Title
Genomic influences on self-reported childhood maltreatment.

Permalink
https://escholarship.org/uc/item/7v35b96p

Journal
Translational psychiatry, 10(1)

ISSN
2158-3188

Authors
Dalvie, Shareefa
Maihofer, Adam X
Coleman, Jonathan RI
et al.

Publication Date
2020-01-27

DOI
10.1038/s41398-020-0706-0

Peer reviewed
Genomic influences on self-reported childhood maltreatment

Shareefa Dalvie, Adam X. Malhofer, Jonathan R. I. Coleman, Bekh Bradley, Gerome Breen, Leslie A. Brick, Chia-Yen Chen, Karmel W. Choi, Laramie E. Duncan, Guia Guffanti, Magali Haas, Supriya Hamal, Israel Liberson, Nicole R. Nugent, Allison C. Provost, Kerry J. Ressler, Katy Torres, Ananda B. Amstadter, Dewleen G. Baker, Adam X. Maihofner, Supriya Harnal, Israel Liberzon, Nicole R. Nugent, Allison C. Provost, Kerryn J. Ressler, Karim W. Choi, Laramie E. Duncan, Angela G. Junglen, Milissa L. Kaufman, Ronald C. Kessler, Karen C. Koenen and Caroline M. Nievergelt

© The Author(s) 2020

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 ($h^2_{SNP}$), (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) ($n = 26,290$). $h^2_{SNP}$ 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. $h^2_{SNP}$ 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.
Introduction

The lifetime prevalence of childhood physical, sexual, and emotional maltreatment ranges from 8% to 36%1. In addition to being highly prevalent, such childhood abuse is associated with the development of mental disorders, including depression2,3, and physical ill health, including non-communicable diseases4,5. 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 maltreatment6.

A twin-based study found that retrospective reports of childhood maltreatment has a heritability of 6%7. 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 children8). 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)9,10. 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 abuse11,12). The latter two rGEs are sometimes collectively referred to as non-passive correlations9.

While a number of key risk factors for childhood maltreatment, including child behavioral characteristics and parental mental health, have been investigated8, studies have seldom focused on associated genetic variation. The few genetic association studies of childhood maltreatment have only considered variants in candidate genes13 and have had insufficient power to detect the small polygenic effect sizes typically associated with behavioral phenotypes14. 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 outcome9 (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 investigated15–17, 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)18, and the replication sample comprised 26,290 individuals—a subset of the PGC-PTSD Freeze 1.5 dataset (PGC1.5)17. 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 #16097×).

Phenotype harmonization

For the childhood maltreatment phenotype, Childhood Trauma Questionnaire (CTQ) scores on physical, sexual, and emotional abuse19 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-
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. 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. 17. 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/a/broadinstitute.org/ricopili/ and https://github.com/orgs/Nealelab/teams/ricopili) as part of the PGC-PTSD Freeze 2 data analysis. 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 ($\text{fhet} < -0.2$ 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. Pre-phasing and phasing was performed using SHAPEIT2 v2.r83723. Imputation was performed with IMPUTE2 v2.2.224 using the 1000 Genomes (1000G) phase 3 data as the reference.

Details regarding the QC, imputation, and ancestry determination of the UKBB dataset can be found in ref. 26. 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)27 and the UK10K28 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 UKBB29. Additional principal component analysis was conducted on the European-only data subset using flashpca230.

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 1.931 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)17. 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 v1.223 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). Effective sample sizes ($N_{\text{eff}}$) for ordinal logistic regressions were calculated as $N_{\text{eff}} = \frac{1}{\text{logit}(p)}$ levels of childhood maltreatment, and for linear regressions as $N_{\text{eff}} = \left(1 - \frac{1}{p}\right)$ levels of childhood maltreatment. 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 meta235.

Functional mapping and annotation

Genome-wide significant hits identified from the GWAS meta-analysis were annotated using the web-
based tool FUbational Mapping and Annotation (FUMA) v1.3.4c

