

Published in final edited form as:

*Mol Psychiatry*. 2023 March 01; 28(3): 1128–1136. doi:10.1038/s41380-022-01871-6.

## DNA methylation and general psychopathology in childhood: an epigenome-wide meta-analysis from the PACE consortium

Jolien Rijlaarsdam<sup>1,28</sup>, Marta Cosin-Tomas<sup>2,3,4,28</sup>, Laura Schellhas<sup>5,6,28</sup>, Sarina Abrishamcar<sup>7</sup>, Anni Malmberg<sup>8</sup>, Alexander Neumann<sup>9</sup>, Janine F. Felix<sup>10,11</sup>, Jordi Sunyer<sup>2,3,4</sup>, Kristine B. Gutzkow<sup>12</sup>, Regina Grazuleviciene<sup>13</sup>, John Wright<sup>14</sup>, Mariza Kampouri<sup>15</sup>, Heather J. Zar<sup>16,17</sup>, Dan J. Stein<sup>18,19</sup>, Kati Heinonen<sup>20,8</sup>, Katri Räikkönen<sup>8</sup>, Jari Lahti<sup>8</sup>, Anke Hüls<sup>7,21</sup>, Doretta Caramaschi<sup>22,23,29</sup>, Silvia Alemany<sup>2,24,25,29</sup>, Charlotte A. M. Cecil<sup>1,26,27,29</sup>

<sup>1</sup>Department of Child and Adolescent Psychiatry/ Psychology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>2</sup>ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain

<sup>3</sup>Universitat Pompeu Fabra, Barcelona, Spain

<sup>4</sup>Centro de investigación biomédica en red en epidemiología y salud pública (ciberesp), Madrid, Spain

---

This work is licensed under a [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) International license.

Correspondence to: Marta Cosin-Tomas; Charlotte A. M. Cecil.

Correspondence and requests for materials should be addressed to Marta Cosin-Tomas or Charlotte A. M. Cecil.

marta.cosin@isglobal.org; c.cecil@erasmusmc.nl.

<sup>28</sup>These authors contributed equally: Jolien Rijlaarsdam, Marta Cosin-Tomas, Laura Schellhas

<sup>29</sup>These authors jointly supervised this work: Doretta Caramaschi, Silvia Alemany, Charlotte A. M. Cecil

### Author Contributions

JR: analysis plan and study design, Generation R analysis, quality control of data and meta-analyses, interpretation of results, manuscript drafting; MCT: Helix and INMA analyses, quality control of data, shadow meta-analysis, functional and genomic enrichment analyses, results interpretation, manuscript drafting, revision; LS: contributed to analysis plan and study design, ALSPAC analysis, interpretation of results, and manuscript review; SaA: DCHS analysis, and manuscript review; AM: GLAKU analysis and manuscript review; AN: Generation R coordination, manuscript review; JF: study design, Generation R coordination, critical revision of the manuscript; JS: INMA coordination, provided funding, manuscript review; KBG: MoBa Helix funding, MoBa coordination; RG: KANC-Helix funding, KANC coordination; JW: BiB-Helix funding, BiB coordination; MK: Rhea-Helix data acquisition; HJZ: DCHS coordination, and manuscript review; DJS: DCHS coordination, and manuscript review; KH: GLAKU coordination, manuscript review; KR: GLAKU coordination, manuscript review; JL: GLAKU coordination, manuscript review; AH: DCHS analysis, and manuscript review; DC: contributed to analysis plan and review of drafts and manuscript; SiA: analysis plan and study design, INMA analysis, manuscript review; CC: analysis plan and study design, Generation R coordination, funding, manuscript drafting and review.

### Competing Interests

The authors declare no competing interests.

### Ethics Approval

All studies acquired approval from local ethics committees and informed consent was obtained for all participants. Full details are listed in the methods supplement.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

- <sup>5</sup>School of Psychological Science, MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
- <sup>6</sup>Institute for Sex Research, Sexual Medicine and Forensic Psychiatry, University Medical Center Hamburg, Eppendorf, Germany
- <sup>7</sup>Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA
- <sup>8</sup>Department of Psychology & Logopedics, University of Helsinki, Helsinki, Finland
- <sup>9</sup>VIB Center for Molecular Neurology, Antwerp, Belgium
- <sup>10</sup>The Generation R Study Group, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands
- <sup>11</sup>Department of Pediatrics, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands
- <sup>12</sup>Division of Climate and Environmental Health, Norwegian Institute of Public Health (NIPH), Oslo, Norway
- <sup>13</sup>Department of Environmental Science, Vytautas Magnus University, 44248 Kaunas, Lithuania
- <sup>14</sup>Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK
- <sup>15</sup>Department of Social Medicine, University of Crete, Crete, Greece
- <sup>16</sup>Department of Paediatrics and Child Health, Red Cross War Memorial Children's Hospital, University of Cape Town, Cape Town, South Africa
- <sup>17</sup>South African Medical Research Council (SAMRC) Unit on Child and Adolescent Health, University of Cape Town, Cape Town, South Africa
- <sup>18</sup>Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa
- <sup>19</sup>South African Medical Research Council (SAMRC) Unit on Risk and Resilience in Mental Disorders, Neuroscience Institute, University of Cape Town, Cape Town, South Africa
- <sup>20</sup>Psychology/ Welfare Sciences, Faculty of Social Sciences, Tampere University, Tampere, Finland
- <sup>21</sup>Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA
- <sup>22</sup>Medical Research Council Integrative Epidemiology Unit, Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK
- <sup>23</sup>Department of Psychology, University of Exeter, Exeter, UK
- <sup>24</sup>Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain

<sup>25</sup>Biomedical Network Research Centre on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Barcelona, Spain

<sup>26</sup>Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

<sup>27</sup>Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

## Abstract

The general psychopathology factor (GPF) has been proposed as a way to capture variance shared between psychiatric symptoms. Despite a growing body of evidence showing both genetic and environmental influences on GPF, the biological mechanisms underlying these influences remain unclear. In the current study, we conducted epigenome-wide meta-analyses to identify both probe- and region-level associations of DNA methylation (DNAm) with school-age general psychopathology in six cohorts from the Pregnancy And Childhood Epigenetics (PACE) Consortium. DNAm was examined both at birth (cord blood; prospective analysis) and during school-age (peripheral whole blood; cross-sectional analysis) in total samples of  $N = 2178$  and  $N = 2190$ , respectively. At school-age, we identified one probe (cg11945228) located in the Bromodomain-containing protein 2 gene (*BRD2*) that negatively associated with GPF ( $p = 8.58 \times 10^{-8}$ ). We also identified a significant differentially methylated region (DMR) at school-age ( $p = 1.63 \times 10^{-8}$ ), implicating the SHC Adaptor Protein 4 (*SHC4*) gene and the EP300-interacting inhibitor of differentiation 1 (*EID1*) gene that have been previously implicated in multiple types of psychiatric disorders in adulthood, including obsessive compulsive disorder, schizophrenia, and major depressive disorder. In contrast, no prospective associations were identified with DNAm at birth. Taken together, results of this study revealed some evidence of an association between DNAm at school-age and GPF. Future research with larger samples is needed to further assess DNAm variation associated with GPF.

## Introduction

Psychiatric disorders or symptoms co-occur more often than would be expected by chance alone [1, 2]. In light of the negative clinical and functional outcomes associated with psychiatric co-occurrence [3, 4], it is important to identify early indicators of risk and underlying biological mechanisms. There is accumulating evidence that, as early as in childhood, the shared variance between psychiatric disorders or symptoms may be usefully represented by a general psychopathology factor (GPF) [5–8]. This GPF in childhood has been found to show temporal stability [6] and to predict long-term functional and psychiatric outcomes in adolescence throughout adulthood [9, 10]. Although previous research has found evidence for both genetic and environmental influences on GPF [7, 11–16], the biological mechanisms underlying these influences remain unclear.

