Genomic Approaches to Posttraumatic Stress Disorder: the Psychiatric Genomic Consortium Initiative

Caroline M. Nievergelt\textsuperscript{1,2}, Allison E. Ashley-Koch\textsuperscript{3}, Shareefa Dalvie\textsuperscript{4}, Michael A. Hauser\textsuperscript{3}, Rajendra A. Morey\textsuperscript{5}, Alicia K. Smith\textsuperscript{6,7}, and Monica Uddin\textsuperscript{8,9}

\textsuperscript{1}University of California San Diego, Department of Psychiatry and Department of Family Medicine and Public Health
\textsuperscript{2}Veterans Affairs San Diego Healthcare System and Veterans Affairs Center of Excellence for Stress and Mental Health
\textsuperscript{3}Department of Medicine, Duke University Medical Center, Durham, NC, United States
\textsuperscript{4}Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa, 7925
\textsuperscript{5}Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham NC 27710, Durham VA Medical Center, Durham, NC 27705
\textsuperscript{6}Emory University, Department of Gynecology and Obstetrics
\textsuperscript{7}Emory University, Department of Psychiatry & Behavioral Sciences
\textsuperscript{8}University of Illinois Urbana-Champaign, Carl R. Woese Institute for Genomic Biology
\textsuperscript{9}University of Illinois Urbana-Champaign, Department of Psychology

Abstract

Post-traumatic stress disorder (PTSD) after exposure to a traumatic event is a highly prevalent psychiatric disorder. Heritability estimates from twin studies as well as from recent molecular data (\(h^2\text{ SNP}\)) indicate moderate to high heritability, yet robust genetic variants for PTSD have not yet been identified and the genetic architecture of this polygenic disorder remains largely unknown.

To date, less than ten large-scale genome-wide association studies (GWAS) of PTSD have been published, with findings that highlight the unique challenges for PTSD genomics, including a complex diagnostic entity with contingency of PTSD diagnosis on trauma exposure, and the large genetic diversity of the study populations.

Corresponding Author: Caroline Nievergelt, Ph.D., Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0667, Tel. 858-246-1895, cnievergelt@ucsd.edu.

Publisher’s Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures

The authors report no biomedical financial interests or potential conflicts of interest.
The Psychiatric Genomics Consortium PTSD group (PGC-PTSD) has brought together over 200 scientists with the goal to increase sample size for GWAS and other genomic analyses to sufficient numbers where robust discoveries of molecular signatures can be achieved. The sample currently includes over 32,000 PTSD cases and 100,000 trauma-exposed controls and collection is ongoing. First results found a significant shared genetic risk of PTSD with other psychiatric disorders, and sex-biased heritability estimates with higher heritability in females compared to males.

This review describes the scope and current focus of the PGC-PTSD, and its expansion from the initial GWAS group to nine working groups, including epigenetics, gene expression, imaging, and integrative systems biology. We further briefly outline recent findings and future directions of ‘omics-based’ studies of PTSD, with the ultimate goal of elucidating the molecular architecture of this complex disorder to improve prevention and intervention strategies.

Keywords
Psychiatric genomics consortium (PGC); genetics; epigenetics; gene expression; imaging; systems biology

Genetics of PTSD in the context of other psychiatric disorders

PTSD is a debilitating psychiatric disorder precipitated by traumatic experience, with subsequent pathological re-experiencing, avoidance, negative alterations in cognitions and mood, and hyperarousal symptoms (DSM-5 (1)). While most individuals exposed to trauma are resilient, PTSD prevalence is directly related to the severity and type of trauma, with rape or direct combat conferring very high risk, and lifetime risk in women is twice that in men (2). PTSD prevalence varies by country, but lifetime prevalence in the US is over 7% – making it among the most common psychiatric disorders (3).

Genetic factors influence who develops PTSD; family and twin studies have estimated heritability of PTSD from ~40-70% following trauma (4–7). However, despite over a decade of research on genetic candidate genes, robust genetic variants for PTSD have yet to be identified and the genetic architecture of this polygenic disorder remains largely unknown.

