Associations between childhood adversity, DNA methylation, and later onset of depression

Tara McDonnell, under the supervision of Dr Andrew Simpkin

Introduction

Globally, an estimated 5.7% of adults suffer from depression, and individuals who have experienced adverse life events are much more likely to develop depression (World Health Organization: WHO, 2025). Specifically, 58.59% of depression and anxiety cases worldwide are potentially attributable to childhood maltreatment (Li, D'Arcy and Meng, 2016).

Biologically, there is little understood about the underlying mechanisms behind this link. One potential biological mechanism underlying this link is DNA methylation (DNAm), an epigenetic mechanism that involves the addition of a methyl group to cytosine to form 5-methylcytosine, influencing gene expression without altering DNA sequence.

This project aims to investigate whether DNA methylation is associated with childhood adversity and later depression-related outcomes, and to identify CpG sites showing evidence of association across both domains. Using existing DNA methylation datasets, linear regression models and other analysis methods will be applied to identify the CpG sites with DNAm levels significantly associated with childhood adversity, and examine their relationship to levels of depressive symptoms. The project also aims to explore the biological relevance of identified CpG sites, particularly in relation to gene regulation and pathways implicated in mental health. The results of this project may provide further insight into the biological processes potentially linking early-life adversity and trauma with later mental health outcomes.

Methods

The main method throughout my study was EWAS (Epigenome-Wide Association Study), a research method that explores the relationship between epigenetic modifications, like DNA methylation, and various traits or diseases. This was conducted by fitting linear regression models between various exposures and up to 480,000 CpG sites in R. The main R packages used were easyEWAS and limma.

A total of 4 DNA methylation datasets were used in the project. The first dataset had features of both maternal and parental stress scores at infancy and preschool-age respectively. Parental stress and psychopathology have been associated with an increased risk of emotional and mental health problems in offspring (Ribas et al., 2024; Dachew et al., 2023), as well as potentially representing a source of stress for the child. Stress scores were calculated across five domains: depression symptoms, family expressed anger, parenting stress, role overload, and financial stress. The dataset was obtained from Essex et al. (2013), 'Epigenetic Vestiges of Early Developmental Adversity: Childhood Stress Exposure and Adolescent DNA Methylation' and the DNA methylation data was obtained using Illumina Human

Methylation 27K. Included were 109 samples and 27577 CpG sites, obtained from Buccal cell tissues, and sex was included as a covariate across the four EWAS. The EWAS were run using the easyEWAS R package and linear regression models on beta values.

The second dataset had features of childhood trauma scores and cortisol stress reactivity. The first analysis on this dataset used the childhood trauma scores, assessed using the short version of the Childhood Trauma Questionnaire (CTQ), as the exposure to predict the DNA methylation and the second used the DNA methylation as the exposure to predict cortisol stress reactivity. Cortisol stress reactivity is strongly associated with increased risk of depression. The dataset was obtained from Houtepen et al. (2016), 'Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans' and the methylation data was obtained using Illumina Human Methylation 450K. Included were 85 samples and 385882 CpG sites, obtained from whole blood tissues. Sex, age, and blood cell composition (including b cell proportion, cd4 t cell proportion, cd8 t cell proportion, monocytes cell proportion, and natural killer cells proportion) were included as covariates along with sentrix and position to account for batch effects. Granulocytes cell proportion was not included as a covariate to avoid collinearity. The EWAS was run using the limma R package and linear regression models on beta values. Limma was used instead of easyEWAS as it ran more effectively on the largely increased size of the dataset due to use of Illumina Human Methylation 450K instead of 27K.

The third dataset had a categorical variable dividing participants into 4 groups, two of which had a history of depression and two that didn't. This variable was converted into a binary history of depression variable which was used as the exposure. The dataset was obtained from Crawford et al. (2018), 'DNA methylation and inflammation marker profiles associated with a history of depression' and the methylation data was obtained using Illumina Human Methylation 450K. Included were 194 samples and 430574 CpG sites, obtained from whole blood tissue. Sex and age were included as covariates, but unfortunately blood cell composition data was not available. The EWAS was run using the limma R package and linear regression models on beta values.

The fourth dataset had features of maternal smoking and drinking during pregnancy and later depression scores at age 10 and 18. DNA methylation data was available at both birth and age 7. Maternal smoking and drinking during pregnancy were each used as exposures to predict DNA methylation at both ages, and methylation at both ages were used to predict depression scores at both ages. The methylation data was obtained using Illumina Human Methylation 450K and included 983 samples and 482855 CpG sites, obtained from whole blood tissue. Sex, age (to account for small variances), and blood cell composition (including b cell proportion, cd4 t cell proportion, cd8 t cell proportion, monocytes cell proportion, and natural killer cells proportion) were included as covariates. Granulocytes cell proportion was not included as a covariate to avoid collinearity. The EWAS was run using linear regression models on beta values.

Results from all EWAS were compared to identify common CpG sites of interest.

Results

Dataset 1

Several CpG sites were nominally associated with the exposures but no associations survived FDR correction. Top 5 genes for each analysis can be seen below:

Maternal Stress During Infancy

probe	gene	BETA	PVAL	FDR
cg12880200	LSM10	0.0032	3e-04	1
cg19103609	PKN1	0.0119	3e-04	1
cg27157038	DNTT	-0.0062	3e-04	1
cg18424538	CNO	0.0014	4e-04	1
cg24340926	RAB33A	0.0109	5e-04	1

Maternal Stress During Preschool

probe	gene	BETA	PVAL	FDR
cg27157038	DNTT	-0.0060	2e-04	0.9999
cg27519424	HLCS	0.0230	3e-04	0.9999
cg20885089	AMPD2	0.0179	6e-04	0.9999
cg17261830	NEK9	0.0022	7e-04	0.9999
cg02936468	C9orf41	0.0209	7e-04	0.9999

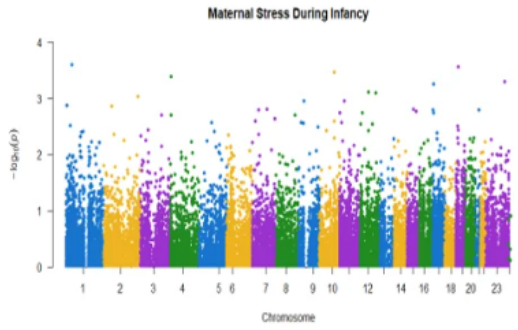
Paternal Stress During Infancy

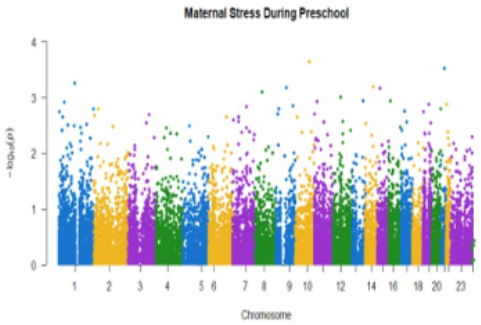
probe	gene	BETA	PVAL	FDR
cg23749046	GPR61	-0.0054	2e-04	1
cg24264578	OR2C1	-0.0068	4e-04	1
cg08198102	PTPLA	-0.0048	5e-04	1
cg17015234	RABL2A	-0.0014	5e-04	1
cg19014419	ZNF300	0.0055	6e-04	1