One of the ways by which genetic and environmental factors might contribute to disease susceptibility is through epigenetic mechanisms that regulate gene expression, such as DNA methylation (DNAm) [17]. Studies have shown that variation in DNAm is influenced by a dynamic interplay of genetic and environmental factors [18]. In turn, alterations in DNAm

patterns across the genome in peripheral tissues including cord blood, and peripheral blood have been found to associate with a wide range of child and adult mental health outcomes, such as conduct problems, attention deficit hyperactivity disorder (ADHD) symptoms, major depressive disorder (MDD), and schizophrenia [19–22]. Despite a growing body of research implicating an involvement of DNAm in individual mental health outcomes, much less work has focused on the relationship between DNAm and general psychopathology [23]. To the best of our knowledge, only one study examined the association between genome-wide DNAm patterns and GPF in childhood. In this study, data were analyzed cross-sectionally in one cohort, focusing on wider biological networks (so called ‘modules’) of co-methylated DNAm probes across the genome [23]. As such, we still lack knowledge on how GPF relates to single DNAm probes and the extent to which associations vary across time (i.e., both cross-sectional and prospective associations). Large multi-cohort epigenome-wide studies, which allow for increased power and generalizability, are needed to improve our understanding of the biological mechanisms underlying shared variance across mental health problems.

We conducted epigenome-wide meta-analyses to investigate both probe-level and region-based associations of DNAm with school-age GPF in the Pregnancy And Childhood Epigenetics (PACE) Consortium. Because it is unclear at which time point differential DNAm may be most relevant to GPF, we examined DNAm both at birth (cord blood; prospective study; pre-symptom manifestation) and at school-age (peripheral whole blood; cross-sectional study) in pooled samples of  $N = 2178$  and  $N = 2190$  children, respectively.

## Methods

### Participants

The prospective analyses included four cohorts from PACE, using complete data on DNAm at birth, general psychopathology in childhood and covariates: the Avon Longitudinal Study of Parents and Children (ALSPAC), Drakenstein Child Health Study (DCHS), Generation R (GENR), and Infancia y Medio Ambiente (INMA). These cohorts have a combined sample size of 2178 (see Table 1). All prospective cohorts included participants of European ancestry, except for DCHS, which included participants of predominantly Black African ancestry or mixed ancestry. See Supplementary Methods for full cohort descriptions.

The cross-sectional analyses included four cohorts from the PACE consortium, using complete data on DNAm and general psychopathology in childhood, as well as covariates; ALSPAC, GENR, Glycyrrhizin in Licorice (GLAKU), and Human Early Life Exposome (HELIX; including six jointly analyzed sub cohorts). These cohorts have a combined sample size of 2190 (see Table 1). All cross-sectional cohorts included participants of European ancestry, except for HELIX, which included participants of European ancestry and participants with a Pakistani background living in the United Kingdom, which were analyzed as separate cohorts in our meta-analysis.

## Measures

**DNA methylation**—DNAm was assessed with the Illumina Infinium HumanMethylation450 (ALSPAC, DCHS, GENR, HELIX, INMA) or the Infinium HumanMethylationEPIC (DCHS, GLAKU) BeadChip assays in cord blood and in peripheral whole blood at ages 7–12 years. The cohorts performed sample processing, quality control (QC) and normalization based on their preferred protocols as described in the Supplementary methods. We used normalized, untransformed beta values, ranging from 0 (fully unmethylated) to 1 (fully methylated). Methylation levels that fell outside of the lower quartile minus 3×interquartile or upper quartile plus 3×interquartile range were removed.

We excluded probes with a call rate <90%, control probes, and probes that mapped to X/Y chromosomes. Following Zhou et al. [24], we further excluded probes with poor base pairing quality (lower than 40 on 0–60 scale), probes with non-unique 30 bp 3'-subsequence (with cross-hybridizing problems), Infinium II probes with SNPs of global MAF over 1% affecting the extension base, and probes with a SNP in the extension base that causes a color channel switch from the official annotation. We also excluded a subset of probes ( $n = 69$ ) that have shown to be unreliable in a recent comparison of the Illumina 450 K and EPIC BeadChips [25]. At the meta-analysis level, we excluded probes which were available in <50% of the cohorts and <50% of the participants. After QC, 404,017 probes remained at birth and 413,497 probes remained at school-age.

**General psychopathology factor**—Mental health symptoms were assessed when children were aged 6–12 years, depending on the cohort. Parent-rated instruments were used, including (i) the Child Behavior Checklist 6–18 (CBCL/6–18) in DCHS, GENR, GLAKU, HELIX, and INMA, and (ii) the Development and Well-being Assessment (DAWBA) in ALSPAC. Instruments are described in more detail in the Supplementary methods. Whereas a single general factor loaded on all CBCL or DAWBA problem subscales, two specific factors loaded on internalizing (CBCL: anxious/ depressed, withdrawn/ depressed, somatic complaints; DAWBA: generalized anxiety disorder, major depressive disorder, social phobia, separation anxiety, specific phobia) versus externalizing (CBCL: rule-breaking behavior, aggressive behavior, attention problems; DAWBA: attention deficit hyperactivity disorder, oppositional defiant disorder, conduct disorder) subscales. For the CBCL, three subscales (social problems, thought problems, other problems) were indicators of the general factor but were not part of the specific internalizing or externalizing factors. Of note, GLAKU included only two of these three CBCL subscales as the 'other problems' subscale was not available.

The internalizing and externalizing factors were allowed to correlate with each other but not with the general factor. As such, the general factor represents the shared variance among all problem scales that is independent of the more specific internalizing and externalizing factors. Previous research reported negative correlations between the GPF and cognitive outcomes [10, 14, 26]. To support the criterion validity of the GPF, we estimated the correlation between the GPF and child cognition across the cohorts.

**Covariates**—We adjusted for the following potential confounders: child sex, gestational age at birth, child age at the assessment of outcome, maternal age, maternal educational level, prenatal maternal smoking status, cell-type proportions estimated using standard algorithms for DNAm at birth [27] or childhood [28], ancestry (depending on the specific cohort), and technical covariates (e.g., batch) (see Supplementary methods). To test the robustness of findings when using a different method to estimate cell-type proportions, we re-ran the cross-sectional EWAS analyses within the cohort with the largest sample size (HELIX, itself comprised of different participating cohorts) with newly estimated cell proportions using IDOL (Salas et al., 2018 [29]) instead of Houseman’s approach [28].

## Statistical analyses

**General psychopathology factor**—We used confirmatory factor analysis (CFA) to fit a general psychopathology model in the full samples with mental health data available (see Supplementary Information 1). Each cohort ran the CFA according to a predefined script, using the Lavaan statistical package [30] in R (<https://www.r-project.org/>). GPF scores were extracted, winsorized at  $\pm 3SD$ , and standardized.

**Cohort-specific EWAS**—Each cohort ran the EWAS according to a predefined analysis plan, using robust linear regression (rlm; *MASS* R-package) to account for potential heteroscedasticity and non-normality. Cohorts excluded all multiple births and chose one random sibling per non-twin sibling pair.

**Meta-analysis**—The cohort-specific results were meta-analyzed at Erasmus MC Rotterdam. A shadow meta-analysis was conducted independently at the Barcelona Institute for Global Health. We performed an inverse-variance weighted fixed effects approach using R and METAL [31]. Probes were annotated using *meffil* [32]. Genome-wide significance was defined at the Bonferroni-corrected threshold of  $p < 1 \times 10^{-7}$ , and suggestive significance at  $p < 1 \times 10^{-5}$ . We included  $p < 1 \times 10^{-4}$  specifically for pathway and enrichment analyses to allow a sufficient number of genes to be included.

We ran two sensitivity meta-analyses. First, we included only cohorts of predominantly European ancestry to check if the results of the main analysis were influenced by ancestry. Second, we performed leave-one-out meta-analyses for hits showing genome-wide significant associations with GPF to ensure that associations were not driven by a single cohort.

**Differentially methylated regions**—Differentially methylated regions (DMRs) were identified using the dmrff package [33] in R. This method first identifies candidate DMRs by screening the meta-level EWAS results for genomic regions each covered by a sequence of CpG sites with EWAS effects in the same direction, EWAS  $p$ -values  $< 0.05$ , and  $< 500$  bp gaps between consecutive CpG sites. Then, summary statistics are calculated for each candidate DMR within each of the cohorts by meta-analyzing the cohort-level EWAS summary statistics of the CpG sites in the region. Meta-analysis is performed by a variation of inverse weighted fixed effects meta-analysis that accounts for non-independence between



CpG sites. Finally, for each candidate DMR, the summary statistics from each cohort are meta-analyzed to obtain a crosscohort meta-analyzed DMR statistic and *p*-value.

