Stephan, Risbrough, Victoria B., Roberts, Andrea L., Rothbaum, Alex O.,
Rothbaum, Barbara O., Roy-Byrne, Peter, Ruggiero, Ken, Rung, Ariane, Rutten,
Bart P. F., Saccone, Nancy L., Sanchez, Sixto E., Schijven, Dick, Seedat,
Soraya, Seligowski, Antonia V., Seng, Julia S., Sheerin, Christina M., Silove,
Derrick, Smith, Alicia K., Smoller, Jordan W., Sponheim, Scott R., Stein, Dan J.,
Stevens, Jennifer S., Sumner, Jennifer A., Teicher, Martin H., Thompson,
Wesley K., Trapido, Edward, Uddin, Monica, Ursano, Robert J., van den
Heuvel, Leigh Luella, Van Hooff, Miranda, Vermetten, Eric, Vinkers, Christiaan
H., Voisey, Joanne, Wang, Yunpeng, Wang, Zhewu, Werge, Thomas, Williams,
Michelle A., Williamson, Douglas E., Winternitz, Sherry, Wolf, Christiane, Wolf,
Erika J., Wolff, Jonathan D., Yehuda, Rachel, Young, Ross McD., Young, Keith
A., Zhao, Hongyu, Zoellner, Lori A., Liberzon, Israel, Ressler, Kerry J., Haas,
Magali and Koenen, Karestan C. 2019. International meta-analysis of PTSD
genome-wide association studies identifies sex- and ancestry-specific genetic
risk loci. Nature Communications 10 , 4558. 10.1038/s41467-019-12576-w file

Publishers page: http://dx.doi.org/10.1038/s41467-019-12576-w
<http://dx.doi.org/10.1038/s41467-019-12576-w>

Please note:
Changes made as a result of publishing processes such as copy-editing,
formatting and page numbers may not be reflected in this version. For the
definitive version of this publication, please refer to the published source. You
are advised to consult the publisher’s version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.
See
http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights
for publications made available in ORCA are retained by the copyright
holders.
International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci

Caroline M. Nievergelt et al.

The risk of posttraumatic stress disorder (PTSD) following trauma is heritable, but robust common variants have yet to be identified. In a multi-ethnic cohort including over 30,000 PTSD cases and 170,000 controls we conduct a genome-wide association study of PTSD. We demonstrate SNP-based heritability estimates of 5–20%, varying by sex. Three genome-wide significant loci are identified, 2 in European and 1 in African-ancestry analyses. Analyses stratified by sex implicate 3 additional loci in men. Along with other novel genes and non-coding RNAs, a Parkinson’s disease gene involved in dopamine regulation, PARK2, is associated with PTSD. Finally, we demonstrate that polygenic risk for PTSD is significantly predictive of re-experiencing symptoms in the Million Veteran Program dataset, although specific loci did not replicate. These results demonstrate the role of genetic variation in the biology of risk for PTSD and highlight the necessity of conducting sex-stratified analyses and expanding GWAS beyond European ancestry populations.

A full list of authors and their affiliations appears at the end of the paper.
Post-traumatic stress disorder (PTSD) is a commonly occurring mental health consequence of exposure to extreme, life threatening stress, and/or serious injury/harm. PTSD is frequently associated with the occurrence of comorbid mental disorders such as major depression and other adverse health sequelae including type 2 diabetes and cardiovascular disease. Given this high prevalence and impact, PTSD is a serious public health problem. An understanding of the biological mechanisms of risk for PTSD is therefore an important goal of research ultimately aimed at its prevention and mitigation.

Exposure to traumatic stress is, by definition, requisite for the development of PTSD, but individual susceptibility to PTSD (conditioned on trauma exposure) varies widely. Twin studies over the past two decades provide persuasive evidence for at least some genetic influence on PTSD risk, and the last decade has witnessed the beginnings of a concerted effort to detect specific genetic susceptibility variants for PTSD.

The Psychiatric Genomics Consortium—PTSD Group (PGC-PTSD) published results from a large GWAS on PTSD, involving a trans-ethnic sample of over 20,000 individuals, approximately 5000 (25%) of whom were cases. With this limited sample size, no individual variants exceeded genome-wide significance; however, significant estimates of SNP heritability and genetic correlations between PTSD and other mental disorders such as schizophrenia were demonstrated for the first time.

It is apparent from previous PGC work on other mental disorders that sample size is paramount for GWAS to discern common genome-wide significant variants of small effect that are replicable. Subsequent to the publication of data from the first freeze, the PGC-PTSD has continued to acquire additional PTSD cases and controls through partnerships with an expanding network of investigators, such that we now have accrued a sample size that has enabled us to turn the corner on genome-wide risk discovery. Presented here are the results of our latest GWA studies that include over 23,000 European and over 4000 African ancestry PTSD cases, now involving a total trans-ethnic sample of over 200,000 individuals. In achieving this sample size, we identify sex- and ancestry-specific findings. GWAS and gene-based analyses across our cohorts indicate genome-wide significant associations, involving genes related to dopamine and immune pathways. We show high genetic correlations between PTSD and related psychiatric disorders such as major depressive disorder, but present evidence that some of the identified loci are likely specific to PTSD. In addition, we construct a highly significant polygenic risk score for PTSD, which is predictive of re-experiencing symptoms (REX), a core feature of PTSD, in the independent Million Veteran Program cohort.

Results

Meta-analysis strategy across ancestries and sex. We report meta-analyses of GWAS from the PGC-PTSD Freeze 2 (PGC2), comprised of an ancestrally diverse group of 206,655 participants (including 32,428 cases) from 60 different PTSD studies, ranging from clinically deeply characterized, small patient groups to large cohorts with self-reported PTSD symptoms (Supplementary Data 1). Trauma exposure included both civilian and/or military events, often with pre-existing exposure to childhood trauma, and the majority of controls were trauma-exposed. First ancestry groups were consistently defined across studies (Supplementary Fig. 1). Primary GWAS were then performed separately in the three largest ancestry groups (European (EUA), African (AFA), and Native American Ancestry (AMA)), then meta-analyzed across studies and ancestry groups. Given the previously observed differences between male and female heritability estimates in PGC-PTSD Freeze 1, we also performed sex-stratified analyses. Quantile-quantile plots showed low inflation across analyses (Supplementary Fig. 2), which was mostly accounted for by polygenic SNP effects with little indication of residual population stratification (see Supplementary Note 1 for additional information).

Heritability of PTSD. We estimated heritability of PTSD in the EUA studies (Table 1) based on information captured by genotyping arrays ($h^2_{SNP}$) from summary statistics using LDSC. Assuming a population prevalence of 30% after trauma exposure, overall $h^2_{SNP}$ in PGC2 was 0.05 on the liability scale ($P = 3.18 \times 10^{-8}$). However, female heritability was highly significant ($h^2_{SNP} = 0.10$, $P = 8.03 \times 10^{-11}$), while male heritability was not significantly different from zero ($h^2_{SNP} = 0.01$, $P = 0.63$).