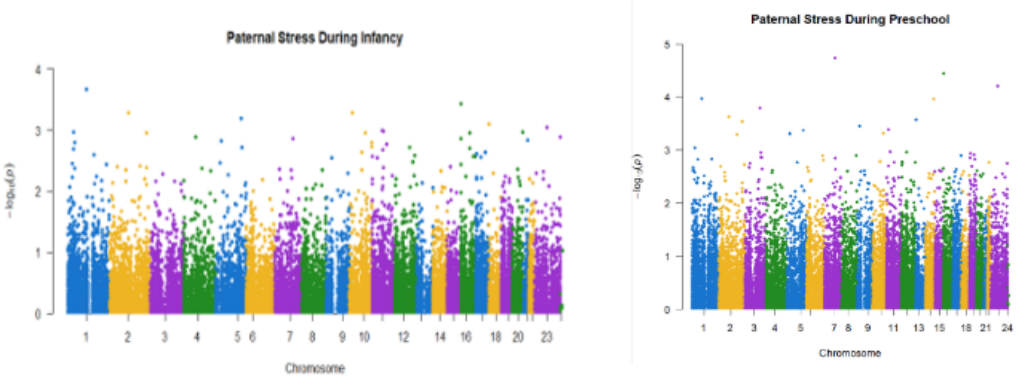
Paternal Stress During Preschool

probe	gene	BETA	PVAL	FDR
cg20002248	ACN9	0.0037	0e+00	0.4986
cg16657615	TNFRSF12A	-0.0040	0e+00	0.4986
cg24034992	YIPF6	-0.0172	1e-04	0.5693
cg26783856	LRRC8D	0.0044	1e-04	0.6033
cg22981461	OTUB2	0.0065	1e-04	0.6033

Manhattan plots for each analysis can be seen below:







Dataset 2

Again, several CpG sites were nominally associated with the exposures but no associations survived FDR correction. Top 5 genes for each analysis can be seen below:

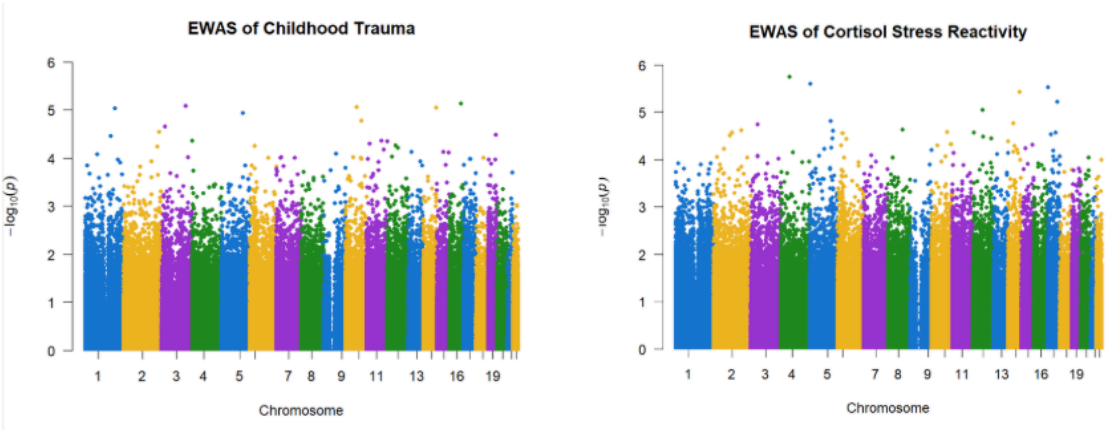
Childhood Trauma

probe	gene	BETA	PVAL	FDR
cg04017036	Na	0.0021	0	0.7169
cg11002227	GMPS	0.0009	0	0.7169
cg17340810	ADAMTS14	-0.0020	0	0.7169
cg16730509	ZNF839	-0.0009	0	0.7169
cg23627149	B3GALT2;CDC73	-0.0016	0	0.7169

Cortisol Stress Reactivity

probe	gene	BETA	PVAL	FDR
cg24852548	HOPX	-20580.6	1.79e-06	0.357667
cg03633948	Na	14586.13	2.47e-06	0.357667
cg23771985	Na	14703.74	2.95e-06	0.357667
cg01270241	Na	10286.62	3.71e-06	0.357667
cg21977075	CACNG1	11569.58	6.05e-06	0.466811

Manhattan plots for each analysis can be seen below:



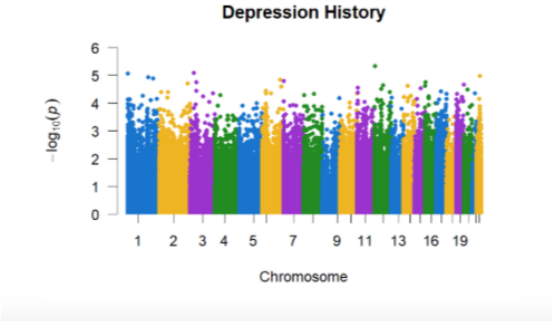
Dataset 3

Again, several CpG sites were nominally associated with the exposure but no associations survived FDR correction. Top 5 genes for the analysis can be seen below:

Depression History

probe	gene	BETA	PVAL	FDR
cg09360912	Na	-0.01151	4.86e-06	0.484766
cg20752683	Na	0.011915	8.32e-06	0.484766
cg12222277	SKI	-0.01362	8.87e-06	0.484766
cg24977886	FAM109B	0.02735	1.10e-05	0.484766
cg23082632	B4GALT3	-0.01097	1.21e-05	0.484766

Manhattan plot can be seen below:



Dataset 4

Dataset 4 had a total of 8 EWAS analyses; 2 exposures and 2 outcomes each for the 2 DNA methylations. Several CpG sites were associated with the exposures and outcomes, with 1 site surviving FDR correction for 'depression scores at 10' ~ 'DNA methylation at 7', 25 sites surviving FDR correction for 'depression scores at 18' ~ 'DNA methylation at 0', 28 sites surviving FDR correction for 'drinking during pregnancy' ~ 'DNA methylation at 7', 198 sites surviving FDR correction for 'smoking during pregnancy' ~ 'DNA methylation at 0', and 20 sites surviving FDR correction for 'smoking during pregnancy' ~ 'DNA methylation at 7'. No sites survived FDR correction for the other analyses.