**Follow-up analyses**—Individual probes showing genome-wide or suggestive significance were looked up in the EWAS catalog [34] and EWAS atlas [35] to examine potential associations with exposures and health outcomes based on existing studies. To further characterize potential environmental - as well as genetic – influences on these sites, we used two different tools: 1) a heritability tool quantifying additive genetic influences as opposed to shared and non-shared environmental influences on DNAm, based on data from monozygotic and dizygotic twins; [36] and 2) the GoDMC database (<http://mqtl.db.godmc.org.uk/>) as a more specific tool for identifying genetic influences on DNAm levels via mQTL mapping. GoDMC is a large-scale collaborative effort including 36 cohorts (4 of which participated in this study: INMA, ALSPAC, GENR, GLAKU), based on whole blood from over 27,000 European samples. We characterized cross-tissue correspondence of DNAm using the Blood Brain DNA Methylation Comparison Tool by Hannon et al [37]., the Blood-Brain Epigenetic Concordance (BECon) [38], and the Iowa Methylation Array Graphing for Experimental Comparison of Peripheral tissue & Gray matter (IMAGE-CpG) [39]. To assess whether methylation levels of CpGs were associated with the expression levels of nearby genes in child blood, we consulted the HELIX Expression Quantitative Trait Methylation (eQTM) catalogue (<https://helixomics.isglobal.org/>), generated from samples overlapping with those included in this study (from the HELIX cohort). Finally, chromatin states associated to the most significant CpGs were assessed using ROADMAP blood 15 reference chromatin states (annotation and enrichment analysis conducted using the Enrichment module of the EASIER R package). Genome Browser (UCSC) was used to further explore the genomic context of the identified DMR.

To identify broader pathways and enrichment for molecular functions, we used the gene ontology (GO-biological processes, GO-molecular functions and GO-cellular components), the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, and the Molecular Signature Database (MSigDB) enrichment methods from the missMethyl R package [40], as implemented in the Functional Enrichment module of the EASIER R package [41]. We ran GWAS enrichment analyses for EWAS using the GenomicRanges Package [42], to identify genomic regions of EWAS suggestive hits ( $p < 1 \times 10^{-4}$ ) that overlapped with the 378 genome-wide significant loci previously reported in GWASs on general psychopathology [16], schizophrenia [43], neuroticism [44], ADHD [45] or anxiety [46] (0.5 Mb window centered to the genomic locus indicated in the original studies).

## Results

### General psychopathology factor

All mental health subscales had significant loadings on the general factor across all cohorts, with all loadings >0.30. For full details on the GPF loadings, correlations, and model fit, see Supplementary Table 1. The loadings of the mental health subscales on the specific internalizing and externalizing factors tended to be lower and were less consistent across the cohorts, as were the correlations between these specific factors. In INMA and HELIX,

a model including the correlation between the specific internalizing and externalizing factors did not fit the data well (see Supplementary Information 1). Therefore, in both INMA and HELIX, the specific internalizing and externalizing factors were not allowed to correlate (i.e., completely orthogonal model; see Supplementary Table 1). In line with previous research [5, 7, 14], GPF consistently negatively correlated with child cognition (see Supplementary methods) across the cohorts (mean  $r = -0.12$ , range =  $-0.08$  to  $-0.13$ ).

### Epigenome-wide meta-analysis

Descriptive statistics across the cohorts are shown in Supplementary Table 2. We note that some differences were observed in GPF levels and sociodemographic characteristics between the full cohort samples and analytical subsamples used in the present study (see Supplementary Table 2). These differences varied depending on the specific cohort and variable examined. We prospectively examined associations of DNAm at birth ( $n = 2178$ ) at 404,017 CpG sites with GPF at school-age. There was no evidence of genomic inflation in the cohort-specific EWASs (range  $\lambda = 0.95$ – $1.14$ ), nor in the meta-analysis ( $\lambda = 1.08$ , see also Fig. 1). As can be seen in Fig. 2, no CpG reached genome-wide significance at  $p < 1 \times 10^{-7}$ , with four CpGs showing  $p < 1 \times 10^{-5}$  (see Table 2). For the top hit (cg02084087), annotated to *TNFRSF25* (TNF Receptor Superfamily Member 25), a 10-point increase in percentage methylation was related to a 0.43 SD increase in general psychopathology symptoms ( $p = 5.54 \times 10^{-6}$ ).

In the cross-sectional meta-analysis of DNAm at school-age ( $n = 2,190$ ) at 413,497 sites, one CpG reached genome-wide significance (see Fig. 2). For this CpG probe (cg11945228), mapped to *BRD2* (Bromodomain-containing protein 2 gene), a 10-point increase in percentage methylation was related to a 3.70 SD decrease in general psychopathology symptoms ( $p = 8.58 \times 10^{-8}$ ). Of note, there was a negative association between DNA methylation at this CpG and GPF in all cohorts except for the HELIX-Pakistani cohort. It is not possible based on the present data however to establish whether this may reflect an ancestry-specific association pattern or the influence of other cohort-specific factors (Supplementary Fig. 2). Twenty other CpGs showed  $p < 1 \times 10^{-5}$  (see Table 3). These 21 top hits identified at school-age did not overlap with the ones observed at birth. Furthermore, as shown in Supplementary Table 3, the significant hit identified at school-age did not reach nominal significance ( $p < 0.05$ ) at birth ( $B = 5.28$ ,  $SE = 3.76$ ,  $p = 0.16$ ). Nominally significant probes identified in childhood correlated at  $r = 0.004$ ,  $p = 0.55$  ( $n = 23,764$ ) with respective probes at birth.

**Sensitivity analyses**—Restricting the meta-analysis to children with European ancestry did not change the overall pattern of results for both prospective ( $n=2027$ ) and cross-sectional ( $n=2125$ ) studies, as evidenced by cross-meta-analysis correlations of effect estimates ( $r_{\text{prospective}} = 0.99$ ,  $r_{\text{cross-sectional}} = 0.99$ ) and consistent directions (95% and 96%, respectively) of effect estimates. The top hit identified at schoolage remained genome-wide significant ( $B = -38.02$ ,  $SE = 6.95$ ,  $p = 4.47 \times 10^{-8}$ ).

Leave-one-out meta-analyses showed that the significant top hit identified during childhood (cg11945228) was robust to excluding all individual cohorts, except GENR (for a leave-one-



out plot, see Supplementary Fig. 1). Furthermore, when looking at the cohort-level EWAS results, the cross-sectional association between cg11945228 and GPF was statistically significant in GENR ( $B=-41.87$ ,  $SE=7.78$ ,  $p=7.39 \times 10^{-8}$ ) but not in the other cohorts (all  $p > 0.25$ , see Supplementary Fig. 2 for a forest plot).

Finally, we re-ran the cross-sectional EWAS analyses within HELIX using a different method to estimate cell-type proportions (i.e. based on Salas et al., 2018 [29] instead of Houseman et al., 2012 [28]). We found that the correlation between the regression beta coefficients for all CpGs was very high ( $r = 0.97$ ) (Supplementary Fig. 3), indicating that results are highly concordant when using these two different methods.

### Differentially methylated regions

In the prospective analyses, there was no evidence of DMRs at birth associated with GPF. In the cross-sectional analyses, one DMR at childhood was associated with GPF (estimate = 10166.54,  $SE = 1800.19$ ,  $p = 1.63 \times 10^{-8}$ ). As shown in Supplementary Table 4, this DMR included 6 CpGs mapped to the gene body of the SHC Adaptor Protein 4 gene (*SHC4*) and close to the Transcription Start Site of the EP300 Interacting Inhibitor Of Differentiation 1 gene (*EIDI*) at chromosome 15. From the 6 CpGs, 2 showed positive associations and 4 showed negative associations between methylation level and GPF.

