional harm by an adult. An unusual feature of our assessments is that we repeatedly interviewed mothers over the years, allowing them to build confidence in the research team. Also, we were able to reassure mothers that if harm to the child was ongoing and had to be reported by us, reporting would be managed through a trusted person, namely the family's registered GP. As the children got older, some

mothers who were initially reluctant to reveal abuse, revealed details of severe abuse at a later interview. At age 5 we used the standardized clinical protocol from the MultiSite Child Development Project (Dodge, Bates, & Pettit, 1990; Lansford et al., 2002). At ages 7, 10, and 12 this interview was modified to expand its coverage of contexts for child harm. Interviews were designed to enhance mothers' comfort with reporting valid child maltreatment information, while also meeting researchers' responsibilities for referral under the U.K. Children Act. Specifically, mothers were asked whether either of their twins had been intentionally harmed (physically or sexually) by an adult or had contact with welfare agencies. If caregivers endorsed a question, research workers made extensive notes on what had happened, and indicated whether physical and/or psychological harm had occurred. Under the U.K. Children Act, our responsibility was to secure intervention if maltreatment was current and ongoing. Such intervention on behalf of E-Risk families was carried out through the family's registered physician with parental cooperation in all but one case. No families left the study following intervention. Over the years of data collection, the study developed a cumulative profile for each child, comprising the caregiver reports, recorded debriefings with research workers who had coded any indication of maltreatment at any of the successive home visits, recorded narratives of the successive caregiver interviews, and information from clinicians whenever the Study team made a child-protection referral. The profiles were reviewed at the end of the age–12 phase by two clinical psychologists. Inter-rater agreement between the coders was 90% of cases for whom maltreatment was identified (100% for cases of sexual abuse), and discrepantly coded cases were resolved by consensus review. These were coded as: 0 = no physical harm at any age; 1 = probable physical harm at any age; and 2 = definite physical harm at any age. There were 15.0% of children coded as probably being exposed to physical harm and 5.1% as definitely physically harmed by 12 years of age. There were 1.5% of the children coded as being exposed to sexual abuse.

Emotional abuse and neglect were coded from research workers' narratives of the home visits at ages 5, 7, 10, and 12. We coded quite severe examples of parental behavior observed. For example, a mother who had schizophrenia screamed and swore at the children throughout the home visit. As another example, a father who was drunk during the home visit repeatedly spoke abusively to the children in front of the research workers. We found that coders could not empirically separate emotional abuse and emotional neglect in a reliable way and thus such experiences were coded together as emotional abuse/neglect. Inter-rater agreement between the coders exceeded 85% for cases with emotional abuse and neglect, and discrepant cases

were resolved by consensus review. Children with no evidence of emotional abuse/neglect were coded as 0 (88.3%), those where there was some indication of emotionally inappropriate/potentially abusive or neglectful behavior were coded as 1 (8.7%), and where there was evidence of severe emotional abuse/neglect the children were coded as 2 (3.0%).

Physical neglect. The cumulative observations of the physical state of the home environment documented by the research workers during home visits to the twins at ages 5, 7, 10 and 12 were reviewed by two raters for evidence of physical neglect. This was defined as any sign that the caretaker was not providing a safe, sanitary, or healthy environment for the child. This included the child not having proper clothing or food, as well as grossly unsanitary home environments. (However, this did not include a family living in a crime-ridden neighborhood for economic reasons.) Inter-rater agreement between the coders was 85%, and discrepantly coded cases were resolved by consensus review. Children with no evidence of physical neglect were coded as 0 (90.9%), those for whom there was an indication of minor physical neglect were coded as 1 (7.1%), and where there was evidence of severe physical neglect the children were coded as 2 (2.0%).

Assessment of victimization in adolescence.

We have previously reported evidence on the reliability and validity of our measurement of adolescent victimization (Fisher et al., 2015). Here we summarize the method.

Participants were interviewed about experiences between 12-18 years using the Juvenile Victimization Questionnaire (JVQ) (Finkelhor, Hamby, Turner, & Omrod, 2011; Hamby SL, 2004), adapted as a clinical interview. The JVQ has good psychometric properties (Finkelhor, Hamby, Ormrod, & Turner, 2005) and was used in the U.K. National Society for the Prevention of Cruelty to Children national survey (L. Radford, Corral, Bradley, & Fisher, 2013; L. C. Radford, S. Bradley, C., et al, 2011), thereby providing benchmark values for comparisons with our cohort.

Within each pair of twins in our cohort, co-twins were interviewed separately by a different research worker and were assured of the confidentiality of their responses. The participants were advised that confidentiality would only be broken if they told the research worker that they were in immediate danger of being hurt, and in such situations the project leader would be informed and would contact the participant to discuss a plan for safety.

We assessed 7 different forms of victimization: maltreatment, neglect, sexual victimization, family violence, peer/sibling victimization, internet/mobile phone victimization, and crime victimization. Each JVQ question was asked for the period ‘since you were 12’.

Participants were given the option to say “yes” or “no” as to whether each type of victimization had occurred in the reporting period. Research workers could rate each item “maybe” if the participant seemed unsure or hesitant in their response or they were not convinced that the participant understood the question or was paying attention. Items rated as “maybe” were recoded as “no” or “yes” by the rating team based on the notes provided by the research workers. When insufficient notes were available, these responses were recoded conservatively as a “no”. Consistent with the JVQ manual (Finkelhor et al., 2011; Hamby SL, 2004), participants were coded as 1 if they reported any experience within each type of victimization category, or 0 if none of the experiences within the category were endorsed. If an experience was endorsed within a victimization category, follow-up questions were asked concerning how old the participant was when it (first) happened, whether the participant was physically injured in the event, whether the participant was upset or distressed by the event; and how long it went on for (by marking the number of years on a Life History Calendar; (Caspi et al., 1996)). In addition, the interviewer wrote detailed notes based on the participant’s description of the worst event. If multiple experiences were endorsed within a victimization category, the participant was asked to identify and report about their worst experience.

All information from the JVQ interview was compiled into victimization dossiers. Using these dossiers, each of the seven victimization categories was rated by an expert in victimology and 3 other members of the E-Risk team who were trained on using the rating criteria. Ratings were made using a 6-point scale: 0 = not exposed, then 1-5 for increasing levels of severity. The anchor points for these ratings were adapted from the coding system used for the Childhood Experience of Care and Abuse interview (CECA; (Bifulco, Brown, & Harris, 1994), which has good inter-rater reliability (Bifulco et al., 1994). The CECA is a comprehensive semi-structured interview whose standardized coding system attempts to improve the objectivity of ratings by basing them on the coder’s perspective (rather than relying on the participant’s judgment) and focusing on concrete descriptions rather than perceptions or emotional responses to the questions, together with considering the context in which the adverse experience occurred.

In our adapted coding scheme, the anchor points of the scale differ for each victimization category, with some focused more on the severity of physical injury that is likely to have been incurred during victimization exposure (crime victimization, family violence, maltreatment), while others are more focused on the frequency of occurrence of victimization (peer/sibling victimization and internet/mobile phone victimization), the physical intrusiveness of the event (sexual victimization), or the pervasiveness of the effects of victimization (neglect). This reflects

the different ways in which severity has previously been defined for different types of victimization (Bifulco et al., 1994). (Given that our sample comprises twins, we also coded if any of the victimization events experienced by each twin had been perpetrated by their co-twin, as it is possible that growing up with a genetically related, same-age child could increase or decrease sibling victimization rates.) Each twin's dossier was evaluated separately and we did not use information provided in the co-twin's dossier about their own or shared victimization experiences to rate direct or witnessed violence exposure for the target twin. High levels of inter-rater reliability were achieved for the severity ratings for all forms of victimization: crime victimization (*intra-class correlation coefficient [ICC]* = 0.89, $p < 0.001$), peer/sibling victimization (*ICC* = 0.91, $p < 0.001$), internet/mobile phone victimization (*ICC* = 0.90, $p < 0.001$), sexual victimization (*ICC* = 0.87, $p < 0.001$), family violence (*ICC* = 0.93, $p < 0.001$), maltreatment (*ICC* = 0.90, $p < 0.001$), and neglect (*ICC* = 0.74, $p < 0.001$).

The ratings for each type of victimization were then grouped into three classes: 0 – no exposure (score of 0), 1 – some exposure (score of 1, 2 or 3), and 2 – severe exposure (score of 4 or 5) due to small numbers for some of the rating points. Combining ratings of 4 and 5 is also consistent with previous studies using the CECA, which have collapsed comparable scale values to indicate presence of 'severe' abuse (e.g., (Bifulco et al., 1994) (Fisher, Bunn, Jacobs, Moran, & Bifulco, 2011)

Cumulative victimization across childhood and adolescence.

We performed a Latent Class Analysis (LCA) using longitudinal data about victimization. LCA is a person-centered technique that classifies individuals into groups based on a profile of variables, in this case the degree of each participant's exposure (i.e., none, moderate, or severe) to the six types of childhood and seven types of adolescent victimization. The LCA was performed using only participants who experienced at least one form of victimization, and was conducted in MPlus v7.4 accounting for clustering of twins within families. Supplementary Table S1 provides detailed results and model-fit statistics. We identified three victimized groups: Individuals who were (a) exposed to domestic violence in childhood (n=254); (b) victimized by peers and 'street crime' (n=412); or (c) exposed to multiple types of violence in both childhood and adolescence (n=158). N=834 individuals were not exposed to childhood or adolescent victimization).

Retrospectively-recalled child abuse and neglect, as reported on the Childhood Trauma Questionnaire (CTQ).

In addition to the above measures of victimization, we were able to look at the recall of victimization assessed through the Childhood Trauma Questionnaire (CTQ) (D. P. Bernstein et al., 1994), completed by Study members at the age–18 follow-up. The CTQ is a popular instrument used to collect retrospective information about the history of 5 categories of childhood maltreatment: emotional, physical, and sexual abuse and emotional and physical neglect. The validity of the instrument has been previously demonstrated in clinical and community samples. Participants reported on their personal experiences of emotional, physical, sexual abuse, and physical and emotional neglect for the period before they were 12 years old (corresponding to the reporting period for prospective reports about childhood maltreatment). Based on CTQ scoring guidelines (D.P. Bernstein & Fink, 1998), a specific category of maltreatment was considered present if the Study member had a moderate to severe score. A polyvictimization score was created by summing reports about these five experiences. 1,507 (91%) participants recalled zero moderate/severe victimization experiences; 97 (5.9%) 1; 25 (1.5%) 2; and 25 (1.5%) recalled 3 or more moderate/severe victimization experiences.

