An epigenome-wide association study of posttraumatic stress disorder in US veterans implicates several new DNA methylation loci

Mark W. Logue1,2,3,4*, Mark W. Miller1,2, Erika J. Wolr1,2, Bertrand Russ Huber1,2, Filomene G. Morrison1,2, Zhenwei Zhou4, Yuanchao Zheng4, Alicia K. Smith5,6, Nikolaos P. Daskalakis7,8,9,10, Andrew Ratanatharathorn11, Monica Uddin12,13, Caroline M. Nieverget14,15,16, Allison E. Ashley-Koch17, Dewleen G. Baker14,15,18, Jean C. Beckham19,20,21, Melanie E. Garrett17, Marco P. Boks22, Elbert Geuze22,23, Gerald A. Grant24, Michael A. Hauser17, Ronald C. Kessler25, Steven A. Schichman39, Regina E. McGlinchey40,7, William P. Milberg40,7, Jasmeet P. Hayes1,2,41, Mieke Verfaellie2,42 and the Traumatic Stress Brain Study Group

Abstract

Background: Previous studies using candidate gene and genome-wide approaches have identified epigenetic changes in DNA methylation (DNAm) associated with posttraumatic stress disorder (PTSD).

Methods: In this study, we performed an EWAS of PTSD in a cohort of Veterans (n = 378 lifetime PTSD cases and 135 controls) from the Translational Research Center for TBI and Stress Disorders (TRACTS) cohort assessed using the Illumina EPIC Methylation BeadChip which assesses DNAm at more than 850,000 sites throughout the genome. Our model included covariates for ancestry, cell heterogeneity, sex, age, and a smoking score based on DNAm at 39 smoking-associated CpGs. We also examined in EPIC-based DNAm data generated from pre-frontal cortex (PFC) tissue from the National PTSD Brain Bank (n = 72).
(Continued from previous page)

**Results:** The analysis of blood samples yielded one genome-wide significant association with PTSD at cg19534438 in the gene G0S2 ($p = 1.19 \times 10^{-7}$, $\rho_{adj} = 0.048$). This association was replicated in an independent PGC-PTSD-EWAS consortium meta-analysis of military cohorts ($p = 0.0024$). We also observed association with the smoking-related locus cg05575921 in AHRR despite inclusion of a methylation-based smoking score covariate ($p = 9.16 \times 10^{-9}$), which replicates a previously observed PGC-PTSD-EWAS association (Smith et al. 2019), and yields evidence consistent with a smoking-independent effect. The top 100 EWAS loci were then examined in the PFC data. One of the blood-based PTSD loci, cg04130728 in CHST11, which was in the top 10 loci in blood, but which was not genome-wide significant, was significantly associated with PTSD in brain tissue (in blood $p = 1.19 \times 10^{-5}$, $\rho_{adj} = 0.060$, in brain, $p = 0.00032$ with the same direction of effect). Gene set enrichment analysis of the top 500 EWAS loci yielded several significant overlapping GO terms involved in pathogen response, including “Response to lipopolysaccharide” ($p = 6.97 \times 10^{-6}$, $\rho_{adj} = 0.042$).

**Conclusions:** The cross replication observed in independent cohorts is evidence that DNA methylation in peripheral tissue can yield consistent and replicable PTSD associations, and our results also suggest that some PTSD associations observed in peripheral tissue may mirror associations in the brain.

**Introduction**

Genetic studies of posttraumatic stress disorder (PTSD) diatheses conducted to date have focused primarily on identifying DNA variants (e.g., single-nucleotide polymorphisms, SNPs) that confer risk for the development of the disorder through candidate gene or genome-wide association studies (GWASs; see, e.g., [1, 2]). More recently, studies have also examined differences between PTSD cases and controls in patterns of gene expression [3] and/or DNA methylation (DNAm [4–6]). DNAm studies involve measurement of a methyl group on the DNA strand at a cytosine-phosphate-guanine (CpG) site, and when this is present in the promoter region of a gene, DNAm tends to be negatively correlated with the expression of the gene. DNAm across the genome can be influenced by a host of genetic and developmental mechanisms, health conditions, and environmental factors ranging from toxin exposure to stress, and it is widely hypothesized to be a mechanism that mediates the effects of trauma exposure on gene expression [7, 8]. Hypothesis and mechanism-focused candidate gene studies have identified PTSD-related differences in DNAm levels in genes associated with the hypothalamic-pituitary-adrenal (HPA) axis (e.g., ADCYAP1 [9], FKBPs [10], and NR3C1 [11]), inflammation (e.g., BDNF [4], HTR2A [12] and IL-18 [13]), and neurotransmission (e.g., BDNF [4], HTR2A [14], and HTR3A [15]). Epigenome-wide association studies (EWASs), on the other hand, take a hypothesis-free approach to identifying DNAm loci from across the genome that are statistically associated with the phenotype of interest. To date, only five published PTSD EWASs have reported single-site associations that survived epigenome-wide multiple-testing correction. First, using a DNAm bead chip that interrogated ~27K loci in a sample of 100 subjects from an urban community cohort, Smith et al. (2011) reported false discovery rate (FDR)-corrected differences between PTSD cases and controls at loci in 5 genes (ACP5, ANXA2, CLEC9A, TLR8, and TPR) [4]. Second, in a study that used a more comprehensive platform measuring methylation at ~850K loci in samples from 96 Australian Vietnam veterans, Mehta et al. (2017) found genome-wide significant associations between DNAm and PTSD in four genes (BRSK1, DOCK2, LCNB8, and NGF) and in one intergenic locus [5]. Rutten et al. (2018) examined pre- to post-deployment changes in DNAm in a cohort of 93 soldiers using a ~450K platform and found 17 loci in or near 8 genes that were associated with increasing PTSD symptoms over time [6]. In that study, replication analyses in a similar pre- and post-deployment cohort of 98 soldiers also showed nominal support for findings in 3 genes (HIST1H2APS2, RNF39, and ZFP57). In an EWAS of methylation in sperm cells from a cohort of Veterans (16 with PTSD and 22 controls), Mehta et al. identified three loci reaching genome-wide significance: two intergenic loci and a CpG in CCDC88C [16]. This could point to a possible role of these loci in the inter-generational transmission of the effects of trauma. Finally, Smith et al. (2017) recently reported results of the largest PTSD EWAS conducted to date based on a meta-analysis of $n = 1896$ participants from 10 cohorts with methylation assessed at ~450K loci. Ten loci achieved genome-wide significance, the most significant of which, cg05575921 ($p = 4.27 \times 10^{-11}$, FDR = 2.15 $\times 10^{-4}$), was located in the smoking-associated gene AHRR. Many of the EWASs of PTSD have also used Gene Set Enrichment Analysis or Functional Network Analysis of genome-wide DNAm as a follow up to their genome-wide association analyses. Significant enrichment has been observed for a number of pathways/biological processes with plausible relevance to PTSD including, most notably, inflammation and...
immortalizing, HPA axis and glucocorticoid signaling, neurogenesis and neurotransmission, circadian rhythms, and cell adhesion [5, 6, 18–20].

To summarize, prior studies of DNAm associations with PTSD have yielded potentially important insights into the epigenetics of PTSD; however, with several noteworthy limitations. All of them were based on DNAm from peripheral samples (usually blood). Though potentially useful for the development of diagnostic biomarkers, blood samples provide only indirect evidence of epigenetic processes in the brain. In addition, most of the sample sizes studied to date have been modest, and the few EWAS-significant associations that have been reported have been accompanied by limited evidence of replication.

Another major consideration involves the influence of potential confounding variables such as cigarette smoking in PTSD studies. Numerous studies have shown the methylene to be exquisitely sensitive to the effects of smoking. The largest EWAS of cigarette smoking conducted to date by Joehanes et al. (N ~ 16k), identified 18760 CpGs on 7201 genes that were differentially methylated in current versus never smokers [21]. Results of that study confirmed numerous prior reports of associations between smoking and DNAm in several genes including, most notably, AHR, F2RL3, and RARA. Because cross-sectional epidemiological studies do not permit inferences about whether DNAm associations are causes, consequences, or effects of third variables, Li et al. (2018) used a genetically informative twin cohort to examine the heritability of the smoking-associated DNAm loci and found strong evidence that most of the observed epigenetic associations were attributable to the effects of cigarette use [22]. These findings are highly relevant to the epigenetics of PTSD because PTSD samples, especially from veteran cohorts, tend to have an elevated prevalence of cigarette smoking relative to the general population [23]. The AHR locus cg05575921 that was associated with PTSD in the Smith et al. EWAS [17] was also highly significant in both the Joehanes et al. EWAS of smoking [21] and the Li et al EWAS of smoking [22], and in the latter, was the most significant locus. Based on this, Smith et al. also performed analyses stratified by smoking status and found that the association between PTSD and cg05575921 was strongest among non-smokers suggesting an association between PTSD and AHR independent of smoking.