### Follow-up analyses

All probes showing significant or suggestive associations with DNAm had twin heritability estimates available, showing mean additive genetic influences of  $r_{\text{birth}} = 0.16$  and  $r_{\text{childhood}} = 9.44 \times 10^{-2}$  (see Supplementary Table 5).

Of the four suggestive probes identified at birth, three were associated with at least one known methylation quantitative trait locus (mQTL) (see Supplementary Table 5a) and one (cg09437808) showed a concordant DNAm pattern ( $r > 0.28$ ,  $p < 0.01$ ) between blood and several brain regions (the prefrontal cortex, entorhinal cortex, and the superior temporal gyrus) according to Hannon et al. tool (see Supplementary Table 6a). This positive correlation between blood and brain is also reported by IMAGE CPG tool ( $r = 0.35$ ,  $p = 0.31$ ) (Supplementary Table 7a) but not identified in BECon.

The genome-wide significant probe identified during childhood (cg11945228) was unrelated to known mQTLs and showed non-significant correlations between blood and brain DNAm (data only available in one of the three online tools used to assess this concordance) (Supplementary Tables 5b and 6b). Of the 20 suggestive probes identified in childhood, ten were associated with at least one known mQTL (see Supplementary Table 5b) and four (cg22691524, cg09040034, cg25182716, cg18436008) showed a significant correlation between blood and at least one brain region DNAm ( $r > 0.25$ ,  $p < 0.04$ ; see Supplementary Table 6b) according to Hannon et al tool. These sites also showed positive correlations in the BECon (3 out of 4 CpGs; Supplementary Table 8b) and IMAGE CPG tool (3 of the 4 CpGs; Supplementary Table 7b). None of the suggestive probes identified at birth or childhood showed links to an eQTM. According to EWAS Atlas and EWAS Catalogue, methylation levels at these top CpGs seem to be variable and sensitive to age, sex, tissue, or substance exposure (smoking, alcohol, polychlorinated biphenyls), and/or associated to

several traits such as inflammatory and neurological diseases (rheumatoid arthritis, Behcet's disease, myalgic encephalomyelitis, multiple sclerosis, among others) (see Supplementary Table 9).

The six probes in the DMR within the childhood analyses showed low evidence of genetic effects, as indicated by both twin-based estimates (mean variance explained by additive genetic influences  $r = 0.007$ ) and the lack of associations with known mQTLs. One probe, cg05867423, was positively correlated between blood and brain according to data from Hannon et al. tool ( $r = 0.36$ ,  $p < 0.002$ ), with positive correlations also identified in the BECon and IMAGE CPG tools (see Supplementary Tables 6c, 7c, 8c). In addition, cg08455700 showed high correlations ( $r = 0.71$ ) between blood and Brodmann area 20 according to the BECon tool (Supplementary Table 7c). None of these probes was related to eQTLs in blood.

Regarding chromatin states, we found that the genome-wide significant probe (cg11945228) and the DMR found at childhood were associated with active states (active transcription start site (TSS)-proximal promoter state and a transcribed state at the 5' and 3' end of genes showing both promoter and enhancer signatures) (Supplementary Table 10). In fact, an enrichment analysis for chromatin states revealed an overrepresentation of active states associated with zinc finger protein genes (ZNF/Rpts) within the most significant CpGs ( $p < 1 \times 10^{-4}$ ) detected in the prospective meta-analysis (Supplementary Table 11a). In contrast, no consistent enrichment for active states vs repressed states was found based on the most significant CpGs detected in the cross-sectional meta-analysis ( $p < 1 \times 10^{-4}$ ). However, we observed a significant underrepresentation of active transcription start site (TSS)-proximal promoter states (TssAFlnk), and an overrepresentation of actively-transcribed states (Tx, TxWk) together with inactive quiescent states (Quies) (Supplementary Table 11b). Moreover, according to ENCODE data on several cell lines, including different blood cell types, the DMR (chr15:49,170,042-49,170,244, GRCh37/hg19) is enriched by H3K27Ac histone marks and overlaps with DNase hypersensitive areas, which are usually associated to active regulatory elements (Supplementary Fig. 4a). Hence, according to Chromatin interaction data (in situ Hi-C Chromatin Structure from a lymphoblastoid cell line), the genomic elements comprised in the region involving the two genes associated with the DMR seem to strongly interact with each other (Supplementary Fig. 4b).

GO, KEGG, and MSigDB analyses revealed no significantly enriched common biological processes, cellular components, molecular functions or pathways for the genes mapped to the probes at  $p < 1 \times 10^{-4}$  in the meta-analyses at birth ( $n = 56$ ) and during childhood ( $n = 104$ ) (see Supplementary Table 12a-f).

Results of the GWAS enrichment analyses for EWAS are presented in Supplementary Table 13. Of the 56 probes at  $p < 1 \times 10^{-4}$  in the prospective EWAS meta-analysis, six overlapped with genomic loci previously linked to general psychopathology [16], schizophrenia [43], neuroticism [44], ADHD [45] or anxiety [46] based on GWASs. Of the 104 probes at  $p < 1 \times 10^{-4}$  in our cross-sectional EWAS meta-analysis, 13 (12.5%) overlapped with genomic loci previously linked to these psychiatric outcomes. Most notably, this cross-sectional enrichment analysis prioritized cg08514304 (*TAOK2*), which

was among the top 10 suggestive hits identified in our cross-sectional EWAS meta-analysis and showed a consistent direction of effect in all cohorts. Finally, regarding the DMR, SNPs within the region comprising the associated genes *SHC4* and *EID1* have been related with psychiatric disorders such as major depressive disorder [47, 48], bipolar disorder [49], mood and psychotic disorders [50], obsessive compulsive disorder [51], and schizophrenia [52] (Supplementary Table 14) (Supplementary Fig. 4b). Interestingly, both *SHC4* and *EID1* genes are highly expressed in the brain according to GTEx data (Supplementary Fig. 4).

## Discussion

We conducted the largest epigenome-wide meta-analysis of GPF in childhood, using DNAm assessments at two different time points (birth and childhood). The analyses revealed little evidence for probe-specific associations between DNAm in cord blood or peripheral blood and GPF. However, we did identify a significant DMR in childhood, implicating two relevant genes.

On the basis of probe-level genome-wide meta-analyses, we found that lower DNA methylation at cg11945228 at school-age was significantly associated with higher levels of GPF. Cg11945228 is located within the *BRD2* gene, a BET (bromodomains and extra terminal domain) family chromatin adaptor that controls the transcription of a wide range of pro-inflammatory genes [53] and is involved in neural tube closure [54], neurogenesis [55], and neuroinflammation [56]. DNAm of the *BRD2* promoter has been implicated in juvenile myoclonic epilepsy, a common adolescent-onset genetic generalized epilepsy syndrome [57]. However, we advise caution when interpreting this specific site because, despite having low variation attributable to heterogeneity across the cohorts, its genome-wide significant association with GPF seems to be driven by one single cohort.

With regards to genes with probes at suggestive significance at school-age (*WDR20*, *MOV10*, and *TAOK2*), these have previously been linked to neurodevelopmental and psychiatric risk, such as autism spectrum disorder (ASD) and schizophrenia [58–65]. Pleiotropy was supported by our cross-sectional GWAS enrichment analyses for EWAS, showing that *TAOK2* overlapped with genomic loci previously linked to schizophrenia [16, 43], as well as obsessive compulsive disorder and bipolar disorder [16]. However, despite these previously established links with mental health outcomes, annotated genes of our overall top hits identified in the EWAS meta-analyses were not enriched for common biological processes, cellular components, molecular functions, or pathways.