We further examined sex differences in heritability in different subsets of the data: the PGC2 data without the UK Biobank (referred to as PGC1.5) and the UK Biobank (UKB) by itself, which comprises a large proportion of PGC2. Similar to the overall PGC2, heritability in PGC1.5 was high in women and not significant in men. In contrast, in the UKB, male heritability was significantly different from zero ($h^2_{SNP} = 0.08$, $P = 3.96 \times 10^{-1}$). The heritability estimates were highest in the UKB and lowest in the PGC1.5. In contrast, the heritability estimates for women were highest in the PGC2 and lowest in the UKB.

Table 1 Heritability estimates in subjects of European ancestry (EUA)

| Sample | N | N | 10% prev | 30% prev | 50% prev |
|--------|---|---|----------|----------|----------|
|        | Cases | Controls | $h^2_{SNP}$ | 95% CI | $h^2_{SNP}$ | 95% CI | $h^2_{SNP}$ | 95% CI |
| All    | 23,212 | 151,447 | 0.04 | 0.02-0.05 | 0.05 | 0.03-0.07 | 0.06 | 0.04-0.08 | 3.2 $\times 10^{-8}$ |
| PGC2   | 12,823 | 35,648 | 0.03 | 0.01-0.06 | 0.05 | 0.01-0.08 | 0.05 | 0.01-0.09 | 0.01 |
| Men    | 10,389 | 115,799 | 0.13 | 0.1-0.15 | 0.17 | 0.14-0.21 | 0.19 | 0.15-0.23 | 2.1 $\times 10^{-18}$ |
| UKB    | 9908   | 75,605 | 0.01 | -0.02 to 0.03 | 0.01 | -0.03 to 0.05 | 0.01 | -0.03 to 0.05 | 0.63 |
| PGC2   | 6364   | 23,905 | 0.01 | -0.04 to 0.05 | 0.01 | -0.05 to 0.07 | 0.01 | -0.05 to 0.08 | 0.69 |
| Women  | 3544   | 51,700 | 0.11 | 0.04-0.17 | 0.15 | 0.05-0.24 | 0.16 | 0.05-0.26 | 1.4 $\times 10^{-3}$ |
| UKB    | 12,973 | 73,627 | 0.07 | 0.05-0.09 | 0.10 | 0.07-0.13 | 0.11 | 0.07-0.14 | 8.0 $\times 10^{-11}$ |
| PGC1.5 | 6128   | 9528   | 0.15 | 0.08-0.22 | 0.21 | 0.11-0.31 | 0.23 | 0.12-0.33 | 2.7 $\times 10^{-5}$ |
| UKB    | 6845   | 64,099 | 0.14 | 0.1-0.18 | 0.19 | 0.13-0.25 | 0.21 | 0.14-0.27 | 2.0 $\times 10^{-10}$ |

Estimates are calculated using LD-score regression (LDSC) at different population prevalences after trauma exposure for the combined PGC freeze 2 samples, and separately for PGC1.5 (without the UK biobank), the UK biobank, and for men and women. Number of SNPs ranges from 1,160,174 to 1,175,791. P-value is testing if $h^2_{SNP}$ is different from zero and applies to all prevalences. PGC2 all European ancestry subjects of PGC freeze 2 (including the UK biobank), PGC1.5 European ancestry subjects in the PGC1.5 EUA (not including the UK Biobank subjects), UKB UK Biobank European subjects, $h^2_{SNP}$ mean SNP-based heritability, 95% CI 95% confidence interval, prev prevalence.
was significant ($h^2_{SNP} = 0.15, P = 1.38 \times 10^{-3}$) and not significantly different ($\delta = 0.23, P = 0.41$) from heritability in women ($h^2_{SNP} = 0.19, P = 2 \times 10^{-10}$). Sensitivity analyses in UKB using different case and control definitions further confirm these results (Supplementary Table 1 and Supplementary Note 1).

We also compared heritability across ancestries using the GCTA GREML method, which allows analysis of admixed populations when individual genotypes are available. GCTA estimates in the smaller EUA data remained similar to LDSC on the full data (Table 2). Heritability in AFA was comparable to estimates for EUA in the overall sample and when stratified by sex.

**Comparability of PGC2 studies and sex-specific analyses.** A comparison of heritability estimates in subsets of PGC2 studies stratified by sex, ancestry, and characteristics of study (i.e. PGC1.5 vs. the large UKB cohort) show significant estimates for PTSD in the range of 10–20% (Tables 1–2). This is with the notable exception of PGC1.5 males (in EUA and AFA analyses), which fail to show significant $h^2_{SNP}$, despite similar numbers of PTSD cases compared to PGC1.5 women. To further evaluate the comparability of PGC2 studies we estimated genetic correlations ($r_g$) between subsets with different characteristics (Supplementary Table 2).

As numerous small studies contribute to PGC1.5 (24 EUA studies with $N < 200$ cases, Supplementary Table 3), we first investigated the contribution of small studies to the overall genetic signal and genetic similarity to the larger PGC1.5 cohorts. The combined subset of small studies showed significant overall heritability ($h^2_{SNP} = 0.12, P = 0.046$) and close to significant genetic correlation with large studies ($r_g = 0.45, P = 0.08$), indicating a meaningful genetic contribution in aggregate.

Subsetting PGC1.5 and UKB by sex showed a high genetic correlation between women and men for UKB ($r_g = 1.03, P = 1.6 \times 10^{-5}$), but no significant genetic correlation between women and men in PGC1.5, which was expected, since $h^2_{SNP}$ in PGC1.5 men is not significant. Next, focusing on PGC1.5 women only, a comparison to UKB showed significant genetic correlations with the overall UKB ($r_g = 0.46, P = 0.0004$) and UKB women ($r_g = 0.46, P = 0.0008$), and a slightly lower, but marginally significant correlation with the smaller UKB male data ($r_g = 0.44, P = 0.052$).

Overall, these findings of significant heritability estimates for PTSD and moderate to high genetic correlations between most PGC2 subsets, including PGC1.5 to UKB ($r_g = 0.73, P = 0.0005$), are promising for GWAS in these data.

**GWAS in subjects of European and African ancestry.** Our largest PTSD meta-analysis in subjects of EUA (maximum number of subjects included in EUA meta-analyses: $N = 23,212$ cases, 151,447 controls, see Table 3 for details) identified two independent, genome-wide significant loci ($P < 5 \times 10^{-8}$), both mapping to chromosome 6, and sex-stratified analyses in men identified two additional loci (Fig. 1a, b, respectively). The smaller meta-analyses in AFA ($N = 4363$ cases, 10,976 controls) identified one genome-wide significant locus, and an additional locus was found in men when stratified by sex (Fig. 1c, d, respectively). No genome-wide significant associations were found in meta-analyses of EUA or AFA women (Supplementary Fig. 3).

Regional plots of the six genomic regions can be found in Supplementary Figs. 4–9. The six leading markers show odds ratios of 1.12–1.33 and no significant heterogeneity across studies (Table 3). This is supported by PM-plots (posterior probability that a SNP effect exists in a given study) showing a high consistency of effects among the studies predicted to have an effect (Supplementary Figs. 10–15). A z-test on the effect sizes confirmed similar effects for men and women for the three leading variants in the joint-sex analyses, and significant sex-specificity for the three male hits identified in the sex-stratified analyses (Supplementary Table 4).