The aim of this study was to identify individual CpG sites and/or sets of genes associated with PTSD using DNAm data from a veteran cohort and a newly-established United States (US) Department of Veterans Affairs (VA) National PTSD Brain Bank [24]. Given the robust associations between smoking and DNAm observed in prior studies, we addressed this important confound by computing a DNAm smoking score based on the top loci from the Li et al. (2018) study and including it as a covariate in our analyses. The primary analysis was a “Discovery” EWAS of lifetime PTSD diagnosis in DNA from whole blood drawn from a cohort of veterans of the post-9/11 conflicts in Iraq and Afghanistan. We then evaluated evidence for replication of the top EWAS results in (a) an analysis of veterans performed as part of the Smith et al. EWAS, and (b) postmortem pre-frontal cortical tissue from the National PTSD Brain Bank. Finally, we examined candidate gene regions and candidate CpGs implicated in previous PTSD epigenetic studies.

Materials and methods
Discovery Cohort
The Discovery Cohort was comprised of veterans recruited by the Translational Research Center for TBI and Stress Disorders (TRACTS), a Department of Veterans Affairs Rehabilitation Research and Development (RR&D) Traumatic Brain Injury Center of Excellence at VA Boston Healthcare System. After applying a quality control pipeline and excluding participants with missing data (see Additional file 1), 541 veterans (n = 378 cases and 135 controls) with DNAm data were available for analyses. Demographic characteristics for all three cohorts are listed in Additional file 1: Table S1. The collection of Discovery Cohort data was performed with the approval of a Department of Veterans Affairs human subjects review board and all subjects provided written informed consent.

Consortium Replication Cohort
We examined results from the recent Smith et al. Consortium EWAS of current PTSD. As it was the closest match to our Discovery Cohort, we requested results from a military cohort meta-analysis. It included 1351 subjects from 7 cohorts: Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS) [25], Marine Resilience Study (MRS) [26, 27], Prospective Research in Stress-related Military Operations (PRISMO) [28, 29], The African American and the European American cohorts from the Mid-Atlantic Mental Illness Research Education and Clinical Center PTSD Study (VA-M-AA and VA-M-EA) [30], the Injury and Traumatic Stress study (INTRuST; see e.g., [31–33]), and the VA Boston Healthcare System National Center for PTSD (VA-NCPTSD) cohort [1]. In the replication cohort analysis, 42% were current PTSD cases, all of which were assessed with Illumina Infinium HumanMethylation450 BeadChip (450K BeadChips; see Smith et al 2019 for details). As it was not available for many participating cohorts, lifetime PTSD was not analyzed.
PTSD Brain Bank
DNA was extracted from the ventromedial prefrontal cortex (vmPFC; Brodmann area 12/32) and dorsolateral prefrontal cortex (dlPFC, Brodmann area 9/46) from 42 PTSD cases and 30 controls. These regions were selected based on findings from functional and structural imaging studies of PTSD suggesting their involvement in the neurobiology of the disorder (see e.g., [34]). A detailed description of the donor identification, post-mortem diagnostic assessment procedures, and tissue extraction and processing is presented elsewhere [24, 35] and in Additional file 1.

Generation of genotype and DNAm data
Genome-wide genotype data were used to generate principal components (PCs) to control for ancestral heterogeneity in each analysis. Genotypes for the Discovery and Brain Bank Cohorts were based on data from Illumina (San Diego CA) HumanOmni2.5-8 BeadChips as described previously [1, 36] and in the Supplementary Methods (Additional file 1). DNAm was assessed in the Discovery and the Brain Bank Cohorts using the Illumina EPIC 850K BeadChips. An EWAS consortium-derived pipeline was used to clean the DNAm data from both cohorts [37]. Additional information about the quality control (QC) pipeline is available in Additional file 1.

Data analyses
Analyses examining associations between DNAm and lifetime PTSD case/control status in the Discovery Cohort were performed with linear models as automated in the Bioconductor limma (Linear Models for Microarray Data) package [38] with the base 2 logit-transformed methylated proportion as the response and PTSD diagnosis as a predictor. Note that we will avoid the use of the term “beta” to refer to coefficient estimates from linear models throughout to avoid confusion with the term beta as it applies to DNAm studies, where it is used to refer to the proportion of DNAm at a given locus. Each limma model included principal components for ancestry, age, sex, estimates of white blood cell proportions, and a DNAm-based “smoking score” as covariates. The latter was based on effect-size estimates for the top-39 probes from a recent smoking EWAS [22]. This score showed highly significant association with self-reported smoking in both the Discovery Cohort and in the VA-NCPTSD cohort \( (p < 2.2 \times 10^{-16}) \) in both cohorts, and was significantly correlated with the number of cigarettes per day in the Discovery Cohort \( (r = 0.52, \ p < 2.2 \times 10^{-16}) \). See Supplementary Methods (Additional file 1) for details about computation of the smoking score.

For the EWAS in the Discovery Cohort, we computed false discovery rate [39] corrected \( p \)-values, also known as Q-values, to control for multiple testing (denoted \( p_{\text{adj}} \)). We then examined the top 100 associated sites from the EWAS in the 450K consortium results and also in the Brain Bank Cohort. In cases in which the observed top associated EPIC locus was not assayed by the 450K BeadChip, when available, we examined correlated “proxy sites” for evidence of replication. A proxy site was defined as a CpG assayed by both EPIC and 450K BeadChips within 5000 bp of a peak EPIC locus that was significantly correlated \( (p < 0.05) \) to the peak EPIC site. When we observed multiple correlated sites to a peak EPIC locus, we took the one with the highest correlation as the proxy. The genes corresponding to the top 500 sites from the Discovery EWAS were examined for enrichment of specific gene ontology (GO) term categories using the gometh function from the R missMethyl package [40]. This function is an extension of the GOseq method [41] which explicitly models the relationship between the number of CpG sites measured within a gene and the probability of that gene appearing within the target list, hence avoiding one of the sources of bias which can influence gene set enrichment analyses. Next, we examined candidate genes and CpG sites previously implicated in studies of DNAm and PTSD. We began by evaluating the significance of the 41 previously implicated CpG sites in blood-based studies of PTSD [4–6, 9, 12, 19, 36, 42, 43] (listed in Additional file 1: Table S2) in the limma EWAS results. Then, we performed a candidate gene examination of 36 previously implicated genes, by examining all sites within a gene from the limma output for association with PTSD using a gene-wide FDR correction based on the number of probes in each gene.

Post-hoc analyses, including evaluation of the role of smoking in the top EWAS associations and also the effects of other potential confounders (depression, depression severity, selective serotonin reuptake inhibitor use, and alcohol use) in the Discovery Cohort, were performed in R using the standard package for linear modeling (lm). As DNAm values can be influenced by nearby SNPs, we also examined the possibility of SNP effects for our top-associated loci. We ran a model similar to what was performed in limma above, but which included as a covariate imputed SNP dosage of any nearby genomic variants noted in the Illumina Annotation with minor allele frequency > 5%. See the Supplementary Materials (Additional file 1) for a more complete description of these analyses.

When analyzing the data from the PTSD brain bank, both regions (vmPFC and dlPFC) were analyzed jointly to increase power, reduce multiple testing, and to focus our attention on methylation differences that are consistent across the PFC. A linear mixed model was used which included a random effect to adjust for the correlations of DNAm between the two brain regions in the
same subject. The model for the PFC included covariates for age, sex, ancestry PCs, and proportion of neurons as estimated from the methylation data [44]. The smoking score when applied to DNAm data from the PFC was found not to correlate with smoking, and indeed, an investigation of smoking in the PFC did not yield any EWAS-significant smoking loci (see Additional file 1 for details). This is consistent with a smaller effect of smoking on methylation in the PFC compared to blood. However, given the modest sample size, we could not exclude such an effect, and smoking as determined by family report and medical history was included as a covariate in the PFC analyses. Although the Brain Bank Cohort was used here primarily to replicate the results from the EWAS performed in blood, we computed the results for the entire genome, so that we could examine the genome-wide distribution of \( p \)-values for inflated significance. Follow-up analyses in limma examined the dlPFC and vmPFC regions separately to explore whether significant associations were jointly observed across regions.