Dunedin Longitudinal Study

Participants were members of the Dunedin Multidisciplinary Health and Development Study, a longitudinal investigation of health and behavior in a representative birth cohort (Poulton, Moffitt, & Silva, 2015). Study members (n = 1,037; 91% of eligible births; 52% male) were all individuals born between April 1972 and March 1973 in Dunedin, New Zealand, who were eligible for the longitudinal study based on residence in the province at 3 years of age and who participated in the first follow-up assessment at 3 years of age. The cohort represented the full range of socioeconomic status on NZ's South Island. On adult health, the cohort matches the NZ National Health and Nutrition Survey (e.g., BMI, smoking, GP visits; (Poulton et al., 2015)). Cohort members are primarily white; approximately 7% self-identify as having partial non-Caucasian ancestry, matching the South Island. Assessments were carried out at birth and at ages 3, 5, 7, 9, 11, 13, 15, 18, 21, 26, 32, and 38 years, when 95% of the 1,007 study members still alive took part. The Otago Ethics Committee approved each phase of the study and informed consent was obtained from all study members.

Our epigenetic study used DNA from a single tissue: blood. At age 38, whole blood was collected from 90% (N=857) of the non-Maori participants in 10mL K_2EDTA tubes. DNA was

extracted from the buffy coat using standard procedures (Bowtell, 1987; Jeanpierre, 1987). Study members who did not provide blood provided buccal swabs, but these were not included in our methylation analysis to avoid tissue-source confounds.)

Genome-wide quantification of DNA methylation

We assayed 833 blood samples (out of 857); 22 samples were not useable (e.g., due to low DNA concentration**).** ~500ng of DNA from each sample was treated with sodium bisulfite using the EZ-96 DNA Methylation kit (Zymo Research, CA, USA). DNA methylation was quantified using the Illumina Infinium HumanMethylation450 BeadChip ("Illumina 450K array") run on an Illumina iScan System (Illumina, CA, USA).

Data were processed and normalized using the *methylumi* (v2.14.0) Bioconductor package from the R statistical programming environment, and subjected to quality control analyses. Samples were removed if the average detection p-value was >= 0.001. To confirm genetic identity of the DNA samples, we assessed genotype concordance between SNP probes on the 450K array and data generated using Illumina OmniExpress12v1.1 genotyping BeadChips. Principal components analysis was performed on the full, normalized dataset and the first two components plotted. Samples formed two major clusters separating on the 1st component, which corresponded to recorded sex. This was used to confirm sex assignment. Samples from 819 participants passed our QC pipeline. For the purpose of this study, only probes that passed QC standards for the E-Risk methylation analysis were used in the analysis.

To permit control for technical variation, we used methylation-array control-probe principal components (Lehne et al., 2015). 32 principal components were needed to explain 90% of the variance. These principal components were used as covariates in all individual-level analyses. To control for cell type composition, we used as covariates white cell-type counts measured using flow cytometry (Sysmex Corporation, Japan) in whole blood samples taken concurrently with the DNA sample.

Assessment of victimization

Study members also completed the Childhood Trauma Questionnaire (CTQ) (D. P. Bernstein et al., 1994) when they were 38 years old. The CTQ inquires about the history of 5 categories of childhood maltreatment: emotional, physical, and sexual abuse and emotional and physical neglect. Based on CTQ scoring guidelines (Bernstein and Fink, 1998), a specific category of maltreatment was considered present if the Study member had a moderate to severe score. A polyvictimization score was created by summing reports about these five experiences. 599 (73.2%) participants recalled zero moderate/severe victimization experiences;

163 (19.9%) 1; 47 (5.7%) 2; and 9 (1.1%) recalled 3 or more moderate/severe victimization experiences.

Data analysis

Linear regression was used to test the association between retrospective reports of childhood victimization and DNA methylation variation. The model included the following covariates: sex, methylation-array control-probe principal components indexing technical variation, and measured cell-type counts. To control for effects of smoking in methylation data, the model was re-fitted by adding information about smoking pack-years as a covariate.

References

Bernstein, D. P., & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report manual*.

Bernstein, D. P., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., Wenzel, K., . . . Ruggiero, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Am J Psychiatry, 151*(8), 1132-1136. doi:10.1176/ajp.151.8.1132

Bifulco, A., Brown, G. W., & Harris, T. O. (1994). Childhood Experience of Care and Abuse (Ceca) - a Retrospective Interview Measure. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35*(8), 1419-1435. doi:DOI 10.1111/j.1469-7610.1994.tb01284.x

Bowes, L., Maughan, B., Ball, H., Shakoor, S., Ouellet-Morin, I., Caspi, A., . . . Arseneault, L. (2013). Chronic bullying victimization across school transitions: the role of genetic and environmental influences. *Dev Psychopathol, 25*(2), 333-346. doi:10.1017/S0954579412001095

Bowtell, D. D. (1987). Rapid isolation of eukaryotic DNA. *Anal Biochem, 162*(2), 463-465.

Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., . . . Silva, P. A. (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research, 6*(2), 101-114. doi:Doi 10.1002/(Sici)1234-988x(199607)6:2<101::Aid-Mpr156>3.3.Co;2-E

Chen, Y. A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., . . . Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics, 8*(2), 203-209. doi:10.4161/epi.23470

Danese, A., Moffitt, T. E., Arseneault, L., Bleiberg, B. A., Dinardo, P. B., Gandelman, S. B., . . . Caspi, A. (2017). The Origins of Cognitive Deficits in Victimized Children: Implications for Neuroscientists and Clinicians. *Am J Psychiatry, 174*(4), 349-361. doi:10.1176/appi.ajp.2016.16030333

Davis, S., Bilke, S., Tim Triche, J., & Bootwalla, M. (2015). methylumi: Handle Illumina methylation data.

Dodge, K. A., Bates, J. E., & Pettit, G. S. (1990). Mechanisms in the cycle of violence. *Science, 250*(4988), 1678-1683.

Finkelhor, D., Hamby, S. L., Ormrod, R., & Turner, H. (2005). The Juvenile Victimization Questionnaire: reliability, validity, and national norms. *Child Abuse Negl, 29*(4), 383-412. doi:10.1016/j.chiabu.2004.11.001

Finkelhor, D., Hamby, S. L., Turner, H. A., & Omrod, R. K. (2011). The Juvenile Victimization Questionnaire: 2nd Revision (JVQ-R2). *Durham, NH: Crimes Against Children Research Center*.

Fisher, H. L., Bunn, A., Jacobs, C., Moran, P., & Bifulco, A. (2011). Concordance between mother and offspring retrospective reports of childhood adversity. *Child Abuse Negl, 35*(2), 117-122. doi:10.1016/j.chiabu.2010.10.003

Fisher, H. L., Caspi, A., Moffitt, T. E., Wertz, J., Gray, R., Newbury, J., . . . Arseneault, L. (2015). Measuring adolescents' exposure to victimization: The Environmental Risk (E-Risk)

Longitudinal Twin Study. *Dev Psychopathol, 27*(4 Pt 2), 1399-1416. doi:10.1017/S0954579415000838

Hamby SL, F. D., Ormrod RK, Turner HA. (2004). The juvenile victimization questionnaire (JVQ): Administration and scoring manual. *Durham NH*. *Crimes Against Children Research Center*.

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., . . . Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics, 13*, 86. doi:10.1186/1471-2105-13-86

Jeanpierre, M. (1987). A rapid method for the purification of DNA from blood. *Nucleic Acids Res, 15*(22), 9611.

Lansford, J. E., Dodge, K. A., Pettit, G. S., Bates, J. E., Crozier, J., & Kaplow, J. (2002). A 12-year prospective study of the long-term effects of early child physical maltreatment on psychological, behavioral, and academic problems in adolescence. *Arch Pediatr Adolesc Med, 156*(8), 824-830.

Lehne, B., Drong, A. W., Loh, M., Zhang, W., Scott, W. R., Tan, S. T., . . . Chambers, J. C. (2015). A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol, 16*, 37. doi:10.1186/s13059-015-0600-x

Moffitt, T. E. (2002). Teen-aged mothers in contemporary Britain. *J Child Psychol Psychiatry, 43*(6), 727-742.

Moffitt, T. E., Caspi, A., Krueger, R. F., Magdol, L., & et al. (1997). Do partners agree about abuse in their relationship?: A psychometric evaluation of interpartner agreement. *Psychological Assessment, 9*(1), 47-56.

Pidsley, R., CC, Y. W., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics, 14*, 293. doi:10.1186/1471-2164-14-293

Poulton, R., Moffitt, T. E., & Silva, P. A. (2015). The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Soc Psychiatry Psychiatr Epidemiol, 50*(5), 679-693. doi:10.1007/s00127-015-1048-8

Price, M. E., Cotton, A. M., Lam, L. L., Farre, P., Emberly, E., Brown, C. J., . . . Kobor, M. S. (2013). Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin, 6*(1), 4. doi:10.1186/1756-8935-6-4

Radford, L., Corral, S., Bradley, C., & Fisher, H. L. (2013). The prevalence and impact of child maltreatment and other types of victimization in the UK: Findings from a population survey of caregivers, children and young people and young adults. *Child Abuse & Neglect, 37*(10), 801-813. doi:10.1016/j.chiabu.2013.02.004

Radford, L. C., S. Bradley, C., et al. (2011). Child Abuse and neglect in the UK today. *London: NSPCC*.

Shakoor, S., Jaffee, S. R., Andreou, P., Bowes, L., Ambler, A. P., Caspi, A., . . . Arseneault, L. (2011). Mothers and children as informants of bullying victimization: Results from an epidemiological cohort of children. *Journal of Abnormal Child Psychology, 39*, 379-387.

Straus, M. A., Straus, M. A., & Gelles, R. J. (1990). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. In *Physical violence in American families: Risk factors*

and adaptions to violence in 8,145 families (pp. 29). New Brunswick, NJ: Transaction Books.