**Analyses across ancestry groups.** In order to study whether the genetic associations with PTSD vary across different ancestries, we first tried to replicate our six EUA and AFA top hits in the other main ancestry groups (EUA, AFA and AMA, respectively). No evidence of replication was found by directly comparing the other main ancestry groups (EUA, AFA and AMA, respectively). We then tried to replicate our six EUA and AFA top hits in the six main ancestry groups (N = 29,556 cases and 166,145 controls) under fixed- and random-effect models (Supplementary Fig. 22).

While lack of replication of the 4 EUA hits may not be conclusive due to limited power in the smaller AFA and AMA data (Supplementary Tables 5–6), the 2 hits in AFA provided an opportunity for a more detailed analysis of ancestry effects. GWAS in the AFA subjects included standard PCs to control for average admixture across the genome. However, to precisely infer local ancestry across the genome of admixed subjects, we further implemented a local ancestry inference (LAI) pipeline (Supplementary Note 1 and Supplementary Fig. 23). We confirmed the AFA top hit rs115539978 to be specific to the African ancestral background (8% MAF on the African, and <1% MAF on the African, and <1% MAF on the African).
European and Native American backgrounds, respectively; (Supplementary Table 7). Conversely, LAI analyses of the male-specific hit indexed by rs142174523 showed no evidence for ancestry-specific effects that would explain the lack of replication in the EUA meta-analyses (Supplementary Table 8); however, the LD-structure in the MHC locus is complex (14).

Integration with functional genomic data. Functional mapping and annotation of the 6 GWAS hits using the FUMA pipeline conservatively predicted five genes ZDHHC14, PARK2, KAZN, TMEM51-ASI and ZNF813 located in EUA risk loci, and five distinct genes LINCO2335, MIR5007, TUC338, LINCO2571 and HLA-B in AFA risk loci (Table 4). In addition, gene-based analyses on 18,222 protein-coding genes based on the EUA and AFA GWAS summary data identified two additional gene-wide significant loci, represented by SH3RF3 (\(P = 4.28 \times 10^{-57}\)) and PODXL (\(P = 2.37 \times 10^{-56}\)) in the EUA analysis. Gene-based analyses in AFA did not result in genome-wide significant loci. The biological function and potential psychiatric relevance of these 12 genes predicted by FUMA are detailed in the Supplementary Note 1 and discussed below.

We next performed gene-set analyses to understand implicated genes in the context of pathways and found four significant, Bonferroni-corrected gene sets (Supplementary Data 2). Of note, the two gene sets identified in the EUA data point towards a role for the immune system in PTSD. This is supported by a number of TNF-related genes summarized in a significant gene-set in AFA.

Annotation of variants in risk loci showed limited evidence of functionality (Table 4 and Supplementary Note 1). Most notably for the AFA top hit on chromosome 13, when testing for chromatin interactions using Hi-C data in neural progenitor cells, significant chromatin conformation interactions were observed between the risk region and a region \(\sim 1100\) kb upstream harboring additional non-coding RNAs including LINCO0458, hsa-mir-1297 and LINCO0558 as well as a region approximately 820 kb downstream harboring the pseudogene HNF4GP1 (Supplementary Fig. 24). eQTL analyses did not show significant associations with gene expression. However, the lack of functional data for this region may be explained by the African ancestry specificity of the GWAS findings since databases available within the FUMA framework, including GTEx and BrainSpan for eQTL analyses, are predominantly based on European populations.

Thus, we expanded our analyses for the AFA top hit through cell culture experiments in lymphoblastoid cell lines (LCLs) from African subjects (see Methods and Supplementary Note 1). We show evidence that the African-ancestry-specific SNP rs115539978 seems to capture a genomic region that may influence the expression of non-coding RNAs from this PTSD risk locus in response to increased glucocorticoid receptor signaling, thus linking this African-specific genetic variant to stress response and non-coding RNA expression (Supplementary Fig. 25).

We further characterized the AFA signal (rs115539978) using psychophysiology and imaging datasets available through the Grady Trauma Project (GTPC) and found evidence that this lead SNP captures a genomic region that is also associated with increased amygdala volume and fear psychophysiology in a traumatized population (Supplementary Note 1 and Supplementary Fig. 26).

Replication of findings in the external MVP cohort. In an attempt to replicate our genomics findings in an adequately-powered external study we used the large MVP cohort, including 146,660 EUA and 19,983 AFA participants assessed for
re-experiencing symptoms (REX), a core feature of PTSD. We first compared the genetic signals between PGC2 PTSD and MVP REX EUA and found a highly significant genetic correlation ($r_g = 0.80$, SE = 0.096, $P = 2.85 \times 10^{-17}$). No evidence of replication was found for the six leading PTSD markers identified in PGC2 EUA and AFA GWAS for the MVP REX-specific symptoms (Supplementary Table 9). However, two of these markers were not directly genotyped and had to be assessed by proxy markers in only moderately high (~75%) LD, and sex-stratified analyses were not available for MVP.

**Polygenic risk scores (PRS) for PTSD.** PRS generated from well-powered GWAS have recently become a tool of high relevance for polygenic disorders and traits (e.g. ref. 15,16). We assessed the predictive value of PRS for PTSD, using our largest cohort, the UKB, as a training sample (Fig. 2a). Our analyses were strongest at a $p$-value threshold $P_T = 0.3$ and showed a highly significant increase in odds to develop PTSD across PRS quintiles in the PGC1.5 EUA target sample, with a variance explained on the liability scale of $r^2 = 0.0015$ (likelihood ratio test $P = 5.44 \times 10^{-7}$). Analyses within the UKB show even stronger PRS predictability, with the highest OR for UKB men with a PRS trained on UKB women, reaching an OR of 1.39 in the 5th quintile, with an overall variance explained of $r^2 = 0.012$ ($P = 4.19 \times 10^{-10}$).

We also tested the overall PGC2 PTSD PRS in the external MVP replication cohort, using REX as the target for predictions. PRS predictions were strongest at $P_T = 0.3$ and highly significant (likelihood ratio test $P = 5.4 \times 10^{-62}$, Supplementary Fig. 27). Mean REX symptoms in MVP EUA participants was 8.48 (4.59 SD), and participants in the 5th quintile of genetic risk had significantly higher REX scores than subjects in the 1st quintile (beta = 0.58, $P = 1.41 \times 10^{-48}$, Fig. 2b).

**Genetic correlation of PTSD with other traits and disorders.** Analysis of shared heritability across common disorders of the brain and specific genetic correlations of psychiatric disorders with cognitive, anthropomorphic and behavioural measures has been facilitated greatly by the development of a centralized database of GWAS results including a web interface for LDSC (LD Hub21). We estimated pairwise genetic correlations ($r_g$) between PTSD and 235 disorders/traits and found 21 significant correlations after conservative Bonferroni correction (Fig. 3, panel A and Supplementary Data 3). Genetic variation associated with PTSD was positively correlated with PRS from other psychiatric traits including depressive symptoms, schizophrenia and neuroticism, as well as epidemiologically comorbid traits such as insomnia, smoking behavior, asthma, hip-waist ratio and coronary artery disease. In contrast, negative $r_g$ with PTSD include subjective...
well-being, education, and a strong correlation with parents’ age at death ($r_g = -0.70$). Significant positive correlations were also found for reproductive traits such as the number of children ever born, and, as previously reported for women\(^{22}\), PTSD was negatively correlated with age at first birth.