As we have elsewhere shown that a large proportion of low variation probes have poor correlation across chips due to a low signal-to-noise ratio [45], we excluded probes from the EWAS and candidate gene analyses when the range of the proportion of DNAm was < 0.10; these were primarily sites where the DNAm proportion was < 0.10 for all subjects or the DNAm proportion was > 0.90 for all subjects. This criterion excluded 417,270 of the 819,877 probes passing quality filters (see Additional file 1 for details), leaving 402,607 sites in the blood-based EWAS. However, these low variation probes were analyzed in limma and retained for the purposes of replication of previously reported loci.

**Results**

**Discovery Cohort EWAS**

There was no evidence of inflation for Discovery EWAS analysis of 402,607 probes (lambda = 1.066; Additional file 1: Fig. S1). Manhattan plots (−log10 \( p \)-values for each locus plotted relative to their genomic position) are presented in Fig. 1, and the top 10 most strongly associated loci are listed in Table 1. One epigenome-wide significant association was found involving probe cg19534438 in the G0/G1 Switch 2 gene (G0S2) on chromosome 1 (coefficient = 0.34, \( p = 1.19 \times 10^{-7} \), \( \hat{p}_{\text{adj}} = 0.048 \); see Fig. 2a). The other top 10 results included probes in BBS9, RCCD1, NCK1, CHST11, TMLHE, 3 intergenic loci, and AHRR (see Additional file 1: Fig. S2). The AHRR result, involving cg05575921 (Fig. 2b), was particularly noteworthy because this probe was one of the 39 used to compute the DNAm smoking score that was included as a covariate in the EWAS. To clarify the nature of this association and determine if it was an artifact of comorbid cigarette use in the PTSD cases, we performed post-hoc analyses examining the association between this AHRR probe, self-reported smoking, the smoking score, and PTSD under various conditions including (a) a model of PTSD with the smoking score excluded, and (b) a model of PTSD with the smoking score included but with cg05575921 excluded from score calculation. In both cases, the association between cg05575921 and PTSD remained robust (\( p < 10^{-4} \); see Supplementary Results and Additional file 1: Tables S3 and S4 for details). As Smith et al. reported that the association between cg05575921 and PTSD was more significant in non-smokers than in smokers, we performed a similar follow-up analysis of subjects segregated by smoking status. Although cg05575921 was not significant in either smokers or non-smokers, perhaps due to the reduction of sample size, there is a more substantial (negative) effect size estimate for cg05575921 in the non-smokers than the smokers (in non-smokers, coefficient = −0.13, \( p = 0.059 \); in smokers, coefficient = −0.048, \( p = 0.79 \)). Post-hoc analyses also showed that the EWAS-significant association involving cg19534438 in G0S2, was not affected by the inclusion or exclusion of self-reported smoking or the DNAm-based smoking score in the model (Additional file 1: Table S4).

**Consortium Military Replication Cohort**

Of the top 100 loci from our EWAS, consortium 450K meta-analysis results were available for 54, and proxies

![Fig. 1 Manhattan plot of an epigenome-wide association study of PTSD in US Veterans, the one EWAS significant locus at G0S2 is highlighted in green](image-url)
were available for an additional 6. We observed 8 nominally significant associations in the replication cohort, more than would be expected under the null ($p = 0.0098$ based on a binomial distribution with 5% chance of success). Across the 60 loci, the effect size estimates were significantly correlated ($r = 0.39$, $p = 0.0022$, see Fig. 3a). The correlation estimate was higher for the $n = 14$ loci with $p < 0.10$ in the Replication Cohort (orange and red points in Fig. 3a; $r = 0.58$, $p = 0.030$) and in the 8 loci with $p < 0.05$ in the Replication Cohort (red points in Fig. 3a; $r = 0.72$, $p = 0.043$). All loci with $p < 0.10$ in the Replication Cohort had the same direction of effect as the Discovery Cohort, which is highly unlikely under the null hypothesis of a 50% chance of agreement ($p = 0.00012$). Five of the loci assessed in the Replication Cohort remained significant after correcting for the 60 loci examined (Table 2). Unsurprisingly, the AHRR locus noted in our analysis was

### Table 1  Top 10 strongest associations from the Discovery Cohort EWAS ($n = 378$ cases and $135$ controls)

| Gene ID | Coefficient | p-value  | p_adj |
|---------|-------------|----------|-------|
| G0S2 cg19534438 | 0.34 | $1.19E-7$ | 0.048 |
| BBS9 cg20152234 | 0.18 | $1.83E-6$ | 0.28 |
| Intergenic cg11504264 | 0.24 | $2.09E-6$ | 0.28 |
| Intergenic cg08000207 | −0.35 | $6.64E-6$ | 0.60 |
| AHRR cg05575921 | −0.13 | $9.16E-6$ | 0.60 |
| RCCD1 cg25526519 | −0.16 | $9.66E-6$ | 0.60 |
| NCK1 cg09423651 | −0.53 | $1.13E-5$ | 0.60 |
| CHST11 cg04130728 | 0.15 | $1.19E-5$ | 0.60 |
| TMLHE cg12111516 | 0.25 | $1.53E-5$ | 0.68 |
| Intergenic cg20974659 | 0.11 | $1.68E-5$ | 0.68 |

Fig. 2 Box/Scatter plot of the proportion of DNAm (Beta) for (a) cg19534438, the genome-wide significant locus in G0S2, (b) cg05575921, the smoking- and PTSD-associated locus in AHRR, and (c) the CHST11 peak locus from the EWAS of whole-blood samples that was corrected significant in the analysis of tissue from the PFC, with blood, dlPFC, and vmPFC methylation plotted separately.
significant in the Consortium Military analysis. Two related genes, \textit{APBA1} and \textit{APBA2} also replicated. The \textit{APBA1} association was based on cg13580827, a proxy for the EPIC-only probe cg06826552, while the \textit{APBA2} locus was present on both chips. The \textit{G0S2} locus, which was genome-wide significant in our EWAS, was likewise associated with PTSD in the Consortium Military Replication Cohort ($p = 9.18 \times 10^{-4}$, $p_{\text{adj}} = 0.014$). Finally, cg25526519 in the \textit{LOC285696} gene also replicated ($p = 0.0033$, $p_{\text{adj}} = 0.040$).

**Brain Bank Cohort results**

When we compared the effect size estimates across the Discovery Cohort top 100 EWAS loci in the Brain Bank PFC samples, we did not observe the same correspondence observed between the two blood-based cohorts. In fact, we observed a significant negative correlation ($r = -0.30$, $p = 0.0022$, see Fig. 3b). This was driven primarily by cg12186981 in OR2AG1, which, of the top 100 blood results we attempted to replicate in brain, had the largest effect size estimated in both blood and brain, but with an opposite direction (in blood: effect size = 1.12, $p = 8.49 \times 10^{-5}$, in brain effect size = −1.15, $p = 0.033$). When this point was excluded, the correlation in effect size estimates across blood and brain was still negative, but not significant ($r = -0.096$, $p = 0.334$). Of the 100 loci examined in the PFC analysis, one, cg04130728 in the \textit{CHST11} gene, survived multiple-testing correction (Table 3). This probe was among the top 10 results in the Discovery Cohort, and the direction of effect was the same across the two cohorts. Analysis of the dIPFC and vmPFC separately in limma indicated that the associations observed with \textit{CHST11} were consistent across both brain regions (in dIPFC, coefficient = 0.32, $p = 0.0034$; in vmPFC, coefficient = 0.30, $p = 0.011$, see Fig. 2c). Information on this locus in the Replication Cohort was not available, as it was not assessed on the 450K chip, nor were there any correlated proxy sites. This analysis also revealed two nominally significant associations with the same direction of effect across the blood and brain samples, namely, cg11339964 in \textit{FBXL7}, and cg12186981 in \textit{PHACTR} (Additional file 1: Fig. S3). Three other nominally significant associations had opposite directions of effect across the discovery and brain bank samples.

Table 2 Probes from the top 100 results in the Discovery EWAS that were significantly associated with PTSD in the Consortium Military Meta-Analysis Replication Cohort

| Gene | ID (proxy) | Discovery: $n = 513$ | Replication: $n = 1351$ |
|------|------------|----------------------|------------------------|
|      | Coef.      | $p$                  | Coef.                  | $p$ | $p_{\text{adj}}$ |
| AHRR | cg05575921 | −0.13                | 9.16E − 6               | −0.23 | 1.15E − 10 | 6.90E − 9 |
| APBA2| cg27184903 | 0.063                | 0.00023                 | 0.044 | 4.51E - 5  | 0.0013 |
| APBA1| cg06826552/ proxy cg13580827 | −0.15/ -0.060 | 9.43E − 5/ 0.0070 | NA/ NA/ -0.042 | 6.37E − 5 | NA/ NA/ |
| G0S2 | cg19534438 | 0.34                 | 1.19E − 7               | 0.067 | 0.00092 | 0.014 |
| LOC285696 | cg23987134 | 0.093                | 8.82E − 5               | 0.039 | 0.0033 | 0.040 |
Genome-wide analysis of the PFC samples did not yield any evidence of inflated significance (Additional file 1: Fig. S4).

