

ARTICLE

Open Access

# Genomic influences on self-reported childhood maltreatment

Shareefa Dalvie<sup>1</sup>, Adam X. Maihofer<sup>2,3,4</sup>, Jonathan R. I. Coleman<sup>5,6</sup>, Bekh Bradley<sup>7,8</sup>, Gerome Breen<sup>5,6</sup>, Leslie A. Brick<sup>9</sup>, Chia-Yen Chen<sup>10,11,12</sup>, Karmel W. Choi<sup>12,13,14</sup>, Laramie E. Duncan<sup>15</sup>, Guia Guffanti<sup>16,17</sup>, Magali Haas<sup>18</sup>, Supriya Harnal<sup>12</sup>, Israel Liberzon<sup>19</sup>, Nicole R. Nugent<sup>9,20</sup>, Allison C. Provost<sup>18</sup>, Kerry J. Ressler<sup>8,16,17</sup>, Katy Torres<sup>2,3,4</sup>, Ananda B. Amstadter<sup>21</sup>, S. Bryn Austin<sup>22,23,24,25</sup>, Dewleen G. Baker<sup>2,3,26</sup>, Elizabeth A. Bolger<sup>17,16</sup>, Richard A. Bryant<sup>27</sup>, Joseph R. Calabrese<sup>28</sup>, Douglas L. Delahanty<sup>29,30</sup>, Lindsay A. Farrer<sup>31</sup>, Norah C. Feeny<sup>32</sup>, Janine D. Flory<sup>33</sup>, David Forbes<sup>34</sup>, Sandro Galea<sup>35</sup>, Aarti Gautam<sup>36</sup>, Joel Gelernter<sup>37,38,39</sup>, Rasha Hammamieh<sup>40</sup>, Marti Jett<sup>40</sup>, Angela G. Junglen<sup>29</sup>, Milissa L. Kaufman<sup>16,17</sup>, Ronald C. Kessler<sup>41</sup>, Alaptagin Khan<sup>17,41</sup>, Henry R. Kranzler<sup>42,43</sup>, Lauren A. M. Lebois<sup>16,17</sup>, Charles Marmar<sup>44</sup>, Matig R. Mavissakalian<sup>28</sup>, Alexander McFarlane<sup>45</sup>, Meaghan O' Donnell<sup>34</sup>, Holly K. Orcutt<sup>46</sup>, Robert H. Pietrzak<sup>47,48</sup>, Victoria B. Risbrough<sup>2,3,4</sup>, Andrea L. Roberts<sup>49</sup>, Alex O. Rothbaum<sup>32</sup>, Peter Roy-Byrne<sup>50</sup>, Ken Ruggiero<sup>51</sup>, Antonia V. Seligowski<sup>17,16</sup>, Christina M. Sheerin<sup>21</sup>, Derrick Silove<sup>52</sup>, Jordan W. Smoller<sup>10,12,14</sup>, Murray B. Stein<sup>2,26,53</sup>, Martin H. Teicher<sup>16,17</sup>, Robert J. Ursano<sup>54</sup>, Miranda Van Hooff<sup>45</sup>, Sherry Winternitz<sup>16,17</sup>, Jonathan D. Wolff<sup>17</sup>, Rachel Yehuda<sup>33,55</sup>, Hongyu Zhao<sup>56</sup>, Lori A. Zoellner<sup>57</sup>, Dan J. Stein<sup>1</sup>, Karestan C. Koenen<sup>11,12,58</sup> and Caroline M. Nievergelt<sup>2,3,4</sup>

## Abstract

Childhood maltreatment is highly prevalent and serves as a risk factor for mental and physical disorders. Self-reported childhood maltreatment appears heritable, but the specific genetic influences on this phenotype are largely unknown. The aims of this study were to (1) identify genetic variation associated with self-reported childhood maltreatment, (2) estimate SNP-based heritability ( $h^2_{\text{SNP}}$ ), (3) assess predictive value of polygenic risk scores (PRS) for childhood maltreatment, and (4) quantify genetic overlap of childhood maltreatment with mental and physical health-related phenotypes, and condition the top hits from our analyses when such overlap is present. Genome-wide association analysis for childhood maltreatment was undertaken, using a discovery sample from the UK Biobank (UKBB) ( $n = 124,000$ ) and a replication sample from the Psychiatric Genomics Consortium-posttraumatic stress disorder group (PGC-PTSD) ( $n = 26,290$ ).  $h^2_{\text{SNP}}$  for childhood maltreatment and genetic correlations with mental/physical health traits were calculated using linkage disequilibrium score regression. PRS was calculated using PRSice and mtCOJO was used to perform conditional analysis. Two genome-wide significant loci associated with childhood maltreatment ( $rs142346759$ ,  $p = 4.35 \times 10^{-8}$ , *FOXP1*;  $rs10262462$ ,  $p = 3.24 \times 10^{-8}$ , *FOXP2*) were identified in the discovery dataset but were not replicated in PGC-PTSD.  $h^2_{\text{SNP}}$  for childhood maltreatment was ~6% and the PRS derived from the UKBB was significantly predictive of childhood maltreatment in PGC-PTSD ( $r^2 = 0.0025$ ;  $p = 1.8 \times 10^{-15}$ ). The most significant genetic correlation of childhood maltreatment was with depressive symptoms ( $r_g = 0.70$ ,  $p = 4.65 \times 10^{-40}$ ), although we show evidence that our top hits may be specific to childhood maltreatment. This is the first large-scale genetic study to identify specific variants associated with self-reported childhood maltreatment. Speculatively, *FOXP* genes might influence externalizing traits and so be relevant to childhood maltreatment. Alternatively, these variants may be associated with a greater likelihood of reporting maltreatment. A clearer understanding of the genetic relationships of childhood maltreatment, including particular abuse subtypes, with a range of phenotypes, may ultimately be useful in developing targeted treatment and prevention strategies.

Correspondence: Shareefa Dalvie (s.dalvie@uct.ac.za)

<sup>1</sup>SA MRC Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa

<sup>2</sup>Department of Psychiatry, University of California San Diego, La Jolla, CA, USA  
Full list of author information is available at the end of the article.

© The Author(s) 2020



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

## Introduction

The lifetime prevalence of childhood physical, sexual, and emotional maltreatment ranges from 8% to 36%<sup>1</sup>. In addition to being highly prevalent, such childhood abuse is associated with the development of mental disorders, including depression<sup>2,3</sup>, and physical ill health, including non-communicable diseases<sup>4,5</sup>. Although these associations are now well established, estimates of effect size vary considerably across epidemiological studies, likely reflecting methodological challenges, including uncertainty about how best to assess childhood maltreatment<sup>6</sup>.

A twin-based study found that retrospective reports of childhood maltreatment has a heritability of 6%<sup>7</sup>. Although the idea that childhood maltreatment is heritable may seem counter-intuitive, work on behavior genetics has long documented the heritability of many exposures perceived as environmental. Such heritability is referred to as gene–environment correlation (rGE), and three potential rGE mechanisms to explain the heritability of childhood maltreatment may be posited. First, a “passive” rGE: parental genes affecting parental behavior may influence the childhood environment (e.g. aggressive parents may be more likely to physically punish their children<sup>8</sup>). Second, an “active” rGE: individuals with genetic variants associated with certain behavioral phenotypes may be more at risk of selecting or creating adverse situations (e.g. risk-taking is heritable and children who are high in risk-taking may be exposed to more trauma)<sup>9,10</sup>. Third, an “evocative” rGE: genetic variation may influence child behavior, which in turn is associated with responses to the child (e.g. genetic factors may influence infant “difficultness”, which in turn is associated with maternal hostile-reactive behavior that is correlated with child abuse<sup>11,12</sup>). The latter two rGEs are sometimes collectively referred to as non-passive correlations<sup>7</sup>.