With the notable exception of asthma, our findings on PTSD correspond closely with genetic correlations between these traits and other psychiatric disorders such as MDD\(^{18}\), SCZ\(^{23}\), BPD\(^{19}\), and ADHD\(^{20}\) (Fig. 3, panels B-E). These findings are not surprising, as pleiotropic effects (i.e. SNPs impacting multiple traits) have been widely reported for psychiatric disorders.

In order to test to what degree genome-wide significant findings from our GWAS meta-analyses were specific to PTSD rather than driven by correlated traits as identified above, we adjusted the top hits from our analyses for the effects of genetically correlated psychiatric traits. Since the strongest correlations were found between PTSD and depressive symptoms ($r_g = 0.80$) and MDD ($r_g = 0.62$), summary statistics from PGC MDD\(^{18}\), as well as MDD plus BPD and SCZ, were included in the analyses. Using a recently implemented method applicable to external GWAS summary data (mtQCOJO\(^{24}\)) to approximate an analysis where these traits are regressed out, we found that effect sizes for the four EUA top hits were not markedly reduced when adjusted for the effects of MDD, or all three psychiatric traits tested simultaneously (Supplementary Tables 10, 11). These findings indicate that the genetic variants identified here are specific to PTSD when tested in the context of the three psychiatric disorders genetically most significantly correlated with PTSD.

**Discussion**

PTSD is a common and debilitating condition influenced by genetic factors, yet common genetic risk variants for PTSD have not been robustly identified. The PGC-PTSD combined data from 60 multi-ethnic cohorts (PGC1.5) and the UK Biobank (PGC2) achieved a sample size of 206,655 participants with 32,428 PTSD cases, over ten times that of any previous analysis\(^{10,23}\). As has been demonstrated in GWAS of SCZ\(^{23}\), BPD\(^{26}\), and recently in MDD\(^{18,27}\) and ADHD\(^{20}\), sample size is critical to produce robust genome-wide significant hits that inform foundational knowledge on the neurobiology of complex psychiatric conditions. These results show this is also true of PTSD. This increased power has led us to draw several major conclusions.

First, our genetic findings squarely place PTSD among the other psychiatric disorders in terms of heritability and genetic relationship with other disorders. While this statement may seem obvious to some, there remains debate about whether PTSD is entirely a social construction\(^{28}\). We found substantial SNP-based heritability (i.e. phenotypic variation explained by genetic differences) at 5–20%, similar to that for major depression\(^{18}\) across methods, studies and ancestries. The heritability results and pattern of genetic correlations are also consistent with our initial findings\(^{10}\) and with those from twin studies. PTSD shares common variant risk with other psychiatric disorders, which show substantial sharing of common variant risk with one another\(^{29}\). PTSD was most significantly (genetically) correlated with major depression, but also with schizophrenia, both of which have genome-wide significant loci implicated in brain function.

Second, our GWAS analyses identified several genetic loci not previously associated with PTSD. These loci pointed to a number of different target genes that merit further investigation. With PARK2, there is a posited role of dopaminergic systems in PTSD. The dopaminergic system has a critical role in fear conditioning which is important in the development and maintenance of PTSD\(^{30}\). There is also epidemiological evidence for association of
Looking deeper into the meaning of GWAS value of deep phenotyping in conducting functional investigations into the biology of PTSD. However, preliminary work from related disorders has now thought to influence risk, in part, through pruning of synapses using immune machinery rather than through classical immune pathways. These ancestry-specific results are preliminary, and even larger PTSD GWAS will facilitate the identification of plausible neurobiological targets for PTSD.

Third, our results also illustrate that there may be genetic contributions to the well-documented association between PTSD and dysregulation in inflammatory and immune processes. It has been widely recognized that PTSD is associated with a broad range of adverse physical health conditions over the life course ranging from type-2 diabetes and cardiovascular disease to dementia and rheumatoid arthritis. Less is known about the biological mechanisms driving the relationship between PTSD and these outcomes. Our genetic correlation analyses may provide some initial clues for further investigation. For example, we found a high genetic correlation ($r_g = 0.49, P = 0.0002$) between PTSD and asthma. Our subsequent gene-set and pathway analyses provide some clues further implicating the immune system. Of note, these genetic results converge with evidence from epidemiologic cohort studies documenting the role of stress-related disorders such as PTSD in autoimmune diseases, case-control studies showing elevations of immune-related biomarkers in women with PTSD, and epigenetic studies pointing to the role of the immune system in PTSD etiology. Further work is needed to determine whether PTSD has genetic overlap with immune disorders broadly and the causal direction between disorders. At minimum, the emerging genetic evidence presented here suggests that association between PTSD and health conditions may, in some cases, have some genetic origin.

Fourth, PGC-PTSD is distinct in relation to current genomics consortia due to its high proportion of data from participants of diverse ancestries. For example, a recent review found that only three percent of all samples in genetic studies were from African ancestry. This contrasts sharply with the 10% of AFA Parkinson disease and synapse formation. $PODXL$ is involved in neural development and synapse formation. $SH3RF3$ is associated with neurocognition and dementia. $ZDHHC14$ is associated with regulation of $\beta$-adrenergic receptors, and $KAZN$ is expressed in the brain, where it has been found to be underexpressed in parvalbumin neurons of the superior temporal cortex of schizoid cases, and overexpressed in the substantia nigra of Parkinson’s cases. Finally, the $HLA-B$ complex may be related through the known role of immunity and inflammation in stress-related disorders. Less is known about the role of the identified RNAs $LINC02335$, $MIRS007$, $TUC338$ and $LINC02571$ in regards to the biology of PTSD. However, preliminary work from our group including imaging and psychophysiology highlights the value of deep phenotyping in conducting functional investigations into the meaning of GWAS findings. Extensive follow-up work is needed to replicate our findings and to determine the function of identified genes and their relationship to putative pathological processes. For example, in SCZ the MHC locus is now thought to influence risk, in part, through pruning of synapses using immune machinery rather than through classical immune pathways. These ancestry-specific results are preliminary, and even larger PTSD GWAS will facilitate the identification of plausible neurobiological targets for PTSD.

Fourth, PGC-PTSD is distinct in relation to current genomics consortia due to its high proportion of data from participants of diverse ancestries. For example, a recent review found that only three percent of all samples in genetic studies were from African ancestry. This contrasts sharply with the 10% of AFA
participants in our consortium. We have the first heritability estimates for PTSD in African ancestry: they are similar to EUA, highly significant in women and lower in men. Our GWAS in subjects of African ancestry indicated at least one ancestry-specific locus using local ancestry methods developed for this analysis. We note the sample size in the AFA analysis has only about 15,000 participants, which is small and under-powered, increasing the chance for false positives. However, other work has shown that genetic studies of underrepresented populations afford the opportunity to discover novel loci that are invariant in European populations. As others have noted, there are major limitations in our knowledge of the genetic and environmental risk architecture of psychiatric disorders in persons of African descent. Our findings provide further evidence of the need to invest in research that includes diverse ancestral populations, to expand reference data, and to continue to develop methods to analyze data from such populations. Until such an investment is made, we are limited in our ability to understand biological mechanisms, predict genetic risk, and produce optimal treatments for non-European populations. African genomes are characterized by shorter haplotype blocks and contain many millions more variants per individual than populations outside Africa. Further, including data from African populations in genetic studies of PTSD and other neuropsychiatric disorders may accelerate genetic discovery and could be useful for fine mapping disease causing alleles.