**Gene set enrichment analysis**

The gene set enrichment analyses of the top 500 loci from the EWAS of the Discovery Cohort revealed five FDR-corrected significant GO terms (see Table 4). Three of the five (GO:0032496: response to lipopolysaccharide, GO:0002237: response to molecule of bacterial origin, and GO:0071216: cellular response to biotic stimulus) involve pathogen response. The genes in these three categories overlap substantially, with each containing PPARGC1A, TICAM1, TICAM2, IL1B, MEF2C, ABCB4, PRKCE, HAMP, RARA, RAR, TFPI, VDR, CDC73, and TNIP3. The other two significant GO terms were GO:0003416: endochondral bone growth and GO:0071216: cellular response to biotic stimulus.

**Candidate probes and candidate gene regions**

Of the 51 CpG sites with prior reports of association with PTSD (Additional file 1: Table S5), only the AHRR locus reported by Smith et al. survived correction for multiple testing ($p_{adj} = 0.00047$), which was not surprising given the fact that cg05575921 was near genome-wide significant in our EWAS. Two additional loci showed nominally significant associations in the Discovery Cohort: cg20098659 in CLEC9A [4] (coefficient $= 0.087$, $p = 0.0098$, $p_{adj} = 0.25$) and cg02357741 in BRSK1 [5] (coefficient $= 0.054$, $p = 0.045$, $p_{adj} = 0.63$). When we examined all of the probes in each of the previously implicated genes (Additional file 1: Table S6), 3 genes showed associations with PTSD that withstood corrections for the number of probes in the respective gene: (a) AHRR (coefficient $= -0.13$, $p = 9.16 \times 10^{-6}$, $p_{adj} = 0.00082$); (b) CLEC9A, where the most significant association was with cg02930518 (coefficient $= 0.089$, $p = 0.0067$, $p_{adj} = 0.040$); and (c) COL9A3, where the most significant locus was cg08021508 (coefficient $= -0.16$, $p = 0.00066$, $p_{adj} = 0.025$). However, only the AHRR association would survive a further adjustment for the 41 previously implicated candidate genes examined.

**Potential confounders**

See the Supplementary Results and Additional file 1: Table S7 for an examination of potential confounders including depression, anti-depressant use, alcohol-use disorders, and SNP effects. All of the CpGs in Table 1 remain significant after inclusion of each of these covariates. Similarly, the three loci in Tables 1, 2, 3 with nearby >5% MAF SNPs remained significant after adjustment for SNP effects. We noted that rs3817870 was a methylation quantitative trait locus (meQTL) for cg19534438 in GOS2 ($p = 0.023$) and rs1550637 was a meQTL for cg25526519 in RCCD1, but that these SNPs

**Table 3** Probes from the top 100 results in the Discovery EWAS that were nominally associated with PTSD in a combined PFC analysis of dIPFC and vmPFC samples from the Brain Bank Cohort

| Gene ID  | Coefficient | p       | p_{adj} |
|---------|-------------|---------|---------|
| CGST11  | 0.15        | 1.19E-05| 0.32    |
| FBXL7   | 0.11        | 1.97E-04| 0.14    |
| PHACTR  | -0.15       | 0.00017 | -0.31   |
| OR2AG1  | 1.12        | 8.49E-05| -1.15   |
| C12orf34| 0.087       | 5.18E-05| -0.12   |
| F9      | -0.43       | 4.32E-05| 0.27    |

**Table 4** Significant enrichment of GO terms in the top 500 sites from the Discovery Cohort EWAS of PTSD

| Term                                      | Ont | N | DE | p   | p_{adj} | Genes                                      |
|-------------------------------------------|-----|---|----|-----|---------|--------------------------------------------|
| GO:00032496: response to lipopolysaccharide| BP   | 322| 19 | 6.97E-06 | 0.042     | PPARGC1A, CNR2, CPS1, TICAM1, ELANE, FGF2, FOXP1, TICAM2, IL1B, MEF2C, ABCB4, PRKCE, HAMP, RARA, RAR, TFPI, VDR, CDC73, TNIP3 |
| GO:00030374: ligand-dependent nuclear receptor transcription coactivator activity| MF   | 69 | 9  | 1.12E-05 | 0.042     | PPARGC1A, NCOA7, PPARG, PRXCB, RAR, RAR, RORA, VDR, MED12 |
| GO:0002237: response to molecule of bacterial origin| BP   | 339| 19 | 1.42E-05 | 0.042     | PPARGC1A, CNR2, CPS1, TICAM1, ELANE, FGF2, FOXP1, TICAM2, IL1B, MEF2C, ABCB4, PRKCE, ERBIN, HAMP, RARA, TFPI, VDR, CDC73, TNIP3 |
| GO:00071216: cellular response to biotic stimulus| BP   | 202| 14 | 1.71E-05 | 0.042     | PPARGC1A, TICAM1, TICAM2, IL1B, MEF2C, ABCB4, PRKCE, HAMP, RARA, TFPI, VDR, WCS1, CDC73, TNIP3 |
| GO:0003416: endochondral bone growth| BP   | 22 | 6  | 1.75E-05 | 0.042     | FGF2, POClA, MSX2, BNC2, RARA, RARB |
were not confounded with the PTSD associations observed at these two loci.

**Discussion**

We performed an EWAS of lifetime PTSD diagnosis using whole blood samples from a cohort of trauma-exposed Veterans of the post-9/11 conflicts. We then evaluated evidence for replication using (a) results from a consortium-based meta-analysis of PTSD using military cohorts and (b) DNAm from PFC from a newly established PTSD brain bank. We observed one epigenome-wide significant association involving cg19534438 in the G0/G1 Switch 2 (G0S2) gene. Methylation at this locus showed a positive association with PTSD diagnosis. This was replicated in the consortium meta-analysis with the same direction of effect \( p = 0.0024, \ p_{adj} = 0.037 \). There is some prior evidence for a G0S2-PTSD link. In vivo and in vitro studies have shown cortisol to suppress G0S2 [46]. Two prior studies have linked G0S2 expression to PTSD. Specifically, using genome-wide expression profiling in a predator-scent stress animal model of PTSD, Daskalakis et al. (2014) found that G0S2 expression was downregulated in both the amygdala and hippocampus of female rats [47]. Similarly, Bam et al. (2016) identified G0S2 as the most downregulated gene in blood from a genome-wide analysis of RNAseq data from a small cohort of PTSD patients \( (n = 5) \) and controls \( (n = 5) \) [48]. The G0S2 protein is well-known for its role in regulating lipid metabolism where it serves as a negative regulator of lipolysis [49, 50]. It has been implicated in mechanisms of obesity, diabetes, aging, and cancer, and linked to gene networks involved in apoptosis, cell communication, and cell death [51]. Because of this, future investigations could examine a potential role of G0S2 in the well-established link between PTSD and metabolic disorders [52, 53].

Although we did not see significant evidence of association between cg19534438 and PTSD in the brain bank data, we note that there is evidence in two online databases that DNAm levels at this locus are correlated between blood and brain. The first database is from the University of Essex (http://epigenetics.essex.ac.uk/blood-brain/) [54] and includes dual assessment of blood and tissue from four different brain regions: PFC, entorhinal cortex, superior temporal gyrus, and cerebellum with tissue from four different brain regions: PFC, entorhinal cortex, superior temporal gyrus, and cerebellum with estimated epigenetic age of methylation sites and regulatory effects in peripheral blood mononuclear cells (PBMCs) indicates that this locus is regulatory for the adjacent gene LAMB3 in CD4T cells \( (\ p = 0.0091, \ p = 4.55 \times 10^{-7}) \). However, both LAMB3 and G0S2 were upregulated in an amyloid β peptide stimulation study of microglia, which suggests that they could be co-regulated [58]. Although we are unable to specifically confirm a regulatory effect for this locus, or confirm association in the brain, cg19534438 and G0S2 remain an interesting locus for further study in relationship to PTSD based on the strong convergent evidence of association from two independent sources of data and prior studies associating G0S2 expression to PTSD-related traits.