While a number of key risk factors for childhood maltreatment, including child behavioral characteristics and parental mental health, have been investigated<sup>6</sup>, studies have seldom focused on associated genetic variation. The few genetic association studies of childhood maltreatment have only considered variants in candidate genes<sup>13</sup> and have had insufficient power to detect the small polygenic effect sizes typically associated with behavioral phenotypes<sup>14</sup>. Also, there are no studies of the genetic overlap of childhood maltreatment with mental and physical health-related traits, using genome-wide single nucleotide polymorphism (SNP) data. Knowledge of specific genetic variation for childhood maltreatment, the heritability of this phenotype, the polygenic risk, and the genetic overlap with other traits may be useful in informing our understanding of the risk factors, the etiology, and the outcomes of childhood maltreatment. This, in turn, may have implications for the design of prevention and treatment programs for adverse health outcomes. For example,

environmental exposures that play a causal role in impacting health outcomes are likely to mediate any observed associations between genetic variants and that health outcome (e.g. early loss of a parent may lead to depression, with such loss then mediating the association between heritability of early parental loss and depression). Thus, preventative strategies would focus on decreasing the risk conferred by the environmental exposure without needing to specifically consider the genetic influences on the health outcome<sup>9</sup> (e.g. development of programs for children who have experienced early loss).

The PGC-PTSD has collaborated to obtain access to well-powered genetic studies of trauma and PTSD that have allowed a number of key genetic questions in this field to be investigated<sup>15–17</sup>, providing a unique opportunity to address knowledge gaps in the area of childhood maltreatment. This study aims to: (1) identify genetic variants associated with childhood maltreatment using a genome-wide association study (GWAS) design, (2) quantify the heritability of childhood maltreatment using SNP-based methods, (3) assess the predictive value of polygenic risk scores (PRS) for childhood maltreatment, and (4) assess the degree of genetic overlap of childhood maltreatment with mental and physical health-related phenotypes, and condition the top genome-wide hits from our analyses when such overlap is present.

## Materials and methods

### Participating studies

Nineteen studies, comprising subjects of European ancestry only, were used in this analysis. The discovery dataset consisted of 124,711 individuals with available childhood maltreatment data from the UK Biobank (UKBB)<sup>18</sup>, and the replication sample comprised 26,290 individuals—a subset of the PGC-PTSD Freeze 1.5 dataset (PGC1.5)<sup>17</sup>. The details of these studies, including the demographics and instruments used to assess maltreatment can be found in Supplementary Table 1. We have complied with relevant ethical regulations for work with human subjects. All subjects provided written informed consent and studies were approved by the relevant institutional review boards and the UCSD IRB (protocol #16097x).

### Phenotype harmonization

For the childhood maltreatment phenotype, Childhood Trauma Questionnaire (CTQ) scores on physical, sexual, and emotional abuse<sup>19</sup> were obtained from the participating studies. From this, an overall childhood maltreatment count score of 0–3 was constructed, based on a count of the three abuse categories listed above. An individual was considered to have endorsed a childhood abuse category if they scored in the moderate to extreme range for that particular category, per established cut-

offs<sup>20</sup> (Supplementary Table 2). If CTQ data were not available, the event assessment during childhood (occurring before 18 years of age) that was most validated for that particular study was obtained, providing a count of the total number of different categories of reported childhood events (e.g. physical, sexual, or severe emotional abuse) along with the range of possible scores for the measure. The reported maltreatment exposure from the UKBB dataset comprised a score of three items where participants were asked whether they were (i) “physically abused by family as a child”, (ii) “sexually molested as a child”, and whether they (iii) “felt hated by family member as a child”. The childhood maltreatment count score, whether it was generated from the CTQ or another instrument, was used as the main outcome measure in the association analysis. The range and mean of maltreatment count scores for each study can be seen in Supplementary Table 1.

#### Global ancestry determination, genotyping quality control, and imputation

Study participants from the PGC-PTSD were genotyped with a number of different arrays (Supplementary Table 1). Genotype data were quality controlled and processed using the standard PGC pipeline, Ricopili-MANC (<https://sites.google.com/a/broadinstitute.org/ricopili/> and <https://github.com/orgs/Nealelab/teams/ricopili>) as part of the PGC-PTSD Freeze 2 data analysis<sup>17,21</sup>. This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. A detailed outline of these methods can be found in ref. <sup>17</sup>. Briefly, ancestry was determined with pre-QC genotypes using a SNPweights panel of 10,000 ancestry informative markers from a reference panel comprising 2911 subjects from 71 diverse populations and six continental groups ([https://github.com/nievergeltlab/global\\_ancestry](https://github.com/nievergeltlab/global_ancestry)). Samples with estimated > 90% European ancestry were classified as European. Samples were excluded if they had call rates < 98%, deviated from the expected inbreeding coefficient ( $f_{het} < -0.2$  or  $> 0.2$ ), or had a sex discrepancy between reported and genotypic sex (based on inbreeding coefficients calculated from SNPs on the X chromosome). Markers were excluded if they had call rates < 98%, > 2% difference in missing genotypes between PTSD cases and controls, or were monomorphic. Markers with a Hardy–Weinberg equilibrium (HWE)  $p < 1 \times 10^{-6}$  in controls were excluded from all subjects. Principal components (PCs) were calculated using the smartPCA algorithm in EIGENSTRAT<sup>22</sup>. Pre-phasing and phasing was performed using SHAPEIT2 v2. r837<sup>23</sup>. Imputation was performed with IMPUTE2 v2.2.2<sup>24</sup> using the 1000 Genomes (1000G) phase 3 data<sup>25</sup> as the reference.

Details regarding the QC, imputation, and ancestry determination of the UKBB dataset can be found in ref. <sup>26</sup>.

Briefly, study participants were genotyped with two custom genotyping arrays (with ~800,000 markers). A two-stage imputation was performed using the Haplotype Reference Consortium (HRC)<sup>27</sup> and the UK10K<sup>28</sup> as the reference panels. Variants in the UKBB dataset were filtered to include only those with a minor allele frequency (MAF) of > 1% and an INFO score of > 0.4. Related individuals (third degree and closer) and those with a genotyping call rate < 98% were excluded. Ancestry was determined by 4-means clustering on the first two PCs provided by the UKBB<sup>29</sup>. Additional principal component analysis was conducted on the European-only data subset using flashpca2<sup>30</sup>.

#### Main GWAS

GWAS analysis was conducted separately for each study. Best-guess genotypes were tested for association to self-reported childhood maltreatment using an ordinal logistic regression model with age, sex, and the first five PCs included as covariates. Variants with a MAF < 0.5% and a genotyping rate < 98% were excluded, for all studies except the UKBB. These analyses were implemented in PLINK 1.9<sup>31</sup> using the plug-in *Rserve*. To ensure computational efficiency, linear regression models were run for 4 of the larger contributing studies (NSS1; NSS2; PPDS; and UKBB,  $N = 143,392$  subjects)<sup>17</sup>. For the NSS1; NSS2; and PPDS studies, age, sex, and 5 PCs were included as covariates in the regression model. For the UKBB dataset, the regression analysis was implemented in BGenie v1.2<sup>32</sup> with age, sex, 6 PCs, batch, and site included as covariates. All tests performed were two-sided.

#### Meta-analysis

As both linear and ordinal logistic models were implemented in the GWASs, which resulted in different effect statistics, fixed effects meta-analysis was conducted across studies using  $p$ -values and direction of effect, weighted according to the effective sample size as the analysis scheme, in METAL (v. March 25 2011)<sup>33</sup>. Effective sample sizes ( $N_{eff}$ ) for ordinal logistic regressions were calculated as  $N_{eff} = \text{harmonic mean} * n$  levels of childhood maltreatment, and for linear regressions as  $N_{eff} = ((1 - \text{probability of having a zero score}) \times \text{mean of nonzero data})$ <sup>34</sup>. Heterogeneity across datasets was tested using the Cochran’s Q-test for heterogeneity, also implemented in METAL. Only variants with an INFO score of > 0.8 and a conservative MAF of > 5% were included in the meta-analysis, except where otherwise indicated in the results. Forest plots were generated for genome-wide significant hits using the R package *meta*<sup>35</sup>.