The significant DMR identified at school-age included 6 CpGs mapped to *SHC4* and *EID1* genes, which are highly expressed in the brain. *SHC4* regulates BDNF-induced MAPK activation [66] and *EID1* plays an important role in the central nervous system [67], being involved in cell proliferation in the brain, synaptic plasticity and memory function. Interestingly, genetic variants in these genes have been previously implicated in multiple psychiatric disorders according to several studies (mostly GWAS), including bipolar disorder [49], obsessive-compulsive disorder [51], mood and psychotic disorders [50], schizophrenia [52], or MDD [47, 48, 68] (Supplementary Table 14) (Supplementary Fig. 4b). The fact that the DMR overlaps with active regulatory elements of these genes and shows evidence of

blood-brain concordance for some of the CpGs supports the potential functional relevance of this region. Mechanistic studies will be needed in future to elucidate biological processes underlying the observed link between DNAm in this region and increased risk for multiple psychiatric outcomes. Of interest, despite similar sample sizes and measures (i.e., almost exclusively the CBCL), the top signals were very different between the prospective and cross-sectional EWASs, as evidenced for example by the lack of a correlation between nominally significant sites for these analyses. This low overlap might be due to the temporally dynamic nature of the methylome. DNAm patterns vary substantially over time [69] and can show temporally specific associations with outcomes, including psychiatric symptoms [20]. Unlike an existing EWAS meta-analysis on ADHD symptoms, which showed the strongest signal prospectively at birth compared to childhood [20], we did not detect any significant prospective associations. This is particularly interesting given the use of largely overlapping samples, suggesting that cord blood DNAm may capture risk for specific psychiatric problems (in this case ADHD) rather than a broader liability to psychopathology.

Strengths of this study include the large sample size and the use of DNAm at two different time points (birth and childhood), enabling the assessment of both prospective and cross-sectional associations with GPF. Another important strength is the use of standardized protocols and scripts to fit GPF to the data in a multicohort setting. The GPF scores we analyzed were previously found to associate with a module of co-methylated DNAm probes across the genome [23], suggesting that it is possible to detect biological correlates of GPF using this study's measure. Furthermore, the current study showed that GPF consistently negatively correlated with child cognition across the cohorts as expected based on existing evidence [7], suggesting that it is capturing a similar, valid construct across the cohorts.

However, the current findings should also be interpreted in the context of several limitations. First, given the possibility of residual confounding and reverse causality, the direction of the observed associations cannot be inferred. DNAm might be a marker for unmeasured environmental factors that could affect GPF via independent pathways. Furthermore, children with higher levels of mental health problems may evoke a particular environment [70], which might affect DNAm. Second, our top hits were unrelated to eQTMs. Future research integrating transcriptomic data will be important for assessing the functional relevance of DNAm changes to gene expression in the brain. Third, because DNAm is tissue specific, our observations in peripheral blood may not reflect DNAm levels in other, potentially more relevant, tissues such as the brain. Despite potential sex differences in mental health problems [71] the current study did not examine sexspecificity for power reasons. Further, participating cohorts used different normalization pipelines, which may have contributed to cohort differences and influenced our results. In future, it would be optimal for meta-analytic studies to utilize a standardized processing pipeline across all samples. Furthermore, we found heterogeneity in CFA parameters, particularly for the specific internalizing and externalizing factors (especially in the GLAKU cohort). This precluded us from investigating whether, aside from the GPF, DNA methylation patterns also associate with variance that is unique to these symptom domains – an interesting question for future research. In future, it would be optimal for meta-analytic studies to utilize a standardized processing pipeline across all samples. Finally, the present findings

are based on a predominantly European population and the cohorts are sampled from settings which largely have socialized healthcare, access to mental health services, and different cultural stigma surrounding mental health than other population groups. Future genome-wide studies with larger sample sizes are needed to replicate our findings in other ancestries and in more diverse settings to further characterize DNAm sites associated with GPF.

In summary, this large EWAS meta-analysis identified one probe (Cg11945228) for which lower DNAm in childhood was associated with higher levels of GPF. Furthermore, one DMR in childhood was associated with GPF. This DMR included 6 CpGs mapped to the *SHC4* gene that has previously been implicated in multiple types of psychiatric disorders in adulthood. In contrast, no prospective associations were identified with DNAm patterns at birth. The current findings call for a more integrative approach to the study of GPF, using multiple omics sources, including the genome, epigenome, and transcriptome, to achieve a more comprehensive understanding of its biological underpinnings.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Data Availability

Site-level meta-analytical results will be made publicly available (Supplementary data file) upon acceptance for publication. For access to cohort-level data, requests can be sent directly to individual studies.

## Code Availability

Analytical codes can be requested from authors.

## References

1. Angold A, Costello EJ, Erkanli A. Comorbidity. *J Child Psychol Psychiatry*. 1999; 40: 57–87. [PubMed: 10102726]
2. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005; 62: 617–27. [PubMed: 15939839]
3. Cuffe SP, Visser SN, Holbrook JR, Danielson ML, Geryk LL, Wolraich ML, et al. ADHD and Psychiatric Comorbidity: Functional Outcomes in a School-Based Sample of Children. *J Atten Disord*. 2020; 24: 1345–54. [PubMed: 26610741]
4. Roy-Byrne PP, Stang P, Wittchen HU, Ustun B, Walters EE, Kessler RC. Lifetime panic-depression comorbidity in the National Comorbidity Survey. Association with symptoms, impairment, course and help-seeking. *Br J Psychiatry J Ment Sci*. 2000; 176: 229–35.
5. Lahey BB, Rathouz PJ, Keenan K, Stepp SD, Loeber R, Hipwell AE. Criterion validity of the general factor of psychopathology in a prospective study of girls. *J Child Psychol Psychiatry*. 2015; 56: 415–22. [PubMed: 25052460]
6. Rijlaarsdam J, Cecil CAM, Buil JM, van Lier PAC, Barker ED. Exposure to Bullying and General Psychopathology: A Prospective, Longitudinal Study. *Res Child Adolesc Psychopathol*. 2021; 49: 727–36. [PubMed: 33481128]