Top 5 genes for each analysis can be seen below:

Depression at 10 ~ Methylation at 0

probe	gene	BETA	PVAL	FDR
cg24112221	ASB2	-11.562015	5.407073e-07	0.2610832
cg25057372	CCNE2	32.416900	1.227874e-06	0.2964426
cg19745314	NPM1	16.599530	1.864711e-06	0.3001283
cg04033559	PDK1	-2.662639	5.268596e-06	0.635992
cg18928683	AGPAT1;RNF5	36.689140	7.256579e-06	0.7007751

Depression at 10 ~ Methylation at 7

probe	gene	BETA	PVAL	FDR
cg00124774	DPH2	93.450165	2.427248e-08	0.01172009
cg25859141	Na	101.313063	2.427248e-08	0.14288049
cg01888767	Na	51.395864	8.877230e-07	0.14288049
cg14128098	ABLIM2	67.574280	1.183703e-06	0.14288918
cg17779456	RRM2	92.495420	1.347818e-05	0.99999874

Depression at 18 ~ Methylation at 0

probe	gene	BETA	PVAL	FDR
cg18255614	NEK11	10.2879250	1.498031e-08	0.004576536
cg20977426	Na	-4.0086723	1.895615e-08	0.004576536
cg27405400	BIN1	13.4110228	1.930658e-07	0.027551920
cg20668447	LRP1	-2.7720359	2.513970e-07	0.027551920
cg07039423	KCP	12.0251638	2.853022e-07	0.027551920

Depression at 18 ~ Methylation at 7

probe	gene	BETA	PVAL	FDR
cg04133132	Na	-3.7916774	2.177525e-07	0.07084851
cg23690444	Na	9.5390968	4.325908e-07	0.07084851
cg14531560	FBLIM1	30.1847563	4.731648e-07	0.07084851
cg02054619	C10orf79	29.2191977	5.869133e-07	0.07084851
cg27405400	BIN1	10.8057320	8.574502e-07	0.08280482

Smoking ~ Methylation at 0

probe	gene	BETA	PVAL	FDR
cg05575921	AHRR	-0.090836011	4.698075e-43	2.349038e-38
cg09935388	GFI1	-0.124426692	5.991843e-15	2.995921e-10
cg12876356	GFI1	-0.139762694	4.096390e-13	1.024097e-08
cg05549655	CYP1A1	0.040526012	6.998679e-13	3.499339e-08
cg18146737	GFI1	-0.143917608	1.966087e-12	9.830437e-08

Smoking ~ Methylation at 7

probe	gene	BETA	PVAL	FDR
cg12803068	MYO1G	0.1053596256	6.513525e-18	3.256762e-13
cg22132788	MYO1G	0.0665389230	6.391079e-15	3.195539e-10
cg04180046	MYO1G	0.0711787819	5.712702e-14	1.428175e-09
cg19089201	MYO1G	0.0584275620	1.251775e-11	6.258875e-07
cg05549655	CYP1A1	0.0312192761	5.454444e-11	2.727222e-06

Drinking ~ Methylation at 0

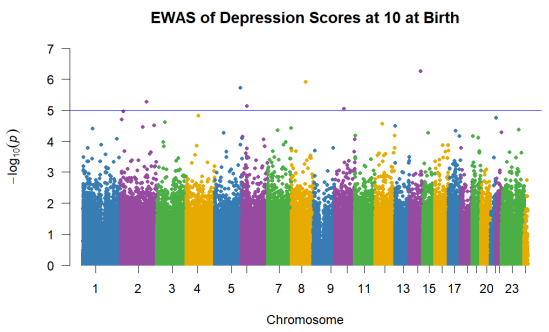
probe	gene	BETA	PVAL	FDR
cg16817435	Na	0.0435640168	3.156458e-06	0.1578229
cg20904024	Na	-0.0263785219	9.841097e-06	0.4920548
cg26224499	TIMM44	0.0005892556	1.196082e-05	0.5980411
cg10243196	RPSAP58	0.0227739459	1.398617e-05	0.6993085
cg07116393	MUL1	0.0024151774	2.916548e-05	0.7291369

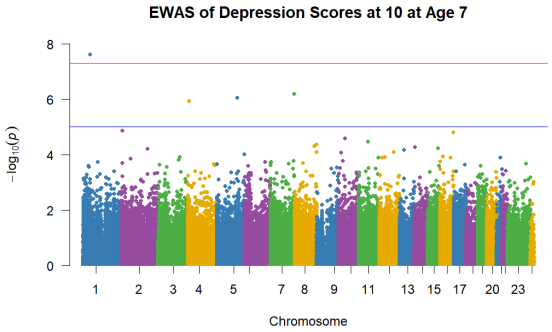
Drinking ~ Methylation at 7

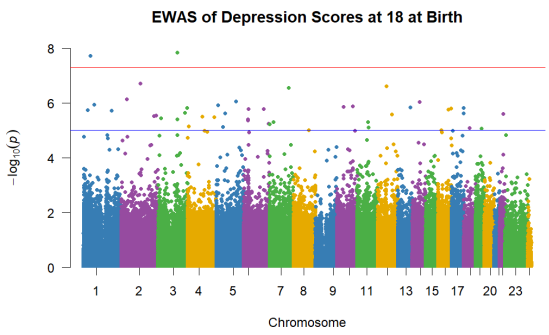
probe	gene	BETA	PVAL	FDR
cg01473643	CNOT4	-0.0137330138	1.912893e-07	0.009564465
cg21253551	Na	-0.0245945964	2.642458e-07	0.013212292
cg10677933	CABIN1	-0.0227732152	2.934530e-07	0.009641398

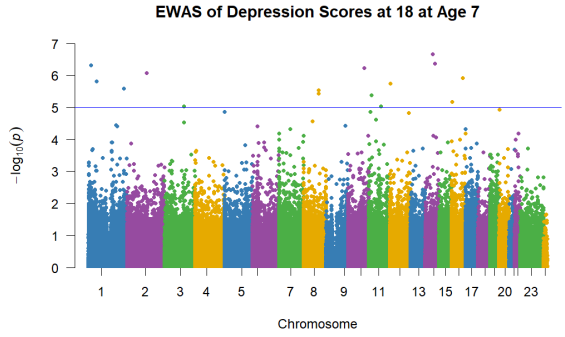
cg01794812	ZNF517	-0.0243019859	5.232279e-07	0.009232575
cg00763223	DAB1;MIR548D2	-0.0128961081	5.307257e-07	0.026536287