#### Functional mapping and annotation

Genome-wide significant hits identified from the GWAS meta-analysis were annotated using the web-

based tool **F**unctional Mapping and Annotation (**F**UMA) v1.3.4c<sup>36</sup>. Default settings were used and annotations were based on the human genome assembly GRCh37 (hg19). The *SNP2GENE* module was used to identify genomic risk loci and these were mapped to protein-coding genes within a 10 kb window. An  $r^2$  of  $\geq 0.6$  was used to identify variants in LD with lead SNPs. The 1000G European Phase 3 was used as the reference dataset. Variants were functionally annotated using ANNOVAR, combined dependent depletion (CADD), RegulomeDB (RDB), and chromatin states (only tissues/cells from brains were included). The NHGRI-EBI GWAS catalog was used to determine any previous associations with the identified risk variants. The GTEx v7 brain tissue, RNAseq data from the CommonMind Consortium and the BRAINEAC database were used to perform eQTL mapping for significant SNP–gene pairs (FDR  $q < 0.05$ ).

A gene-based analysis was performed within FUMA using MAGMA whereby SNPs were mapped to 18,989 protein-coding genes. Genome-wide significance was set at a Bonferroni-corrected threshold  $p < 2.63 \times 10^{-6}$ . In addition, gene-based test statistics were used to determine whether specific biological pathways are associated with childhood maltreatment. This was performed for 10,678 curated gene sets and GO terms obtained from MsigDB, using MAGMA. The significance threshold was set at a Bonferroni-corrected threshold of  $p = 4.68 \times 10^{-6}$  (0.05/10,678).

#### Heritability estimation

Linkage disequilibrium score regression (LDSR) is a technique for quantifying polygenicity and confounding, such as population stratification, in GWAS summary statistics<sup>37</sup>. This is accomplished by evaluating the relationship between linkage disequilibrium (LD) scores (the average squared correlation of a SNP with all neighboring SNPs) and SNP test statistics. Using this approach, the LDSR intercept was used to estimate the proportion of inflation in test statistics due to polygenic signal (rather than inflation due to population stratification and cryptic relatedness), with the Eq. (1)—(LDSR intercept–1)/(mean observed chi-square–1)<sup>17</sup>. Using GWAS summary statistics, SNP-based heritability ( $h^2_{\text{SNP}}$ ) was calculated, which is one of the applications of LDSR.

#### Polygenic risk scoring

Using PRSice v2.1.3.beta<sup>38</sup>, PRS were calculated in target samples (PGC1.5) based on SNP effect sizes from childhood maltreatment GWAS in non-overlapping discovery/training samples (UKBB). Multiple  $p$ -value thresholds ( $P_T$ ) (0.001, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 1) were generated using the best guess genotype data of target samples. Variants with a MAF  $< 5\%$  were excluded from

the discovery dataset. As a default in PRSice, LD pruning was performed whereby variants were pruned if they were nearby (within 250 kb) and in LD ( $r^2 > 0.1$ ) with the leading variant (lowest  $p$ -value) in a given region. For this analysis, a rescaled childhood maltreatment phenotype was generated whereby the childhood maltreatment score for each individual was divided by the theoretical maximum score for a given study. Best-fit PRS (at  $P_T = 0.0354$ ) were used to predict childhood maltreatment status as a quantitative trait, adjusting for five PCs and dummy study indicator variables. As women in PGC1.5 experienced significantly more childhood maltreatment than men, we generated PRS in women and men separately. The proportion of variance explained by PRS was estimated as the difference in Nagelkerke's  $R^2$  between the full model (which includes PRS plus covariates) and the null model (which only has the covariates). PRS prediction plots were based on quantiles of PRS, with effect sizes calculated in reference to the lowest quantile.  $p$ -values for PRS were derived from a likelihood ratio test comparing the two models. The significance threshold was set at a Bonferroni-corrected threshold of  $p = 0.006$  (0.05/8).

#### Genetic correlation

Another application of LDSR is the measurement of genetic correlation, i.e. the degree and direction of shared genetic effects between different traits<sup>37,39</sup>. Cross-cohort genetic correlation ( $r_g$ ) was calculated using LDSR. The web-based interface for LDSR, **LD Hub**, was used to further calculate pairwise genetic correlations between childhood maltreatment and 247 non-UKBB traits of interest including psychiatric, anthropomorphic, smoking behavior, reproductive, aging, education, autoimmune, and cardio-metabolic categories.

#### Conditional analyses of childhood maltreatment top hits

To evaluate if the effects of top variants in the UKBB GWAS and meta-analysis were specific to childhood maltreatment, we conditioned childhood maltreatment on genetically correlated traits using the multi-trait conditional and joint analysis (mtCOJO)<sup>40</sup> feature in GCTA<sup>41</sup>. Data for major depressive disorder (MDD)<sup>42</sup> (from <https://www.med.unc.edu/pgc/results-and-downloads>) was used to minimize sample overlap with the UK Biobank data. The effect of the correlated trait on childhood maltreatment was estimated using a generalized summary-data based Mendelian randomization analysis of significant LD independent SNPs ( $r^2 < 0.05$ , based on 1000G Phase 3 CEU samples). The threshold for significance was set at  $p < 5 \times 10^{-6}$ , due to having less than the required 10 significant independent SNPs at the program default of  $p < 5 \times 10^{-8}$ , for the correlated trait.

## Results

### GWAS and meta-analysis

We report GWAS results from our discovery dataset (UKBB) ( $n = 124,711$ ) and meta-analysis ( $n = 151,001$ ). In our UKBB discovery dataset, we identified two genome-wide significant loci ( $p < 5 \times 10^{-8}$ ) associated with childhood maltreatment (Table 1 and Fig. 1), rs142346759 (chr3,  $\beta = 0.015$ ,  $p = 4.35 \times 10^{-8}$ ) and rs10262462 (chr7,  $\beta = -0.016$ ,  $p = 3.24 \times 10^{-8}$ ). These variants remained significant in the meta-analysis (Table 1, Supplementary Figs. 1 and 2). Additional variants on chromosome 7 (rs1859100,  $\beta = 0.015$ ,  $p = 3.91 \times 10^{-8}$ ) and chromosome 12 (rs917577,  $\beta = 0.017$ ,  $p = 2.64 \times 10^{-8}$ ) (Supplementary Fig. 3), also achieved genome-wide significance in the meta-analysis. Running an ordinal regression on these hits in the UKBB led to similar results (data not shown). None of these hits were replicated in PGC1.5 (Table 1 and Supplementary Fig. 4).

Quantile–quantile (qq) plots indicate minimal inflation of  $p$ -values across studies (Supplementary Figs. 5–7). Using the LDSR intercept method, polygenic effects account for 93% and 94% of the observed inflation in test statistics for the UKBB dataset (intercept = 1.0096, SE = 0.0064) and meta-analysis (intercept = 1.0095, SE = 0.0077), respectively (Supplementary Figs. 5 and 7), consistent with minimal population stratification and cryptic relatedness.