Fifth, although PTSD heritability remained relatively stable across methods, studies, and ancestries, sex differences in heritability were observed in the overall cohort analyses as well as in the AFA analyses. It is of note that the sex differences in heritability were not evident in the UK Biobank data, which we hypothesize is due to differences between the PGC1.5 cohorts and the UK Biobank. PTSD is conditional on trauma exposure, which is also highly heterogeneous across individuals and populations. Unlike PGC1.5, the UK Biobank cohort is comprised of few to no subjects who experienced military-related trauma. In contrast, a substantial proportion of men in the PGC1.5 cohorts were from military cohorts, while virtually all women were civilians. The environmental experiences (e.g. military versus civilian) and index traumatic events leading to PTSD in male subjects versus female subjects could explain observed lower heritability estimates in males in the PGC1.5 cohorts. In future work, we aim to investigate this empirically by pooling detailed trauma and PTSD phenotypic information on both males and females and by modeling the effects of measurement variability on heritability estimates. Future work could also aim to increase samples of civilian men and military women to allow for analyses stratified by military trauma and sex.

Lastly, we report a significant polygenic risk score for PTSD, which also significantly predicts re-experiencing symptoms in independent data from the MVP. However, larger sample sizes are needed to achieve sensitivity and specificity at levels of clinical utility. In the future, polygenic risk may eventually be useful in algorithms developed to identify vulnerable persons after exposure to trauma. PTSD is one of the most theoretically preventable mental disorders, as many people exposed to trauma come to clinical attention in first response settings such as emergency rooms, intensive care units, and trauma centers. Controlled clinical trials show that PTSD risk can be significantly reduced by early preventive interventions. However, these interventions have nontrivial costs, making it infeasible to offer them to all persons exposed to trauma, given that only a small minority goes
on to develop PTSD\textsuperscript{61,62}. They are also unnecessary for many survivors who recover spontaneously\textsuperscript{35}. To be cost-effective, risk prediction rules are needed to identify which exposed persons are at high risk of PTSD. Such risk prediction tools have been developed\textsuperscript{63,64}, but none to date has included polygenic risk as a predictor.

These findings advance our understanding of the genetic basis of PTSD, but they also demonstrate that PGC-PTSD remains under-powered for the detection of most risk loci and associated pathways. PTSD is similar to major depression in both prevalence (among trauma-exposed persons) and in heritability. There are now 100 genome-wide significant signals for major depression; notably, that level of discovery required 246,363 cases and ancestrally diverse studies from Europe, Africa and the Americas. Of these, 12 were Participating studies that spanned the global diversity of PTSD and were nested within the modiﬁed pipeline (PGC-PTSD). The ancestry pipeline was shared with external sites in order to ensure consistency in ancestry calling across cohorts.

Genotype quality control. The standard PGC pipeline RICOPILI was used to perform QC, but modifications were made to allow for ancestrally diverse data. In the modified pipeline, each dataset was processed separately, including subjects of all ancestries. Sample exclusion criteria: using SNPs with call rates >95%, samples were excluded with call rates <98%, deviation from expected inbreeding coefficient (\(f_{het} < 0.2\) or >0.2), or a sex discrepancy between reported and estimated sex based on inbreeding coefficients calculated from SNPs on X chromosomes. Marker exclusion criteria: SNPs were excluded for call rates <98%, a >2% difference in missing genotypes between cases and controls, or being monomorphic. Hardy-Weinberg equilibrium (HWE) was assessed using the modiﬁed pipeline (PGC-PTSD) and highest homozygous ancestry group in the data, identified SNPs with a HWE P-value <1×10\(-6\) in controls, and excluded these SNPs in all subjects of the speciﬁc datasets, irrespective of ancestry.

Relatedness within studies. Within-study relatedness was estimated using the IBS function in PLINK \textsuperscript{19,99}. From each pair with relatedness \(r\) >0.2, one individual was removed from further analysis, retaining cases where possible.

Calculation of principal components (PC’s) for GWAS. For each dataset, unrelated subjects were subset into the three ancestry groups (EUA, AFA; AMA; Supplementary Tables 3, 5, 6) for analysis. SNPs were excluded that had a MAF <5%, HWE P >1×10\(-5\), call rate <98%, were ambiguous (A/T, G/C), or due to being located in the MHC region (chr. 6, 25–35 MB) or chromosome 8 inversion (chr. 8, 7–13 MB). SNPs with pairwise LD-pruned \((r^2 > 0.2)\) and a random set of 65,555 markers was used for each subset to calculate PC’s based on the smartPCA algorithm in EIGENSTRAT\textsuperscript{79}.

Imputation. Imputation was based on the 1000 Genomes phase 3 data (1KGP phase 3\textsuperscript{71}). Any dataset using a human genome assembly version prior to GRCh37 (hg19) was lifted over to GRCh37 (hg19). SNP alignment proceeded as follows: for each dataset, SNPs were aligned to the same strand as the 1KGP phase 3 data. For ambiguous markers, the largest ancestry group was used to calculate allele frequencies and only SNPs with MAF >40% and ≤15% difference between matching 1KGP phase 3 ancestry data were retained. Pre-phasing was performed using default settings in SHAPEIT2 v2.r837\textsuperscript{72} without reference subjects, and phasing was done in 3 megabase (MB) blocks, where an additional 1 MB of buffer was added to either end of the block. Haplotypes were then imputed using default settings in IMPUTE2 v2.2.2\textsuperscript{73}, with 1KGP phase 3 reference data and genome map, a 1 MB buffer, and effective population size set to 20,000. RICOPILI default filters for MAF and Info were removed since analyses were run across ancestry groups at this step. Imputed datasets were deposited with the PGC DAC and are available for approved requests.