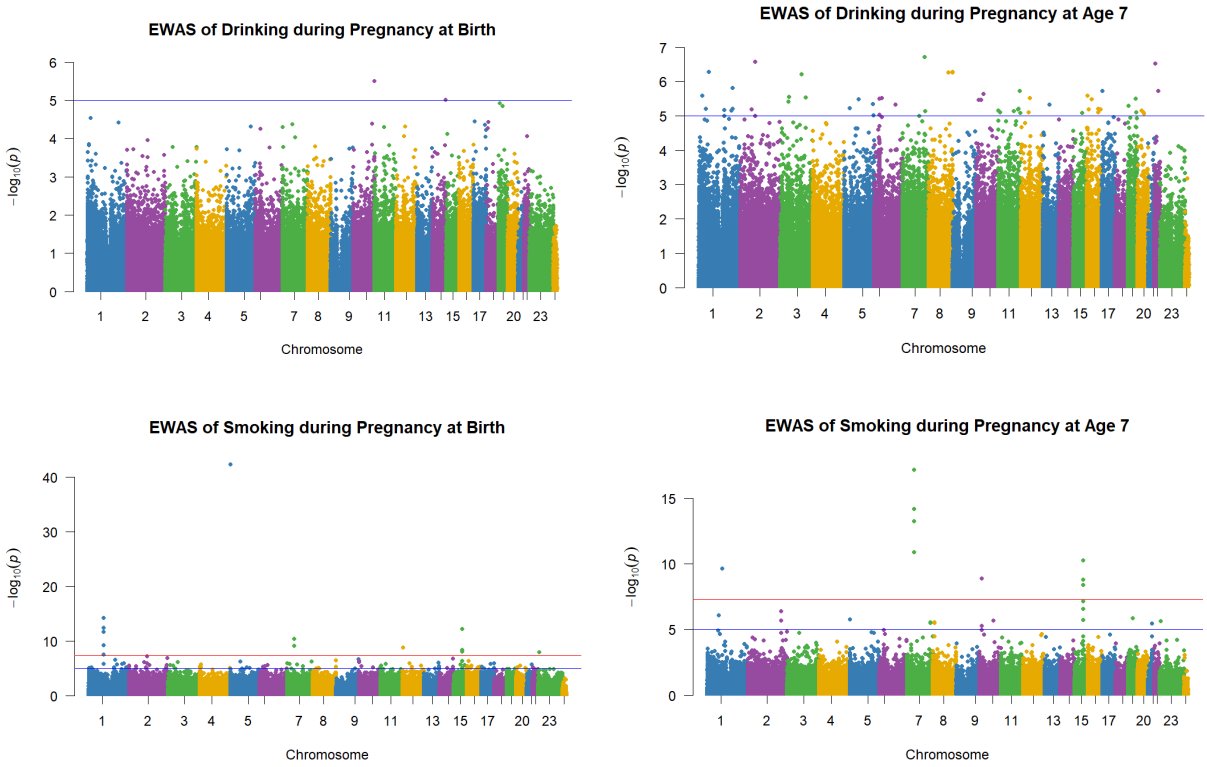
Manhattan plots for each analysis can be seen below:











Comparison across EWAS

Comparison of the CpG sites with the lowest p-values from each EWAS identified a number of sites that overlapped across analyses. Sites were considered of particular interest where they showed nominal significance ($p < 0.05$) in both a childhood adversity-related exposure and a mental health-related outcome. This included parental stress, childhood trauma, and maternal smoking or drinking during pregnancy as childhood adversity-related exposures, and depression history, cortisol stress reactivity, and depression scores as mental health-related outcomes.

A total of 14 CpG sites showed nominally significant associations with both an adversity-related exposure and a mental health-related outcome. These 14 sites should therefore be considered candidate CpGs showing cross-domain evidence rather than 14 independently FDR-significant EWAS findings. Four of these sites also had an FDR-significant association with a mental health outcome. These were cg20977426, cg24826867 (IRF8), cg02189786 (THAP11/CENPT), and cg13473356 (PEX5L). All four were FDR significant in the analysis of depression scores at age 18 using methylation measured at birth.

cg20977426 showed an association with maternal drinking during pregnancy at age 7 ($p = 0.0498$) and a strong association with depression scores at age 18 using methylation measured at birth ($p = 1.90 \times 10^{-8}$). cg24826867, located in IRF8, was associated with maternal stress during infancy ($p = 0.0312$) and preschool ($p = 0.00349$), as well as depression scores at age 18 using methylation measured at birth ($p = 1.64 \times 10^{-6}$). cg02189786, located in THAP11/CENPT, was associated with maternal smoking during pregnancy at age 7 ($p = 0.0194$) and depression scores at age 18 using methylation measured at birth ($p = 1.78 \times 10^{-6}$). cg13473356, located in PEX5L, was associated with childhood trauma scores ($p = 0.0209$) and depression scores at age 18 using methylation measured at birth ($p = 2.28 \times 10^{-6}$).

Several additional sites showed nominal associations across multiple adversity-related and mental health-related EWAS. cg01107031 (TGFB1) showed associations with maternal stress during preschool ($p = 0.00769$), paternal stress during infancy ($p = 0.00714$), paternal stress during preschool ($p = 0.00150$), childhood trauma ($p = 0.00889$), and cortisol stress reactivity ($p = 0.000708$). cg02920604 (MYO1E) was associated with maternal drinking during pregnancy at age 7 ($p = 0.00215$), maternal smoking during pregnancy at birth ($p = 0.00688$), depression history ($p = 0.00382$), and cortisol stress reactivity ($p = 0.00863$). cg03002480 (UBR1) was associated with paternal stress during infancy ($p = 0.00977$), maternal smoking during pregnancy at birth ($p = 0.00877$), depression scores at age 10 ($p = 0.00439$), and depression scores at age 18 ($p = 0.000507$).

cg13725825 showed associations with childhood trauma ($p = 0.00286$) and maternal drinking during pregnancy at age 7 ($p = 0.00468$), as well as cortisol stress reactivity ($p = 0.00223$) and depression scores at age 18 using methylation measured at age 7 ($p = 0.000943$). cg23596567 (ZNF18) was associated with maternal smoking during pregnancy at age 7 ($p = 0.00814$), depression history ($p = 0.00111$), cortisol stress reactivity ($p = 0.00976$), and depression scores at age 18 using methylation measured at age 7 ($p = 0.00936$). cg03363175 (NLE1) showed associations with maternal stress during preschool ($p = 0.0226$), paternal stress during infancy ($p = 0.00278$), paternal stress during preschool ($p = 0.00230$), and depression history ($p = 0.00870$).

cg13701109 (ADAMTS19) was associated with maternal stress during infancy ($p = 0.00905$), maternal stress during preschool ($p = 0.00872$), maternal smoking during pregnancy at birth ($p = 0.0312$), and cortisol stress reactivity ($p = 0.00347$). cg21246431 (GK5) was associated with paternal stress during infancy ($p = 0.00697$), paternal stress during preschool ($p = 0.000161$), cortisol stress reactivity ($p = 0.00486$), and depression scores at age 18 using methylation measured at birth ($p = 0.0394$).

The sites cg23876203 (HOXA3) and cg20562102 showed the greatest number of nominally significant associations, with six EWAS showing $p < 0.05$ for each site. cg23876203 was associated with maternal smoking and drinking during pregnancy at both birth and age 7, and with depression scores at age 10 using methylation measured at both birth and age 7. cg20562102 was associated with maternal smoking during pregnancy at both birth and age 7, and with depression scores at both age 10 and age 18 using methylation measured at both birth and age 7.