Methylation at cg05575921 in the Aryl-hydrocarbon Receptor Repressor (AHRR) gene, which was associated with PTSD in Smith et al. 2019, was among the top 10 results from the Discovery EWAS. In this study and in the consortium meta-analysis, PTSD cases showed reduced methylation at this locus compared to controls. The AHRR gene, and cg05575921 in particular, has a well-established association with smoking. The AHR protein is primarily known for its role in xenobiotic metabolism (i.e., metabolism of foreign chemicals generally and aromatic hydrocarbons, specifically). Smoking reduces methylation of AHRR, which in turn is associated with increased expression of the gene and enhanced negative feedback inhibition of AHR signaling. DNAm at cg05575921 is one of the strongest and most reliable indicators of smoking in the epigenome [59], but AHRR methylation at this locus, and at other CpG sites within the gene, has also been associated with other phenotypes above and beyond the effects of cigarette smoking, including epigenetic age acceleration (e.g., [60]) and C-reactive protein levels (e.g., [61]), both of which have been linked to PTSD in other recent studies [12]. According to iMethyl (accessed Jan 21, 2020), adjacent loci \( (\text{chr}5:373,355 \text{ and chr}5:373,398) \) are negatively associated with expression of the AHRR gene in monocytes \( (p < 5 \times 10^{-4}) \). Additionally, as reported in Smith et al., methylation at this locus is associated with lower kynurenine and kynurenic acid [17]. Based on the examination of the impact of the smoking score and our follow-up analyses using the score and smoking, we conclude that our data are consistent with a PTSD effect which is independent of the association with smoking. However, we
note that this will be hard to establish conclusively using human case/control cohorts given the sensitivity of this locus to smoking exposure. For example, even if one were to look at pediatric PTSD in a cohort of children who have never smoked, it is possible that trauma exposure in childhood is associated with increased exposure to second-hand smoke. Therefore, molecular methods (including animal models) may be key in disentangling the association between smoking and AHRR methylation from any putative associations with PTSD and/or trauma exposure. The ImageCpG and Essex databases indicate some level of correlation between blood and brain tissue at this locus (in ImageCpG \( r = 0.51, p = 0.019 \); in the Essex database PFC \( r = 0.28, p = 0.016 \), for other regions \( p > 0.05 \)). However, we note that this locus was not associated with smoking in the PFC, and the smoking score was not predictive of a history of smoking in the PFC samples (in dIPFC \( p = 0.44 \), in vmPFC \( p = 0.38 \)).

When we examined the Consortium Military Cohort replication data, two additional loci from our top 100 associated sites were implicated. These are Amyloid beta A4 precursor protein-binding family A member 1 (APBA1) locus cg08826552, in association observed with the 450K BeadChip proxy locus cg13580827, and the Amyloid beta A4 precursor protein-binding family A member 2 (APBA2) locus cg27184903. APBA1 on chromosome 9 and APBA2 on chromosome 15 are expressed in neurons and are involved in protein transport and synaptic function [62]. They play a role in the trafficking of APP, and hence potentially play a role in Alzheimer’s disease [62]. APBA1 (also known as MINT1 and X11) has been implicated in a GWAS of cognitive performance and educational attainment from the UK Biobank [63]. It was also differentially expressed in the brains of schizophrenia cases vs controls [64]. Analyses of genomic deletions have implicated APBA2 (also known as MINT2 and X11L) in autism [65] and schizophrenia [66] pathogenesis. Teschler et al. observed increased methylation at two APBA2 CpGs cg21917349 and cg12044210 in female bipolar cases (\( n = 24 \)) compared to controls (\( n = 11 \)) [67]. While APBA1 and APBA2 are intriguing PTSD candidate genes, these loci were not genome-wide significant in either the Discovery Cohort or the Consortium Military EWAS, and hence further replication is warranted before these can be considered PTSD loci.

Finally, cg04130728 in the carbohydrate (chondroitin 4) sulphotransferase 11 gene (CHST11) was among the top 10 most strongly associated CpG sites in the Discovery Cohort and this probe associated with PTSD in the PFC at a level that withstood correction for the 100 probes that were examined for replication in the Brain Bank Cohort. CHST11 is an enzyme that is part of a family of carbohydrate sulfotransferases that modify carbohydrate scaffolds involved in mechanisms of extracellular signaling and adhesion. In the brain, CHST11 is involved in generating chondroitin-4-sulfate (C4S), an important component of the brain’s extracellular matrix that regulates neuronal plasticity [68], fear learning [69], and neuroinflammation [70]. The cg04130728 locus is not assessed on the 450K chip, and we could identify no correlated proxies to examine it in the Smith et al. consortium replication data. We were also therefore unable to examine blood/brain correlation in the Essex database. This locus was available in ImageCpG, which did not indicate significant correlation (\( r = 0.11, p = 0.62 \)). However, we note that ImageCpG is limited in the number of samples assessed (\( n = 21 \)), is not brain-region specific, and only contains a small proportion of frontal cortex samples. Therefore, it is quite possible that methylation at this probe displays substantial blood/brain correlation in the PFC, but that more data will be needed from specific Brodmann areas before this can be confirmed. As the iMethyl database was generated using 450K data, we were unable to find evidence that this CpG was associated with CHST11 expression in blood, and we were unable to identify a suitable database for a similar examination of regulatory effects of CHST11 in brain tissue.

Our examination of candidate probes and gene regions implicated in prior studies of the epigenetics of PTSD replicated the prior observed association with the AHRR locus, which was compelling as a candidate locus (\( p_{adj} = 0.00047 \)) despite its lack of genome-wide significance. None of the other previously implicated probes (listed in Additional file 1: Table S2) yielded effect sizes that survived multiple-testing correction. Several factors may explain why AHRR was the only locus to replicate. First, our primary analyses adjusted for smoking through the use of a DNAm-based smoking score and this was not done in many of the prior PTSD EWASs. Secondly, our clinical cohorts were comprised exclusively of Veterans, the majority of whom were male and most of whom had chronic PTSD. For this reason, our findings may not generalize to other PTSD populations (and vice versa). It is also noteworthy that 12 of the 51 previously implicated loci had low methylation ranges (range < 0.10; see Additional file 1: Table S6), which could make them unreliable and/or difficult to replicate [45].

Conclusions

The replication of the AHRR locus implicated in the PTSD-consortium EWAS in the discovery cohort, the replication of G0S2, the top locus from this EWAS, in the Consortium Military Meta-Analysis, and the broader evidence of agreement in effect size direction for the Discovery Cohort and the Replication Cohort are cause for renewed optimism in the search for reliable blood
methylational based PTSD biomarkers. In many large-scale consortium GWAS studies from the PGC, there was an inflection point past which increasing sample size produced an increasing number of replicable loci. Although it may well be that PTSD DNA/m signatures will differ by sex, trauma type, and other demographic factors, these replicated loci may indicate that a similar tipping point has been reached in large scale EWAS of PTSD. Further, this study adds to the growing body of evidence of disruption to inflammation and immune response processes in the pathophysiology of PTSD.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13148-020-0820-0.

Additional file 1. Supplementary methods and results.

Additional file 2. Complete EWAS results for the Discovery Cohort.

Additional file 3. Replication of top EWAS results in the PFC of the Brain Bank Cohort.

Abbreviations

450K BeadChip: Illumina Infinium HumanMethylation450 BeadChip; CpG: Cytosine-phosphate-guanine site; dIPFC: Dorsolateral prefrontal cortex; DNA: Deoxynucleobnic acid; DNAm: DNA methylation; EPIC: Illumina EPIC Methylation BeadChip; EWAS: Epigenome-wide association study; FDR: False discovery rate; GO: Gene ontology; GWAS: Genome-wide association study; HPA: axis: Hypothalamic-pituitary-adrenal axis; p: p-value; pFDR: False-discovery rate corrected p-value; PFC: Prefrontal cortex; PGC: Psychiatric Genomics Consortium; PTSD: Posttraumatic stress disorder; QC: Quality control; SNP: Single-nucleotide polymorphism; TBI: Traumatic brain injury; TRACTS: Translational Research Center for TBI and Stress Disorders; vmPFC: Ventromedial prefrontal cortex