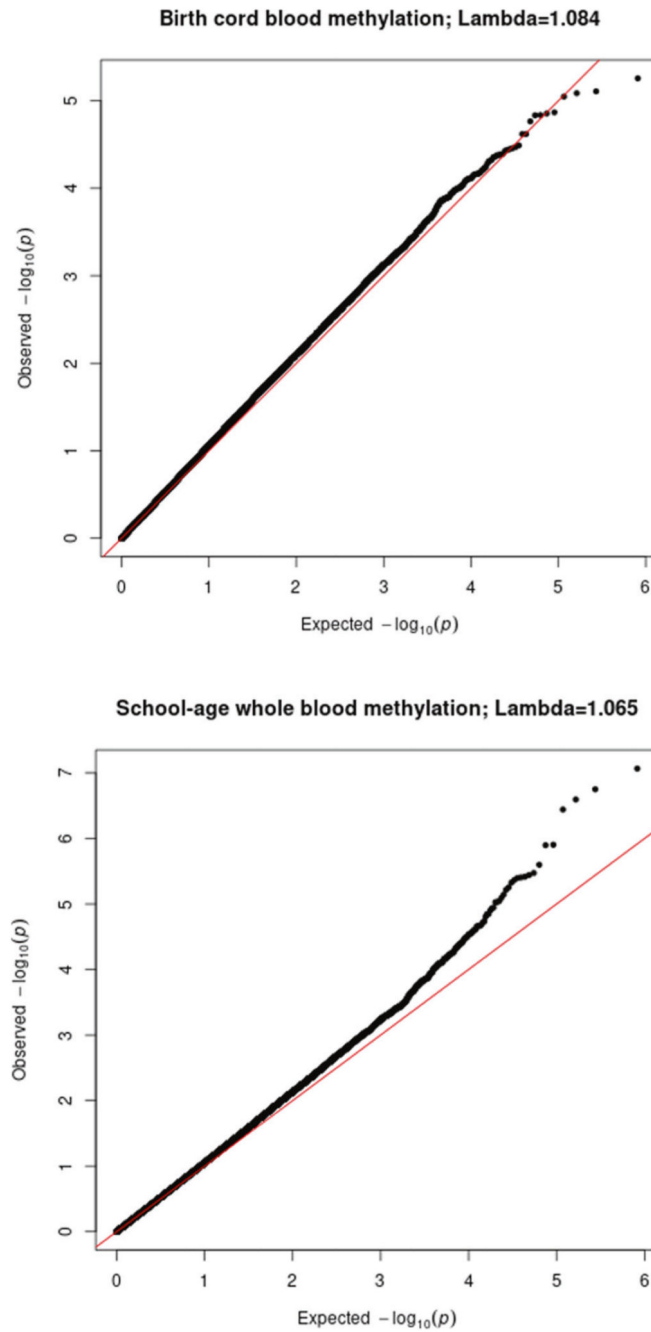
7. Neumann A, Pappa I, Lahey BB, Verhulst FC, Medina-Gomez C, Jaddoe VW, et al. Single Nucleotide Polymorphism Heritability of a General Psychopathology Factor in Children. *J Am Acad Child Adolesc Psychiatry*. 2016; 55: 1038. e4 [PubMed: 27871638]
8. Caspi A, Moffitt TE. All for One and One for All: Mental Disorders in One Dimension. *Am J Psychiatry*. 2018; 175: 831–44. [PubMed: 29621902]
9. Pettersson E, Lahey BB, Larsson H, Lichtenstein P. Criterion Validity and Utility of the General Factor of Psychopathology in Childhood: Predictive Associations With Independently Measured Severe Adverse Mental Health Outcomes in Adolescence. *J Am Acad Child Adolesc Psychiatry*. 2018; 57: 372–83. [PubMed: 29859553]
10. Sallis H, Szekely E, Neumann A, Jolicoeur-Martineau A, van IJendoorn M, Hillegers M, et al. General psychopathology, internalising and externalising in children and functional outcomes in late adolescence. *J Child Psychol Psychiatry*. 2019; 60: 1183–90. [PubMed: 31049953]
11. Brikell I, Larsson H, Lu Y, Pettersson E, Chen Q, Kuja-Halkola R, et al. The contribution of common genetic risk variants for ADHD to a general factor of childhood psychopathology. *Mol Psychiatry*. 2020; 25: 1809–21. [PubMed: 29934545]
12. Riglin L, Thapar AK, Leppert B, Martin J, Richards A, Anney R, et al. Using Genetics to Examine a General Liability to Childhood Psychopathology. *Behav Genet*. 2020; 50: 213–20. [PubMed: 31828458]
13. Brodbeck J, Fassbinder E, Schweiger U, Fehr A, Späth C, Klein JP. Differential associations between patterns of child maltreatment and comorbidity in adult depressed patients. *J Affect Disord*. 2018; 230: 34–41. [PubMed: 29407536]
14. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clin Psychol Sci J Assoc. Psychol Sci*. 2014; 2: 119–37.
15. Campbell M, Jahanshad N, Mufford M, Choi KW, Lee P, Ramesar R, et al. Overlap in genetic risk for cross-disorder vulnerability to mental disorders and genetic risk for altered subcortical brain volumes. *J Affect Disord*. 2021; 282: 740–56. [PubMed: 33601715]
16. Cross-Disorder Group of the Psychiatric Genomics Consortium. plee0@mgh.harvard.edu, Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell*. 2019; 179: 1469. e11 [PubMed: 31835028]
17. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. *Nat Rev Genet*. 2018; 19: 129–47. [PubMed: 29129922]
18. Meaney MJ. Epigenetics and the biological definition of gene x environment interactions. *Child Dev*. 2010; 81: 41–79. [PubMed: 20331654]
19. Cecil CAM, Walton E, Jaffee SR, O'Connor T, Maughan B, Relton CL, et al. Neonatal DNA methylation and early-onset conduct problems: A genome-wide, prospective study. *Dev Psychopathol*. 2018; 30: 383–97. [PubMed: 28595673]
20. Neumann A, Walton E, Alemany S, Cecil C, González JR, Jima DD, et al. Association between DNA methylation and ADHD symptoms from birth to school age: a prospective meta-analysis. *Transl Psychiatry*. 2020; 10: 398. [PubMed: 33184255]
21. Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for colocalization of genetic associations and differential DNA methylation. *Genome Biol*. 2016; 17: 176. [PubMed: 27572077]
22. Zhu Y, Strachan E, Fowler E, Bacus T, Roy-Byrne P, Zhao J. Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: A Monozygotic Discordant Twin Study. *Transl Psychiatry*. 2019; 9: 215. [PubMed: 31477685]
23. Rijlaarsdam J, Barker ED, Caserini C, Koopman-Verhoeff ME, Mulder RH, Felix JF, et al. Genome-wide DNA methylation patterns associated with general psychopathology in children. *J Psychiatr Res*. 2021; 140: 214–20. [PubMed: 34118639]
24. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucl Acids Res*. 2017; 45: e22. [PubMed: 27924034]



25. Solomon O, MacIsaac J, Quach H, Tindula G, Kobor MS, Huen K, et al. Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. *Epigenetics*. 2018; 13: 655–64. [PubMed: 30044683]
26. Patalay P, Fonagy P, Deighton J, Belsky J, Vostanis P, Wolpert M. A general psychopathology factor in early adolescence. *Br J Psychiatry J Ment Sci*. 2015; 207: 15–22.
27. Gervin K, Salas LA, Bakulski KM, van Zelm MC, Koestler DC, Wiencke JK, et al. Systematic evaluation and validation of reference and library selection methods for deconvolution of cord blood DNA methylation data. *Clin Epigenetics*. 2019; 11: 125. [PubMed: 31455416]
28. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinforma*. 2012; 13: 86.
29. Salas LA, Koestler DC, Butler RA, Hansen HM, Wiencke JK, Kelsey KT, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol*. 2018; 19: 64. [PubMed: 29843789]
30. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *J Stat Softw*. 2012; 48: 1–36.
31. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinforma Oxf Engl*. 2010; 26: 2190–1.
32. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinforma Oxf Engl*. 2018; 34: 3983–9.
33. Suderman M, Staley JR, French R, Arathimos R, Simpkin A, Tilling K. dmrff: identifying differentially methylated regions efficiently with power and control. *bioRxiv*. 2018.
34. Battram T, Yousefi P, Crawford G, Prince C, Shekhali Babaei M, Sharp G, et al. The EWAS Catalog: a database of epigenome-wide association studies. *Wellcome Open Res*. 2022; 7: 41. [PubMed: 35592546]
35. Li M, Zou D, Li Z, Gao R, Sang J, Zhang Y, et al. EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res*. 2019; 47: D983–D988. [PubMed: 30364969]
36. Hannon E, Knox O, Sugden K, Burrage J, Wong CCY, Belsky DW, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLoS Genet*. 2018; 14 e1007544 [PubMed: 30091980]
37. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics*. 2015; 10: 1024–32. [PubMed: 26457534]
38. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry*. 2017; 7 e1187 [PubMed: 28763057]
39. Braun P, Han S, Nagahama Y, Gaul L, Heinzman J, Hing B, et al. 28 - IMAGE-CpG: DEVELOPMENT OF A WEB-BASED SEARCH TOOL FOR GENOME-WIDE DNA METHYLATION CORRELATION BETWEEN LIVE HUMAN BRAIN AND PERIPHERAL TISSUES WITHIN INDIVIDUALS. *Eur Neuropsychopharmacol*. 2019; 29: S796.
40. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinforma Oxf Engl*. 2016; 32: 286–8.
41. isglobal-brge/EASIER: EwAS: quality control, meta-analyses and EnRichment version 0128 from GitHub. Accessed May 2022 <https://rdrr.io/github/isglobal-brge/EASIER/>
42. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol*. 2013; 9 e1003118 [PubMed: 23950696]
43. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014; 511: 421–7. [PubMed: 25056061]
44. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J, et al. Metaanalysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet*. 2018; 50: 920–7. [PubMed: 29942085]