To address this gap in knowledge, the field of psychiatric genomics has moved its focus over the last decade from small studies on specific candidate genes (8) to agnostic, genome-wide association studies (GWA studies) and ultimately to well-powered, large-scale meta-analyses made possible through efforts such as the Psychiatric Genomics Consortium (PGC) (9). Recent success in the identification of robustly associated genetic variants in psychiatric disorders such as schizophrenia (10), bipolar disorder (11) and major depressive disorder (MDD) (12) has confirmed that very large sample sizes are necessary to discover loci with the small genotypic relative risks typically seen in psychiatric disorders. Accordingly, the short-term goal of the PGC is to obtain GWAS data on 100,000 cases for each of its nine disorders (9).
**Genome-wide association approaches to PTSD**

PGC-PTSD ([https://pgc-ptsd.com/](https://pgc-ptsd.com/)) was initiated in 2013 by bringing together four groups with published GWA studies of PTSD (13–16), making it a relative latecomer to the PGC (17). To date, four additional association studies of PTSD with genome-wide data have been published (18–21) and at least one more study in Danish soldiers has been completed (Wang et al., pers. com.). Typically, these studies present a finding that meets criteria for genome-wide significance plus evidence of replication in at least one independent cohort (Table 1). While this has been the gold standard for GWA studies, none of the identified genes in these GWA studies has robustly replicated across multiple studies. With the rapid availability of PGC-PTSD summary data on a large number of studies, best practice guidelines for GWAS replication are currently being discussed (e.g. [https://www.cohenveteransbioscience.org/2017/06/29/psychiatric-genomic-consortium-workshop-summary/](https://www.cohenveteransbioscience.org/2017/06/29/psychiatric-genomic-consortium-workshop-summary/)) and can now include the pre-specified selection of specific replication cohorts matched for example on ancestry, gender, and trauma type.

The PGC-PTSD has adopted pipelines and protocols established by the PGC ([https://data.broadinstitute.org/mpg/ricopili/](https://data.broadinstitute.org/mpg/ricopili/)), which facilitates integration of data across disorders (e.g. (22)). However, PGC-PTSD has to consider some unique challenges not faced by other groups, including the contingency of PTSD diagnosis on trauma exposure, a complex and changing diagnostic entity (23), and very diverse genetic ancestry within and across study cohorts, resulting in considerable heterogeneity (17).

To address ancestral diversity, the PGC pipeline was extended to include an ancestry inference module, which allows for stable ancestry determination across studies ([https://github.com/nievergeltlab](https://github.com/nievergeltlab)). It was designed to be portable to enable PGC-PTSD studies that cannot share individual-level genotype data (e.g. some US military and non-US cohorts) to generate summary-level results for meta-analysis. The majority of GWA studies to date have been performed in subjects of European (EA) and African American (AA) ancestry groups, while carefully addressing residual population stratification (Figure 1). With the sustained PGC-PTSD efforts to increase sample size, other ancestries such as Latinos and East Asians are reaching considerable size. Trans-ethnic GWAS have been generated using meta-analytical approaches (16, 24), and alternative mega-analysis methods are currently employed to leverage all available samples, irrespective of ancestry (25).

The first PGC-PTSD publication in 2017 included 11 multi-ethnic studies of 5,000 PTSD cases and 15,000 controls (freeze 1) and is largest published genetic association study of PTSD to date. SNP-based heritability ($h^2_{SNP}$) estimates for PTSD were ~29% in females, but substantially lower in males (24), consistent with lower twin-based heritability estimates in males. In addition, the study found significant shared genetic risk of PTSD with schizophrenia, bipolar disorder and MDD. Investigating the implications of sex-based heritability in PTSD and cross-disorder genetic risks are high priority for future study design.
Expanding the PGC-PTSD scope

Although successful in demonstrating some aspects of the genetic architecture of PTSD, the PGC-PTSD freeze 1 was underpowered to identify genome-wide significant loci (24). Thus, expanding the sample size to sufficient numbers for GWAS remains one of the main goals of the PGC-PTSD. The current freeze 2 includes over 32,000 PTSD cases and 100,000 trauma-exposed controls (Nievergelt et al., manuscript in preparation), approaching the number of cases for which other PGC studies showed first robust discoveries (26).