### Integration with functional genomic data

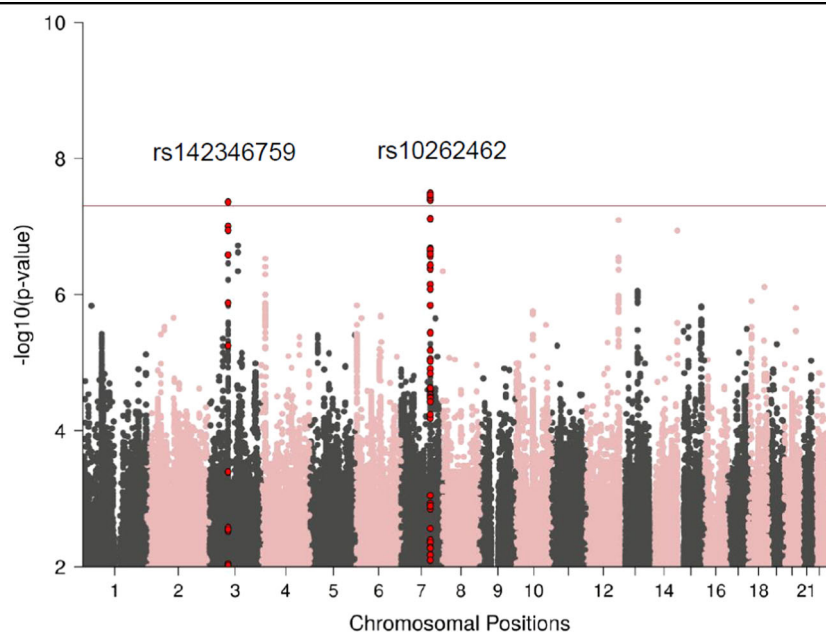
Using the web-based tool FUMA, the two UKBB GWAS hits were each annotated to two genes, *FOXP1* and *FOXP2* (Table 2). Gene-based analysis of the UKBB GWAS summary statistics further identified three gene-wide significant genes, *KIF26B* ( $p = 1.67 \times 10^{-7}$ ), *CNTNAP5* ( $p = 8.89 \times 10^{-7}$ ), and *EXOC2* ( $p = 2.04 \times 10^{-6}$ ) from a total of 18,989 protein-coding genes. Gene-set analysis did not reveal any significant pathways associated with childhood maltreatment. Limited functionality of the two risk variants (rs142346759 and rs10262462) was observed (Table 2). One of the SNPs in LD for the risk variant on chromosome 3, rs142346759, obtained a CADD score of >12.37, indicating that this SNP may be deleterious. Six of the SNPs in LD with the risk variant on chromosome 7, rs10262462, had a CADD score of >12.37. No significant eQTLs were identified for either risk locus.

The chromosome 7 variant identified in the meta-analysis, rs1859100, also mapped to the gene *FOXP2* and is located in the same genomic risk locus (chr7:114,015,707–114,287,116 base pairs) as rs10262462. The other hit observed in the meta-analysis, rs917577, was mapped to an intergenic region on chromosome 12. This variant obtained an RDB categorical score of 2B, indicating that it is likely to affect transcription factor binding. No eQTLs exist in the selected tissue types for this region (Table 2).

**Table 1 Genome-wide significant hits in the UK Biobank, PGC-PTSD Freeze 1.5 (PGC 1.5), and meta-analyses.**

Variant	Chr	Position (bp)	Gene	Discovery: UK Biobank				Replication: Freeze 1.5				Meta-analysis										
				A1	A2	A1 freq	Z-score	Beta	p-Value	n	Z-score	Beta <sup>a</sup>	p-value	r <sup>2</sup> (%)	P-het	n	Z-score	Beta <sup>a</sup>	p-Value	r <sup>2</sup> (%)	P-het	n
rs142346759	3	71,362,232	FOXP1	G	A	0.039	5.476	0.015	4.35E-08	124,711	1.490	0.016	0.136	5	0.394	10,858 <sup>b</sup>	5.674	0.033	1.40E-08	0	0.491	135,569 <sup>b</sup>
rs1859100	7	114,194,615	FOXP2	G	T	0.597	5.517	0.016	3.45E-08	124,711	0.846	0.007	0.397	0	0.978	14,495	5.495	0.015	3.91E-08	0	0.973	139,206
rs10262462	7	114,180,062	FOXP2	A	G	0.403	-5.528	-0.016	3.24E-08	124,711	-0.729	-0.006	0.466	0	0.969	14,495	-5.468	-0.015	4.56E-08	0	0.958	139,206
rs917577	12	126,548,817	Intergenic	C	G	0.265	5.367	0.015	8.02E-08	124,711	1.476	0.016	0.14	0	0.781	11,463	5.564	0.017	2.64E-08	0	0.831	136,174

<sup>a</sup>An approximation was used to transform the Z-statistics from the effective sample-size-weighted meta-analysis (the output of the software METAL) into a beta value<sup>74</sup>. This was calculated on the scale of the MRS dataset  
<sup>b</sup>As this variant has a MAF of < 5% in the UKBB, only studies with a minor allele count of at least five alleles were included in the meta analysis



**Fig. 1** Manhattan plot of UKBB GWAS for childhood maltreatment, showing the top variants. The horizontal line represents genome-wide significance at  $p < 5 \times 10^{-8}$ .

### Heritability of reported childhood maltreatment

GWAS summary statistics were used to estimate the  $h^2_{\text{snp}}$  of childhood maltreatment with the tool LDSR (Table 3). The  $h^2_{\text{snp}}$  was estimated at 0.057 ( $p = 1.60 \times 10^{-32}$ ) for the UKBB discovery dataset and 0.123 ( $p = 0.002$ ) for PGC1.5. The  $h^2_{\text{snp}}$  for the meta-analysis was 0.057 ( $p = 4.48 \times 10^{-46}$ ).

### Polygenic risk scoring

We assessed the predictive value of PRS for childhood maltreatment, using our largest cohort, the UKBB, as a training sample. Our analyses showed a highly significant increase in effect size to develop childhood maltreatment across PRS quantiles in the PGC1.5 target sample, with a variance explained of  $r^2 = 0.0025$  ( $p = 1.8 \times 10^{-15}$ ). Participants in the 5th quantile of genetic risk had significantly higher childhood maltreatment scores than subjects in the 1st quantile (beta = 0.042,  $p = 4.78 \times 10^{-16}$ ; Supplementary Fig. 8). Since women reported significantly more childhood maltreatment than men (PGC1.5 mean childhood maltreatment: women = 0.32, men = 0.127,  $p < 1 \times 10^{-80}$ ), PRS were also calculated separately for women and men. When stratified by sex, PRS had significantly higher explanatory power in women ( $r^2 = 0.0053$ ) relative to men ( $r^2 = 0.0015$ ) ( $p = 0.0002$ , Supplementary Fig. 8).

### Genetic correlations of reported childhood maltreatment with other traits and disorders

All pairwise genetic correlations are listed in Supplementary Table 3. The  $r_g$  for childhood maltreatment

between the UKBB and PGC1.5 datasets was 0.63 ( $p = 3.28 \times 10^{-6}$ ). To determine whether there is significant genetic overlap between childhood maltreatment and other traits and disorders, pairwise genetic correlations were calculated using the web-based tool LD Hub. A total of 27 significant correlations (Bonferroni-corrected  $p$ -value threshold =  $0.05/247 = 0.0002$ ) were found between childhood maltreatment in the meta-analysis and 247 non-UKBB traits. The top 10 highest genetic correlations are plotted in Fig. 2 with depressive symptoms ( $r_g = 0.70$ ,  $p = 4.65 \times 10^{-40}$ ) having the most significant correlation with childhood maltreatment. There were also positive genetic correlations with “MDD” ( $r_g = 0.71$ ,  $p = 4.13 \times 10^{-11}$ ), “PGC cross-disorder analysis” ( $r_g = 0.47$ ,  $p = 1.62 \times 10^{-14}$ ) and “neuroticism” ( $r_g = 0.44$ ,  $p = 1.14 \times 10^{-17}$ ). Significant negative genetic correlations between childhood maltreatment and “age of first birth” ( $r_g = -0.47$ ,  $p = 2.61 \times 10^{-27}$ ), “subjective well-being” ( $r_g = -0.46$ ,  $p = 1.00 \times 10^{-18}$ ), and “mother’s age at death” ( $r_g = -0.36$ ,  $p = 7.42 \times 10^{-6}$ ) were also observed.

### Conditional analyses of childhood maltreatment top hits

As depressive symptoms ( $r_g = 0.70$ ,  $p = 4.65 \times 10^{-40}$ ) and MDD ( $r_g = 0.71$ ,  $p = 4.13 \times 10^{-11}$ ) were the most genetically correlated with childhood maltreatment, we conditioned the top hits from our meta-analysis for the effects of MDD using publicly available summary statistics for MDD<sup>42</sup>. We found that effect sizes for the four top hits for childhood maltreatment remained similar when adjusted for the effects of MDD (Supplementary Table 4).