Overall, these comparisons identified 14 CpG sites showing associations across both childhood adversity-related exposures and mental health-related outcomes. Of these, four sites — cg20977426, cg24826867, cg02189786, and cg13473356 — had FDR-significant associations with depression scores at age 18 using methylation measured at birth. The remaining sites showed nominal associations across multiple EWAS, with cg23876203 and cg20562102 showing the greatest number of cross-domain associations. The results from the four datasets included different measures of parental stress, childhood trauma, maternal smoking and drinking, cortisol stress reactivity, depression history, and depression scores, allowing these overlapping CpG sites to be identified across several measures of adversity and mental health.

Discussion

The present study identified 14 CpG sites showing nominal associations with both a childhood adversity-related exposure and a mental health-related outcome. Four sites — cg20977426, cg24826867 (IRF8), cg02189786 (THAP11/CENPT), and cg13473356 (PEX5L) — also showed FDR-significant associations with depression scores at age 18 when methylation was measured at birth. The remaining sites showed nominal associations across multiple EWAS, with cg23876203 (HOXA3) and cg20562102 each showing associations in six EWAS.

These findings fit within a growing literature suggesting that early-life adversity may become biologically embedded through epigenetic processes. Reviews of childhood maltreatment and DNA methylation have identified associations across a range of candidate genes and genome-wide studies, including genes involved in stress regulation, neurodevelopment and immune function, although considerable heterogeneity exists between studies in the adversity measures, tissues, developmental stages and CpGs identified (Cecil et al., 2020; Parade et al., 2021; Rubens et al., 2023). Similarly, reviews of DNA methylation and depression have identified associations between methylation and depression-related phenotypes, while highlighting substantial methodological heterogeneity and relatively limited replication of individual loci (Li et al., 2019; Szyf et al., 2016).

The current approach extends this literature by focusing on overlap between adversity and mental-health EWAS. Rather than considering associations with adversity or depression in isolation, the 14 CpGs were prioritised because the same loci showed evidence of association across both domains. This is particularly interesting for the four FDR-significant sites, because their associations with depression at age 18 were observed using methylation measured at birth, providing temporal separation between the methylation measurement and the later mental-health outcome.

IRF8 and neuroimmune pathways

Among the 14 sites, cg24826867, annotated to IRF8, is particularly compelling biologically. In the present analyses, cg24826867 was associated with maternal stress during infancy and preschool, while methylation at birth was strongly associated with depression scores at age 18 ($p = 1.64 \times 10^{-6}$) and survived FDR correction.

IRF8 is a transcription factor with an important role in microglial development and function. Recent experimental work demonstrated that IRF8 binds enhancer regions during postnatal microglial development and contributes to chromatin accessibility, histone modifications and DNA-methylation patterns that are characteristic of microglia. Loss of IRF8 altered microglial identity and transcriptional programmes, demonstrating that IRF8 is involved not simply in immune-cell differentiation but in establishing the epigenetic landscape of microglia (Saeki et al., 2024).

This is potentially important in the context of childhood adversity and depression because microglia are increasingly implicated in the effects of stress on the developing and mature brain. The relevance of IRF8 to behaviour is supported by experimental evidence showing that IRF8 deficiency in mice produces sex-dependent anxiety-like and repetitive behaviours, with altered microglial oxidative-stress responses also observed (Zhu et al., 2025). Although these findings concern IRF8 function rather than the specific

human cg24826867 site, they provide a plausible biological context for the current association between maternal stress, IRF8-associated methylation and later depression.

An important limitation is that cg24826867 being annotated to IRF8 does not demonstrate that methylation at this CpG regulates IRF8 expression. Nevertheless, the combination of an adversity association in the present data, an FDR-significant prospective association with depression, and independent evidence for a role of IRF8 in microglial epigenetic regulation makes this one of the more compelling candidates for further investigation.

PEX5L and neuronal excitability

cg13473356, annotated to PEX5L, provides another particularly interesting finding. This site was associated with childhood trauma ($p = 0.0209$) and showed an FDR-significant association with depression scores at age 18 using methylation measured at birth ($p = 2.28 \times 10^{-6}$).

PEX5L encodes the protein TRIP8b, an auxiliary subunit of hyperpolarisation-activated cyclic nucleotide-gated (HCN) channels. These channels contribute to neuronal excitability, particularly in hippocampal and cortical neurons. Experimental deletion of TRIP8b in mice alters HCN channel localisation and function in hippocampal pyramidal neurons and produces behavioural changes interpreted as having antidepressant-like properties (Lewis et al., 2011).

The relevance of this pathway to the current results is therefore notable. Childhood trauma was associated with methylation at cg13473356, while the same CpG showed one of the strongest associations with depression at age 18. The biological literature provides a plausible connection between PEX5L/TRIP8b and neuronal excitability and depression-related behaviour, although the existing experimental evidence does not demonstrate that methylation of cg13473356 changes PEX5L expression or HCN-channel function. In addition, cg13473356 has previously been identified as an age-associated CpG in an independent methylation study, showing that this site is not unique to the current dataset (Florath et al., 2014)

Thus, PEX5L represents a particularly interesting candidate because the exact CpG is associated with both childhood trauma and later depression in the present study, while gene-level evidence provides a plausible neurobiological mechanism. Further work examining whether cg13473356 is an expression quantitative trait methylation site for PEX5L would be particularly informative.

THAP11/CENPT and developmental regulation

cg02189786, annotated to THAP11/CENPT, was associated with maternal smoking during pregnancy at age 7 and showed an FDR-significant association with depression scores at age 18 using methylation measured at birth ($p = 1.78 \times 10^{-6}$).

THAP11 is a DNA-binding transcription factor involved in regulation of gene expression through interactions with HCFC1 and ZNF143. It regulates genes involved in fundamental cellular processes

including protein biosynthesis and metabolism, and pathogenic THAP11 variants have been associated with neurodevelopmental phenotypes (Quintana et al., 2017).

The relevance to the current findings is therefore potentially developmental: the CpG was associated with prenatal smoking, an exposure occurring during a period of substantial epigenetic and neurodevelopmental change, and methylation at birth was subsequently associated with depression at age 18. There is not currently strong evidence that THAP11 methylation specifically mediates the effects of prenatal smoking or contributes to depression. Nevertheless, its role as a transcriptional regulator involved in development makes it a plausible candidate through which an early environmental exposure could potentially influence later biological phenotypes.

The CENPT annotation also provides an alternative biological context because CENPT is involved in kinetochore and centromere function and therefore cell division. The presence of two gene annotations means that the biological interpretation of cg02189786 should not be restricted to THAP11 alone. Further annotation and functional analyses would be required to establish which gene, if either, is most directly relevant to this CpG.

TGFB1 and stress-related signalling

cg01107031 (TGFB1) showed one of the strongest patterns of convergence across the present EWAS. The CpG was associated with maternal stress during preschool, paternal stress during infancy and preschool, childhood trauma, and cortisol stress reactivity.