The SNP2GENE module was used to identify genomic risk loci and these were mapped to protein-coding genes within a 10 kb window. An \( r^2 > 0.6 \) 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. 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)\(^{-1}\). 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, 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 (\( P_1 \)) (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 \( P_1 = 0.0354 \)) 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. 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) feature in GCTA. Data for major depressive disorder (MDD) (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 \( p < 5 \times 10^{-8} \), 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 \( (\{chr}3, \beta = 0.015, \ p = 4.35 \times 10^{-8}) \) and rs10262462 \( (\{chr}7, \beta = -0.016, \ p = 3.24 \times 10^{-8}) \). 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, \( \text{FOXP1 and FOXP2} \) (Table 2). Gene-based analysis of the UKBB GWAS summary statistics further identified three genome-wide significant genes, \( \text{KIF26B} \ (p = 1.67 \times 10^{-7}), \text{CNTNAP5} \ (p = 8.89 \times 10^{-8}) \), and \( \text{EXOC2} \ (p = 2.04 \times 10^{-8}) \) 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 \( (rs142346759 \text{ and } rs10262462) \) was observed (Table 2). One of the SNPs in LD for the risk variant on chromosome 3, \( rs142346759 \), obtained a CADDS score of \( >12.37 \), indicating that this SNP may be deleterious. Six of the 18 SNPs in LD with the risk variant on chromosome 7, \( rs10262462 \), had a CADDS 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 \( \text{FOXP2} \) and is located in the same genomic risk locus \( (\{chr}7:114,015,707–114,287,116 \text{ base pairs}) \) 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 | A1 | A2 | A1 freq | Discovery: UK Biobank | Replication: Freeze 1.5 | Meta-analysis |
|---------|--------------------|------|----|----|---------|-----------------------|-----------------------|--------------|
| rs142346759 | 3 | 114,015,707 | \( \text{FOXP1} \) | G | A | 0.009 | 5.467 | 0.015 | 4.35E-08 | 1.490 | 0.016 | 0.146 | 124,711 |
| rs10262462 | 7 | 114,180,062 | \( \text{FOXP2} \) | A | G | 0.403 | 5.537 | 0.017 | 3.24E-08 | 1.490 | 0.017 | 0.140 | 124,711 |
| rs917577 | 12 | 126,548,817 | Intergenic | C | G | 0.305 | 5.517 | 0.017 | 0.140 | 124,711 |

\( \text{A1 and A2 represent risk alleles for either risk locus.} \)

\( \text{An approximation was used to transform the Z-statistics from the effective sample-size-weighted meta-analysis (the output of the software METAL) into a beta value.} \)

\( \text{This was calculated on the scale of the MRS dataset.} \)

\( \text{The chromosome 7 variant identified in the meta-analysis, } rs1859100, \text{ also mapped to the gene } FOXP2 \) and is located in the same genomic risk locus \( (\{chr}7:114,015,707–114,287,116 \text{ base pairs}) \) 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).
Heritability of reported childhood maltreatment

GWAS summary statistics were used to estimate the $h^2_{snp}$ of childhood maltreatment with the tool LDSR (Table 3). The $h^2_{snp}$ 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 $h^2_{snp}$ 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 $= 0.127$, $p = 1 \times 10^{-30}$), 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” ($r_g = -0.36$, $p = 7.42 \times 10^{-10}$) 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. 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).
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. FOXP2 contains highly conserved genomic sites, including an intronic region within this gene, located about 107 kb downstream from our risk variant. FOXP1 and FOXP2 have ~60% homology at the amino acid level and both proteins have been implicated in cognitive disorders, including expressive language impairment. 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...
symptoms, and externalizing symptoms. FOXP2 has mainly been investigated in regards to speech and language development, but has also been found to be associated with depression and attention deficit hyperactivity disorder (ADHD). Further, an intronic variant in the FOXP2 gene, rs727644, has been associated with risk-taking behavior. 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, consistent with a non-passive rGE mechanism. Alternatively, phenotypes such as depression or neuroticism may increase the likelihood of individuals recalling childhood maltreatment.

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. 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. Indeed, a recent systematic review found very low concordance between prospective and retrospective measures of childhood maltreatment and those who retrospectively report childhood adversity were at greater risk for having psychopathology than those who prospectively reported childhood maltreatment.