**Table 2 Functional mapping and annotation of UKBB GWAS and meta-analysis.**

Sample	GWAS hit lead variant	GWAS p-value	Position (hg19)	#SNPs in LD ( $r^2 > 0.6$ )	Genomic coordinated risk locus (hg19)	Predicted genes in risk locus	SNPs in LD with CADD scores > 12.37	SNPs in LD with RegulomeDB scores < 5	Chromatin state analysis (Roadmap Epigenomics) in neuronal cell lines/tissues <sup>a</sup>	eQTL (28 neuronal tissue/cell lines from CommonMind Consortium, BRAINEAC or GTEx v7)
UKBB GWAS	rs142346759	4.35E-08	3: 71362232	5	chr3:71355932-71417539	FOXP1 intronic	rs34936081	rs13083684 (intronic)	Mainly quietcent	None
	rs10262462	4.36E-08	7: 114180062	77	chr7:114015707-114287116	FOXP2 intronic	rs7785701, rs6466488, rs10249234, rs9332390, rs66823671, rs12533005	rs2396753 (intronic), rs12536335 (intronic)	Mainly quietcent	None
Meta-analysis	rs1859100	3.91E-08	7: 114194615	74	chr7:114015707-114287116	FOXP2 intronic	rs66823671, rs71149745, rs7785701, rs12533005, rs9332390, rs10249234, rs1476535, rs6466488	rs2396753 (intronic), rs12536335 (intronic), rs9332390 (intronic)	Mainly quietcent	None
	rs917577	2.64E-08	12:126548817	27	chr12:126548816-126585545	Intergenic	None	None	Mainly quietcent	None

<sup>a</sup>In neuronal cell lines/tissues E053, E054, E067, E068, E069, E070, E071, E072, E073, E074, E081, E082, E125

**Table 3 Heritability estimates based on LD-score regression (LDSR).**

Sample	n	$h^2_{\text{snp}}$	SE	Z	p-value
UKBB	124,711	0.057	0.005	11.40	1.60E-32
PGC1.5	26,290	0.123	0.040	3.08	2.00E-03
Meta-analysis	151,001	0.057	0.004	14.25	4.48E-46

Estimates are calculated for the UK biobank (UKBB), the PGC-PTSD Freeze 1.5 (PGC1.5), and meta-analysis

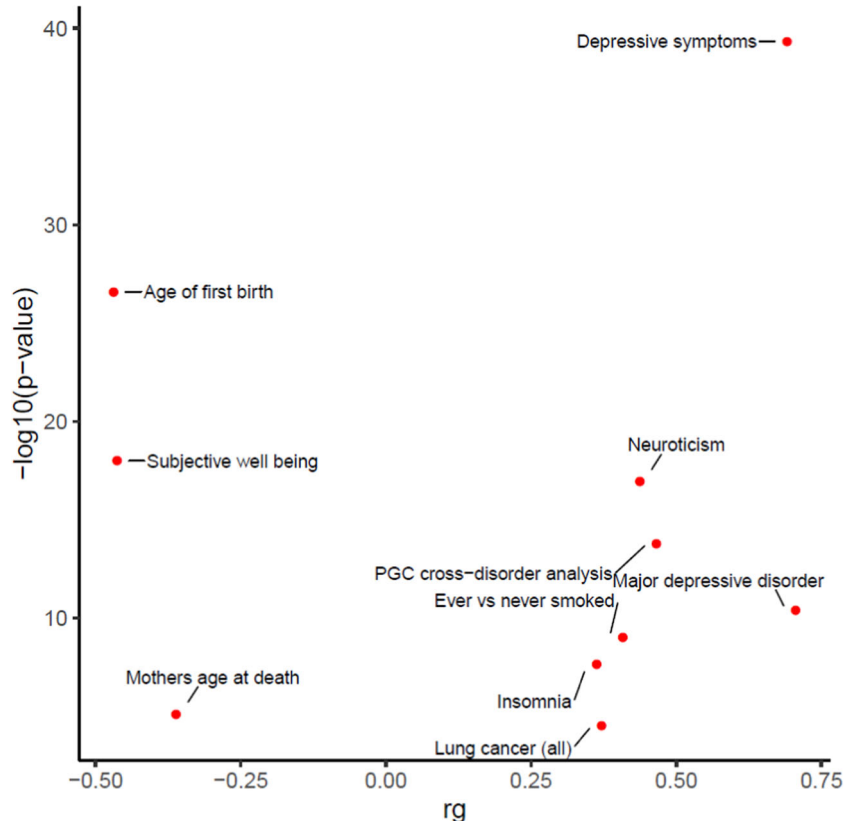
These findings indicate that the genetic variants identified here are specific to childhood maltreatment when tested in the context of MDD, the disorder genetically most significantly correlated with childhood maltreatment.

**Discussion**

The main findings of this study were that (1) variants located in the genes *FOXP1* and *FOXP2* and on chromosome 12 are significantly associated with childhood maltreatment, (2) the SNP-based estimate of childhood maltreatment is ~6%, (3) PRS of self-reported childhood maltreatment derived from a discovery cohort can significantly predict this phenotype in a target cohort, with 0.25% of variance explained, and (4) childhood maltreatment is significantly genetically correlated with “depressive symptoms” and “MDD”, “neuroticism”, “age of first birth”, and “subjective well-being”, despite showing evidence that our top hits may be specific to childhood maltreatment when conditioning on MDD.

Two genome-wide loci for childhood maltreatment identified in our discovery dataset were also significant in the meta-analysis: rs142346759 (chr3p13), an intronic variant in *FOXP1* and rs10262462 (chr7q31.1) an intronic variant located in *FOXP2*. Both genes form part of the forkhead box superfamily of transcription factors which are widely expressed, and which play important roles during development and adulthood. *FOXP1* and *FOXP2* fall under the FOXP sub-family (also comprising *FOXP3* and *FOXP4*) which has functions in oncogenic and tumor suppressive pathways<sup>43</sup>. *FOXP2* contains highly conserved genomic sites, including an intronic region within this gene, located about 107 kb downstream from our risk variant<sup>44</sup>. *FOXP1* and *FOXP2* have ~60% homology at the amino acid level (<https://www.ncbi.nlm.nih.gov/books/NBK7023/>) and both proteins have been implicated in cognitive disorders, including expressive language impairment<sup>45</sup>. In the meta-analysis, we observed an additional genome-wide variant, located in an intergenic region on chromosome 12, but as this variant does not map to a particular gene, its possible biological mechanism is unclear.

Notably, variation within *FOXP1* has been found to have associations with language impairment, internalizing



**Fig. 2** Top ten genetic correlations between several groups of traits (from psychiatric, anthropomorphic, smoking behavior, reproductive, aging, education, autoimmune, and cardio-metabolic categories) and childhood maltreatment (meta-analysis).

symptoms, and externalizing symptoms<sup>46</sup>. *FOXP2* has mainly been investigated in regards to speech and language development<sup>47</sup>, but has also been found to be associated with depression<sup>48</sup> and attention deficit hyperactivity disorder (ADHD)<sup>49</sup>. Further, an intronic variant in the *FOXP2* gene, rs727644, has been associated with risk-taking behavior<sup>50,51</sup>. While most work on childhood maltreatment has emphasized subsequent risk for mental and physical disorders, it is possible that externalizing behaviors increase risk for childhood trauma<sup>52</sup>, consistent with a non-passive rGE mechanism. Alternatively, phenotypes such as depression or neuroticism may increase the likelihood of individuals recalling childhood maltreatment<sup>53,54</sup>.