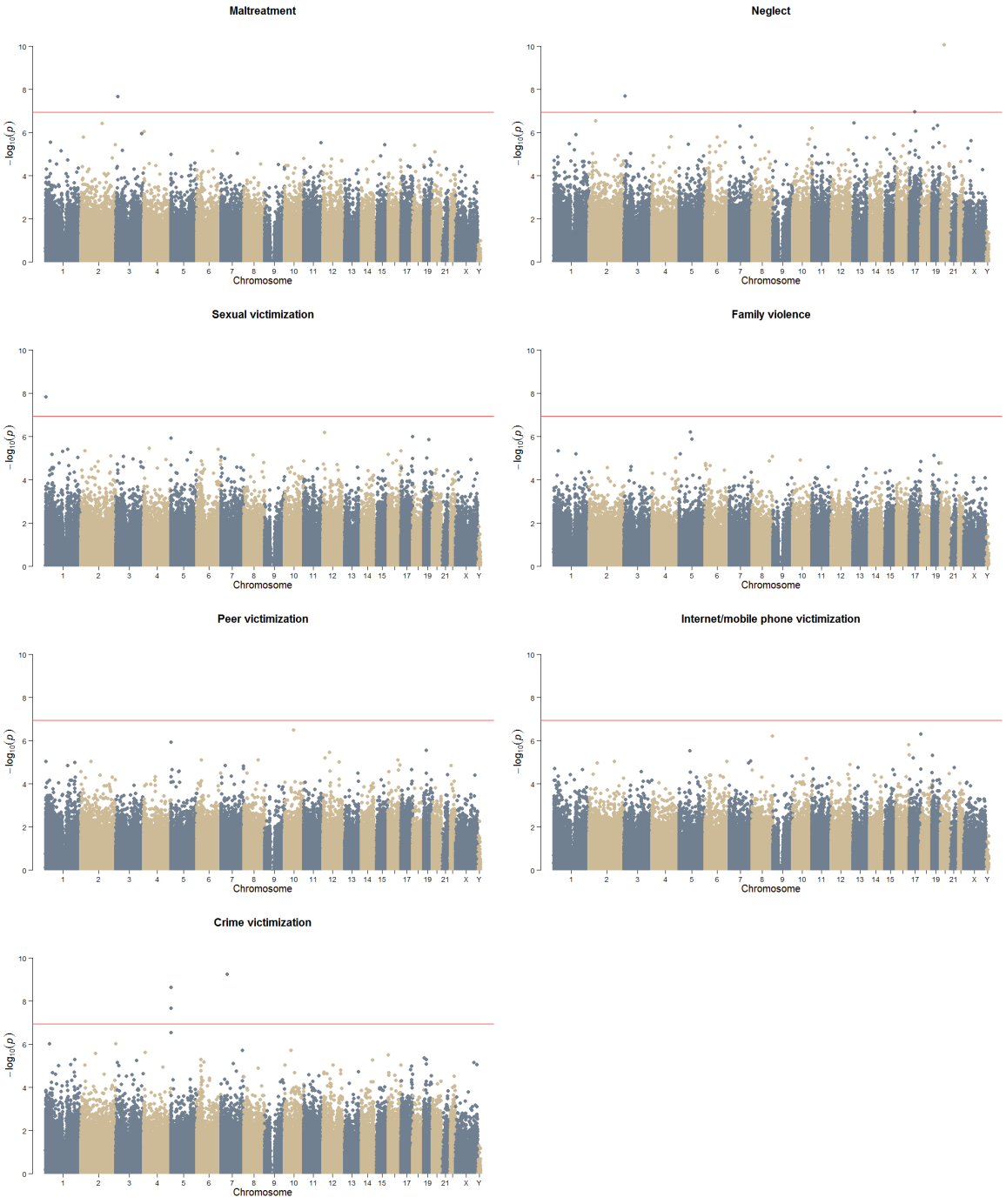

Figure S1. Few novel associations and sparsely distributed significant findings were observed for the association between seven individual types of adolescent victimization and DNA methylation. For the seven individual victimization types constituting the adolescent polyvictimization measure (i.e., maltreatment, neglect, sexual victimization, family violence, peer/sibling victimization, internet/mobile phone victimization, and crime victimization) a total

of eight array-wide significant associations were observed across four victimization types (maltreatment, neglect, sexual victimization and crime victimization). Two of these associations (probes cg05575921 and cg21161138, both annotated to *AHRR*) had been identified in the EWAS of adolescent polyvictimization (**Figure 1**).

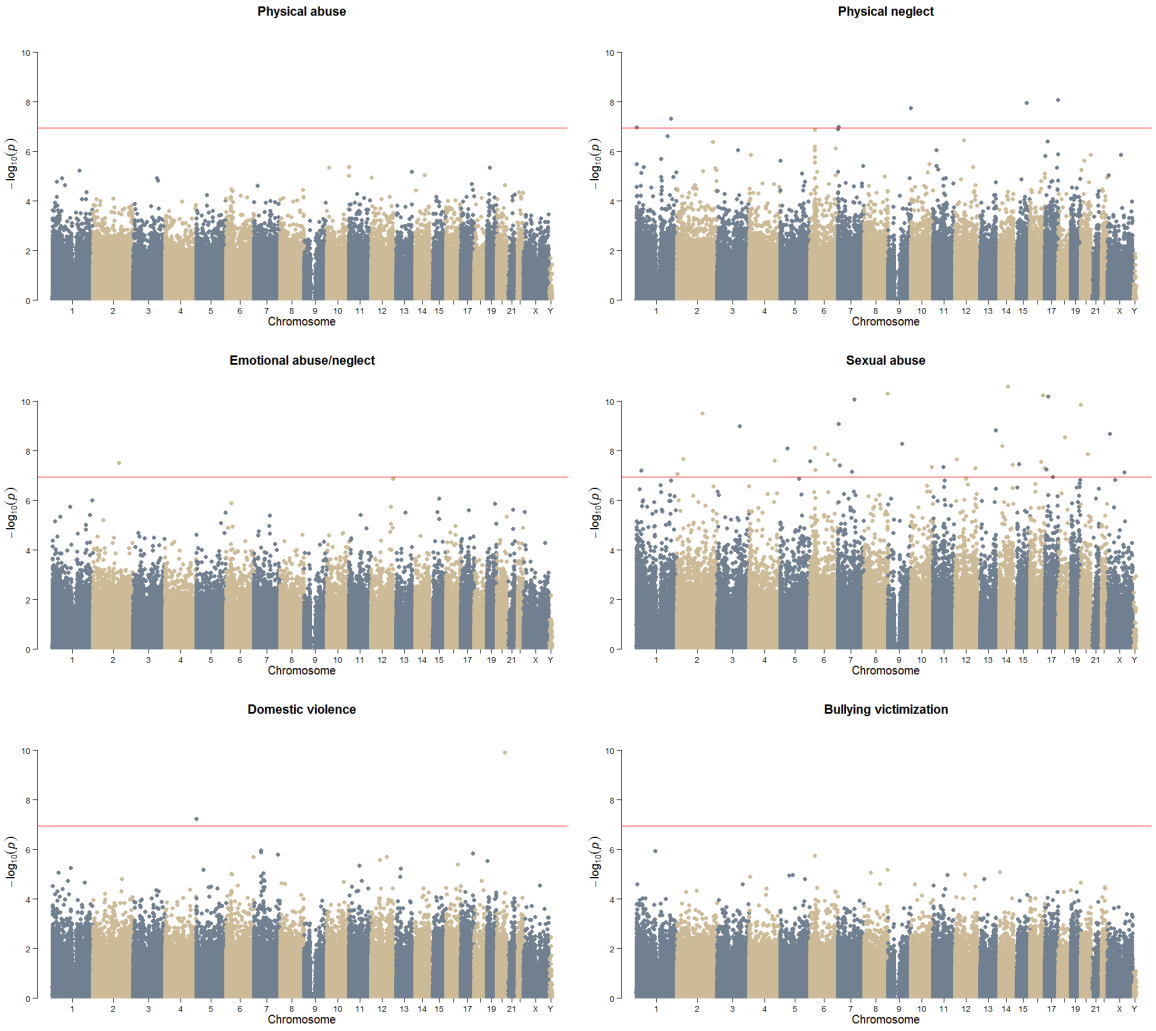

Figure S2. Few novel associations and sparsely distributed significant findings were observed for the association between six individual types of childhood victimization and DNA methylation. For the six individual victimization types constituting the childhood polyvictimization measure (i.e. physical abuse, physical neglect, emotional neglect, sexual victimization, domestic violence and peer victimization), a total of 48 array-wide significant associations were observed across four of the victimization types (emotional neglect, domestic violence, physical neglect and sexual victimization). None of these probes were shared between victimization types, nor were they identified in the EWAS of childhood polyvictimization (**Figure 3**). Interestingly, of these 48 probes, 39 were associated with childhood sexual victimization. These probes are listed in **Table S3**.

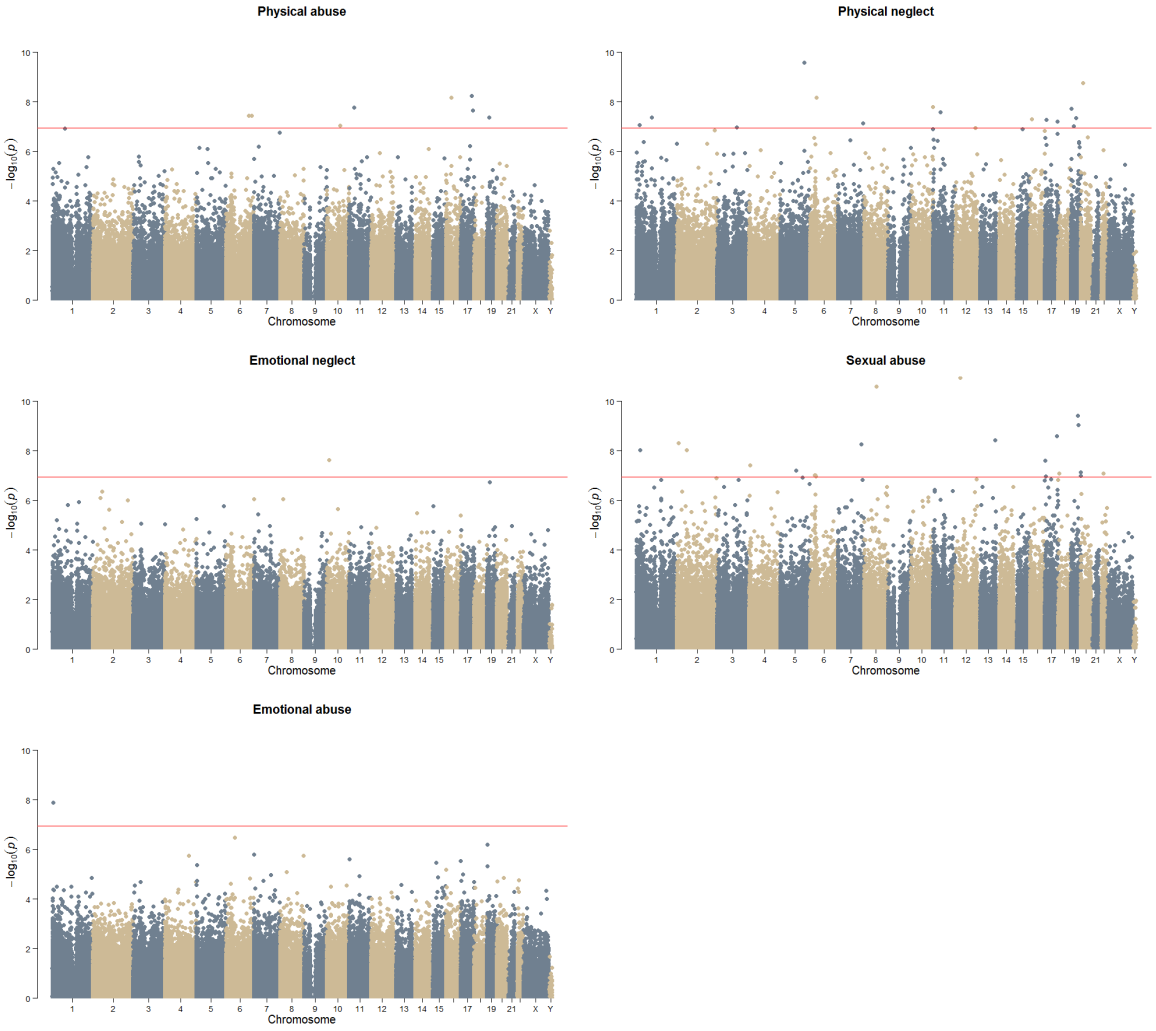

Figure S3. **Few novel associations and sparsely distributed significant findings were observed for the association between five individual types of retrospective childhood victimization measures and DNA methylation.** For the five individual victimization types constituting the retrospective childhood trauma questionnaire (CTQ) measure (i.e., physical abuse, physical neglect, emotional neglect, sexual abuse and emotional abuse), a total of 48 array-wide significant associations were observed across the five victimization types. None of these 48 probes were shared between victimization types and none were identified in the EWAS of childhood polyvictimization (**Figure 3**). Of these 48 probes, 22 were associated with sexual abuse; these probes are listed in **Table S3**. Additionally, those probes associated with sexual abuse were not shared with sexual victimization measured in adolescence (**Figure S1)** or childhood (**Figure S2).**