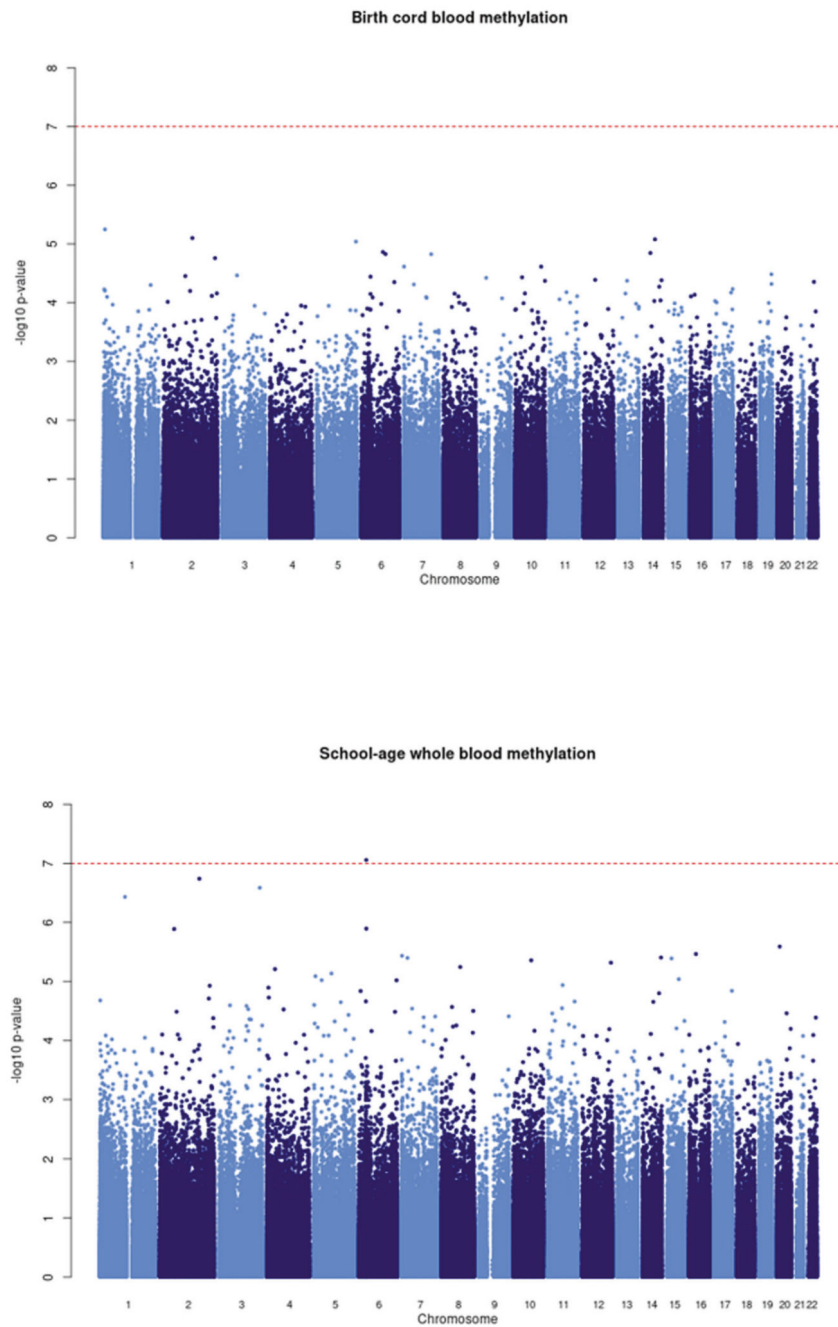
45. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet.* 2019; 51: 63–75. [PubMed: 30478444]
46. Levey DF, Gelernter J, Polimanti R, Zhou H, Cheng Z, Aslan M, et al. Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *Am J Psychiatry.* 2020; 177: 223–32. [PubMed: 31906708]
47. Aragam N, Wang K-S, Pan Y. Genome-wide association analysis of gender differences in major depressive disorder in the Netherlands NESDA and NTR population-based samples. *J Affect Disord.* 2011; 133: 516–21. [PubMed: 21621269]
48. Howard DM, Hall LS, Hafferty JD, Zeng Y, Adams MJ, Clarke T-K, et al. Genomewide haplotype-based association analysis of major depressive disorder in Generation Scotland and UK Biobank. *Transl Psychiatry.* 2017; 7: 1263. [PubMed: 29187746]
49. Budde M, Friedrichs S, Alliey-Rodriguez N, Ament S, Badner JA, Berrettini WH, et al. Efficient region-based test strategy uncovers genetic risk factors for functional outcome in bipolar disorder. *Eur Neuropsychopharmacol J Eur Coll Neuropsychopharmacol.* 2019; 29: 156–70.
50. Blokland GAM, Grove J, Chen C-Y, Cotsapas C, Tobet S, Handa R, et al. Sex-Dependent Shared and Nonshared Genetic Architecture Across Mood and Psychotic Disorders. *Biol Psychiatry.* 2022; 91: 102–17. [PubMed: 34099189]
51. Boraska V, Davis OSP, Cherkas LF, Helder SG, Harris J, Krug I, et al. Genome-wide association analysis of eating disorder-related symptoms, behaviors, and personality traits. *Am J Med Genet Part B Neuropsychiatr Genet Publ Int Soc Psychiatr Genet.* 2012; 159B: 803–11.
52. Goes FS, McGrath J, Avramopoulos D, Wolyniec P, Pirooznia M, Ruczinski I, et al. Genome-wide association study of schizophrenia in Ashkenazi Jews. *Am J Med Genet Part B Neuropsychiatr Genet Publ Int Soc Psychiatr Genet.* 2015; 168: 649–59.
53. Wang N, Wu R, Tang D, Kang R. The BET family in immunity and disease. *Signal Transduct Target Ther.* 2021; 6: 23. [PubMed: 33462181]
54. Gyuris A, Donovan DJ, Seymour KA, Lovasco LA, Smilowitz NR, Halperin ALP, et al. The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis. *Biochim Biophys Acta.* 2009; 1789: 413–21. [PubMed: 19362612]
55. Garcia-Gutierrez P, Juarez-Vicente F, Wolgemuth DJ, Garcia-Dominguez M. Pleiotrophin antagonizes Brd2 during neuronal differentiation. *J Cell Sci.* 2014; 127: 2554–64. [PubMed: 24695857]
56. DeMars KM, Yang C, Candelario-Jalil E. Neuroprotective effects of targeting BET proteins for degradation with dBET1 in aged mice subjected to ischemic stroke. *Neurochem Int.* 2019; 127: 94–102. [PubMed: 30872008]
57. Pathak S, Miller J, Morris EC, Stewart WCL, Greenberg DA. DNA methylation of the BRD2 promoter is associated with juvenile myoclonic epilepsy in Caucasians. *Epilepsia.* 2018; 59: 1011–9. [PubMed: 29608786]
58. Wockner LF, Noble EP, Lawford BR, Young RM, Morris CP, Whitehall VLJ, et al. Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl Psychiatry.* 2014; 4: e339. [PubMed: 24399042]
59. McKinney B, Ding Y, Lewis DA, Sweet RA. DNA methylation as a putative mechanism for reduced dendritic spine density in the superior temporal gyrus of subjects with schizophrenia. *Transl Psychiatry.* 2017; 7 e1032 [PubMed: 28195572]
60. Skariah G, Seimetz J, Norsworthy M, Lannom MC, Kenny PJ, Elrakhawy M, et al. Mov10 suppresses retroelements and regulates neuronal development and function in the developing brain. *BMC Biol.* 2017; 15: 54. [PubMed: 28662698]
61. Vissers LELM, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. *Nat Rev Genet.* 2016; 17: 9–18. [PubMed: 26503795]
62. Hanson E, Bernier R, Porche K, Jackson FI, Goin-Kochel RP, Snyder LG, et al. The cognitive and behavioral phenotype of the 16p11.2 deletion in a clinically ascertained population. *Biol Psychiatry.* 2015; 77: 785–93. [PubMed: 25064419]

63. Zufferey F, Sherr EH, Beckmann ND, Hanson E, Maillard AM, Hippolyte L, et al. A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *J Med Genet.* 2012; 49: 660–8. [PubMed: 23054248]
64. Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med.* 2008; 358: 667–75. [PubMed: 18184952]
65. Chang H, Li L, Li M, Xiao X. Rare and common variants at 16p11.2 are associated with schizophrenia. *Schizophr Res.* 2017; 184: 105–8. [PubMed: 27889382]
66. You Y, Li W, Gong Y, Yin B, Qiang B, Yuan J, et al. ShcD interacts with TrkB via its PTB and SH2 domains and regulates BDNF-induced MAPK activation. *BMB Rep.* 2010; 43: 485–90. [PubMed: 20663410]
67. Liu R, Lei JX, Luo C, Lan X, Chi L, Deng P, et al. Increased EID1 nuclear translocation impairs synaptic plasticity and memory function associated with pathogenesis of Alzheimer's disease. *Neurobiol Dis.* 2012; 45: 902–12. [PubMed: 22186421]
68. Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T, et al. Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry.* 2009; 14: 359–75. [PubMed: 19065144]
69. Mulder RH, Neumann A, Cecil CAM, Walton E, Houtepen LC, Simpkin AJ, et al. Epigenome-wide change and variation in DNA methylation in childhood: trajectories from birth to late adolescence. *Hum Mol Genet.* 2021; 30: 119–34. [PubMed: 33450751]
70. Singham T, Viding E, Schoeler T, Arseneault L, Ronald A, Cecil CM, et al. Concurrent and Longitudinal Contribution of Exposure to Bullying in Childhood to Mental Health: The Role of Vulnerability and Resilience. *JAMA Psychiatry.* 2017; 74: 1112–9. [PubMed: 28979965]
71. Bale TL, Epperson CN. Sex as a Biological Variable: Who, What, When, Why, and How. *Neuropsychopharmacol Publ Am Coll Neuropsychopharmacol.* 2017; 42: 386–96.