In addition, future analyses will also explore analytical models that are potentially stronger than conventional case-control analyses, including quantitative symptom scores and clusters as well as trauma exposure, similarly to a recent meta-analysis of GWA studies on anxiety disorder (27). The grouping of sub-clinical PTSD cases with trauma-exposed controls is a potential limitation of the conventional PGC case-control studies.

A promising development in the PGC-PTSD is the expansion from the initial GWAS group to nine integrated working groups (see Figure 2). While some working groups such as the physical health (28), psychophysiology, and imaging groups have extended the phenotype (i.e., PTSD diagnosis) with highly relevant additional phenotypes, other working groups have assembled complimentary ‘omic’ type data such as copy-number variants (CNV), methylation, and gene-expression. A microbiome group has recently been initiated. Finally, the systems biology group is charged with integration of these multiple data types to maximally leverage data resources for discovery. A rational for some of these efforts is discussed below.

Gene Expression analysis in PTSD

Analysis of gene expression in PTSD presents challenges, as the underlying molecular events causing this disorder likely occur in the central nervous system (CNS), and archival of postmortem brain tissue from individuals affected by PTSD lags behind other CNS disorders such as Parkinson’s and Alzheimer’s Disease. However, considerable data also point to PTSD being characterized by systemic immune and metabolic perturbations caused by stress-responsive changes in the hypothalamic–pituitary–adrenal (HPA) axis (29–31). These systemic changes give rise to differential gene expression signatures in peripheral blood of PTSD cases versus controls that can serve as biomarkers for disease, and can also provide insight into disease-associated systemic effects on immune function and organ pathology. Moreover, the changes occurring in the periphery may be promoting or exacerbating changes in the brain (32). Thus, gene expression analysis of peripheral blood of PTSD cases and controls is being performed by the PGC-PTSD not merely as a matter of convenience, but because it is likely to illuminate critical disease processes and potentially identify individuals most at-risk for PTSD.

A number of gene expression studies in peripheral blood have already been reported. While statistical power of some of these studies has been limited by small sample size (33–38), others have reported gene expression changes that were statistically significant after rigorous correction for multiple testing (39–41). These studies have replicated a few differentially
expressed genes: \textit{USP48} was identified as differentially expressed in PTSD cases versus controls by two studies (39, 40), and \textit{DICER1} was similarly identified by two studies (40, 41). While in general there is little concordance of specific genes between studies, pathway and gene network analyses have consistently and reproducibly identified differential expression of transcripts involved in innate immunity, interferon signaling, and wound healing (42, 43). These studies have provided valuable insights into the pathobiology of PTSD. With a combined sample size of almost 5000 samples, future peripheral blood gene expression studies by the PGC-PTSD should refine and extend these findings. As CNS tissue samples become increasingly available, gene expression data from post-mortem brains can be compared and integrated with findings from peripheral blood.

**PTSD Epigenetics**

PTSD is unique among psychiatric disorders in that it requires a traumatic environmental event as part of its diagnosis. Among the different epigenetic modifications, DNA methylation has received the most attention by researchers studying psychosocial stress, childhood trauma, and PTSD due to its relative stability and its ability to be assessed with microarrays that facilitate replication within and between studies (40, 44–50). Immune dysregulation figured prominently among the biologic pathways associated with PTSD and are replicable between studies (44, 46). A recent study examined DNA methylation along with microRNA (miRNA), another epigenetic modification, in a small group of PTSD cases and controls. The authors noted reductions in miRNA levels in PTSD cases and proposed that epigenetic changes may contribute to systemic inflammation in PTSD (51). These studies echo those of other psychiatric disorders that emphasize the cross-talk between the peripheral immune system and the brain (32, 52).