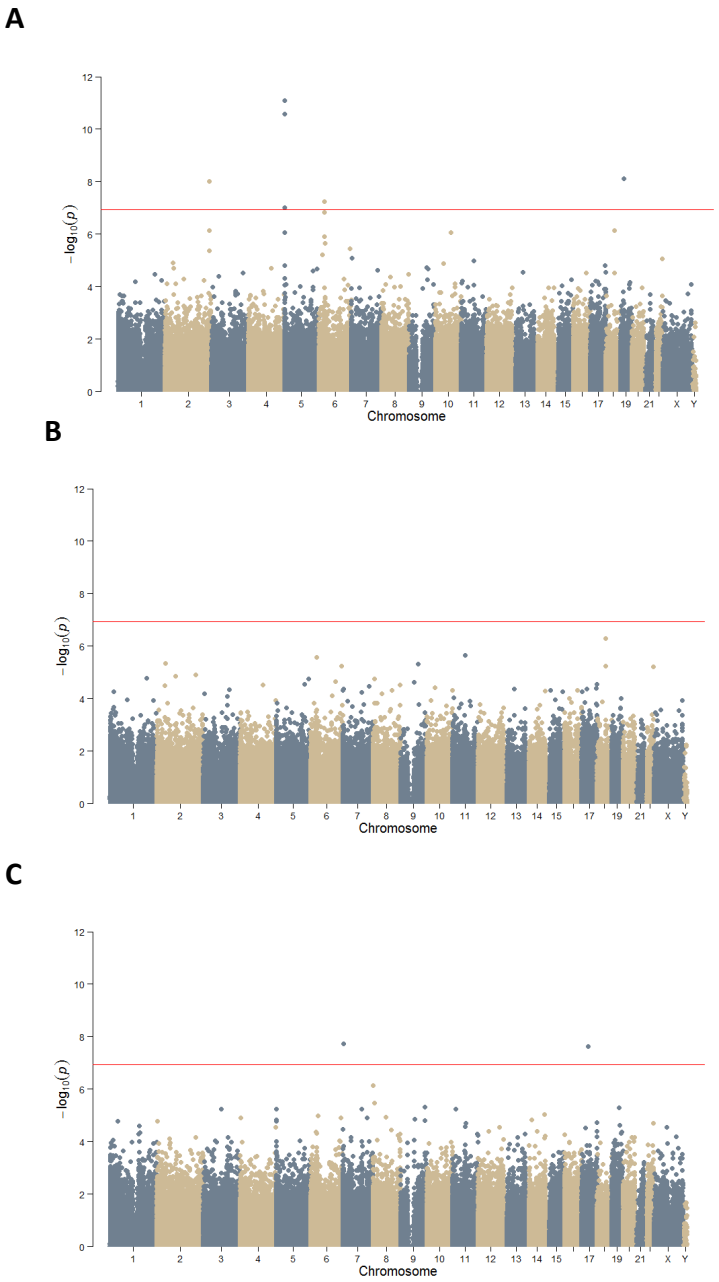

Figure S4. The association between retrospective reports of childhood victimization on the Childhood Trauma Questionnaire and DNA methylation in the Dunedin Longitudinal Study. **Panel A:** Six probes passed the array-wide multiple testing threshold ($P < 1.16 \times 10^{-7}$; red line). **Panel B**: We identified no significant associations when adding smoking pack-years as a further covariate, suggesting that the association between retrospective reports of victimization and DNA methylation is confounded by smoking. **Panel C:** Two probes were associated with retrospective reports of childhood sexual abuse; these probes are listed in **Table S3**. Neither of these probes were amongst the six identified in **Panel A**, nor those identified in analyses of

sexual victimization measured in adolescence (**Figure S1**), sexual victimization measured in childhood (**Figure S2**), or retrospective reports of childhood sexual abuse (**Figure S3**) in the E-Risk study (cf. **Table S3**).

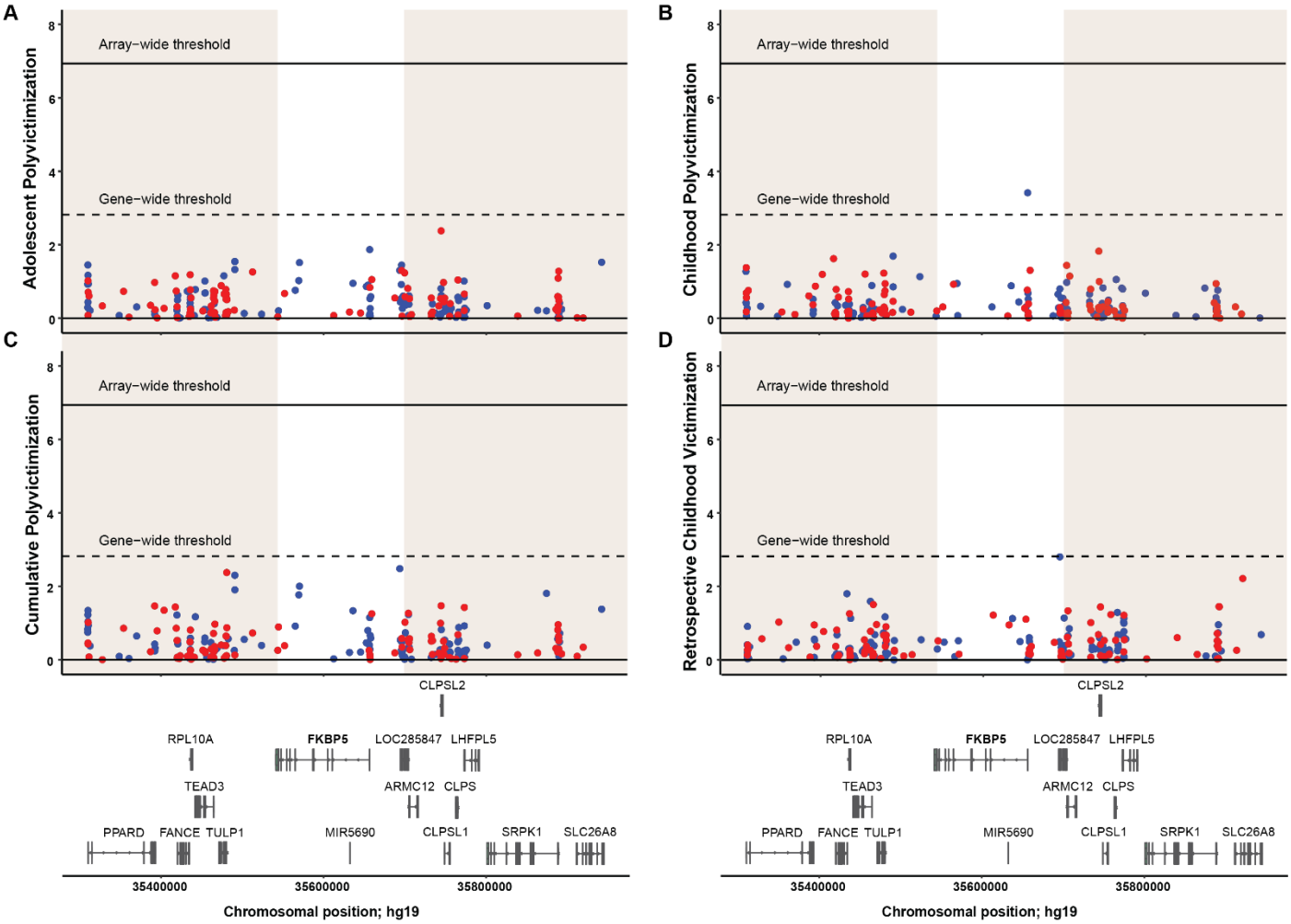

Figure S5. **The association between victimization and DNA methylation across probes annotated to *FKBP5*.** Panels **A-D** show regional Manhattan plots of $-\log_{10} P$ values for all probes annotated to the genomic region surrounding *FKBP5* in analyses of **(A)** adolescent polyvictimization, **(B)** childhood victimization, **(C)** cumulative victimization, and **(D)** retrospective reports of childhood victimization. The Panels show both the array-wide multiple testing threshold ($P<1.16\times10^{-7}$) and the gene-wide significance threshold based on a Bonferroni correction for the number of probes annotated to the gene (N=33). Effect directions are indicated in blue (• hypomethylated) and red (• hypermethylated) circles. The white areas highlight probes localized to the genomic coordinates for *FKBP5*. The shaded areas highlight probes annotated to the *FKBP5* gene region (±250 Kb up- and downstream of the gene), with diagrammatic gene annotations below for illustrative purposes. Of the 33 probes annotated to *FKBP5* one probe (cg00140191) passed the gene-wide threshold for childhood polyvictimization.