Acknowledgements

The MRS study would like to acknowledge Victoria B. Ristrouth Ph.D (VA San Diego Healthcare System & UCSD), Mark A. Geyer (UCSD), Daniel T. O’Connor (UCSD), all MRS investigators, as well as the MRS administrative core and data collection staff listed in the Methods article (Baker et al, Prev Chronic Dis. 2012;9(10):E97), and to thank the Marine and Navy Corpsmen volunteers for their military service and participation in MRS. The Army STARRS Team consists of the following Co-Principal Investigators, Robert J. Ursano, MD (Uniformed Services University of the Health Sciences and) Murray B. Stein, MD, MPH (University of California San Diego and VA San Diego Healthcare System); Site Principal Investigators, Steven Heeringa, PhD (University of Michigan), James Wagner, PhD (University of Michigan), and Ronald C. Kessler, PhD (Harvard Medical School); Army liaison/consultant, Kenneth Cox, MD, MPH (US Army Public Health Center); other team members, Pablo A. Alliaga, MS (Uniformed Services University of the Health Sciences), COL David M. Benedek, MD (Uniformed Services University of the Health Sciences); Susan Borja, PhD (NIH); Tianxu Cai, ScD (Harvard School of Public Health); Laura Campbell-Sills, PhD (University of California San Diego); Chia-Yen Chen, ScD (Harvard Medical School); Carol S. Fullerton, PhD (Uniformed Services University of the Health Sciences); Nancy Geletar, MA (University of Michigan); Joel Geleman, MD (Yale University); Robert K. Gilford, PhD (Uniformed Services University of the Health Sciences); Feng He, MS (University of California San Diego); Paul E. Hurwitz, MPH (Uniformed Services University of the Health Sciences); Sonia Jain, PhD (University of California San Diego); Kevin Jensen, PhD (Yale University); Kristen Jepson, PhD (University of the California San Diego); Tzu-Chyng Kao, PhD (Uniformed Services University of the Health Sciences); Lisa Lewandowski-Romps, PhD (University of Michigan); Holly Herberman Mash, PhD (Uniformed Services University of the Health Sciences); James E. McCarroll, PhD, MPH (Uniformed Services University of the Health Sciences); Adam X. Malmoher (University of California San Diego); Colter Mitchell, PhD (University of Michigan); James A. Nafteh, PhD (Uniformed Services University of the Health Sciences); Tsz Hin Hinz Ng, MPH (Uniformed Services University of the Health Sciences); Caroline M. Nievergelt, PhD (University of California San Diego); Matthew K. Nock, PhD (Harvard University); Stephen Ripke, MD (Harvard Medical School); Nancy A. Sampson, BA (Harvard Medical School); CDR Patcho Santiago, MD, MPH (Uniformed Services University of the Health Sciences); Rosan Segman, MD (Hadassah University Hospital, Israel); Jordan W. Smoller, MD, ScD (Harvard Medical School); Xiaoying Sun, MS (University of California San Diego); Erin Ware PhD (University of Michigan); LTC Gary H. Wynn, MD (Uniformed Services University of the Health Sciences); Alan M. Zaslavsky, PhD (Harvard Medical School); and Lee Zhang, MD (Uniformed Services University of the Health Sciences). Members of the INTRuST Consortium Biorepository Working Group include Gerald A. Grant, MD, Christine E. Marx, MD, Mark S. George, MD, Thomas W. Allister, MD, Norberto Andaluz, MD, Lori Shutter, MD, Paul Cumbra, MD, Ross D. Zafonte, DO, Sonia Jain, PhD, Yue-Jun Qiu, and Michael Hauser, PhD. The Mid-Atlantic MIRECC Workgroup contributors include Mira Branu, PhD, Patrick S. Calhoun, PhD, Eric Dierdt, PhD, Eric B. Bologen, PhD, John A. Fairbank, PhD, Robin A. Hurley, MD, Jason D. Khits, PhD, Angela Kirby, MS, Christine E. Mann, MD, Scott D. McDonald, PhD, Scott D. Moore, MD, PhD, Rajendra A. Morey, MD, MS, Jennifer C. Naylor, PhD, Steven C. Pickett, PsyD, Jared Rowland, PhD, Cindy Swinkels, PhD, Steven T. Szabo, MD, PhD, Katherine H. Taber, PhD, Larry A. Tupil, PhD, Elizabeth E. Van Voorhees, PhD, H. Ryan Wagner, PhD, Ruth E. Yoash-Gantz, PsyD.

“The Traumatic Stress Brain Research Group (exclusive of previously named authors) are: Matthew J. Friedman MD, PhD (Director); Victor E. Alvarez MD, MPH, PhD (Project Director); David Beneked MD, MPH, Christopher Brady, PhD (PhD, MD); Dianni Cruz MS 46, David A Davis PhD 49, Ronald S. Duman PhD 43,50,51, Matthew J. Girgenti 50,51, PhD, Melanie Hardegree RN 50, Paul E. Holtzheimer MD 43,44, Terence M. Keane PhD 43,45,46, Neil Kowell MD 45,46, John H. Krystal MD 43,50,51, Ann McKee MD 45,46, Brian Max PhD 45,46, Deborah Mash PhD 52, William K. Scott PhD 40, Thor Stein MD 45,46, PhD, Robert Ursano MD 43, Douglas E. Williamson PhD 48,51, Keith A. Young PhD 40, 43, National Center for PTSD, 44 Department of Psychiatry, Geisel School of Medicine at Dartmouth. 45. Departments of Psychiatry, Neurology, and/or Pathology & Laboratory Science, Boston University. 46. VA Boston Healthcare System. 47. Department of Psychiatry, Uniformed Services University of Health Science. 48. Department of Psychiatry & Behavioral Sciences, Duke University Medical Center. 49. Departments of Human Genetics, Public Health Sciences and/or Neurology, University of Miami Miller School of Medicine. 50. Department of Psychiatry, Yale University School of Medicine. 51. VA Connecticut Healthcare System. 52. Department of Biomedical Sciences, Nova Southeastern University. 53. Durham VA Healthcare System. 54. Department of Psychiatry & Behavioral Science, Texas A&M University System Health Science Center.

Contributors’ contributions

Authors MWL, MWM, EAW, BRH, FGM, AKS, and NPD contributed to study design, data interpretation, and drafting of the manuscript. MWL and ZY analyzed data and/or prepared figures and tables for the manuscript. MUI, CMM, AEA-K, DGB, JCB, MEG, MBP, EP, GG, GAG, MAH, RCK, NAK, AXM, CEM, X-JQ, VBR, BFBS, MBR, RUI, EU, CHV, and EBW contributed to the consortium meta-analysis used for replication. AS and SAS supplied samples and phenotype data used in the study. All authors reviewed and approved the manuscript prior to submission.

Funding

This work was funded by 1I01BX003477, a Department of Veterans Affairs BLRD grant to MWL, 1R03AG051877, 2R21AG063427-01, and 1ID1CX001276-01A2 to EJW, 5T32MH019836-18 to FGM, R21MH102834 to MWM, 1U01MH108826 to Smith/Logue/Nievergelt/Uddin, the National Center for PTSD Behavioral Sciences Division, and the Translational Research Center for TBI and Stress Disorders (TRACTS), a Department of Veterans Affairs Rehabilitation Research and Development (RR&D) Traumatic Brain Injury Center of Excellence (B3001-D) at VA Boston Healthcare System. Genotype and methylation data for the TRACTS study and Brain Bank Cohort was generated with the support of resources and/or facilities of the Pharmacogenomics Analysis Laboratory (Research Service, Central Arkansas Veterans Health System, Little Rock, Arkansas), a core research laboratory funded by the Cooperative Studies Program, Research and Development, Department of Veterans Affairs. Marine Resiliency Study (MRS) funding from The Marine Corps, Navy Bureau of Medicine and Surgery (BUMED) and VA Health Research and Development (HSRD) who provided funding for MRS data collection and analysis (PI DGB).
and NIH R01MH093500 which funded the GWAS assays and analysis (PI CMN). Army STARSS was sponsored by the Department of the Army and funded under cooperative agreement number U01MH087981 (2009-2015) with the National Institutes of Health, National Institute of Mental Health (NIH/NIMH). Subsequently, STARSS-L5 was sponsored and funded by the Department of Defense (USLHS grant no. HU0001-15-2-0004). Data collection of PRSMO was funded by the Dutch Ministry of Defence, and DNA methylation analyses in PRSMO were funded by the VENI Award fellowship from the Netherlands Organisation for Scientific Research (NWO, grant no. 916.11.086). The PTSD and TBI Injury and Traumatic Stress Clinical Consortium (INTRuST) was funded by a grant from the US Department of Defense (PI: Stein, MB; W81XWH08-2-0159). Mid-Atlantic Mental Illness Research Education and Clinical Center the study of Post-Deployment Mental Health Study funding included a Clinical Sciences Research and Development (CSRD) Research Career Scientist Award (ID 15-RCS-009) to Dr. Beckham, a CSRD Career Development Award (#K2 CX00052S) to Dr. Kimbrel, and Biomedical and Laboratory Research and Development (BLR&D) Merit Award to Dr. Beckham from the US Department of Veterans Affairs (VA). This work was also supported by the VA Mid-Atlantic Mental Illness Research, Education and Clinical Center (MIRECC), the Durham Veterans Affairs Medical Center, the VA Office of Mental Health Services, and the VA Office of Research and Development.