**Fig. 1. Quantile-quantile plot of the meta-analytic associations of DNA methylation at birth and DNA methylation at school-age with general psychopathology.**

The diagonal line represents the distribution of the expected p-values under the null. Points above the diagonal refer to p-values that are lower than expected.



**Fig. 2.** Manhattan plot of  $-\log_{10}$  p-values versus CpG position (base pair and chromosome) showing meta-analytic associations of DNA methylation at birth and DNA methylation at school-age with general psychopathology.

The red line indicates genome-wide significance ( $p < 1 \times 10^{-7}$ ).

**Table 1**  
**Cohort characteristics.**

Cohort	Ancestry/ethnicity	n	DNAm age	GPF age	GPF instrument	Effect sizes			$\lambda$
						1st Qu	Median	3rd Qu	
<i>Birth EWAS</i>									
ALSPAC	European	643	0	10	DAWBA	-1.10	0.02	1.18	0.98
DCHS	Black African, mixed	151	0	6	CBCL	-4.35	-0.77	2.75	0.95
Generation R	European	1092	0	10	CBCL	-0.77	0.08	1.00	1.01
INMA	European	292	0	9	CBCL	-1.53	0.39	3.03	1.14
Meta		2178				-0.51	0.07	0.77	1.08
<i>Childhood EWAS</i>									
ALSPAC	European	697	7	10	DAWBA	-1.26	0.02	1.28	1.00
Generation R	European	434	10	10	CBCL	-1.51	0.08	1.84	1.03
GLAKU	European	215	12	12	CBCL	-1.66	0.31	2.24	0.91
HELIX	European	779	8	8	CBCL	-1.36	0.32	2.22	0.99
HELIX	Pakistani	65	7	7	CBCL	-4.47	2.52	10.95	1.20
Meta		2190				-0.54	0.18	1.00	1.07



**Table 2**  
**DNA methylation at birth and general psychopathology: meta-analytic associations with  $p < 1 \times 10^{-5}$**

CpG	Gene	Chr	Position	n	B	SE	p	Direction	I <sup>2</sup>	Heterogeneity p-value
cg02084087	<i>TNFRSF25</i>	chr1	6526049	2175	4.31	0.95	$5.54 \times 10^{-6}$	++++	46.2	0.13
cg11777523	<i>GPR148</i>	chr2	131485418	2019	9.26	2.07	$7.80 \times 10^{-6}$	+?++	23.6	0.27
cg14358879	<i>SLC8A3</i>	chr14	70655920	2174	-7.82	1.75	$8.20 \times 10^{-6}$	----	33.7	0.21
cg09437808	-	chr5	176107069	2177	2.46	0.55	$8.96 \times 10^{-6}$	++++	0	0.47

*Chr* chromosome, *n* number of participants, *SE* standard error, *Direction* direction of the effect per study (ALSPAC, DCHS, GENR, INMA) in alphabetical order (+ = positive direction, - = negative direction, ? = not present); *I*<sup>2</sup> heterogeneity statistic reflecting the variation attributable to heterogeneity across studies (high values suggest high heterogeneity)

Effect estimates (B) represent the SD increase in GPF for each increase of 100% in DNAm.

**Table 3**  
**DNA methylation at school-age and general psychopathology: meta-analytic associations**  
**with  $p < 1 \times 10^{-5}$**

CpG	Gene	Chr	Position	n	B	SE	p	Direction	I <sup>2</sup>	Heterogeneity p-value
cg11945228	<i>BRD2</i>	chr6	32940368	2173	-37.00	6.91	$8.58 \times 10^{-8}$	----+	0	0.53
cg18862005	-	chr2	177940863	1972	3.78	0.72	$1.78 \times 10^{-7}$	++?++	0	0.64
cg22691524	-	chr3	185300576	2180	7.84	1.52	$2.54 \times 10^{-7}$	+++++	0	0.47
cg00719568	<i>MOV10</i>	chr1	113239645	2184	4.20	0.83	$3.62 \times 10^{-7}$	+++++	0	0.67
cg09040034	<i>KIFC1</i>	chr6	33362567	1966	4.89	1.01	$1.25 \times 10^{-6}$	++?+-	40	0.17
cg24514921	<i>VPS54</i>	chr2	64246311	2184	12.17	2.51	$1.27 \times 10^{-6}$	++++-	0	0.45
cg25182716	-	chr20	13622875	2178	6.42	1.36	$2.52 \times 10^{-6}$	+++++	0	0.76
cg08514304	<i>TAOK2</i>	chr16	29994437	2175	6.88	1.48	$3.36 \times 10^{-6}$	+++++	0	0.59
cg00470277	-	chr7	2669915	2185	4.55	0.98	$3.62 \times 10^{-6}$	+++++	0	0.99
cg26353764	<i>WDR20</i>	chr14	102660055	2187	6.82	1.48	$3.86 \times 10^{-6}$	+++++	43.7	0.13
cg27009703	<i>HOXA9</i>	chr7	27204894	2121	-8.58	1.86	$3.93 \times 10^{-6}$	-----	0	0.52
cg12492087	<i>ZFP106</i>	chr15	42749885	2178	6.39	1.39	$4.01 \times 10^{-6}$	+++++	0	0.92
cg18436008	-	chr10	80535327	2176	5.48	1.19	$4.31 \times 10^{-6}$	+++++	0	0.56
cg17281031	-	chr12	128223216	2186	4.50	0.98	$4.73 \times 10^{-6}$	+++++	14.6	0.32
cg08327106	<i>RALYL</i>	chr8	85094842	2175	8.60	1.89	$5.60 \times 10^{-6}$	+++++	0	0.66
cg11236841	-	chr4	35567978	2175	4.67	1.03	$6.08 \times 10^{-6}$	+++++	54.3	0.07
cg21525176	<i>LHFPL2</i>	chr5	77906752	2182	6.09	1.36	$7.19 \times 10^{-6}$	+++++	59.2	0.04
cg26420013	<i>NSUN2;SRD5A1</i>	chr5	6632020	2189	2.44	0.55	$8.03 \times 10^{-6}$	+++++	56.1	0.06
cg17087232	<i>MAN2C1</i>	chr15	75651821	2187	3.45	0.78	$9.00 \times 10^{-6}$	+++++	0	0.98
cg16360861	<i>RAI14</i>	chr5	34684597	2175	5.12	1.16	$9.36 \times 10^{-6}$	+++++	0	0.53
cg00737264	<i>SMOC2</i>	chr6	169049498	2188	1.88	0.42	$9.40 \times 10^{-6}$	+++++	41.3	0.15

*Chr* chromosome, *n* number of participants, *SE* standard error, *Direction* direction of the effect per study (ALSPAC, GENR, GLAKU, HELIX, HELIX-Pakistani) in alphabetical order (+ = positive direction, - = negative direction, ? = not present); *I*<sup>2</sup> heterogeneity statistic reflecting the variation attributable to heterogeneity across studies (high values suggest high heterogeneity)

Effect estimates (B) represent the SD increase in GPF for each increase of 100% in DNAm.