Other studies suggest more widespread mechanisms for epigenetic dysregulation in PTSD. For example, Maddox and colleagues reported DNA methylation differences in \textit{HDAC4}, a histone deacetylase, in the blood of women with PTSD and went on to show that variation in genetic and epigenetic predictors of \textit{HDAC4} expression associated with fear-potentiated startle response and functional connectivity differences in the amygdala (53). Similarly, lower expression of \textit{DICER1}, which is required for processing mature miRNAs, is associated with PTSD cases with comorbid depression and increased amygdala activation in response to fearful stimuli (41, 54), a neural phenotype strongly associated with risk for PTSD even prior to trauma exposure (reviewed in (55)).

In addition to these genome-scale approaches, epigenetic summary measures may be particularly informative. The most widely used measure is the ‘epigenetic clock’ (56, 57) for assessing age acceleration, which is associated with psychosocial stress and higher mortality risk (57, 58). The phenomenon describes methylation-based prediction of age that exceeds chronological age. Of note, a relatively high proportion – almost 25% of epigenetic clock-related CpG sites are located in glucocorticoid response elements (GREs) – a genomic region in which methylation levels vary in relation trauma exposure (59) and dexamethasone suppression (57). These environmentally sensitive genomic sites have been explicitly linked to traumatic stress, neural integrity and mortality (60), discussed in more detail below,
further demonstrating the utility of using epigenetic summary measures as an index of the biologic impact of lived experience, including trauma exposure.

There are numerous challenges in conducting epigenetic studies of PTSD. Similar to gene expression discussed above, DNA methylation patterns vary by tissue type, with the majority of studies to date conducted in blood, whereas the most relevant tissue is brain. As new postmortem samples and single-cell technologies become available, there will be substantial advancement in identification of genes whose regulation is altered in those with PTSD. This may further support the identification of peripheral biomarkers or may restrict the scope of peripheral tissues. A second challenge is limited platforms, and thus limited coverage, for DNA methylation or other arrays that are widely used for population-scale studies. Epigenome-wide investigations require cost-effective and highly reproducible methods to achieve the sample sizes required to detect associations that withstand multiple testing correction. The PGC-PTSD EWAS group (Figure 2) has the goal to assemble such data to perform meta-analyses across cohorts with a common multi-site analysis pipeline (61). In some ways, these platform limitations increase opportunities for replication, but also complicate linking genome-wide discoveries with those based on sequencing or targeted assays.

PTSD Imaging Genetics

Elevated risk of psychopathology may be more powerfully investigated with intermediate phenotypes (or endophenotypes) than clinical diagnoses. Brain measures from MRI may have a simpler underlying genetic architecture involving fewer individual genes or pathways than the polygenicity driving overall risk for psychopathology (62), and offer a more precise and reproducible phenotype than clinical diagnostic scales (63). A GWAS of continuous brain measures may be statistically more powerful and more efficient than binary traits (diagnosis), which may disguise complexities such as co-morbidity and syndromal heterogeneity (64). Furthermore, brain phenotypes may provide common pathways for the combined effects of environmental and genetic risk factors that may underlie multiple diagnoses (62). However, there are two important caveats (1) the effect sizes for gene effects on neuroimaging phenotypes are unlikely to be greater than for behavioral traits or psychiatric disorders (65), and (2) genetics of brain phenotypes may reveal common mechanistic pathways for a number of psychiatric disorders resulting in a loss of specificity when moving from psychiatric disorder to brain phenotype – different disorders may possess nearly identical brain abnormalities (66). This loss of specificity may prove advantageous for drug development, by facilitating the design of a target-specific intervention that is effective for multiple neuropsychiatric disorders.