Main GWAS. The analysis strategy for the main association analyses is shown in Supplementary Tables 3, 5 and 6. Analyses were performed separately for each study and ancestry group, unless otherwise indicated. The minimum number of subjects per analysis unit was set at 50 cases and 50 controls, or a total of at least 200 subjects, and subsets of smaller size were excluded. Smaller studies of similar composition were genotyped jointly in preparation for joint analyses (e.g. PSY1, PSY3). For studies with unrelated subjects, imputed SNP dosages were tested for association with PTSD under an additive model using logistic regression in PLINK 1.9, including the first five PCs as covariates. For family and twin studies (e.g. EVA, QIMR), analyses were performed using linear mixed models in GEMMA v0.96\textsuperscript{74} including a genetic relatedness matrix (GRM) as a random effect to account for population structure and relatedness, and the first five PCs as covariates. The UKB data (UKB) were analyzed with BGenie v1.2 (https://www.biorxiv.org/content/ early/2017/07/20/166298) using a linear regression with 6 PCs, and batch and center indicator variables as covariates (see Supplementary Methods for additional details). In addition, all GWAS analyses were also performed stratified by sex.
**Meta-analyses.** Summary statistics on the linear scale (from GEMMA and BGenie) were converted to a logistic scale prior to meta-analysis (for formula see [23]). Within each dataset and ancestry group, summary statistics were filtered to MAF ≥1% and PLINK INFO score ≥0.6. Meta-analyses across studies were performed within each of the three ancestry groups and across all ancestry groups. Inverse variance weighted fixed effects meta-analysis was performed with METAL (v. March25 2011)[32]. Heterogeneity between datasets was tested with a Cochran test and for nominally significant Q-values, a Han-Eskin random effects model (RE–HE) meta-analysis was performed with METAFAST v.2.0.1[17]. Markers with summary statistics in less than 25% of the total effective sample size or present in less than three studies were removed from meta-analyses. Quantile-quantile (QQ) plot of the effect estimates from the meta-analyses was visually inspected to confirm the presence of outliers. To account for multiple comparisons in analyses stratified by sex, genome-wide significance was also considered at P < 1.67 × 10⁻⁸. For genome-wide significant hits, Forest plots and PM-Plots were generated using the programs METASOFT with default settings and M-values were generated using the MCMC option[17,28]. For a given study and SNP, the M-value is the posterior probability that there is a SNP effect that study. Studies with M-values ≤0.1 are genotype data with no effect, values ≥0.1 and ≤0.9 are ambiguous, and values >0.9 are predicted to have an effect. In PM-plots, M-values are plotted against log₁₀ P-values. Regional association plots were generated using LocusZoom[29] with 400KB windows around the index variant and compared to the corresponding windows in the other ancestry groups, including the 1000 Genomes Nov. 2014 reference populations EUR, AFR, and AMR, respectively. No sex-specific analyses were performed as the difference in the effect estimates from male and female sex-stratified analyses was negligible.

**Comparability of PGC2 studies.** To compare the genetic signal between specific PGC2 subsets, LDSC[22] was used to estimate heritability and genetic correlations. Small EUA studies with N ≥200 cases and total effective sample size of N > 500 were selected (N = 24 studies; GWAS including 2102 cases and 7366 controls, effective N = 5162) and compared to larger studies. To reduce standard error given this relatively small sample, we estimated LD-score regression intercepts from the corresponding 1KGP phase ≥1% in the linear scale using REML in LDAK, including 5 PC chromosomes, informed by the HapMap combined b37 recombination map[30]. Markers with values ≥0.1 are predicted to have no effect, values <0.1 are ambiguous, and values >0.9 are predicted to have an effect. Heterogeneity between datasets was tested with a Cochran test and for nominally signif-

**Estimating PTSD heritability.** SNP-based heritability estimates (h²SNP) in EUA subjects were calculated using LDSC on meta-analysis summary data. Estimates were calculated for the combined PGC freeze 2 samples (PGC2) and separately for PGC1.5 (without UKB), the UK biobank (including alternative subject/phenotype selections), and for men and women. Unconstrained regression intercepts were used to account for potential inclusion of related subjects and residual population stratification, and precomputed LD scores from 1KGP EUR populations were used. For population prevalence we used a range of values (conservative low at 10%, moderate at 20%, high at 50%), based on prevalences reported for subjects exposed to different types of trauma[31]. Sample prevalence was set to the actual proportion of cases in each set of data. To estimate h²SNP in admixed individuals and compare h²SNP across different ancestries, individual-level genotype data was analyzed using an unweighted mixed model[11] as implemented in the LDKA software[31]. For each ancestry group (EUA and AMA), subjects with ≥1% ancestry dosage to databases containing known functional annotations, including Indians living in the region. For primary analyses, genome-wide signif-

**Local ancestry deconvolution.** A pipeline was developed to determine local ancestry in subjects with African and/or Native American admixture (AFA, AMA; Supplementary Fig. 28). Additional QC to consistently prepare cohort data for downstream analysis was performed with a custom script (https://github.com/eatkinson/Post-QC). Post-QC steps in the workflow extracting admiscusual data, removing duplicate loci, updating SNP IDs to dbSNP 144, orienting data to the 1KGP reference (with removal of indels and loci that either were not found in 1KGP or that had different coding alleles), flipping alleles that were on the wrong strand, and removing ambiguous SNPs. Genotyping was accomplished via a 723,305–SNP Affymetrix Axiom biobank array, customized for the MVP. Imputation was performed with Minimac[33] and the 1000 Genomes November 2014 reference (AFR, and Native American, respectively) at a threshold of p > 10⁻⁶ using linear regression with the first 10 principal components, age, and sex included as covariates. The results were filtered with imputation quality score R² ≥ 0.9, MAF > 0.01 and HWE test P-value > 1 × 10⁻⁶. LDSC was used to estimate genetic correlation with the PGEC EUA sample. The PGEC EUA GWAS summary statistics were merged and the merged dataset was then used to test for association between PRS and re-experiencing symptoms.

**Functional mapping and annotation.** We used Functional Mapping and Annotation of genetic associations (FUMA) v1.3.0 (https://fuma.github.io) to annotate GWAS data and obtain functional characterization of risk loci. Annotations are based on human genome assembly GRCh37 (hg19). FUMA was used with default settings unless stated otherwise. The SNPGene module was used to define independent genomic risk loci and variants in LD with lead SNPs (r² > 0.6, calculated using ancestry appropriate 1KGP reference genotypes). SNPs in risk loci were mapped to protein–coding genes with a 10 kb window. Functional consequences for SNPs were obtained by mapping the SNPs on their chromosomal position and reference alleles to databases containing known functional annotations, including ANNOVAR, Combined Annotation Dependent Depletion (CADD), RegulomeDB, and chromatin states (only brain tissues/cell types were selected). Next eQTL mapping was performed on significant (FDR q < 0.05) SNP–gene pairs, mapping to GTeX v7 brain tissue, RNA-seq data from the CommonMind Consortium and the BRAINEAC database. Chromatin interactions mapping was performed using built-in chromatin interaction data from the dorsolateral prefrontal cortex, hippocampus and neuronal progenitor cell line. We used a FDR q < 1 × 10⁻⁵ to define significant interactions, based on previous recommendations.
modified to account for the differences in cell lines used here. SNPs were also checked for previously reported phenotypic associations in published GWAS listed in the NHGRI-EBI catalog.

**Gene-based and gene-set analysis with MAGMA.** Gene-based analysis was performed with the FUMA implementation of MAGMA. SNPs were mapped to 18,222 protein-coding genes. For each gene, its association with PTSD was determined as the weighted mean $\chi^2$ test statistic of SNPs mapped to the gene, where LD patterns were calculated using ancestry appropriate 1KG reference genotypes. Significance of genes was set at a Bonferroni-corrected threshold of $P = 0.05/18,222 = 2.7 \times 10^{-8}$.