Availability of data and materials
A complete set of the results of the Ewas in the TRACTS study, and the results of the top 100 loci examined in the PFC analysis are presented in additional file 2 and additional file 3. Due to the provisions of the informed consent documents for this study, individual-level data cannot be posted publicly; however, those data may be obtained via a data-use agreement with the TRACTS study. Contact Dr. Regina McClinichney (Regina_McClinchey@hms.harvard.edu) to request access.

Ethics approval and consent to participate
This research was performed under the supervision and approval of US Department of Veterans Affairs human subjects review board. All participants supplied written informed consent.

Consent for publication
Not applicable.

Competing interests
FM participated in this study while also an NIMH-funded research post-doctoral fellow at Boston University/VA Boston Healthcare System, but is currently an employee of BlackThorn Therapeutics. MBS has in the past 3 years been a consultant for Actelion, Aptinyx, Bionomics, Dart Neuroscience, Healthcare Management Technologies, Jansen, Neuronx Biosciences, Oxsea Biopharmaceuticals, Pfizer, and Resilience Therapeutics. In the past 3 years, RCK received support for his epidemiological studies from Sandhi Aventis; was a consultant for DataSat, Inc., Johnson & Johnson Wellness and Prevention, Sage Pharmaceuticals, Shire, Takeda; and served on an advisory board for the Johnson & Johnson Services Inc. Lake Nona Life Project. None of the remaining authors have any conflicts to disclose. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, the Department of Defense, NIMH, or the US government.

Author details
1National Center for PTSD, VA Boston Healthcare System, Boston, MA, USA.
2Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA.
3Biomedical Genetics, Boston University School of Medicine, Boston, MA, USA.
4Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA.
5Department of Gynecology and Obstetrics, Emory University, Atlanta, GA, USA.
6Department of Psychiatry and Behavioral Sciences, Emory University, Atlanta, GA, USA.
7Department of Psychiatry, Harvard Medical School, Boston, MA, USA.
8McLean Hospital, Belmont, MA, USA.
9Cohen Veterans Bioscience, Cambridge, MA, USA.
10Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
11Department of Epidemiology, Columbia University, New York, NY, USA.
12Genomics Program, University of South Florida College of Public Health, Tampa, FL, USA.
13Global Health and Infectious Disease Research Program, University of South Florida College of Public Health, Tampa, FL, USA.
14Department of Psychiatry, University of California San Diego, La Jolla, CA, USA.
15Center of Excellence for Stress and Mental Health, Veterans Affairs San Diego Healthcare System, San Diego, CA, USA.
16Research Service, Veterans Affairs San Diego Healthcare System, San Diego, CA, USA.
17Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC, USA.
18Psychiatry Service, Veterans Affairs San Diego Healthcare System, San Diego, CA, USA.
19Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA.
20Research, Durham VA Medical Center, Durham, NC, USA.
21Genetics Research Laboratory, VA Mid-Atlantic Mental Illness Research, Education, and Clinical Center (MIRECC), Durham, NC, USA.
22Department of Psychiatry, UMC Utrecht Brain Center, Utrecht, Vlaardingen, Netherlands.
23Brain Research and Innovation Centre, Netherlands Ministry of Defence, Utrecht, Utrecht, Netherlands.
24Department of Neurosurgery, Stanford University Medical Center, Stanford, CA, USA.
25Department of Health Care Policy, Harvard Medical School, Boston, MA, USA.
26Duke Molecular Physiology Institute, Duke University, Durham, NC, USA.
27Department of Critical Care Medicine, Neurology, and Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA.
28Department of Psychiatry & Behavioral Sciences, Duke University Medical Center, Durham, NC, USA.
29School for Mental Health and Neuroscience, Department of Psychiatry and Neuropsychology, Maastricht Universitair Medisch Centrum, Maastricht, Limburg, Netherlands.
30Million Veteran Program, Veterans Affairs San Diego Healthcare System, San Diego, CA, USA.
31Department of Psychiatry, Uniformed Services University, Bethesda, MD, USA.
32Arq, Psychotrauma Research Expert Group, Diemen, NH, Netherlands.
33Department of Psychiatry, Leiden University Medical Center, Leiden, ZH, Netherlands.
34Netherlands Defense Department, Research Center, Utrecht, UT, Netherlands.
35Department of Psychiatry, New York University School of Medicine, New York, NY, USA.
36Department of Anatomy and Neurosciences, Amsterdm UMC (location VUMc), Amsterdam, Holland, Netherlands.
37Department of Psychiatry, Amsterdam UMC (location VUMc), Amsterdam, Holland, Netherlands.
38Institute for Social Research, Survey Research Center, University of Michigan, Michigan, MI, USA.
39Pharmacogenomics Analysis Laboratory, Research Service, Central Arkansas Veterans Healthcare System, Little Rock, AR, USA.
40Geriatric Research Educational and Clinical Center and Translational Research Center for TBI and Stress Disorders, VA Boston Health Care System, Boston, MA, USA.
41Department of Psychology and Chronic Brain Injury Program, The Ohio State University, Columbus, OH, USA.
42Memory Disorders Research Center, VA Boston Healthcare System and Boston University School of Medicine, Boston, MA, USA.

Received: 31 October 2019 Accepted: 29 January 2020
Published online: 14 March 2020

References
1. Logue MW, Baldwin C, Guffanti G, Melista E, Wolf EJ, Reardon AF, Uddin M, Wildman D, Galea S, Koenen KC, Miller MW. A genome-wide association study of post-traumatic stress disorder identifies the retinoid-related orphan receptor alpha (RORA) gene as a significant risk locus. Mol Psychiatry. 2013;18:937–42.
2. Nievergelt CM, Malholef AR, Kengel T, Atkinson EG, Chen CY, Choi K, Coleman J, Dalvie S, Duncan LE, Gelernter J, et al. International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. Nat Commun. 2019;10:4558.
3. Logue MW, Smith AK, Baldwin C, Wolf EJ, Guffanti G, Ratnatharathorn A, Stone A, Schichman SA, Humphries D, Binder EB, et al. Analysis of gene expression in PTSD implicates genes involved in the glucocorticoid receptor pathway and neural responses to stress. Psychoneuroendocrinology. 2015;57:1–13.
4. Smith AK, Connelly KN, Kiliaru V, Mercer KB, Weiss TE, Bradley B, Tang Y, Gillespie CF, Cubells JF, Ressler KJ. Differential immune system DNA methylation and cytokine regulation in post-traumatic stress disorder. Am J Med Genet B Neuropsychiatr Genet. 2011;158B:700–8.
5. Mehta D, Brueinig D, Camilo-Roa T, Lawford B, Harvey W, Morris CP, Smith AK, Binder EB, Young RM, Volpe I, Guffanti G. DNA methylation analysis in combat veterans reveals a novel locus for PTSD. Acta Psychiatr Scand. 2017;136:493–505.
6. Rutten BP, Vermetten E, Vinkers CH, Ursini G, Daskalakis NP, Pishva E, de Nijs L, Houtepen LC, Eijssen L, Jaffe AE, et al. Longitudinal analyses of the DNA methylation in deployed military servicemen identify susceptibility loci for post-traumatic stress disorder. Mol Psychiatry. 2018;23:1143–56.
7. Szfy M, Tang YY, Hill KG, Musci R. The dynamic epigenome and its implications for behavioral interventions: a role for epigenetics to inform disorder prevention and health promotion. Transl Behav Med. 2016;6:655–62.

8. Zannas AS, Provencal N, Binder EB. Epigenetics of posttraumatic stress disorder: current evidence, challenges, and future directions. Biol Psychiatry. 2015;78:327–35.

9. Ressler KJ, Mercer KB, Bradley B, Jovanovic T, Mahan A, Kerley K, Duvall M, Logue MW, Schichman SA, Stone A, Milberg WP, et al. Predictors of risk and resilience for posttraumatic stress disorder among ground combat Marines: methods of the Marine Resiliency Study. Prev Chronic Dis. 2012;9:E97.

10. Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, et al. Epigenetic biomarkers as predictors of combat stress vulnerability and resilience in U.S. marines: a genome-wide association study across multiple ancestries implicates PRTFDC1 as a potential PTSS gene. Psychoneuroendocrinology. 2015;51:459–71.

11. Baker DG, Nash WP, Litz BT, Geyer MA, Risbrough VB, Nievergelt CM, O’Connor DT, Larson GE, Schork NJ, O’Connor DT, et al. Map of the human genome from the U.S. Department of Energy’s (DOE’s) National Human Genome Research Institute. Am J Med Genet B Neuropsychiatr Genet. 2017;174:619–26.