An international collaboration of investigators (17) within the PGC and Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) plans to investigate the genetic effects of complex brain traits (66). The first major analysis of the PGC-ENIGMA PTSD Working Group with 1,868 samples has demonstrated that PTSD is associated with smaller hippocampus and amygdala volume (67). Exposure to childhood trauma was negatively associated with hippocampal and amygdala volume when adjusting for age, sex, and intracranial volume (67). Both structures have ample a priori evidence implicating their role
in PTSD starting with the report of reduced hippocampal volume in a small PTSD sample over 20 years ago (68). However, we confirmed this finding across a large number of demographically and clinically heterogeneous cohorts analyzed with a standardized segmentation technique, and a harmonized analysis protocol across all sites. Methodological consistency was promoted by using the same statistical models across all samples, making this the largest and most powerful study of subcortical volumes in PTSD to date. The analysis of 12 hippocampal subfield volumes, DTI measures of 26 white matter tracts, the cortical thickness of 78 regions-of-interest, and whole-brain vertex-based analyses are currently underway using ENIGMA pipelines, which can be downloaded at https://pgc-ptsd.com/methods-tools/imaging-pipeline/.

Forthcoming analyses gene-by-environment (GxE) GWAS of relevant structural brain phenotypes with childhood trauma as major risk factor are planned with the long-term goal of identifying genetic modulators of brain structure that help early prediction and treatment for a range of psychiatric disorders, followed by deep sequencing in a subset of samples to identify potential causal variants within the coding and/or regulatory regions of implicated risk loci. Over 40 participating sites have coalesced around the common goal to form ENIGMA-PGC-PTSD, which has already received 4,000 samples among which 3,000 samples have been aggregated and analyzed. Nevertheless, the analyses are expected to be woefully underpowered with the large number of phenotypes available in neuroimaging data. Two approaches address the shortcomings of previous neuroimaging-genetics studies that have been plagued by small sample size due to the large expense of MRI acquisition, and the use of candidate genes that have been criticized for being susceptible to population stratification and fueling information bottlenecks. First, no candidate gene analyses are planned. Second, replication samples of neuroimaging data such as the UK Biobank and the Million Veteran Program (MVP) will be leveraged. Third, we will focus on polygenic risk score (PRS) calculation and PRS x E interaction analyses. Discovery samples for calculating PRS include (1) the GWAS of the PTSD diagnosis from the PGC from 80,000 samples, (2) the GWAS of subcortical volumetry performed from 31,000 normative neuroimaging and genomic samples that provided several SNP associations (69), and (3) the GWAS of cortical thickness and surface area in which preliminary results from 30,000 samples has generated 120 SNPs that show robust genome-wide significant associations after correction for 78 cortical structures. The testing of PRS that are calculated in discovery samples, which have yielded robust genome wide associations, in our sample avoids the criticisms previously leveled against neuroimaging studies (70).

Addressing the inaccessibility of brain tissue

To date, EWAS studies of PTSD have been limited to accessible peripheral tissues, specifically whole blood (40, 44, 46). While potentially informative as biomarkers of the disorder, the extent to which PTSD-associated DNA methylation patterns in blood or other peripheral tissues reflect patterns that may exist in the brain remains unknown. As the target organ of most interest to the disorder, the brain remains a challenge to access in living individuals, and brain-based epigenetic predictors of PTSD have yet to be identified.
Despite these challenges, recent work has attempted to bridge the link between brain and periphery by examining peripherally derived epigenetic biomarkers of neuroimaging-based phenotypes and endophenotypes of PTSD. Much of this work to date has adopted a candidate gene approach (e.g. (71, 72)). A notable exception is a study conducted by Wolf and colleagues (60) which tested the hypothesis that PTSD is associated with accelerated cellular age and degraded neural integrity, as well as reduced performance on executive function tasks, in a sample of U.S. veterans. Lifetime PTSD severity was found to be positively associated with DNA methylation-derived age, and these age estimates were negatively associated with neural integrity in the genu of the corpus callosum and working memory performance. Of note, the mean age in this sample was ~32 years, suggesting that the neurobiological effects of traumatic stress may impair neuropsychological functioning of veteran populations. This genome-scale study represents a genomic approach to PTSD with high public health impact, as it holds the potential to identify individuals who may be most in need of intervention by leveraging peripherally-derived, polygenic epigenetic measurements that are predictive of neural integrity and memory performance. Cohorts within the PGC-PTSD that include both neuroimaging and EWAS data are optimally positioned to combine efforts and pursue similar studies in the future, augmented by the meta-analytic framework currently being employed by both the EWAS and the neuroimaging working groups within the PGC-PTSD.