TGFB1 is involved in TGF-β signalling, a pathway with important roles in immune regulation, cellular differentiation and tissue responses to environmental stress. There is also evidence connecting TGFB1 with depression-related phenotypes. Białek et al. (2020) reported associations between TGFB1 genetic variation and the occurrence, severity and treatment response of major depressive disorder, providing evidence that variation in this pathway may be relevant to depression.

The present finding is particularly interesting because the association was not limited to a single adversity measure. cg01107031 was associated with both maternal and paternal stress measures and childhood trauma, as well as cortisol stress reactivity. This pattern is consistent with the possibility that TGFB1-related signalling could form part of a broader biological response to early environmental stress.

Importantly, cg01107031 itself has also been identified in independent methylation datasets, including as a promoter-associated CpG in TGFB1 (Erbe et al., 2021). This does not establish that the site regulates TGFB1 in the present samples, but it strengthens the rationale for investigating whether methylation at this locus has functional consequences.

HOXA3 and developmental programming

cg23876203 (HOXA3) showed six nominally significant associations, making it one of the most repeatedly associated CpGs in the analysis. It was associated with maternal smoking and drinking during

pregnancy at both birth and age 7, as well as with depression scores at age 10 using methylation measured at both birth and age 7.

HOXA3 is a homeobox transcription factor involved in embryonic development, morphogenesis and differentiation. More directly relevant to the current study, Collender et al. (2023) identified differential methylation involving the HOXA3 region in newborn cord blood in relation to maternal adverse childhood experiences. This provides independent evidence that methylation around HOXA3 may be sensitive to early-life adversity, although the CpG identified in that study was not cg23876203 (Collender et al., 2023).

The present findings therefore provide an interesting extension of previous work. cg23876203 showed associations with prenatal exposures and later depression, while independent studies have implicated the broader HOXA3 region in methylation associated with adversity. This convergence is particularly relevant given the developmental function of HOXA3, as prenatal environmental exposures occur during periods in which developmental gene regulation is highly dynamic.

The association with prenatal smoking is also consistent with the wider literature demonstrating that maternal smoking during pregnancy can alter offspring DNA methylation, with some methylation differences detectable from birth and persisting into later childhood (Richmond et al., 2015). However, the major replicated smoking-associated CpGs reported by Richmond et al. were concentrated in loci such as AHRR, CYP1A1, GFI1 and MYO1G, rather than cg23876203. Thus, the present HOXA3 finding should be considered complementary rather than a direct replication of the established prenatal-smoking methylation signature.

ADAMTS19 and brain expression in depression

cg13701109 (ADAMTS19) was associated with maternal stress during infancy and preschool, maternal smoking during pregnancy at birth, and cortisol stress reactivity. Although there is relatively little literature on ADAMTS19 in relation to childhood adversity, there is evidence linking the gene to depression-related brain biology. Yuan et al. (2021) identified ADAMTS19 as differentially expressed in both the hippocampus and prefrontal cortex in major depressive disorder, with expression reduced in MDD samples compared with controls.

This is noteworthy because the present association involves methylation rather than gene expression, but occurs at a locus annotated to the same gene. ADAMTS19 encodes a member of the ADAMTS family of extracellular metalloproteinases, suggesting a potential connection with extracellular-matrix and cellular signalling processes (Yuan et al., 2021). Although it remains unknown whether cg13701109 influences ADAMTS19 expression, the convergence between stress-related methylation in the present study and altered ADAMTS19 expression in depression-related brain tissue provides a reasonable basis for further investigation.

NLE1 and neurodevelopment

cg03363175 (NLE1) showed associations with maternal stress during preschool, paternal stress during infancy and preschool, and depression history. NLE1 encodes notchless homolog 1 and is expressed in the human brain, including the hippocampal formation and cerebral cortex (Uhlén et al., 2015).

The presence of NLE1 expression in brain regions relevant to emotional regulation provides some biological context for the current finding, particularly because the associations extend across several measures of parental stress and depression history. However, there is currently considerably less evidence directly connecting NLE1 to depression or childhood adversity than exists for genes such as IRF8 or PEX5L. The present finding may therefore represent a potentially novel association involving a gene expressed in brain tissue, warranting replication before stronger conclusions can be drawn.

UBR1 and neurodevelopmental phenotypes

cg03002480 (UBR1) was associated with paternal stress during infancy, maternal smoking during pregnancy at birth, and depression scores at both age 10 and age 18. UBR1 encodes an E3 ubiquitin ligase involved in the N-end rule pathway, which regulates protein degradation. Loss-of-function mutations in UBR1 cause Johanson-Blizzard syndrome, a developmental disorder that can include intellectual disability and neurological abnormalities (Zenker et al., 2005).

Although this does not provide evidence that UBR1 methylation contributes to depression, it does indicate that disruption of UBR1 can have substantial developmental consequences. The repeated association of cg03002480 with both prenatal adversity and depression across two ages is therefore interesting in the context of developmental programming. Again, the relevant limitation is that the present study identifies a methylation association near UBR1 rather than demonstrating altered UBR1 expression.

MYO1E, ZNF18 and GK5

The remaining genes provide additional, although more preliminary, biological context. cg02920604 (MYO1E) was associated with maternal drinking and smoking during pregnancy, depression history and cortisol stress reactivity. MYO1E encodes an unconventional myosin involved in actin-dependent membrane trafficking and has recognised roles in immune-cell biology, including MHC class II trafficking in dendritic cells. This is potentially relevant given the repeated involvement of stress and immune-related phenotypes in the present results, although there is currently no strong evidence that MYO1E methylation itself is involved in depression or childhood adversity.

cg23596567 (ZNF18) was associated with maternal smoking during pregnancy, depression history, cortisol stress reactivity and depression scores at age 18 using methylation measured at age 7. ZNF18 encodes a zinc-finger transcription factor involved in transcriptional regulation and is expressed in the brain (Victor et al., 2023). ZNF18 expression has also been examined in patient-derived neuronal models, although the evidence currently relates to neurodevelopmental disorders rather than depression specifically. Thus, the present finding is consistent with the possibility that a transcriptional regulator expressed in neural tissue may be relevant to adversity-related phenotypes, but further evidence is required.

Finally, cg21246431 (GK5) was associated with paternal stress during infancy and preschool, cortisol stress reactivity and depression scores at age 18. GK5 has previously been identified in a blood-based transcriptomic study of schizophrenia and bipolar disorder, where alternative splicing of the gene differed between diagnostic groups (Mirnics et al., 2011). This provides some evidence that GK5 may be relevant to psychiatric phenotypes, although the previous study examined RNA splicing rather than DNA methylation and involved different disorders. Consequently, the current association provides an additional hypothesis regarding GK5 that requires independent replication.