To see if specific biological pathways were implicated in PTSD, gene-based test statistics were used to perform a competitive set-based analysis of 10,894 predefined curated gene sets and GO terms obtained from MxSigDB using MAGMA. Significance of pathways was set at a Bonferroni-corrected threshold of $P = 0.05/10,894 = 4.6 \times 10^{-6}$. To test if tissue-specific gene expression was associated with PTSD, gene-set-based analysis was also used with expression data from GTex v7 RNA-seq and BrainSpan RNA-seq, where the expression of genes within tissue was used to define the gene properties used in the gene-set analysis model.

**Functional follow-up of the AFA top hit rs115539978.** Cell Culture Experiments, RNA extraction and qPCR. Lymphocytes from cell lines (LCLs) from the superpopulation were obtained from the Coriell Institute, NJ (Supplementary Table 13, N = 6 lines each for the homzygous major and homzygous minor allele). Cells were cultured in RPMI 1640 medium with GlutaMAX (Thermo Scientific, 61870-036) supplemented with 15% FBS (Thermo Scientific, 26140079) and 1X Antibiotic-Antimycotic (Thermo Scientific, 15240-062) at 37°C and 5% CO2 in a humidified incubator. For Dexamethasone (Dex) treatment, a final concentration of 100 nM Dex (Sigma-Aldrich) in 100% Ethanol was added to the medium for a total of 4 hr. All experiments were run in duplicates.

RNA was extracted using the Quick-RNA MiniPrep Kit (Zymo, R2062) according to the manufacturer's instructions. SYBR green qPCR reactions were carried out in duplicates using POWERUP SYBR Green Master Mix (Life Technologies, A25743) using the SuperScript IV First Strand Kit (Life Technologies, 18091200) according to the manufacturer's recommendations. Data were analyzed using the ΔΔCT method$^{48}$ and GAPDH as reference. Between group differences were calculated using one-way ANCOVA with sex as covariate. Significance threshold was set at $P = 0.05$.

**Deep phenotyping of the AFA top hit rs115539978.** Neuroimaging: Scanning of 87 GTPC subjects took place on a 3.0 T Siemens Trio with echo-planar imaging (Siemens, Milwauk, PA). High-resolution T1-weighted anatomical scans were collected using the 3D MP-RAGE sequence, with 176 contiguous 1 mm sagittal slices (Siemens, Malvern, PA). High-resolution T1-weighted anatomical scans were collected using a 3D MP-RAGE sequence, with 176 contiguous 1 mm sagittal slices. Preprocessing of the functional data was performed following the ENIGMA 2 protocol (https://surfer.nmr.mgh.harvard.edu). Data quality checks were performed following the ENIGMA 2 protocol. Subcortical structures were extracted through automated segmentation, and from subcortical structures was extracted through automated segmentation, and from subcortical structures was extracted through automated segmentation. In addition, a 30% prevalence, using the equation found in Lee et al$^{8}$. P-values for PTSD were derived from a likelihood ratio test comparing the two models.

**Genetic correlation of PTSD with other traits and disorders.** Bivariate LD-score regression (LDSC) was used to calculate pairwise genetic correlation ($r_g$) between PTSD and 235 traits with publicly available GWAS summary statistics on LD Hub$^{12}$. Summary statistics for PTSD studies were restricted to the EUA meta-analysis, including UKB subjects (23,212 cases, 151,447 controls) and significance was evaluated based on a conservative Bonferroni correction for 235 phenotypes (i.e., excluded traits and traits measured twice in independent studies were counted independently).

In addition, these phenotypes were compared with genetic correlations reported for PTSD and several psychiatric disorders, including 221 phenotypes and MDD$^{19}$, 172 phenotypes and Schizophrenia (SCZ)$^{19}$, 196 phenotypes and bipolar disorder (BPD)$^{19}$ and 219 phenotypes and attention-deficit/hyperactivity disorder (ADHD)$^{20}$. Due to substantial overlap with other traits, two education, four anthropometric and two cancer phenotypes were omitted.

**Conditional analyses to test for disease specific effects.** To evaluate if the effects of the top variants identified in the PTSD GWAS were specific to PTSD, we conditioned PTSD on MDD, and MDD plus BP D plus SCZ using the multi-trait conditional and joint analysis (mtCOJO)$^{24}$ feature in GCTA to regress out the effects of correlated traits based on external GWAS summary data. MDD was selected here as the main psychiatric trait because of the high co-morbidity and genetic correlation of depressive symptoms and PTSD ($r_g = 0.60$ for depressive symptoms and $r_g = 0.70$ for MDD; see Supplementary Data 3). Publicly available summary statistics were supplied as program inputs: Bipolar cases vs. controls for BP D, and MDD excluding 23andMe for MDD (both from https://www.med. unc.edu/pgc-results-and-downloads); Schizophrenia: CLOZUK + PGCG meta-analysis for SCZ (https://walters.pyscm.cf.ac.uk/). The effect of each psychiatric trait on PTSD was estimated using a generalized summary-data-based Mendelian randomization analysis of significant LD independent psychiatric trait SNPs ($r^2 > 0.5$, based on 1000 G Phase 3 CEU samples), where the threshold for significance was set to $P < 5 \times 10^{-7}$ due to having less than the required 10 significant independent SNPs at the program default $P < 5 \times 10^{-8}$ for MDD. Estimates of heritability, genetic correlation, and sample overlap of psychiatric trait and PTSD GWAS were estimated using precomputed LD scores based on 1000 G Europeans that were supplied with LDSC (https://data.broadinstitute.org/alkesgroup/LDSCORE/eur_w_ld_chr.tar.bz2).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The full meta-analyses 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 authors upon request.

**References**

1. Ben Barnes, J., Hayes, A. M., Contractor, A. A., Nash, W. P. & Litz, B. T. The structure of co-occurring PTSD and depression symptoms in a cohort of Marines pre- and post-deployment. Psychiatry Res. 259, 442–449 (2018).
2. Koenen, K. C. et al. Post-traumatic stress disorder and cardiometabolic disease: improving causal inference to inform practice. Psychol. Med. https://doi.org/10.1017/S0033291716002294 (2016).
3. Pollard, H. B. et al. “Soldier’s Heart”: a genetic basis for elevated cardiovascular disease risk associated with post-traumatic stress disorder. Front. Mol. Neurosci. 9, 87 (2016).
4. Howlett, J. R. & Stein, M. B. Prevention of trauma and stressor-related disorders: a review. *Neuropsychopharmacology* **41**, 357–369 (2016).

5. Shalev, A., Liberman, I. & Marmar, C. Post-traumatic stress disorder. *N. Engl. J. Med.* **376**, 2459–2469 (2017).

6. Kremen, W. S., Koenen, K. C., Afari, N. & Lyons, M. J. Twin studies of posttraumatic stress disorder: differentiating vulnerability factors from sequelae. *Neuropsychopharmacology* **62**, 647–653 (2017).

7. Wu, C. J. et al. A classical twin study of PTSD symptoms and resilience: evidence for a single spectrum of vulnerability to traumatic stress. *Depress Anxiety* https://doi.org/10.1002/da.22712 (2017).

8. Nieweg, C. M. et al. Genomic approaches to posttraumatic stress disorder: the psychiatric genomic consortium initiative. *Biol Psychiatry* **83**, 831–839 (2018).