12. Schechter DS, Maser DA, Pointet VC, Ate T, Stenzel P, Liozis G, Wenzel S, Ladd A, Ginzburg L, Chinnock J, et al. The association of serotonin receptor 3A methylation with maternal violence exposure, neural activity, and child aggression. Behav Brain Res. 2017;325:268–77.

13. Mehta D, Peltzer ES, Brugniet L, Lawford B, Smith AK, Myers AE, et al. Stress exposure and psychopathology alter methylation of the serotonin receptor 2A (HTR2A) gene in preschoolers. Dev Psychopathol. 2017;29:1619–26.

14. Sperbeck E, Piehler ES, Brown EG, McLeay S, Alcock E, Logue MW, et al. Genome-wide association studies of posttraumatic stress disorder in a cohort of Iraq-Afghanistan era veterans. J Affect Disord. 2015;184:225–34.

15. Lepage C, de Pierrefeu A, Koertie IK, Coleman MJ, Pasnarkov G, Grant G, Marx CE, Morey RA, Flashman LA, George MS, et al. White matter abnormalities in mild traumatic brain injury with and without posttraumatic stress disorder: a subject-specific diffusion tensor imaging study. Brain Imaging Behav. 2018;12:870–81.

16. Forzo GA, Goodkind MS, Oathes DJ, Zaitz V, Harvey M, Peng KK, Weiss ME, Thompson AL, Zaduck SE, Finney-Cent N, et al. Selective effects of psychotherapy on frontopolar cortical function in PTSD. Am J Psychiatry. 2017;174:1175–84.

17. Kavelaars A, Symanski E, Young RM, Voisey J, Initiative P. DNA methylation from germline cells in veterans with PTSD. J Psychiatric Res. 2019;116:42–50.

18. Smith AK, Ratnakarathorn A, Mafhoo AX, Navaiaux RK, Aiello AE, Amstadter AB, Ashley-Koch AE, Baker DG, Becker KM, Boks MP, et al. Epigenome-wide meta-analysis of PTSD across 10 military and civilian cohorts identifies novel methylation loci. BioRxiv. 2019; https://doi.org/10.1101/585109.

19. Hammarmeirh R, Chakabory N, Gautham A, Muhie S, Yang Y, Donohue D, Kumar R, Daigle BJ Jr, Zhang Y, Amara DA, et al. Whole-genome DNA methylation status associated with clinical PTSD measures of OIF/OEF veterans. Transl Psychiatry. 2017;7:e169.

20. Kuan PF, Wang JC, Xiong M, Kotov R, Marist CJ, Guldbrand T, Gonzalez A, Yang X, Koenen K, Bremmer L, Lu B. An epigenome-wide DNA methylation study of PTSD and depression in World Trade Center responders. Transl Psychiatry. 2017;7:e171158.

21. Uddin M, Aiello AE, Wildman DE, Koenen KC, Pawelec G, de Los Sar, Goldmann E, Galea S. Epigenetic and immune function profiles associated with posttraumatic stress disorder. Proc Natl Acad Sci U S A. 2010;107:9470–5.

22. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, Guan W, Xu T, Ellis CE, Astleley S, et al. Epigenetic signatures of cigarette smoking. Circ Cardiovasc Genet. 2016;9:436–47.

23. Li S, Wong EM, Bui M, Nguyen TL, Joo JE, Stone J, Dite GS, Giles GS, Saffery R, Southey MC, Hopper JL. Causal effect of smoking on DNA methylation in peripheral blood. Genome Biol. 2018;10:18.

24. Fu SS, McFall M, Saxon AJ, Beckham JC, Carmody TP, Baker DG, Joseph AM. Post-traumatic stress disorder and smoking: a twin and family study. Clin Epigenetics. 2018;10:659–704.

25. Nievergelt CM, Mafhoo AX, Mustapic M, Yung K, Schork NJ, Miller MW, Logue MW, Geyer MA, Risbrough VB, O’Connor DT, Baker DG. Genomic predictors of combat stress vulnerability and resilience in U.S. marines: a genome-wide association study across multiple ancestries implicates PRTFDC1 as a potential PTSD gene. Psychoneuroendocrinology. 2015;51:459–71.

26. Ashley-Koch AE, Garrett ME, Gibson J, Liu Y, Dennis MF, Kimbell NA. Veterans Affairs Mid-Atlantic Mental Illness Research E, Clinical Center W, Beckham JC, Hsu EA. Genome-wide association study of posttraumatic stress disorder in a cohort of Iraq-Afghanistan era veterans. J Affect Disord. 2015;184:225–34.

27. Logue MW, Schichman SA, Stone A, Logue MW, et al. Randomized placebo-controlled trial of methylphenidate or galantamine for persistent emotional and cognitive symptoms associated with PTSD and/or traumatic brain injury. Neuropsychopharmacology. 2016;41:1191–8.

28. Rasmusson AM, Marx CE, Jain S, Farrel GM, Tsai J, Sun X, Geraci TD, Hannen MB, Johr L, Rosse R, et al. A randomized controlled trial of ganaoxide in posttraumatic stress disorder. Psychopharmacology (Berl). 2017;234:2425–57.

29. Beckham JC, Hauser MA: Genome-wide association study of posttraumatic stress disorder: a subject-specific diffusion tensor imaging study. Brain Imaging Behav. 2018;12:870–81.

30. Forzo GA, Goodkind MS, Oathes DJ, Zaitz V, Harvey M, Peng KK, Weiss ME, Thompson AL, Zaduck SE, Finney-Cent N, et al. Selective effects of psychotherapy on frontopolar cortical function in PTSD. Am J Psychiatry. 2017;174:1175–84.

31. Kavelaars A, Symanski E, Young RM, Voisey J, Initiative P. DNA methylation from germline cells in veterans with PTSD. J Psychiatric Res. 2019;116:42–50.

32. Smith AK, Ratnakarathorn A, Mafhoo AX, Navaiaux RK, Aiello AE, Amstadter AB, Ashley-Koch AE, Baker DG, Becker KM, Boks MP, et al. Epigenome-wide meta-analysis of PTSD across 10 military and civilian cohorts identifies novel methylation loci. BioRxiv. 2019; https://doi.org/10.1101/585109.

33. Hammarmeirh R, Chakabory N, Gautham A, Muhie S, Yang Y, Donohue D, Kumar R, Daigle BJ Jr, Zhang Y, Amara DA, et al. Whole-genome DNA methylation status associated with clinical PTSD measures of OIF/OEF veterans. Transl Psychiatry. 2017;7:e169.

34. Kuan PF, Wang JC, Xiong M, Kotov R, Marist CJ, Guldbrand T, Gonzalez A, Yang X, Koenen K, Bremmer L, Lu B. An epigenome-wide DNA methylation study of PTSD and depression in World Trade Center responders. Transl Psychiatry. 2017;7:e171158.

35. Uddin M, Aiello AE, Wildman DE, Koenen KC, Pawelec G, de Los Sar, Goldmann E, Galea S. Epigenetic and immune function profiles associated with posttraumatic stress disorder. Proc Natl Acad Sci U S A. 2010;107:9470–5.

36. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, Guan W, Xu T, Ellis CE, Astleley S, et al. Epigenetic signatures of cigarette smoking. Circ Cardiovasc Genet. 2016;9:436–47.

37. Li S, Wong EM, Bui M, Nguyen TL, Joo JE, Stone J, Dite GS, Giles GS, Saffery R, Southey MC, Hopper JL. Causal effect of smoking on DNA methylation in peripheral blood. Genome Biol. 2018;10:18.

38. Fu SS, McFall M, Saxon AJ, Beckham JC, Carmody TP, Baker DG, Joseph AM. Post-traumatic stress disorder and smoking: a systematic review. Nicotine Tob Res. 2007;9:1017–84.

39. Friedman MJ, Huber BR, Brady CB, Ursano RJ, Benedek DM, Kowall NW, McKeen AC. Traumatic Stress Brain Research G: VA’s National PTSD Brain Bank: a National Resource for Research. Curr Psychiatry Rep. 2017;19:73.

40. Steen MB, Chen CY, Ursano RJ, Cai T, Gudzi, Clinic Epigenetics. 2018;10:18.

41. Koenen K, Bromet E, Luft BJ. An epigenome-wide DNA methylation study of PTSD in female mice and women with PTSD. Mol Psychiatry. 2018;23:658–65.

42. Uddin M, Ratnakarathorn A, Armstrong D, Kuan PF, Aiello AE, Bremot EJ, Galea S, Koenen KC, Lut B, Resler KJ, et al. Epigenetic meta-analysis across three civilian cohorts identifies NRG1 and HGS as birth-based biomarkers for post-traumatic stress disorder. Epigenomics. 2018;10(12):1585–601.