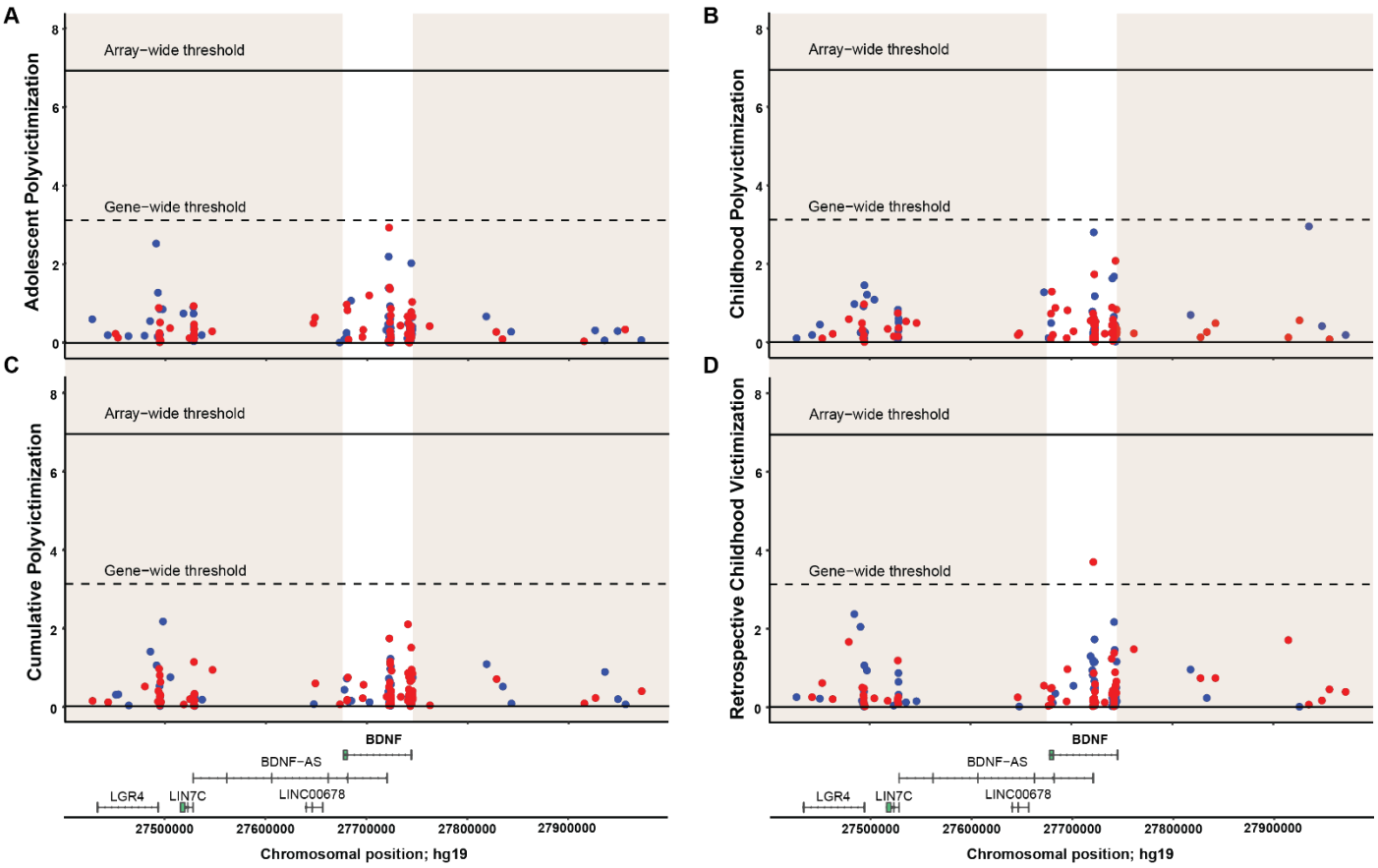

Figure S6. **The association between victimization and DNA methylation across probes annotated to *BDNF*.** Panels **A-D** show regional Manhattan plots of $-\log_{10} P$ values for all probes annotated to the genomic region surrounding *BDNF* in analyses of **(A)** adolescent polyvictimization, **(B)** childhood victimization, **(C)** cumulative victimization, and **(D)** retrospective reports of childhood victimization. The Panels show both the array-wide multiple testing threshold ($P<1.16\times10^{-7}$) and the gene-wide significance threshold based on a Bonferroni correction for the number of probes annotated to the gene (N=66). Effect directions are indicated in blue (● hypomethylated) and red (● hypermethylated) circles. The white areas highlight probes localized to the genomic coordinates for *BDNF*. The shaded areas highlight probes annotated to the *BDNF* gene region (±250 Kb up- and downstream of the gene), with diagrammatic gene annotations below for illustrative purposes. Of the 66 probes annotated to *BDNF* one probe (cg20954537) passed the gene-wide threshold for retrospective childhood victimization.

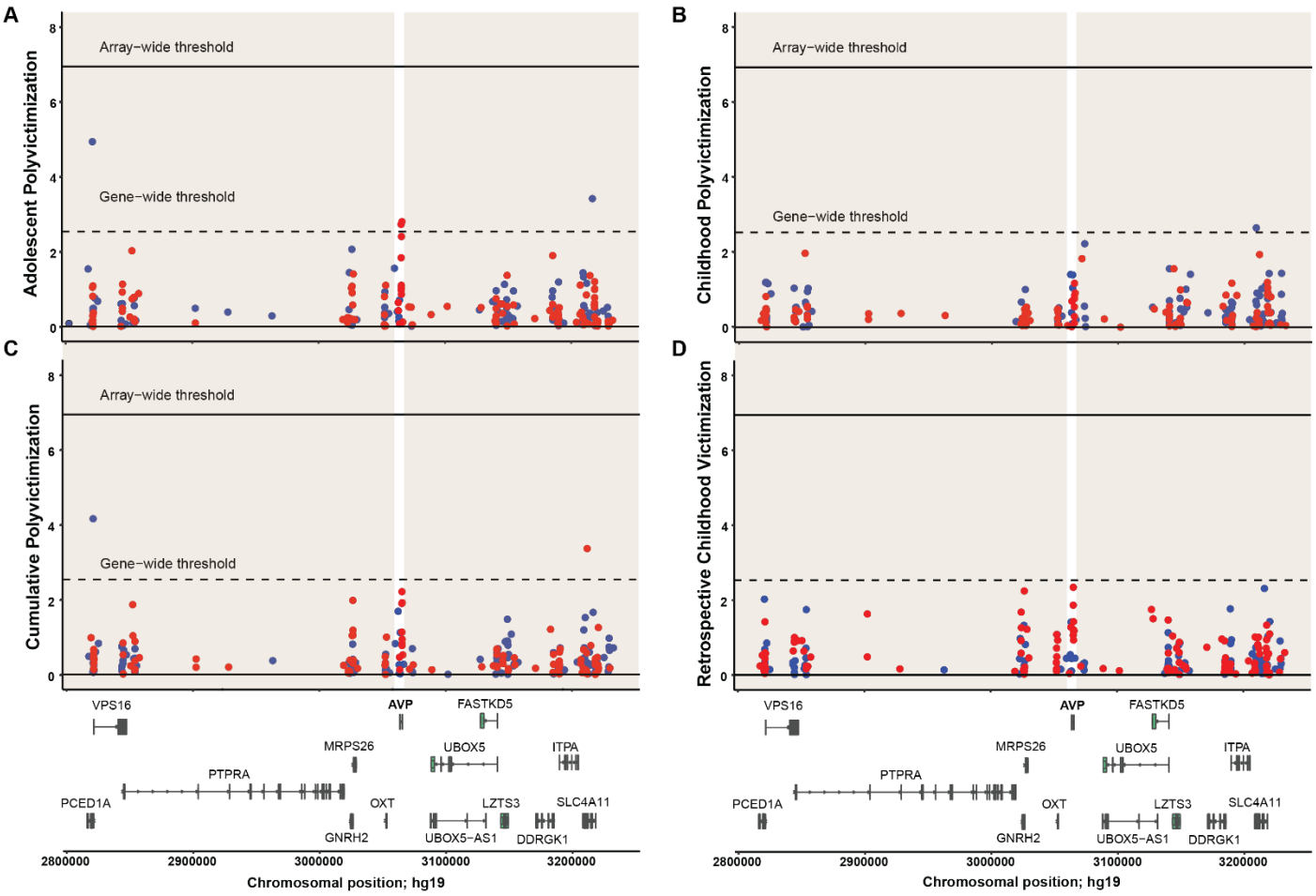

Figure S7. **The association between victimization and DNA methylation across probes annotated to *AVP*.** Panels **A-D** show regional Manhattan plots of $-\log_{10} P$ values for all probes annotated to the genomic region surrounding *AVP* in analyses of **(A)** adolescent polyvictimization, **(B)** childhood victimization, **(C)** cumulative victimization, and **(D)** retrospective reports of childhood victimization. The Panels show both the array-wide multiple testing threshold ($P<1.16\times10^{-7}$) and the gene-wide significance threshold based on a Bonferroni correction for the number of probes annotated to the gene (N=17). Effect directions are indicated in blue (● hypomethylated) and red (● hypermethylated) circles. The white areas highlight probes localized to the genomic coordinates for *AVP*. The shaded areas highlight probes annotated to the *AVP* gene region (±250 Kb up- and downstream of the gene), with diagrammatic gene annotations below for illustrative purposes. Of the 17 probes annotated to *AVP* two probes (cg23035419 and cg25551168) passed the gene-wide threshold for adolescent polyvictimization.