Sites without established gene-level evidence

Three of the 14 sites — cg20977426, cg13725825 and cg20562102 — do not have a strong gene annotation that provides an obvious biological pathway to discuss. Nevertheless, their patterns within the present data make them important candidates.

cg20977426 was associated with maternal drinking during pregnancy at age 7 and showed an exceptionally strong association with depression scores at age 18 using methylation measured at birth ($p = 1.90 \times 10^{-8}$), surviving FDR correction. cg13725825 was associated with childhood trauma, maternal drinking during pregnancy, cortisol stress reactivity and depression scores at age 18 using methylation measured at age 7. cg20562102 showed six nominally significant EWAS, including associations with maternal smoking during pregnancy and depression scores at ages 10 and 18 across both birth and age-7 methylation.

The absence of a well-established gene-level mechanism for these sites does not make them less interesting. Indeed, cg20977426 is particularly noteworthy because it combines an adversity association with one of the strongest depression associations observed among the 14 sites. The lack of previous biological evidence instead makes these loci potentially novel findings, which will require replication and functional annotation.

Convergence across adversity and mental-health phenotypes

Taken together, the 14 sites show several patterns that are consistent with the hypothesis that early adversity and later mental health may share epigenetic correlates. The strongest evidence comes from the four FDR-significant sites, all of which were associated with depression at age 18 using methylation measured at birth. However, the broader set of nominally significant sites is also informative because several were associated with multiple types of adversity or multiple mental-health phenotypes.

For example, cg01107031 showed associations across maternal stress, paternal stress and childhood trauma, in addition to cortisol stress reactivity. cg23876203 and cg20562102 each showed six nominally significant associations, making them the most consistently associated sites across the EWAS. This repeated pattern is potentially more informative than an isolated nominal association because it suggests that some loci may be responsive to multiple related aspects of the early environment or may be associated with multiple manifestations of later mental-health vulnerability.

At the same time, these phenotypes are not independent. Prenatal smoking, prenatal drinking, parental stress and childhood adversity can be correlated, as can depression scores and cortisol-related measures. Consequently, repeated nominal associations cannot by themselves distinguish a genuinely shared biological pathway from correlated exposures, shared confounding or chance findings. This is a recognised challenge in the wider literature on adversity-related DNA methylation (Parade et al., 2021; Rubens et al., 2023).

Interpretation of CpG-to-gene relationships

An important consideration when interpreting these findings is the relationship between CpG sites and their annotated genes. Illumina CpG annotation does not necessarily mean that methylation at a site directly regulates transcription of the nearest or annotated gene. DNA methylation can have different effects depending on its genomic location, and the functional relationship between an individual CpG and gene expression must generally be demonstrated empirically through expression quantitative trait methylation analyses or functional experiments.

Nevertheless, gene-level evidence remains highly relevant to the interpretation of the current findings. For example, the biological relevance of IRF8 to microglial epigenetic regulation, PEX5L to neuronal HCN-channel function, HOXA3 to developmental regulation, TGFB1 to immune and depression-related biology, and ADAMTS19 to gene-expression changes observed in depression provides plausible hypotheses about the biological significance of the corresponding CpGs (Saeki et al., 2024; Lewis et al., 2011; Collender et al., 2023; Białek et al., 2020; Yuan et al., 2021).

Therefore, the absence of evidence that a particular CpG regulates its annotated gene should not prevent discussion of the gene itself. Instead, the distinction should be explicit: the present study demonstrates an association at the CpG, while the existing literature provides biological context at the gene or pathway level. This distinction is particularly important for novel loci such as cg20977426, cg13725825 and cg20562102, for which there is currently less established biological context.

Limitations and future directions

Several limitations should be considered. First, the criterion used to identify the 14 sites included nominal significance, meaning that some associations may represent false-positive findings arising from the large number of CpGs and EWAS examined. Although requiring evidence across both adversity and mental-health domains provides a useful prioritisation strategy, it does not eliminate the possibility of chance overlap. This is consistent with the wider literature, in which replication of individual adversity-associated CpGs has often been inconsistent (Cecil et al., 2020; Parade et al., 2021; Rubens et al., 2023).

Second, the methylation measurements were obtained from blood, and peripheral blood methylation cannot necessarily be assumed to represent methylation in the brain. This is particularly important when interpreting genes such as IRF8 and PEX5L, where much of the relevant biological evidence comes from neural or animal models. Differences in blood-cell composition may also contribute to observed methylation differences, although the EWAS models accounted for estimated cell proportions.

A major strength is the longitudinal structure of the datasets. In particular, the depression-at-18 analysis used methylation measured at birth, allowing the four FDR-significant associations to be examined prospectively. The findings at cg23876203 and cg20562102 are also notable because associations were observed across methylation measured at different developmental stages and later depression outcomes. Nevertheless, temporal precedence does not establish causality, and the present results cannot determine whether methylation is a cause, consequence or correlate of adversity-related biological processes.

Overall, the downstream analysis identifies a set of CpG sites that provide candidate links between childhood adversity and later mental-health phenotypes. The most compelling candidates include cg24826867/IRF8, cg13473356/PEX5L, cg02189786/THAP11/CENPT, cg01107031/TGFB1 and cg23876203/HOXA3, where the current findings can be considered alongside independent evidence implicating the corresponding genes or genomic regions in microglial regulation, neuronal signalling, development, stress-related signalling or adversity-related methylation (Saeki et al., 2024; Lewis et al., 2011; Collender et al., 2023; Białek et al., 2020).

The findings at ADAMTS19, UBR1, NLE1, ZNF18, MYO1E and GK5 also provide potentially useful biological hypotheses, although the evidence connecting these genes to adversity or mental health is less direct. Finally, cg20977426, cg13725825 and cg20562102 represent potentially novel loci for which the strongest evidence currently comes from their cross-EWAS pattern rather than from established gene-level mechanisms.

Further research should therefore focus on replication of these CpGs in independent longitudinal cohorts, assessment of CpG–gene expression relationships, investigation of relevant tissues and cell types, and functional studies of the most strongly supported loci. Such analyses would help establish whether the observed methylation differences are reproducible biological signatures of early adversity, whether they influence the expression or function of the annotated genes, and whether they contribute to the pathway from childhood adversity to later mental-health difficulties.

References

Background / epidemiology

Li, M., D'Arcy, C. & Meng, X. (2016). Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis, and proportional attributable fractions. Psychological Medicine, 46(4), 717–730. https://doi.org/10.1017/S0033291715002743.

World Health Organization (2025). Depressive disorder (depression). WHO Fact Sheet, 29 August 2025.

Ribas, L.H., Montezano, B.B., Nieves, M., Kampmann, L.B. & Jansen, K. (2024). The role of parental stress on emotional and behavioral problems in offspring: a systematic review with meta-analysis. Jornal de Pediatria, 100(6), 565–585. https://doi.org/10.1016/j.jped.2024.02.003.

Dachew, B.A., Ayano, G., Duko, B., Lawrence, B., Betts, K. & Alati, R. (2023). Paternal depression and risk of depression among offspring: a systematic review and meta-analysis. JAMA Network Open, 6(8), e2329159. https://doi.org/10.1001/jamanetworkopen.2023.29159.