Data integration and systems biology approaches

Based on current findings, the underlying etiology of PTSD is likely the result of a complex interplay between various molecular systems (73), and to delineate this, a holistic (systems biology) approach which integrates different ‘omics layers’ amongst PTSD cases and trauma-exposed controls may be the next logical step. One of the strengths of the PGC-PTSD is that multiple forms of genomic data have been generated for many datasets, which will allow cross-platform analyses to be performed (Figure 2). Several methods have been proposed for meta-analysis across platforms (74–77), as well as for imputing data to improve the ability to combine datasets (78–82). Such joint analyses of multiple –omics datasets have been previously termed “genomic convergence” and have great potential to inform the genetic architecture of PTSD (83).

Through the co-ordination of the various PGC-PTSD working groups (Figure 2), there is potential to identify DNA methylation patterns that may be giving rise to altered gene expression, and sufficient power to conduct eQTL and meQTL analyses, which assesses whether particular genetic variants alter the levels of expression and methylation of specific genes, respectively. For example, previous studies for PTSD have shown that the risk variant, rs363276 located within an intronic region of SLC18A2, is an eQTL, significantly associated with decreased expression of the genes SLC18A2 and PDZD8 in the dorsolateral prefrontal cortex (DLPFC) of post-mortem human brains (84). Another PTSD risk variant rs717947, located at chromosome 4p15, was shown to be a meQTL (85). These types of analyses provide clues on the etiology of PTSD; however, identifying definitive causal risk factors requires alternate methods, such as Mendelian Randomization (MR), which quantifies causality between a risk factor and a disease outcome by utilizing SNP data as an instrumental variable (86). This technique has been applied to PTSD whereby a causal
relationship was identified between plasma dopamine beta-hydroxylase (DBH), an enzyme which catalyzes the synthesis of norepinephrine, on symptoms of re-experiencing (87). Another study investigated the relationship between BMI adjusted weight circumference (WCadj) and PTSD in women and found that an increase in WCadj results in a relative decrease in the risk of developing PTSD (88).

Current data integration efforts will be augmented through the increasing availability of publically available data sets, such as GTex (89) and PsychENCODE (90), which can provide additional annotation for the PTSD-specific findings. For example, recently developed methods such as PrediXcan uses genome-wide variation to predict or “impute” gene expression in test datasets, based on tissue-dependent modeling performed on transcriptome data from reference databases such as GTex. The imputed expression can be tested for association to the phenotype of interest enabling the identification of trait associated loci (91). Additionally, as technologies continue to evolve and become more widely-available for single cell RNA sequencing (92, 93), induced pluripotent stem cell (iPSC)-derived neural progenitor cells (94) and brain organoids (95–97), these technologies can also be integrated into the ongoing PGC-PTSD efforts, particularly with respect to understanding the roles of specific genes and variants in PTSD risk.

Studies of this nature fill a significant gap in the available literature on the complex genetic mechanisms and pathways underlying complex psychiatric disorders, including PTSD. Systems biology approaches will lay the groundwork for future development of more accurate diagnostic methods, improved management and the development of more suitable and individualized treatment strategies for patients.

### Conclusion

The promise of finding genetic determinants of psychiatric disorder is identifying etiologic pathways for targeted interventions. However, before this can become a reality, biological validation of genetic findings will be required. Making individual prediction to support the emerging discipline of precision medicine holds the promise of personalized medical decisions driven by an individual’s genetic make-up and environment, other risk factors, and large databases of genotype-phenotype relationships.

### Acknowledgments

We thank the research groups contributing to the PGC-PTSD working groups for sharing their data and time to make this work possible, and especially the over 80,000 research participants world-wide who shared personal experiences and biological samples with these research studies.