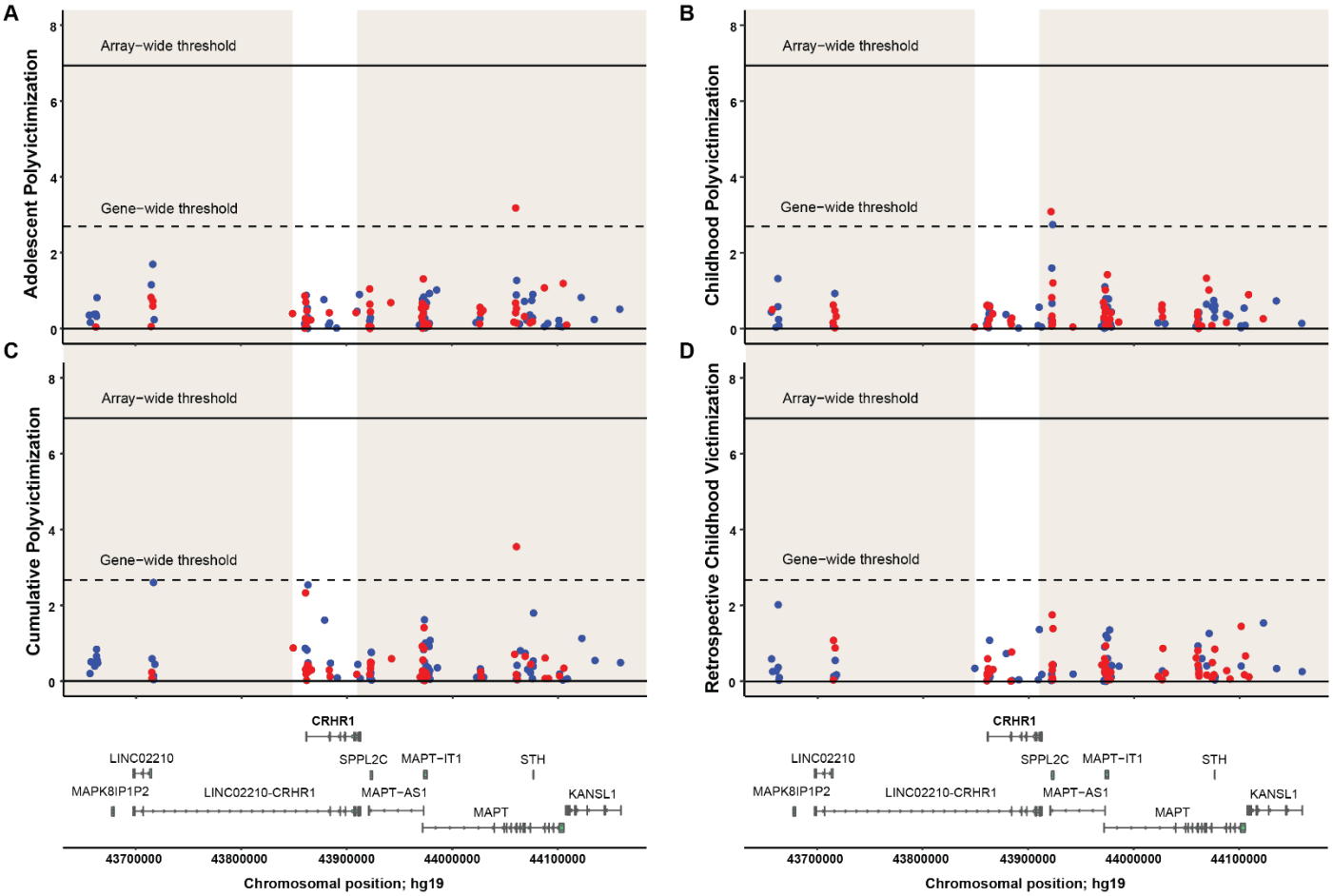

Figure S8. **The association between victimization and DNA methylation across probes annotated to *CRHR1*.** Panels **A-D** show regional Manhattan plots of $-\log_{10} P$ values for all probes annotated to the genomic region surrounding *CRHR1* in analyses of **(A)** adolescent polyvictimization, **(B)** childhood victimization, **(C)** cumulative victimization, and **(D)** retrospective reports of childhood victimization. The Panels show both the array-wide multiple testing threshold ($P<1.16\times10^{-7}$) and the gene-wide significance threshold based on a Bonferroni correction for the number of probes annotated to the gene (N=24). Effect directions are indicated in blue (● hypomethylated) and red (● hypermethylated) circles. The white areas highlight probes localized to the genomic coordinates for *CRHR1*. The shaded areas highlight probes annotated to the *CRHR1* gene region (±250 Kb up- and downstream of the gene), with diagrammatic gene annotations below for illustrative purposes. Of the 24 probes annotated to *CRHR1* none passed the array- or gene-wide thresholds for any of the victimization exposures. Two probes (cg15679139 and cg20995065), located close to the 3' end of *CRHR1* and annotated to *MAPT-AS1* and *SPPL2C*, passed gene-wide threshold for childhood victimization.

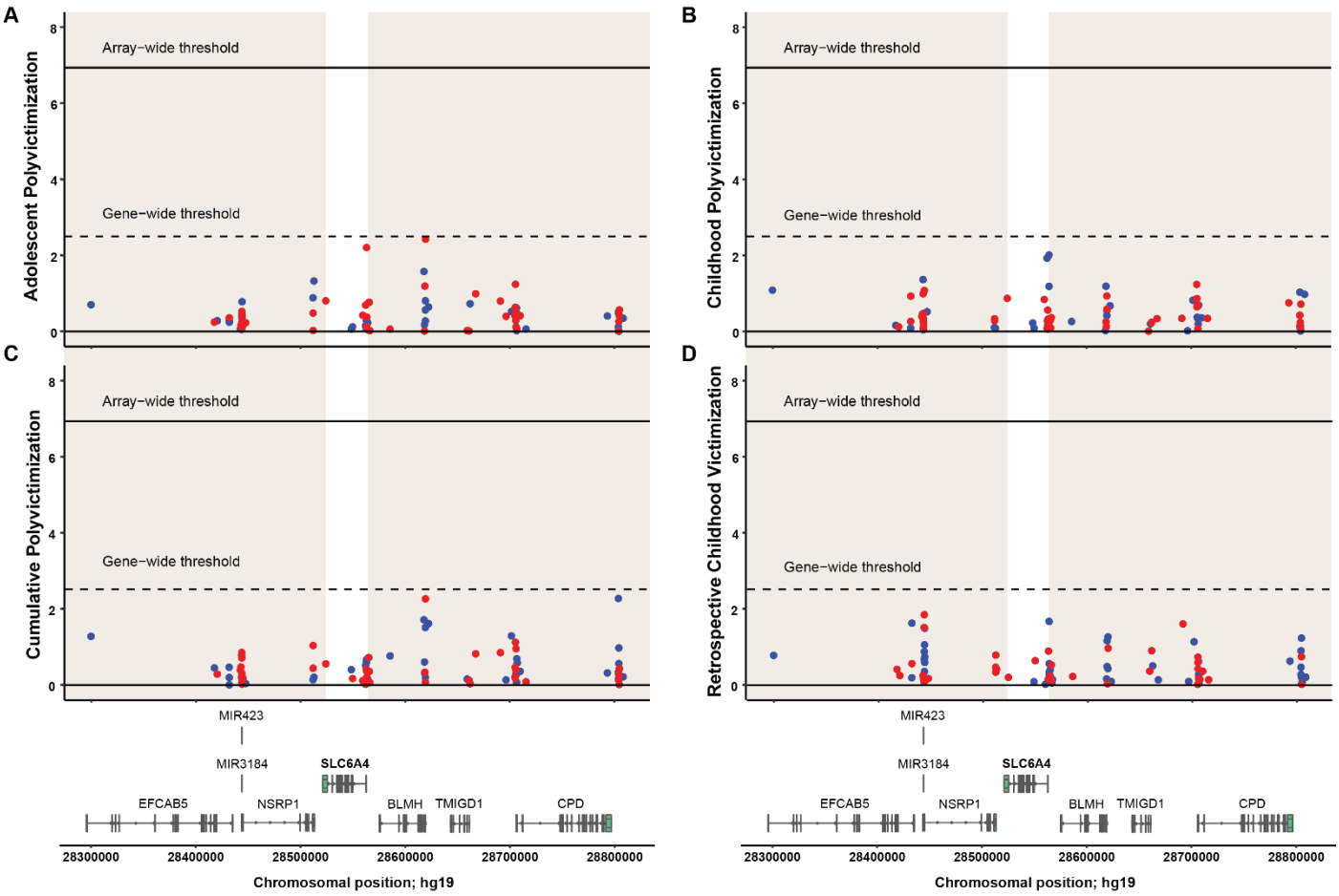

Figure S9. **The association between victimization and DNA methylation across probes annotated to *SLC6A4*.** Panels **A-D** show regional Manhattan plots of $-\log_{10} P$ values for all probes annotated to the genomic region surrounding *SLC6A4* in analyses of **(A)** adolescent polyvictimization, **(B)** childhood victimization, **(C)** cumulative victimization, and **(D)** retrospective reports of childhood victimization. The Panels show both the array-wide multiple testing threshold ($P<1.16\times10^{-7}$) and the gene-wide significance threshold based on a Bonferroni correction for the number of probes annotated to the gene (N=16). Effect directions are indicated in blue (● hypomethylated) and red (● hypermethylated) circles. The white areas highlight probes localized to the genomic coordinates for *SLC6A4*. The shaded areas highlight probes annotated to the *SLC6A4* gene region (±250 Kb up- and downstream of the gene), with diagrammatic gene annotations below for illustrative purposes. Of the 16 probes annotated to *SLC6A4* none passed the array- or gene-wide thresholds for any of the victimization exposures.

Table S1. **Results of a Latent Class Analysis using longitudinal data about victimization**. We examined fit statistics for 2 to 6 groups. The 3-class solution was the preferred solution. Exposure to parental intimate-partner violence in childhood defined the victimization profile of 254 young people. Peer victimization in childhood and adolescence and/or 'street crime' in adolescence defined the victimization profile of 412 young people. Cumulative polyvictimization defined the victimization profile of 158 young people; as children, 85% were exposed to domestic violence; 75% were bullied; 64% were physically abused; 49% were emotionally abused or neglected; and 38% were physically neglected. As adolescents, 97% experienced conventional crime-related violence; 93% were victimized by peers; 62% were maltreated; 58% experienced family violence; 53% experience internet violence; 41% were victims of sexual abuse; and 37% were neglected. For analysis purposes, we compared the 158 young people who were exposed to cumulative polyvictimization across both childhood and adolescence to the 834 young people who were not exposed to childhood or adolescent victimization.

No. Groups	Loglikelihood	AIC	BIC	Entropy	LMR Adjusted LRT Test	*P* value
2	-9718.351	19542.702	19805.044	0.742	608.664	0.0353
3	-9462.985	19085.971	19481.959	0.720	508.024	0.0062
4	-9366.310	18946.621	19476.255	0.768	192.326	0.6988
5	-9291.425	18850.850	19514.131	0.776	149.002	0.2524
6	-9213.723	18749.446	19546.373	0.803	154.625	0.7650

Table S2. Methylation probes significantly associated with smoking pack years by age 18. The table shows summary statistics for probes passing the array-wide significance threshold in an analysis of DNA methylation and pack years by age 18 in E-Risk (**Column A)**. For comparison purposes, summary statistics of significant probes reported in a recently published study of DNA methylation and cigarette smoking (current smoker vs never) are found in **Column B.**

			A. Pack years at age 18		**B. Results reported in Joehanes *et al*., (2016)**	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**
cg21161138	Chr5:399,311	AHRR; Intron	-0.012	1.03E-19	-0.053	4.42E-16
cg25949550	Chr7:145,814,305	CNTNAP2; Intron	-0.003	4.41E-18	-0.020	9.02E-25
cg19572487	Chr17:38,475,975	RARA; Intron	-0.005	1.07E-17	-0.054	3.80E-20
cg05575921	Chr5:373,377	AHRR; Intron	-0.025	8.72E-17	-0.180	4.55E-26
cg01940273	Chr2:233,284,885	AC068134.5; Upstream	-0.011	8.74E-17	-0.081	2.03E-34
cg08709672	Chr1:206,224,285	AVPR1B; 5'UTR	-0.005	1.40E-16	-0.023	1.55E-21
cg14817490	Chr5:392,919	AHRR; Intron	-0.008	1.72E-16	-0.049	1.17E-19
cg05951221	Chr2:233,284,401	AC068134.5; Exon	-0.010	2.90E-16	-0.097	4.35E-27
cg07826859	Chr7:45,020,037	MYO1G; Upstream	-0.004	4.27E-16	-0.022	1.26E-12
cg14753356	Chr6:30,720,059	RN7SL353P; Downstream	-0.008	2.48E-15	-0.038	9.24E-18
cg22132788	Chr7:45,002,437	MYO1G; 5'UTR	0.010	3.79E-15	0.046	8.14E-19
cg21322436	Chr7:145,812,793	CNTNAP2; Upstream	-0.005	6.76E-15	-0.032	3.52E-18
cg23916896	Chr5:368,755	AHRR; Intron	-0.007	2.36E-14	-0.059	6.32E-20
cg26703534	Chr5:377,357	AHRR; Intron	-0.010	3.02E-14	-0.041	1.22E-21
cg20295214	Chr1:206,226,745	AVPR1B; Intron	-0.005	7.99E-14	-0.018	5.49E-15
cg11071448	Chr1:202,584,464	SYT2; Intron	-0.005	9.37E-14	-0.018	1.19E-13
cg02013841	Chr13:49,159,766	LINC00462; Upstream	-0.006	3.92E-13	-0.013	6.08E-11
cg25648203	Chr5:395,395	AHRR; Intron	-0.008	9.08E-13	-0.039	2.12E-14
cg02451831	Chr7:26,578,049	KIAA0087; CDS	-0.006	1.38E-12	-0.024	6.53E-10
cg01899089	Chr5:369,968	AHRR; Intron	-0.004	1.71E-12	-0.032	1.08E-15