In this study we estimated SNP-based heritability for childhood maltreatment to be ~6%. A first possibility, in line with a link between *FOXP* variants and externalizing symptoms, is that genetic factors influence environmental factors indirectly through temperament and behavior<sup>9</sup>. A second possibility, consistent with the link of *FOXP* variants to internalizing symptoms and depression, is that genetic factors influence the recall of childhood maltreatment. In particular, retrospective assessment of childhood maltreatment may be limited by recall bias and

the respondent's subjective assessment of the event<sup>55,56</sup>. Indeed, a recent systematic review found very low concordance between prospective and retrospective measures of childhood maltreatment<sup>57</sup> and those who retrospectively report childhood adversity were at greater risk for having psychopathology than those who prospectively reported childhood maltreatment<sup>58</sup>.

A twin-based study estimated the heritability of reported childhood maltreatment (comprising physical, and sexual maltreatment and neglect) to be 6%<sup>7</sup>, the same as our SNP-based estimate. As twin-based studies capture latent heritability across the entire genome, these heritability estimates are generally higher than SNP-based heritability estimates, which are limited to common variation and by the number of markers present and tagged on the genotyping array used<sup>15</sup>. However, in this twin study, when considering each maltreatment category separately, the heritability of childhood physical maltreatment, sexual maltreatment, and neglect was 28%, 0%, and 24%, respectively. This suggests that only physical abuse and neglect are heritable and that sexual abuse is not genetically influenced. It is notable that these twin data, then, do not support an rGE for some abuse types (i.e. sexual abuse).



We show that PRS derived from the UKBB was significantly predictive of childhood maltreatment in PGC1.5, explaining 0.25% of the variance for this exposure. Although the variance explained is relatively modest, we expect greater prediction accuracy with future larger sample sizes. When stratifying by sex, PRS had significantly higher explanatory power in women relative to men. This is expected as women had higher mean self-reported childhood maltreatment scores than men in PGC1.5.

The PRS results suggests a polygenic architecture for self-reported childhood maltreatment but does not indicate the mechanism by which genetic factors are able to influence this exposure. However, our finding of positive genetic correlations between childhood maltreatment, depressive symptoms, and MDD provides support for the hypothesis that genetic factors predisposing to reporting early life maltreatment overlap with those underlying depression. Genetic correlations between depression, stressful life events, and lifetime trauma have led to the hypothesis that genes increasing risk for the development of depression predispose individuals to entering into adverse environments<sup>59,60</sup>. Depressed individuals with and without trauma exposure differ in associated genetic variation, with trauma-exposed individuals having greater SNP-based heritability, supporting this hypothesis<sup>26,61</sup>. On the other hand, polygenic scores for MDD were associated with greater reporting of stressful life events in individuals with MDD<sup>62</sup>. Indeed, current mood can influence the recall of childhood experiences, and individuals with current depression are at an increased likelihood of reporting early life adversity<sup>63</sup>. Notably, although we show that childhood maltreatment is significantly genetically correlated with depression, results from our conditional GWAS analysis indicates that our top four hits are specific to self-reported childhood maltreatment, favoring a non-passive rGE mechanism for childhood maltreatment.

In addition to depression, we found significant positive genetic correlations between childhood maltreatment and “neuroticism” and “PGC cross-disorder analysis” (comprised of GWAS summary statistics of five psychiatric disorders: autism spectrum disorder, attention deficit-hyperactivity disorder, bipolar disorder, MDD, and schizophrenia). We observed negative genetic correlations of childhood maltreatment with “age of first birth” and “subjective well-being”. Associations between early life maltreatment and each of these phenotypes have previously been observed<sup>61,64–72</sup>. Further investigation is required to delineate the mechanisms that play a role in the relationship between childhood maltreatment and these outcomes.

Our study had a number of limitations that deserve emphasis. First, the genetic correlation between the UKBB

and PGC1.5 datasets was only 0.63, indicating differences between the datasets, which possibly explains the non-replication of our top hit and of greater SNP heritability in PGC1.5. The UKBB dataset comprises healthy volunteers who are typically of a higher socioeconomic status and in better overall health than the general population of comparable age<sup>73</sup>, and the findings reported here may not be generalizable to the general population. However, it is also worth noting that the top hits were significant in the meta-analysis, where additional hits for childhood maltreatment were detected in an intergenic region on chromosome 12. Second, although many of the study sites included in the final meta-analysis utilized the well-validated CTQ, childhood maltreatment was measured in a diversity of ways across the different studies. Thus, our main phenotype was not homogenous and may reflect different aspects of childhood maltreatment in different contributing studies.

This is the first large-scale genetic study to identify specific variants associated with self-reported childhood maltreatment. Variation in *FOXP* genes and the polygenic architecture associated with childhood maltreatment may put individuals at greater risk for maltreatment. Alternatively, however, these variants may be associated with a greater likelihood of reporting maltreatment, given the high genetic correlation between childhood maltreatment and depression, and neuroticism. Using the available data, we are unable to indicate definitively which of these explanations is a better one, and it is possible that different mechanisms have more robust explanatory power in accounting for different abuse subtypes as well as different associated psychopathologies. A clearer understanding of the genetic relationships of childhood maltreatment, including particular abuse subtypes, with a range of different phenotypes, may ultimately be useful in developing targeted treatment and prevention strategies.

## Disclaimer

The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citations of commercial organizations or trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

## Acknowledgements

This work was funded by Cohen Veterans Bioscience, the NIMH/U.S. Army Medical Research and Materiel Command Grant R01MH106595 to C.M.N., I.L., K.J.R. and K.C.K., One Mind, and supported by 5U01MH109539 to the Psychiatric Genomics Consortium. Statistical Analysis were carried out on the NL Genetic Cluster computer (URL) hosted by SURFsara. Genotyping of samples was supported in part through the Stanley Center for Psychiatric Genetics at the Broad Institute of MIT and Harvard. This research has been conducted using the UK biobank resource under application number 16577.

This work would not have been possible without the contributions of the investigators who comprise the PGC-PTSD working group, and especially the more than 206,000 research participants worldwide who shared their life experiences and biological samples with PGC-PTSD investigators. Full acknowledgements for each study are in the *Supplementary Note*.