A twin-based study estimated the heritability of reported childhood maltreatment (comprising physical, and sexual maltreatment and neglect) to be 6%, 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. 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 suggest 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. Depressed individuals with and without trauma exposure differ in associated genetic variation, with trauma-exposed individuals having greater SNP-based heritability, supporting this hypothesis. On the other hand, polygenic scores for MDD were associated with greater reporting of stressful life events in individuals with MDD. 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. 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. 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, 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 Material Command Grant R01MH106959 to CMN, IL, KJR, and KCK; One Mind, and supported by SU01MH09339 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
1SA MRC Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa. 2Department of Psychiatry, University of California San Diego, La Jolla, CA, USA. 3Veterans Affairs San Diego Healthcare System, Center of Excellence for Stress and Mental Health, San Diego, CA, USA. 4Veterans Affairs San Diego Healthcare System, Research Service, San Diego, CA, USA. 5King’s College London, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, London, UK. 6King’s College London, NIHR BRC at the Maudsley, London, UK. 7Atlanta VA Health Care System, Mental Health Service Line, Decatur, GA, USA. 8Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, USA. 9Alpert Medical School of Brown University, Providence, RI, USA. 10Massachusetts General Hospital, Analytic and Translational Genetics Unit, Boston, MA, USA. 11Massachusetts General Hospital, Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Boston, MA, USA. 12Broad Institute of MIT and Harvard, Stanley Center for Psychiatric Research, Cambridge, MA, USA. 13Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, MA, USA. 14Massachusetts General Hospital, Department of Psychiatry, Boston, MA, USA. 15Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. 16Harvard Medical School, Department of Psychiatry, Boston, MA, USA. 17McLean Hospital, Belmont, MA, USA. 18Cohen Veterans Bioscience, Cambridge, MA, USA. 19Department of Psychiatry, University of Michigan Medical School, Ann Arbor, MI, USA. 20Bradley-Hasbro Children’s Research Center of Rhode Island Hospital, Providence, RI, USA. 21Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Richmond, VA, USA. 22Harvard Medical School, Department of Psychiatry, Boston, MA, USA. 23Harvard School of Public Health, Department of Social and Behavioral Sciences, Boston, MA, USA. 24Boston Children’s Hospital, Division of Adolescent and Young Adult Medicine, Boston, MA, USA. 25Harigham and Women’s Hospital, Channing Division of Network Medicine, Boston, MA, USA. 26Veterans Affairs San Diego Healthcare System, Psychiatry Service, San Diego, CA, USA. 27Department of Psychology, University of New South Wales, Sydney, NSW, Australia. 28Department of Psychology, University Hospitals, Cleveland, OH, USA. 29Department of Psychological Sciences, Kent State University, Kent, OH, USA. 30Research and Sponsored Programs, Kent State University, Kent, OH, USA. 31Department of Medicine, Boston University School of Medicine, Boston, MA, USA. 32Department of Psychological Sciences, Case Western Reserve University, Cleveland, OH, USA. 33Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 34Department of Psychiatry, University of Melbourne, Melbourne, VIC, Australia. 35Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA. 36US Army Medical Research and Materiel Command, Fort Detrick, MD, USA. 37US Department of Veterans Affairs, Department of Psychiatry, West Haven, CT, USA. 38VA Connecticut Healthcare Center, West Haven, CT, USA. 39Department of Genetics and Neuroscience, Yale University School of Medicine, New Haven, CT, USA. 40US Army Medical Research and Materiel Command, USACERL, Fort Detrick, MD, USA. 41Harvard Medical School, Department of Health Care Policy, Boston, MA, USA. 42Department of Psychiatry, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. 43Mental Illness Research, Education and Clinical Center, Crescenz VAMC, Philadelphia, PA, USA. 44Department of Psychiatry, New York University School of Medicine, New York, NY, USA. 45University of Adelaide, Centre for Traumatic Stress Studies, Adelaide, SA, Australia. 46Department of Psychology, Northern Illinois University, DeKalb, IL, USA. 47U.S. Department of Veterans Affairs National Center for Posttraumatic Stress Disorder, West Haven, CT, USA. 48Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA. 49Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA. 50Department of Psychiatry, University of Washington, Seattle, WA, USA. 51Department of Nursing and Department of Psychiatry, Medical University of South Carolina, Charleston, SC, USA. 52Department of Psychiatry, University of New South Wales, Sydney, NSW, Australia. 53Veterans Affairs San Diego Healthcare System, Million Veteran Program, San Diego, CA, USA. 54Department of Psychiatry, Uniformed Services University, Bethesda, MD, USA. 55Department of Mental Health, James J. Peters VA Medical Center, Bronx, NY, USA. 56Department of Biostatistics, Yale University, New Haven, CT, USA. 57Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA. 58Harvard 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/.

Conflict of interest
HRK 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. HRK. 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. Kesler 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 Aptinyx, Bioinformatics, Dart Neuroscience, Janssen, Jazz Pharmaceuticals, Neurocine 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
1. Child maltreatment. Global estimates. http://apps.who.int/gho/data/view.main.VIOLENCECHILDMALTESTIMENTv2016) (2016).
2. 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).
3. Copeland, W. E. et al. Association of childhood trauma exposure with adult psychiatric disorders and functional outcomes. JAMA Network Open 1, e184493–e184493 (2018).
4. 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).
5. 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).
6. 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. & Pan, C) Ch. 2. (Cambridge University Press, United Kingdom, 2010).
67. Palmer-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).