			A. Pack years at age 18		B. Results reported in Joehanes *et al*., (2016)	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**
cg03636183	Chr19:17,000,536	F2RL3; CDS	-0.010	4.57E-12	-0.095	1.12E-20
cg19089201	Chr7:45,002,238	MYO1G; 5'UTR	0.006	6.85E-12	0.031	1.56E-15
cg25189904	Chr1:68,299,492	GNG12-AS1; Intron	-0.011	6.87E-12	-0.077	5.22E-22
cg07986378	Chr12:11,898,235	ETV6; Intron	-0.007	8.46E-12	-0.029	9.63E-10
cg06126421	Chr6:30,720,031	RN7SL353P; Downstream	-0.010	1.58E-11	-0.097	1.72E-20
cg02186444	Chr17:73,120,976	ARMC7; Intron	0.004	4.25E-11	0.015	2.64E-15
cg15159987	Chr19:17,003,889	CPAMD8; 5'UTR	-0.005	9.82E-11	-0.019	2.95E-07
cg23973524	Chr19:18,873,173	CRTC1; Intron	0.005	1.04E-10	0.025	1.94E-17
cg10951873	Chr1:25,254,697	RUNX3; Intron	-0.002	1.22E-10	-0.007	1.03E-13
cg05221370	Chr7:110,738,835	IMMP2L; Intron	-0.004	1.63E-10	-0.017	3.36E-09
cg12803068	Chr7:45,002,918	MYO1G; Exon	0.012	3.93E-10	0.063	9.06E-23
cg26764244	Chr1:68,299,510	GNG12-AS1; Intron	-0.005	4.00E-10	-0.032	1.77E-28
cg15542713	Chr1:42,385,532	HIVEP3; Intron	0.008	4.12E-10	0.040	2.44E-13
cg03450842	Chr10:80,834,946	ZMIZ1; Intron	-0.005	4.28E-10	-0.026	1.87E-14
cg03991871	Chr5:368,398	AHRR; Intron	-0.009	4.40E-10	-0.038	2.60E-16
cg24049493	Chr1:42,385,940	HIVEP3; Intron	0.007	4.90E-10	0.046	7.79E-17
cg05460226	Chr17:8,804,278	PIK3R5; Intron	-0.006	7.29E-10	-0.021	1.33E-09
cg07465627	Chr17:53,167,406	STXBP4; Intron	-0.003	9.40E-10	-0.014	3.78E-15
cg16702313	Chr14:74,251,877	ELMSAN1; Intron	-0.002	1.11E-09	-0.008	9.45E-10
cg21606956	Chr1:184,211,885	Y_RNA; Upstream	-0.002	2.17E-09	-0.003	4.48E-05
cg04368724	Chr6:31,760,592	VARS; 3'UTR	0.003	2.42E-09	0.010	8.05E-07
cg04640972	Chr10:8,373,473	RP5-1119O21.2; Downstream	0.003	2.52E-09	0.010	1.45E-05
cg10750182	Chr10:73,497,465	C10orf105; 3'UTR	-0.003	2.89E-09	-0.019	6.25E-19
cg26271591	Chr2:178,125,955	NFE2L2; Intron	-0.005	3.03E-09	-0.032	7.62E-13

			A. Pack years at age 18		B. Results reported in Joehanes *et al*., (2016)	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**
cg08266095	Chr5:123,149,715	KRT18P16; Downstream	-0.005	3.65E-09	-0.023	3.00E-14
cg16744741	Chr4:82,125,976	PRKG2; CDS	0.005	3.65E-09	0.012	4.68E-10
cg09099830	Chr16:30,485,436	ITGAL; 5'UTR	-0.005	3.78E-09	-0.022	6.44E-15
cg26242531	Chr14:104,190,629	ZFYVE21; Intron	0.005	4.34E-09	0.016	7.82E-15
cg26529655	Chr5:424,370	AHRR; Intron	-0.003	4.61E-09	-0.010	8.70E-08
cg16145216	Chr1:42,385,613	HIVEP3; Intron	0.006	6.02E-09	0.030	6.70E-48
cg02228160	Chr5:143,192,066	HMHB1; Intron	0.004	6.88E-09	0.014	1.86E-12
cg04956244	Chr17:38,511,591	RARA; CDS	0.002	7.33E-09	0.012	1.52E-29
cg13185177	Chr3:194,119,836	GP5; Intron	0.005	9.47E-09	0.023	6.49E-15
cg04885881	Chr1:11,123,117	SRM; Upstream	-0.005	1.08E-08	-0.042	2.15E-14
cg20505728	Chr3:43,800,489	RP4-672N11.1; Downstream	0.003	1.23E-08	0.005	1.04E-09
cg15342087	Chr6:30,720,160	RN7SL353P; Downstream	-0.002	1.47E-08	-0.031	1.63E-17
cg19427338	Chr2:42,566,406	COX7A2L; Intron	0.003	1.63E-08	0.013	5.69E-11
cg24090911	Chr5:400,731	AHRR; Intron	-0.005	1.81E-08	-0.024	8.73E-12
cg23126342	Chr13:67,801,076	PCDH9; CDS	0.006	2.23E-08	0.027	3.05E-12
cg09935388	Chr1:92,947,587	GFI1; Intron	-0.009	2.41E-08	-0.083	3.14E-17
cg24859433	Chr6:30,720,154	RN7SL353P; Downstream	-0.003	2.41E-08	-0.031	3.27E-12
cg10420527	Chr11:68,138,504	LRP5; Intron	-0.003	2.56E-08	-0.011	4.63E-08
cg05603985	Chr1:2,161,048	SKI; CDS	-0.002	3.88E-08	-0.012	1.76E-43
cg06679494	Chr17:6,921,246	MIR497HG; Exon	-0.002	3.94E-08	*NA*	*NA*
cg04551776	Chr5:393,365	AHRR; Intron	-0.004	4.79E-08	-0.024	5.77E-31
cg27537125	Chr1:25,349,680	MIR4425; Upstream	-0.003	5.61E-08	-0.019	1.35E-19
cg10965178	Chr1:43,766,703	TIE1; TSS	-0.003	5.76E-08	-0.011	2.53E-10
cg04224247	ChrX:9,984,514	WWC3; Exon	-0.006	5.76E-08	-0.021	9.98E-05

			A. Pack years at age 18		B. Results reported in Joehanes *et al*., (2016)	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**
cg13039251	Chr5:32,018,552	PDZD2; Intron	0.003	6.05E-08	0.030	1.36E-15
cg13399816	Chr1:68,299,467	GNG12-AS1; Intron	-0.003	6.11E-08	-0.021	2.83E-15
cg24688690	Chr5:345,849	AHRR; Intron	-0.002	6.17E-08	-0.007	1.64E-11
cg21461196	Chr18:48,494,909	ELAC1; Exon	-0.001	6.42E-08	*NA*	*NA*
cg14712058	Chr19:16,988,034	SIN3B; Exon	-0.003	7.22E-08	-0.014	9.05E-15
cg00950497	Chr10:116,393,422	ABLIM1; Intron	0.005	7.35E-08	0.012	2.30E-07
cg09554443	Chr1:167,487,761	CD247; TSS	0.004	7.56E-08	0.016	1.42E-11
cg22222281	Chr12:6,308,709	CD9; Upstream	-0.004	7.72E-08	*NA*	*NA*
cg02532700	Chr22:37,257,403	CTA-833B7.2; Intron	-0.004	7.88E-08	-0.024	5.17E-13
cg01513913	Chr14:106,329,109	IGHJ6; Downstream	-0.003	8.64E-08	-0.020	1.31E-11
cg00310412	Chr15:74,724,869	SEMA7A; Intron	-0.002	9.26E-08	-0.023	5.74E-17
cg21611682	Chr11:68,138,268	LRP5; Intron	-0.004	9.84E-08	-0.026	1.35E-18
cg00501876	Chr3:39,193,202	CSRNP1; Intron	-0.003	9.99E-08	-0.022	2.51E-15
cg16624521	Chr5:172,064,551	NEURL1B; Upstream	0.003	1.00E-07	0.011	3.83E-07
cg09762515	Chr7:101,556,539	CUX1; Exon	0.004	1.09E-07	0.015	1.05E-09

Table S3. Methylation probes significantly associated with sexual victimization, as measured in childhood, in adolescence, and retrospectively in young adulthood. The Table shows 39 probes that were significantly associated with childhood sexual victimization (in bold red); 1 with adolescent sexual victimization (in bold blue); 22 with retrospective reports of sexual abuse as reported on the Childhood Trauma Questionnaire (CTQ) (in bold green); and 2 with retrospective reports of sexual abuse as reported on the Childhood Trauma Questionnaire (CTQ) in the Dunedin Longitudinal Study (in bold purple). The Table documents that methylation correlates of sexual victimization are not reproduced across different operationalizations of this stressor. Note: Childhood sexual victimization correlates are shown in red; adolescent sexual victimization correlates are shown in blue; retrospective reports of sexual abuse correlates are shown in green; retrospective reports of sexual abuse correlates in the Dunedin Longitudinal Study are shown in purple. Bolded values pass array-wide significance thresholds; non-bolded values do not pass the significance threshold. Percent DNA methylation change per unit change in victimization can be calculated by *B**100.

			Childhood sexual victimization		adolescent sexual victimization		retrospective reports of sexual abuse (CTQ)		retrospective reports of sexual abuse (CTQ); Dunedin Longitudinal Study	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**
cg01956420	Chr13:110,959,667	MRPL23; Intron	**0.006**	**7.58E-12**	0.000	9.92E-01	0.003	1.50E-02	-0.001	4.71E-02
cg02424139	Chr6:106,957,080	PAPLN; Intron	**-0.022**	**2.62E-11**	-0.003	1.05E-01	-0.003	5.49E-01	0.000	1.00E+00
cg03128860	Chr17:4,893,029	AC104417.1; Upstream	**0.014**	**4.87E-11**	0.000	5.15E-01	0.005	3.76E-02	-0.001	5.89E-01
cg03624528	Chr5:41,261,526	WWOX; Intron	**-0.007**	**5.76E-11**	0.000	9.33E-01	-0.001	6.34E-01	0.000	9.52E-01
cg03643838	Chr11:1,978,107	CCDC144B; Intron	**0.010**	**6.46E-11**	0.000	5.76E-01	0.002	2.32E-01	0.000	7.56E-01
cg04577497	Chr12:10,766,215	FAM200A; Intron	**-0.010**	**8.63E-11**	0.000	5.66E-01	0.002	4.44E-01	0.000	8.80E-01
cg05233289	Chr11:64,120,679	RP4-673D20.3; Intron	**-0.007**	**1.41E-10**	0.000	8.45E-01	0.000	8.21E-01	0.002	1.01E-01
cg05406868	Chr7:5,467,365	AC093375.1; Downstream	**-0.015**	**3.17E-10**	0.000	8.78E-01	-0.001	5.45E-01	-0.001	6.89E-01
cg05499054	Chr8:75,148,757	RP11-1275H24.1; Exon	**-0.007**	**8.37E-10**	0.000	7.12E-01	-0.002	8.73E-02	0.002	3.84E-02
cg06317056	Chr17:7,589,295	MRPS22; Intron	**0.004**	**1.01E-09**	0.000	5.13E-01	0.000	8.58E-01	0.001	2.92E-01
cg06732228	Chr19:41,119,352	ATP11A-AS1; Exon	**0.003**	**1.53E-09**	0.000	9.86E-01	0.000	9.67E-01	0.000	2.09E-01