#### Author details

<sup>1</sup>SA MRC Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa. <sup>2</sup>Department of Psychiatry, University of California San Diego, La Jolla, CA, USA. <sup>3</sup>Veterans Affairs San Diego Healthcare System, Center of Excellence for Stress and Mental Health, San Diego, CA, USA. <sup>4</sup>Veterans Affairs San Diego Healthcare System, Research Service, San Diego, CA, USA. <sup>5</sup>King's College London, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, London, UK. <sup>6</sup>King's College London, NIHR BRC at the Maudsley, London, UK. <sup>7</sup>Atlanta VA Health Care System, Mental Health Service Line, Decatur, GA, USA. <sup>8</sup>Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, USA. <sup>9</sup>Alpert Medical School of Brown University, Providence, RI, USA. <sup>10</sup>Massachusetts General Hospital, Analytic and Translational Genetics Unit, Boston, MA, USA. <sup>11</sup>Massachusetts General Hospital, Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Boston, MA, USA. <sup>12</sup>Broad Institute of MIT and Harvard, Stanley Center for Psychiatric Research, Cambridge, MA, USA. <sup>13</sup>Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA, USA. <sup>14</sup>Massachusetts General Hospital, Department of Psychiatry, Boston, MA, USA. <sup>15</sup>Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. <sup>16</sup>Harvard Medical School, Department of Psychiatry, Boston, MA, USA. <sup>17</sup>McLean Hospital, Belmont, MA, USA. <sup>18</sup>Cohen Veterans Bioscience, Cambridge, MA, USA. <sup>19</sup>Department of Psychiatry, University of Michigan Medical School, Ann Arbor, MI, USA. <sup>20</sup>Bradley/Hasbro Children's Research Center of Rhode Island Hospital, Providence, RI, USA. <sup>21</sup>Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Richmond, VA, USA. <sup>22</sup>Harvard Medical School, Department of Psychiatry, Boston, MA, USA. <sup>23</sup>Harvard School of Public Health, Department of Social and Behavioral Sciences, Boston, MA, USA. <sup>24</sup>Boston Children's Hospital, Division of Adolescent and Young Adult Medicine, Boston, MA, USA. <sup>25</sup>Brigham and Women's Hospital, Channing Division of Network Medicine, Boston, MA, USA. <sup>26</sup>Veterans Affairs San Diego Healthcare System, Psychiatry Service, San Diego, CA, USA. <sup>27</sup>Department of Psychology, University of New South Wales, Sydney, NSW, Australia. <sup>28</sup>Department of Psychiatry, University Hospitals, Cleveland, OH, USA. <sup>29</sup>Department of Psychological Sciences, Kent State University, Kent, OH, USA. <sup>30</sup>Research and Sponsored Programs, Kent State University, Kent, OH, USA. <sup>31</sup>Department of Medicine, Boston University School of Medicine, Boston, MA, USA. <sup>32</sup>Department of Psychological Sciences, Case Western Reserve University, Cleveland, OH, USA. <sup>33</sup>Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. <sup>34</sup>Department of Psychiatry, University of Melbourne, Melbourne, VIC, Australia. <sup>35</sup>Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA. <sup>36</sup>US Army Medical Research and Materiel Command, Fort Detrick, MD, USA. <sup>37</sup>US Department of Veterans Affairs, Department of Psychiatry, West Haven, CT, USA. <sup>38</sup>VA Connecticut Healthcare Center, West Haven, CT, USA. <sup>39</sup>Department of Genetics and Neuroscience, Yale University School of Medicine, New Haven, CT, USA. <sup>40</sup>US Army Medical Research and Materiel Command, USACEHR, Fort Detrick, MD, USA. <sup>41</sup>Harvard Medical School, Department of Health Care Policy, Boston, MA, USA. <sup>42</sup>Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. <sup>43</sup>Mental Illness Research, Education and Clinical Center, Crescenzo VAMC, Philadelphia, PA, USA. <sup>44</sup>Department of Psychiatry, New York University School of Medicine, New York, NY, USA. <sup>45</sup>University of Adelaide, Centre for Traumatic Stress Studies, Adelaide, SA, Australia. <sup>46</sup>Department of Psychology, Northern Illinois University, DeKalb, IL, USA. <sup>47</sup>U.S. Department of Veterans Affairs National Center for Posttraumatic Stress Disorder, West Haven, CT, USA. <sup>48</sup>Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA. <sup>49</sup>Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>50</sup>Department of Psychiatry, University of Washington, Seattle, WA, USA. <sup>51</sup>Department of Nursing and Department of Psychiatry, Medical University of South Carolina, Charleston, SC, USA. <sup>52</sup>Department of Psychiatry, University of New South Wales, Sydney, NSW, Australia. <sup>53</sup>Veterans Affairs San Diego Healthcare System, Million Veteran Program, San Diego, CA, USA. <sup>54</sup>Department of Psychiatry, Uniformed Services University, Bethesda, MD, USA. <sup>55</sup>Department of Mental Health, James J. Peters VA Medical Center, Bronx, NY,

USA. <sup>56</sup>Department of Biostatistics, Yale University, New Haven, CT, USA.

<sup>57</sup>Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA. <sup>58</sup>Harvard School of Public Health, Department of Epidemiology, Boston, MA, USA

#### Data availability

The full meta-analysis summary statistics are available for download from the Psychiatric Genomics Consortium at <https://www.med.unc.edu/pgc/results-and-downloads/>. Access to individual-level data for available datasets may be requested through the PGC Data Access Portal at <https://www.med.unc.edu/pgc/shared-methods/data-access-portal/>. All other data that support the findings of this study are available from the corresponding author upon request.

#### Code availability

GWAS and associated scripts are available at [https://github.com/nievergeltlab/ct\\_gwas/](https://github.com/nievergeltlab/ct_gwas/).

#### Conflict of interest

H.R.K. is a member of the American Society of Clinical Psychopharmacology's Alcohol Clinical Trials Initiative (ACTIVE), which in the last 3 years was supported by AbbVie, Alkermes, Amygdala Neurosciences, Arbor, Ethypharm, Indivior, Lilly, Lundbeck, Otsuka, and Pfizer. H.R.K. and J.G. are named as inventors on PCT patent application #15/878,640 entitled: "Genotype-guided dosing of opioid agonists," filed January 24, 2018. In the past 3 years, D.J.S. has received research grants and/or consultancy honoraria from Lundbeck and Sun. In the past 3 years, R.C.K. received support for his epidemiological studies from Sanofi Aventis; was a consultant for Johnson & Johnson Wellness and Prevention, Sage Pharmaceuticals, Shire, Takeda; and served on an advisory board for the Johnson & Johnson Services Inc. Lake Nona Life Project. Kessler is a co-owner of DataStat, Inc., a market research firm that carries out healthcare research. M.B.S. has in the past 3 years been a consultant for Aptinix, Bionomics, Dart Neuroscience, Janssen, Jazz Pharmaceuticals, Neurocrine Biosciences, Oxeia Biopharmaceuticals, and Pfizer. R.Y. is a co-inventor of the following patent application: "Genes associated with posttraumatic-stress disorder. European Patent# EP 2334816 B1.

#### Publisher's note

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

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41398-020-0706-0>).

Received: 3 August 2019 Revised: 28 November 2019 Accepted: 11 December 2019

Published online: 27 January 2020

#### References

- Child maltreatment. *Global estimates*. <http://apps.who.int/gho/data/view.main.VIOLENCECHILDMALTREATMENTv> (2016).
- McLaughlin, K. A. et al. Childhood adversities and post-traumatic stress disorder: evidence for stress sensitisation in the World Mental Health Surveys. *Br. J. Psychiatry* **211**, 280–288 (2017).
- Copeland, W. E. et al. Association of childhood trauma exposure with adult psychiatric disorders and functional outcomes. *JAMA Network Open* **1**, e184493–e184493 (2018).
- Basu, A., McLaughlin, K. A., Misra, S. & Koenen, K. C. Childhood maltreatment and health impact: the examples of cardiovascular disease and type 2 diabetes mellitus in adults. *Clin. Psychol.* **24**, 125–139 (2017).
- Suglia, S. F. et al. Childhood and adolescent adversity and cardiometabolic outcomes: a scientific statement from the American Heart Association. *Circulation* **137**, e15–e28 (2018).
- Koenen, K. C., Roberts, A. L., Stone, D. M. & Dunn, E. C. in *The Impact of Early Life Trauma on Health and Disease: The Hidden Epidemic*, Vol. 1 (eds Lanius, R. A., Vermetten, E. & Pain, C.) Ch. 2 (Cambridge University Press, United Kingdom, 2010).