Financial support was provided by the Stanley Center for Psychiatric Genetics at the Broad Institute, One Mind, Cohen Veterans Bioscience, and the National Institute of Mental Health (NIMH) grants R01-MH111671-01 (RAM, AAK, MAH), and NIMH/U.S. Army Medical Research and Materiel Command support for R01MH1106595 (CMN) and R01MH108826 (AKS, MU, CMN). AAK and MAH were supported by the Department of Veterans Affairs’ (VA) Mid-Atlantic Mental Illness Research, Education, and Clinical Center (MIRECC) and the Research & Development and Mental Health Services of the Durham Veterans Affairs Medical Center.
References


Biol Psychiatry. Author manuscript; available in PMC 2019 May 15.


Biol Psychiatry. Author manuscript; available in PMC 2019 May 15.
43. Breen MS, Tylee DS, Maihofer AX, Neylan TC, Mehta D, Binder E, et al. PTSD Blood Transcriptome Mega-Analysis: Shared Inflammatory Pathways Across Biological Sex and Modes of Trauma. Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology. 2017


Figure 1. Ancestry composition and main ancestry groups analyzed in the PGC-PTSD
Principal component (PC) plot showing inferred ancestry of > 80,000 subjects from 56
studies. Main population groups analyzed to date include homogeneous subjects of
European ancestry (N>40,000), and one-way admixed African Americans (N≈20,000) and
Latinos, including Native Americans (N>6,500). With expansion of the data collection,
additional populations, such as East Asians, will reach optimal sizes for analysis. The plot is
based on PC1 and PC4 to highlight the Asian populations.
The PGC-PTSD GWAS group has drastically expanded its scope since its initiation in 2013. It currently includes working groups with emphasis on complimentary phenotypes (psychophysiology, physical health, and imaging), working groups contributing complimentary ‘omics’ data (CNV, epigenetics, transcriptome, and microbiome), as well as a systems biology group aiming at integration of the different types of data. Arrows are indicating primary flow of data, but interactions among groups are expanding.
## Table 1

Genome-wide significant hits and replication in individual PTSD studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Hits</th>
<th>Cohort</th>
<th>N PTSD</th>
<th>N controls</th>
<th>Sex</th>
<th>Trauma</th>
<th>P-value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logue et al. (14)</td>
<td>RORA</td>
<td>NCPT</td>
<td>295</td>
<td>196</td>
<td>EA</td>
<td>pred. M</td>
<td>2.50E-08</td>
<td></td>
</tr>
<tr>
<td>replication replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guffanti et al. (13)</td>
<td>LINC01090</td>
<td>DNHS</td>
<td>43</td>
<td>41</td>
<td>AA</td>
<td>pred. M</td>
<td>0.05</td>
<td>gene-base test gene-base test</td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xie et al. (15)</td>
<td>COBL</td>
<td>GSD</td>
<td>300</td>
<td>1,273</td>
<td>EA</td>
<td>M+F</td>
<td>3.96E-08</td>
<td></td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xie et al. (15)</td>
<td>TLL1</td>
<td>GSD</td>
<td>207</td>
<td>1692</td>
<td>EA</td>
<td>M+F</td>
<td>2.99E-07</td>
<td>marginal GWAS hit</td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solovieff et al. (18)</td>
<td>SLC18A2</td>
<td>DNHS</td>
<td>845</td>
<td>1,093</td>
<td>EA</td>
<td>F</td>
<td>2.10E-05</td>
<td>candidate-gene study</td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nievergelt et al. (16)</td>
<td>PRTFDC1</td>
<td>MRS</td>
<td>940</td>
<td>2,554</td>
<td>mixed M M</td>
<td>military</td>
<td>2.04E-09</td>
<td></td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stein et al. (20)</td>
<td>ANKRD55</td>
<td>NSS</td>
<td>497</td>
<td>815</td>
<td>AA</td>
<td>pred. M</td>
<td>2.34E-08</td>
<td></td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stein et al. (20)</td>
<td>ZNF626</td>
<td>NSS</td>
<td>2,140</td>
<td>2,909</td>
<td>EA</td>
<td>pred. M</td>
<td>4.59E-08</td>
<td></td>
</tr>
<tr>
<td>replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Ancestry: EA: European ancestry; AA: African American
2. Trauma: cohorts have been separated by military and civilian; individual type of trauma is not considered here