			Childhood sexual victimization		adolescent sexual victimization		retrospective reports of sexual abuse (CTQ)		retrospective reports of sexual abuse (CTQ); Dunedin Longitudinal Study	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**
cg06863310	Chr16:79,034,834	NHS; CDS	0.008	2.08E-09	0.000	7.90E-01	-0.001	6.60E-01	0.002	2.03E-01
cg07098277	ChrX:106,361,851	SIGLEC15; Intron	-0.019	2.93E-09	-0.002	9.38E-02	-0.004	1.28E-01	-0.003	1.92E-01
cg07173670	Chr16:89,575,360	C9orf64; 3'UTR	-0.005	5.11E-09	0.001	2.55E-01	0.000	8.75E-01	0.001	2.87E-01
cg07415373	Chr22:27,152,966	TTC6; CDS	-0.007	6.54E-09	0.000	5.35E-01	-0.001	7.37E-01	0.000	7.65E-01
cg08099136	Chr6:32,811,250	PPP1R10; 3'UTR	0.016	7.57E-09	0.001	2.35E-01	0.003	2.13E-01	-0.001	5.49E-01
cg08121175	Chr1:25,566,282	C6; TSS	0.017	7.96E-09	0.004	1.10E-02	0.008	1.00E-01	-0.002	3.90E-01
cg08568736	Chr20:1,757,731	AIM1; Upstream	0.014	1.40E-08	0.000	7.47E-01	0.002	1.67E-01	0.000	7.73E-01
cg08755218	Chr16:67,260,980	SERINC3; 5'UTR	0.023	1.40E-08	0.001	7.32E-01	0.002	4.78E-01	-0.001	8.75E-01
cg09396850	Chr13:113,400,629	CEBPZ-AS1; Intron	-0.006	2.13E-08	0.001	2.65E-01	0.000	9.62E-01	0.001	3.42E-01
cg10491628	Chr6:29,521,171	MAGOHB; TSS	0.003	2.26E-08	0.000	4.74E-01	0.000	9.47E-01	0.000	4.56E-01
cg10741369	Chr5:180,633,325	AKAP12; CDS	0.009	2.37E-08	0.001	4.42E-02	0.003	7.66E-02	-0.002	2.70E-02
cg10988336	Chr5:95,767,767	LRAT; Upstream	0.004	2.46E-08	0.000	4.47E-01	0.002	3.34E-02	-0.001	2.06E-01
cg11087819	Chr8:142,838,433	CTC-338M12.1; Upstream	0.004	2.60E-08	0.000	9.28E-01	0.001	2.11E-01	0.000	7.79E-01
cg11788523	Chr19:58,446,721	TMEM208; TSS	-0.002	2.82E-08	0.000	3.77E-01	-0.001	5.02E-02	0.000	5.31E-01
cg12062572	Chr15:35,271,966	ZNF770; 5'UTR	0.009	3.56E-08	0.000	5.01E-01	0.005	2.47E-03	0.000	9.03E-01
cg12200038	Chr14:38,091,469	RP11-736N17.8; Upstream	0.004	3.58E-08	0.000	7.14E-01	0.000	8.20E-01	0.001	1.25E-02
cg12666976	Chr2:1,482,789	PHF14; Upstream	-0.004	3.96E-08	0.000	1.93E-01	-0.002	3.02E-03	0.000	7.18E-01
cg13208584	Chr20:43,128,002	RP11-12J10.3; Intron	0.007	4.46E-08	0.000	8.73E-01	0.002	2.83E-01	0.000	8.41E-01
cg13417559	Chr9:86,571,383	CCDC88B; 5'UTR	0.003	4.49E-08	0.000	5.22E-01	0.000	8.29E-01	0.001	1.98E-02
cg13431226	Chr3:139,062,593	UBC; TSS	-0.005	5.01E-08	0.000	2.32E-01	0.001	2.97E-01	0.000	3.36E-01
cg14834903	Chr14:73,712,901	SPG7; CDS	0.002	5.06E-08	0.000	8.18E-01	-0.001	7.28E-02	0.000	9.21E-01
cg15543489	Chr19:46,527,497	RP11-199F11.2; Exon	-0.005	5.78E-08	0.000	5.82E-01	0.001	6.31E-01	0.000	6.21E-01
cg16030145	Chr17:18,527,434	PPT2; TSS	0.004	5.81E-08	0.000	8.97E-01	0.000	6.91E-01	0.000	7.20E-01

			Childhood sexual victimization		adolescent sexual victimization		retrospective reports of sexual abuse (CTQ)		retrospective reports of sexual abuse (CTQ); Dunedin Longitudinal Study	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**
cg16800461	Chr7:11,013,434	SRSF4; Upstream	**-0.007**	**6.17E-08**	0.001	9.43E-02	-0.001	3.66E-01	0.000	5.38E-01
cg17553592	ChrX:17,394,344	ABCB1; Intron	**-0.006**	**7.19E-08**	0.000	3.31E-01	-0.001	2.01E-01	0.000	9.40E-01
cg18534992	Chr14:103,550,403	RBM41; Intron	**-0.012**	**7.50E-08**	0.000	7.88E-01	-0.002	3.67E-01	-0.001	6.12E-01
cg18673377	Chr19:58,446,849	TPO; Intron	**-0.012**	**8.94E-08**	0.000	9.87E-01	-0.002	4.07E-01	-0.001	4.32E-01
cg18764577	Chr18:7,117,789	TOB1-AS1; Intron	**-0.010**	**1.12E-07**	0.000	7.38E-01	0.002	2.54E-01	0.000	7.03E-01
cg18998670	Chr2:58,273,503	PRDM16; Intron	0.000	8.03E-01	**-0.002**	**1.49E-08**	0.000	6.63E-01	-0.001	3.50E-01
cg19658829	Chr12:125,399,893	FOXN3; Intron	0.001	5.37E-01	0.000	7.30E-01	**0.005**	**6.58E-12**	-0.001	2.86E-01
cg19840494	Chr14:89,606,172	PKP2; Upstream	-0.002	6.74E-02	0.000	9.50E-01	**-0.002**	**1.15E-11**	0.001	2.60E-01
cg19861260	Chr7:87,228,748	JPH1; 5'UTR	-0.001	3.56E-01	-0.001	3.42E-03	**0.003**	**2.57E-11**	-0.001	5.73E-02
cg19975346	Chr10:126,366,082	LTBP4; TSS	-0.001	7.06E-01	0.000	4.11E-01	**-0.003**	**3.80E-10**	0.001	1.70E-01
cg20568219	Chr6:151,671,319	PGLYRP1; Upstream	0.005	1.50E-01	0.001	1.32E-01	**0.009**	**9.44E-10**	0.001	3.45E-01
cg20667796	Chr17:73,720,233	ITGB4; Intron	0.003	8.96E-02	0.000	4.61E-01	**0.005**	**2.56E-09**	-0.001	1.85E-01
cg20785360	Chr4:155,547,951	COL4A2; 5'UTR	-0.002	4.16E-01	-0.001	2.00E-01	**-0.005**	**3.77E-09**	0.000	7.14E-01
cg20862283	Chr2:156,838,907	NOL10; Intron	-0.001	4.00E-02	0.000	5.06E-01	**-0.002**	**5.05E-09**	0.000	4.50E-01
cg21207593	Chr17:33310494	EPHA1-AS1; Intron	-0.003	2.71E-01	0.000	6.61E-01	**-0.005**	**5.46E-09**	0.000	7.61E-01
cg21393124	Chr17:48,981,322	C1orf63; Downstream	-0.004	3.85E-05	0.000	7.65E-01	**-0.003**	**9.25E-09**	0.000	5.54E-01
cg21987316	Chr6:32,121,952	VRK2; Intron	-0.001	3.04E-01	0.000	8.29E-01	**-0.002**	**9.57E-09**	0.000	6.37E-01
cg22697786	Chr7:4939324	PRPF8; Intron	0.001	5.80E-01	0.000	7.69E-01	**-0.003**	**2.47E-08**	0.000	5.18E-01
cg23146534	Chr7:99,149,190	SORCS2; CDS	0.001	5.88E-01	0.000	5.76E-01	**-0.005**	**3.79E-08**	0.002	7.16E-02
cg23181627	Chr17:1,587,711	PCSK1; CDS	-0.001	8.85E-01	-0.001	4.77E-01	**-0.008**	**6.31E-08**	0.000	9.41E-01
cg23687288	Chr2:37,424,014	ZNF418; TSS	-0.006	1.21E-01	0.000	8.94E-01	**-0.008**	**7.29E-08**	0.000	9.88E-01
cg24030138	Chr18:43,419,101	CTA-211A9.5; Intron	-0.003	4.81E-01	0.001	3.86E-01	**0.008**	**8.24E-08**	-0.003	1.80E-01
cg24330922	Chr6:30,584,244	LAMA1; TSS	-0.002	2.24E-01	0.000	2.55E-01	**-0.002**	**8.40E-08**	0.000	8.85E-01

			Childhood sexual victimization		adolescent sexual victimization		retrospective reports of sexual abuse (CTQ)		retrospective reports of sexual abuse (CTQ); Dunedin Longitudinal Study	
Illumina probe ID	**Probe Start Position**	**Closest Gene/position**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**	***B***	***P* value**
cg24940499	Chr1:29,508,758	OR2I1P; Exon	-0.002	5.51E-02	0.000	6.18E-01	**-0.003**	**9.72E-08**	0.000	9.50E-01
cg25150520	Chr1:3,142,481	PSMB8; Intron	-0.002	5.71E-01	0.002	1.94E-02	**0.012**	**9.80E-08**	-0.001	4.81E-01
cg25375420	Chr6:36,009,600	ZNF418; Upstream	-0.004	3.52E-01	0.000	9.04E-01	**-0.006**	**1.04E-07**	0.000	9.58E-01
cg25655234	Chr4:7,194,840	INCA1; Intron	0.000	8.13E-01	-0.001	2.51E-02	**0.004**	**1.07E-07**	-0.001	3.98E-02
cg25804072	Chr2:10,829,676	MAPK14; Intron	0.005	2.60E-02	0.001	1.17E-02	**0.007**	**1.11E-07**	0.000	9.36E-01
cg27347930	Chr7:143,113,812	LIG3; CDS	-0.001	7.17E-01	0.000	5.65E-01	0.005	1.57E-02	**-0.008**	**2.39E-08**
cg27458987	Chr12:33,050,049	MMD2; Downstream	0.001	4.72E-01	0.000	6.18E-01	0.002	1.81E-01	**-0.006**	**1.84E-08**