7. Schulz-Heik, R. J. et al. Investigation of genetically mediated child effects on maltreatment. *Behav. Genet.* **39**, 265–276 (2009).
8. Knafo, A. & Jaffee, S. R. Gene–environment correlation in developmental psychopathology. *Dev. Psychopathol.* **25**, 1–6 (2013).
9. Jaffee, S. R. & Price, T. S. Gene–environment correlations: a review of the evidence and implications for prevention of mental illness. *Mol. Psychiatry* **12**, 432 (2007).
10. Rutter, M. Gene–environment interdependence. *Dev. Sci.* **10**, 12–18 (2007).
11. Boivin, M. et al. The genetic–environmental etiology of parents’ perceptions and self-assessed behaviours toward their 5-month-old infants in a large twin and singleton sample. *J. Child Psychol. Psychiatry* **46**, 612–630 (2005).
12. Berlin, L. J., Dodge, K. A. & Reznick, J. S. Examining pregnant women’s hostile attributions about infants as a predictor of offspring maltreatment. *JAMA Pediatrics* **167**, 549–553 (2013).
13. Rehan, W. et al. Gene–environment correlation between the dopamine transporter gene (DAT1) polymorphism and childhood experiences of abuse. *J. Interpers. Violence* **33**, 2059–2072 (2018).
14. Munafò, M. R. & Flint, J. Dissecting the genetic architecture of human personality. *Trends Cogn. Sci.* **15**, 395–400 (2011).
15. Duncan, L. E. et al. Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Mol. Psychiatry* **23**, 666–673 (2018).
16. van der Merwe, C. et al. Concordance of genetic variation that increases risk for anxiety disorders and posttraumatic stress disorders and that influences their underlying neurocircuitry. *J. Affect. Disord.* **245**, 885–896 (2019).
17. Nievergelt, C. M. et al. International meta-analysis of PTSD genome-wide association studies identifies sex-and ancestry-specific genetic risk loci. *Nat. Commun.* **10**, 1–16 (2019).
18. Allen, N. E., Sudlow, C., Peakman, T. & Collins, R. UK biobank data: come and get it. *Sci. Transl. Med.* **6**, 224 (2014).
19. Bernstein, D. P. et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse Negl.* **27**, 169–190 (2003).
20. Bernstein, D. P. & Finkel, L. *Childhood Trauma Questionnaire: A Retrospective Self-report: Manual* (The Psychological Corporation, Orlando, 1998).
21. Lam, M. et al. RICOPIII: Rapid Imputation for COntortias PipeLine. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btz633> (2019).
22. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
23. Delaneau, O., Zagury, J.-F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6 (2013).
24. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
25. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
26. Coleman, J. et al. Genome-wide gene–environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. *Mol. Psychiatry* <https://doi.org/10.1038/s41380-019-0546-6> (2020).
27. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).
28. Walter, K. et al. The UK10K project identifies rare variants in health and disease. *Nature* **526**, 82–90 (2015).
29. Warren, H. R. et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat. Genet.* **49**, 403–415 (2017).
30. Abraham, G., Qiu, Y. & Inoué, M. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. *Bioinformatics* **33**, 2776–2778 (2017).
31. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
32. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
33. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
34. Zhou, B., Shi, J. & Whittemore, A. S. Optimal methods for meta-analysis of genome-wide association studies. *Genet. Epidemiol.* **35**, 581–591 (2011).
35. Schwarzer, G. meta: an R package for meta-analysis. *R News* **7**, 40–45 (2007).
36. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
37. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
38. Euesden, J., Lewis, C. M. & O’Reilly, P. F. PRSice: polygenic risk score software. *Bioinformatics* **31**, 1466–1468 (2014).
39. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
40. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
41. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
42. Ripke, S. et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry* **18**, 497–511 (2013).
43. Lam, E. W.-F., Brosens, J. J., Gomes, A. R. & Koo, C.-Y. Forkhead box proteins: tuning forks for transcriptional harmony. *Nat. Rev. Cancer* **13**, 482–495 (2013).
44. Atkinson, E. G. et al. No evidence for recent selection at FOXP2 among diverse human populations. *Cell* **174**, 1424–1435. e1415 (2018).
45. Bacon, C. & Rappold, G. A. The distinct and overlapping phenotypic spectra of FOXP1 and FOXP2 in cognitive disorders. *Hum. Genet.* **131**, 1687–1698 (2012).
46. Siper, P. M. et al. Prospective investigation of FOXP1 syndrome. *Mol. Autism* **8**, 57 (2017).
47. Fisher, S. E. & Scharff, C. FOXP2 as a molecular window into speech and language. *Trends Genet.* **25**, 166–177 (2009).
48. Arnau-Soler, A. et al. Genome-wide by environment interaction studies of depressive symptoms and psychosocial stress in UK Biobank and Generation Scotland. *Transl. Psychiatry* **9**, 14 (2019).
49. Demontis, D. et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
50. Clifton, E. A. et al. Genome-wide association study for risk taking propensity indicates shared pathways with body mass index. *Commun. Biol.* **1**, 36 (2018).
51. Strawbridge, R. J. et al. Genetics of self-reported risk-taking behaviour, trans-ethnic consistency and relevance to brain gene expression. *Transl. Psychiatry* **8**, 178 (2018).
52. Cui, N. & Liu, J. Cognitive and behavioral risk factors for child physical abuse among Chinese children: a multiple-informant study. *Child Adolesc. Psychiatry Ment. Health* **10**, 36 (2016).
53. Bernet, C. Z. & Stein, M. B. Relationship of childhood maltreatment to the onset and course of major depression in adulthood. *Depression Anxiety* **9**, 169–174 (1999).
54. Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A. & Livesley, W. J. Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms: a twin study. *Am. J. Psychiatry* **159**, 1675–1681 (2002).
55. Widom, C. S. & Morris, S. Accuracy of adult recollections of childhood victimization, part 2: childhood sexual abuse. *Psychol. Assess.* **9**, 34–46 (1997).
56. Hardt, J. & Rutter, M. Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *J. Child Psychol. Psychiatry* **45**, 260–273 (2004).
57. Baldwin, J. R., Reuben, A., Newbury, J. B. & Danese, A. Agreement between prospective and retrospective measures of childhood maltreatment: a systematic review and meta-analysis. *JAMA Psychiatry* **76**, 584–593 (2019).
58. Newbury, J. B. et al. Measuring childhood maltreatment to predict early-adult psychopathology: comparison of prospective informant-reports and retrospective self-reports. *J. Psychiatr. Res.* **96**, 57–64 (2018).
59. Kendler, K. S., Kessler, R. C., Neale, M. C., Heath, A. C. & Eaves, L. J. The prediction of major depression in women: toward an integrated etiologic model. *Am. J. Psychiatry* **150**, 1139–1139 (1993).
60. Kendler, K. S. & Karkowski-Shuman, L. Stressful life events and genetic liability to major depression: genetic control of exposure to the environment? *Psychol. Med.* **27**, 539–547 (1997).
61. Peterson, R. E. et al. Molecular genetic analysis subdivided by adversity exposure suggests etiologic heterogeneity in major depression. *Am. J. Psychiatry* **175**, 545–554 (2018).
62. Mullins, N. et al. Polygenic interactions with environmental adversity in the aetiology of major depressive disorder. *Psychol. Med.* **46**, 759–770 (2016).
63. Colman, I. et al. Consistency in adult reporting of adverse childhood experiences. *Psychol. Med.* **46**, 543–549 (2016).
64. Rosenman, S. & Rodgers, B. Childhood adversity and adult personality. *Aust. N. Z. J. Psychiatry* **40**, 482–490 (2006).
65. Pesonen, A. K. et al. Reproductive traits following a parent–child separation trauma during childhood: a natural experiment during World War II. *Am. J. Hum. Biol.* **20**, 345–351 (2008).
66. Matheson, S. L., Shepherd, A. M., Pinchbeck, R. M., Laurens, K. R. & Carr, V. J. Childhood adversity in schizophrenia: a systematic meta-analysis. *Psychol. Med.* **43**, 225–238 (2013).

67. Palmier-Claus, J. E., Berry, K., Bucci, S., Mansell, W. & Varese, F. Relationship between childhood adversity and bipolar affective disorder: systematic review and meta-analysis. *Br. J. Psychiatry* **209**, 454–459 (2016).
68. Kerns, C. M., Newschaffer, C. J., Berkowitz, S. & Lee, B. K. Brief report: examining the association of autism and adverse childhood experiences in the National Survey of Children's Health: the important role of income and co-occurring mental health conditions. *J. Autism Dev. Disord.* **47**, 2275–2281 (2017).
69. Bjorkenstam, E., Bjorkenstam, C., Jablonska, B. & Kosidou, K. Cumulative exposure to childhood adversity, and treated attention deficit/hyperactivity disorder: a cohort study of 543 650 adolescents and young adults in Sweden. *Psychol. Med.* **48**, 498–507 (2018).
70. Corcoran, M. & McNulty, M. Examining the role of attachment in the relationship between childhood adversity, psychological distress and subjective well-being. *Child Abuse Negl.* **76**, 297–309 (2018).
71. Mandelli, L., Petrelli, C. & Serretti, A. The role of specific early trauma in adult depression: a meta-analysis of published literature. *Childhood trauma and adult depression. Eur. Psychiatry* **30**, 665–680 (2015).
72. Li, M., D'arcy, C. & Meng, X. Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis, and proportional attributable fractions. *Psychol. Med.* **46**, 717–730 (2016).
73. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
74. Lee, J. J. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).