9. Gelmenter, J. et al. Genome-wide association study of post-traumatic stress disorder reexperiencing symptoms in >165,000 US veterans. *Nat. Neurosci.* **22**, 1394–1401 (2019).

10. Duncan, L. E. et al. Largest GWAS of PTSD (N=20,070) yields genetic overlap with schizophrenia and sex differences in heritability. *Mol. Psychiatry* https://doi.org/10.1038/s41386-017-0020-4 (2017).

11. Sullivan, P. F. et al. Psychiatric genomics: an update and an agenda. *Am. J. Psychiatry* **175**, 15–27 (2018).

12. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *PLoS Genet.* **8**, e1002555 (2012).

13. Han, B. & Eskin, E. Interpreting meta-analyses of genome-wide association studies. *PLoS Genet.* **8**, e1002832 (2012).

14. Matzaraki, V., Kumar, V., Wijmenga, C. & Zhernakova, A. The MHC locus from polygenicity in genome-wide association studies. *Trends Genet.* **32**, 5–9 (2016).

15. Hill, R. & Eskin, E. Interpreting meta-analyses of genome-wide association studies. *PLoS Genet.* **8**, e1002832 (2012).

16. Park, J. H. et al. Estimation of effect size distribution from genome-wide association studies. *Mol. Psychiatry* **18**, 76 (2017).

17. Brainstorm, C. et al. Analysis of shared heritability in common disorders of the brain. *Science* https://doi.org/10.1126/science.aap8757 (2018).

18. Green, R. et al. A classical twin study of PTSD symptoms in 2.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).

19. Park, J. H. et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat. Genet.* **42**, 570–575 (2010).

20. Brainstorm, C. et al. The invention of post-traumatic stress disorder and the social construction of a psychiatric category. *Am. J. Psychiatry* **176**, 1274–1281 (2018).

21. Demontis, D. et al. Discovery of the first genome-wide significant risk locus for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63 (2019).

22. Zheng, J. et al. LD Hub: a centralized database and web interface to perform large-scale genome-wide association analysis. *PLoS ONE* **12**, e0186857 (2017).

23. Sklar, P. et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near OX1A. *Nat. Genet.* **43**, 977 (2011).

24. Howard, D. M. et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343 (2019).

25. Zhub, Z. et al. Cross associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).

26. Stein, M. B. et al. Genome-wide association studies of posttraumatic stress disorder in 2 cohorts of US Army soldiers. *JAMA Psychiatry* **73**, 695–704 (2016).

27. Kalisch, R., Gerlicher, A. M. & Duvarci, S. A dopaminergic basis for fear extinction. *Trends Cogn. Sci.* **23**, 274–277 (2019).

28. Sun, X. et al. Post-traumatic stress disorder and risk of parkinson disease: a nationwide longitudinal study. *Am. J. Geriatr. Psychiatry* **25**, 917–923 (2017).

29. Kaasinen, V. & Vahberg, T. Striatal dopamine in Parkinson disease: a meta-analysis of imaging studies. *Ann. Neurol.* **82**, 873–882 (2017).

30. Lee, J. H., Lee, S. & Kim, J. H. Amygdala circuits for fear memory: a key role for dopamine regulation. *Neuroscientist* https://doi.org/10.1177/1073858416679936 (2016).
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-019-12576-w.

Correspondence and requests for materials should be addressed to C.M.N.

Peer review information Nature Communications thanks Andrew McIntosh, Cathryn Lewis and Noah S. Philip for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019
Emory University, Department of Gynecology and Obstetrics, Atlanta, GA, USA.
Boston University, Dean’s Office, Boston, MA, USA.

New York University School of Medicine, Department of Psychiatry, New York, NY, USA.
United States Army, Command, Fort Sill, OK, USA.

University of Adelaide, Department of Psychiatry, Adelaide, South Australia, AU, Australia.
VA Boston Health Care System, GRECC/TRACTS, Boston, MA, USA.
Harvard University, Department of Psychology, Boston, MA, USA.
UNC Institute for Trauma Recovery, Department of Emergency Medicine, Chapel Hill, NC, USA.
Gallipoli Medical Research Institute, PTSD Initiative, Greenslopes, Queensland, AU, Australia.
Queensland University of Technology, School of Psychology and Counseling, Faculty of Health, Kelvin Grove, QLD, AU, Australia.
Aarhus University Hospital, Psychosis Research Unit, Risskov, DK, Denmark.
Aarhus University, Centre for Integrated Register-based Research, Aarhus, DK, Denmark.
University of Adelaide, Department of Psychiatry, Adelaide, South Australia, AU, Australia.
VA Boston Health Care System, GRECC/TRACTS, Boston, MA, USA.

New York University School of Medicine, Department of Psychiatry, New York, NY, USA.
United States Army, Command, Fort Sill, OK, USA.

University of Texas Health Science Center at San Antonio, Department of Psychiatry, San Antonio, TX, USA.
University of Washington, Department of Psychiatry, Seattle, WA, USA.
U.S. Department of Veterans Affairs National Center for Posttraumatic Stress Disorder, West Haven, CT, USA.
Minneapolis VA Health Care System, Department of Mental Health, Minneapolis, MN, USA.

Harvard T.H. Chan School of Public Health, Department of Environmental Health, Boston, MA, USA.

Medical University of South Carolina, Department of Nursing and Department of Psychiatry, Charleston, SC, USA.
Maastricht Universitair Medisch Centrum, School for Mental Health and Neuroscience, Department of Psychiatry and Neuropsychology, Maastricht, Limburg, NL, Netherlands.
Universidad Peruana de Ciencias Aplicadas Facultad de Ciencias de la Salud, Department of Medicine, Lima, Lima, PE, USA.
University of Michigan, School of Nursing, Ann Arbor, MI, USA.
University of New South Wales, Department of Psychiatry, Sydney, NSW, AU, Australia.
Columbia University Medical Center, Department of Medicine, New York, NY, USA.
Mental Health Centre Sct. Hans, Institute of Biological Psychiatry, Roskilde, DK, Denmark.
Oslo University Hospital, KG Jebsen Centre for Psychosis Research, Norway Division of Mental Health and Addiction, Oslo, NO, USA.
University of South Florida College of Public Health, Genomics Program, Tampa, FL, USA.

Uniformed Services University, Department of Psychiatry, Bethesda, Maryland, USA.

Charité - Universitätsmedizin, Department of Psychiatry and Psychotherapy, Berlin, Germany.
Harvard Medical School, Department of Psychiatry, Boston, MA, USA.

Medical University of South Carolina, Department of Psychiatry and Behavioral Sciences, Charleston, SC, USA.
University of Copenhagen, Department of Clinical Medicine, Copenhagen, Denmark.
James J Peters VA Medical Center, Department of Mental Health, Bronx, NY, USA.
Baylor Scott and White Central Texas, Department of Psychiatry, Temple, TX, USA.
Columbia University, Department of Biostatistics, New Haven, CT, USA.
University of Washington, Department of Psychiatry and Behavioral Sciences, Seattle, WA, USA.

*Email: cnievergelt@ucsd.edu