Dataset papers

Essex, M.J., Boyce, W.T., Hertzman, C., Lam, L.L., Armstrong, J.M., Neumann, S.M.A. & Kobor, M.S. (2013). Epigenetic vestiges of early developmental adversity: childhood stress exposure and DNA methylation in adolescence. Child Development, 84(1), 58–75. https://doi.org/10.1111/j.1467-8624.2011.01641.x.

Houtepen, L.C., Vinkers, C.H., Carrillo-Roa, T., Hiemstra, M., van Lier, P.A., Meeus, W., Branje, S., Heim, C.M., Nemeroff, C.B., Mill, J., Schalkwyk, L.C., Creyghton, M.P., Kahn, R.S., Joëls, M. et al. (2016). Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. Nature Communications, 7, 10967. https://doi.org/10.1038/ncomms10967.

Crawford, B., Craig, Z., Mansell, G., White, I., Smith, A., Spaull, S., Imm, J., Hannon, E., Wood, A., Yaghootkar, H., Ji, Y., Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium, Mullins, N., Lewis, C.M., Mill, J. & Murphy, T.M. (2018). DNA methylation and inflammation marker profiles associated with a history of depression. Human Molecular Genetics, 27(16), 2840–2850. https://doi.org/10.1093/hmg/ddy199.

Richmond, R.C., Simpkin, A.J., Woodward, G., Gaunt, T.R., Lyttleton, O., McArdle, W.L., Ring, S.M., Smith, A.D.A., Timpson, N.J., Tilling, K., Davey Smith, G. & Relton, C.L. (2015). Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). Human Molecular Genetics, 24(8), 2201–2217. https://doi.org/10.1093/hmg/ddu739.

Reviews / general epigenetic literature

Cecil, C.A.M., Zhang, Y. & Nolte, T. (2020). Childhood maltreatment and DNA methylation: A systematic review. Neuroscience & Biobehavioral Reviews, 112, 392–409. https://doi.org/10.1016/j.neubiorev.2020.02.019.

Parade, S.H., Huffhines, L., Daniels, T.E., Stroud, L.R., Nugent, N.R. & Tyrka, A.R. (2021). A systematic review of childhood maltreatment and DNA methylation: candidate gene and epigenome-wide approaches. Translational Psychiatry, 11, 134. https://doi.org/10.1038/s41398-021-01207-y.

Rubens, M., Bruenig, D., Adams, J.A.M., Suresh, S.M., Sathyanarayanan, A., Haslam, D., Shenk, C.E., Mathews, B. & Mehta, D. (2023). Childhood maltreatment and DNA methylation: A systematic review. Neuroscience & Biobehavioral Reviews, 147, 105079. https://doi.org/10.1016/j.neubiorev.2023.105079.

Li, M., D'Arcy, C., Li, X., Zhang, T., Joober, R. & Meng, X. (2019). What do DNA methylation studies tell us about depression? A systematic review. Translational Psychiatry, 9, 68. https://doi.org/10.1038/s41398-019-0412-y.

Szyf, M., Tang, Y.Y., Hill, K.G. & Musci, R. (2016). The dynamic epigenome and its implications for behavioral interventions: a role for epigenetics to inform disorder prevention and health promotion. Translational Behavioral Medicine, 6(1), 55–62. https://doi.org/10.1007/s13142-016-0387-7.

Gene / CpG-specific literature

Saeki, K., Pan, R., Lee, E., Kurotaki, D. & Ozato, K. (2024). IRF8 defines the epigenetic landscape in postnatal microglia, thereby directing their transcriptome programs. Nature Immunology, 25, 1928–1942. https://doi.org/10.1038/s41590-024-01962-2.

Zhu, S., Saeki, K., Karlsson, R.-M., Abebe, D. & Ozato, K. (2025). IRF8 deficiency causes anxiety-like behavior in a sex-dependent manner. bioRxiv [Preprint]. https://doi.org/10.1101/2025.09.02.673764.

Lewis, A.S. et al. (2011). Deletion of the hyperpolarization-activated cyclic nucleotide-gated channel auxiliary subunit TRIP8b impairs hippocampal Ih localization and function and promotes antidepressant behavior in mice. Journal of Neuroscience.

Erbe, R., Wang, Z., Wu, S. et al. (2021). Evaluating the impact of age on immune checkpoint therapy biomarkers. Cell Reports, 36(8), 109599. https://doi.org/10.1016/j.celrep.2021.109599.

Florath, I., Butterbach, K., Müller, H., Bewerunge-Hudler, M. & Brenner, H. (2014). Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. Human Molecular Genetics, 23(5), 1186–1201. https://doi.org/10.1093/hmg/ddt531.

Quintana, A.M. et al. (2017). Mutations in THAP11 cause an inborn error of cobalamin metabolism and developmental abnormalities. Human Molecular Genetics, 26(15), 2838–2849. https://doi.org/10.1093/hmg/ddx157.

Białek, K., Czarny, P., Watała, C., Wigner, P., Talarowska, M., Gałecki, P., Szemraj, J. & Sliwinski, T. (2020). Novel association between TGFA, TGFB1, IRF1, PTGS2 and IKBKB single-nucleotide polymorphisms and occurrence, severity and treatment response of major depressive disorder. PeerJ, 8, e8676. https://doi.org/10.7717/peerj.8676.

Collender, P., Bozack, A.K., Veazie, S., Nwanaji-Enwerem, J.C., Van Der Laan, L., Kogut, K., Riddell, C., Eskenazi, B., Holland, N., Deardorff, J. & Cardenas, A. (2023). Maternal adverse childhood experiences (ACEs) and DNA methylation of newborns in cord blood. Clinical Epigenetics, 15, 162. https://doi.org/10.1186/s13148-023-01581-y.

Yuan, N., Tang, K., Da, X., Gan, H., He, L., Li, X., Ma, Q. & Chen, J. (2021). Integrating Clinical and Genomic Analyses of Hippocampal-Prefrontal Circuit Disorder in Depression. Frontiers in Genetics, 11, 565749. https://doi.org/10.3389/fgene.2020.565749.

Uhlén, M. et al. (2015). Tissue-based map of the human proteome. Science, 347(6220), 1260419. https://doi.org/10.1126/science.1260419.

Zenker, M. et al. (2005). Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome). Nature Genetics, 37, 1345–1350. https://doi.org/10.1038/ng1681.

Victor, A.K., Hedgecock, T., Donaldson, M., Johnson, D., Rand, C.M., Weese-Mayer, D.E. & Reiter, L.T. (2023). Analysis and comparisons of gene expression changes in patient-derived neurons from ROHHAD, CCHS, and PWS. Frontiers in Pediatrics, 11, 1090084. https://doi.org/10.3389/fped.2023.1090084.

Mirnics, K. et al. (2011). Alternatively spliced genes as biomarkers for schizophrenia, bipolar disorder and psychosis: a blood-based spliceome-profiling exploratory study. Molecular Psychiatry.