

Genetic Association Analysis of 300 Genes Identifies a Risk Haplotype in *SLC18A2* for Post-traumatic Stress Disorder in Two Independent Samples

Nadia Solovieff^{1,2,3}, Andrea L Roberts⁴, Andrew Ratanatharathorn⁵, Michelle Haloosim⁵, Immaculata De Vivo^{6,7}, Anthony P King⁸, Israel Liberzon^{8,9}, Allison Aiello¹⁰, Monica Uddin^{11,12}, Derek E Wildman^{11,13}, Sandro Galea⁵, Jordan W Smoller^{1,2,3}, Shaun M Purcell^{1,2,3,14} and Karestan C Koenen^{*5}

¹Department of Center of Human Genetics Research, Massachusetts General Hospital, Boston, MA, USA; ²Department of Stanley Center for Psychiatric Research, the Broad Institute of Harvard and MIT, Cambridge, MA, USA; ³Department of Psychiatry, Harvard Medical School, Boston, MA, USA; ⁴Department of Society, Human Development and Health, Harvard School of Public Health, Boston, MA, USA; ⁵Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA; ⁶Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA; ⁷Channing Division of Network Medicine, BWH, Harvard Medical School, Boston, MA, USA; ⁸Mental Health Service, VA Ann Arbor Health Systems, MI, USA; ⁹Department of Psychiatry, University of Michigan Medical School, Ann Arbor, MI, USA; ¹⁰Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA; ¹¹Center for Molecular Medicine and Genetics, Department of Psychiatry & Behavioral Neurosciences, Wayne State University School of Medicine, Detroit, MI, USA; ¹²Department of Psychiatry and Behavioral Neurosciences, Wayne State University School of Medicine, Detroit, MI, USA; ¹³Department of Obstetrics and Gynecology, Wayne State University School of Medicine, Detroit, MI, USA; ¹⁴Division of Psychiatric Genomics, Mount Sinai School of Medicine, New York, NY, USA

The genetic architecture of post-traumatic stress disorder (PTSD) remains poorly understood with the vast majority of genetic association studies reporting on single candidate genes. We conducted a large genetic study in trauma-exposed European-American women (N = 2538; 845 PTSD cases, 1693 controls) by testing 3742 SNPs across more than 300 genes and conducting polygenic analyses using results from the Psychiatric Genome-Wide Association Studies Consortium (PGC). We tested the association between each SNP and two measures of PTSD, a severity score and diagnosis. We found a significant association between PTSD (diagnosis) and SNPs (top SNP: rs363276, odds ratio (OR) = 1.4, $p = 2.1 \times 10^{-5}$) in *SLC18A2* (vesicular monoamine transporter 2). A haplotype analysis of 9 SNPs in *SLC18A2*, including rs363276, identified a risk haplotype (CGGCGGAAG, $p = 0.0046$), and the same risk haplotype was associated with PTSD in an independent cohort of trauma-exposed African-Americans ($p = 0.049$; N = 748, men and women). *SLC18A2* is involved in transporting monoamines to synaptic vesicles and has been implicated in a number of neuropsychiatric disorders including major depression. Eight genes previously associated with PTSD had SNPs with nominally significant associations ($p < 0.05$). The polygenic analyses suggested that there are SNPs in common between PTSD severity and bipolar disorder. Our data are consistent with a genetic architecture for PTSD that is highly polygenic, influenced by numerous SNPs with weak effects, and may overlap with mood disorders. Genome-wide studies with very large sample sizes are needed to detect these types of effects.

Neuropsychopharmacology (2014) **39**, 1872–1879; doi:10.1038/npp.2014.34; published online 12 March 2014

INTRODUCTION

Post-traumatic stress disorder (PTSD) is a debilitating mental disorder that occurs following exposure to a potentially traumatic life event and is defined by three symptom clusters: reexperiencing, avoidance and numbing, and arousal (American Psychiatric Association, 1994). The

disorder is pervasive in the US general population with a lifetime prevalence of 7.6% (Kessler *et al*, 2005): 1 in 9 women and 1 in 20 men will meet the criteria for diagnosis at some point in their lives (Roberts *et al*, 2011). Although 50–85% of the US population report trauma exposure, only 2–50% develop PTSD (Breslau *et al*, 1999; Kessler *et al*, 1995). In fact, only about half of all those who are exposed to traumatic events of even the most severe interpersonal violence such as a completed rape develop the disorder (Breslau *et al*, 1999; Kessler *et al*, 1995). The identification of risk markers, and particularly biomarkers, that distinguish between persons at high and low risk of developing PTSD following trauma exposure has been identified as a priority research goal by the Institute of Medicine

*Correspondence: Dr KC Koenen, Department of Epidemiology, Mailman School of Public Health, Columbia University, 722 West 168th Street, Room 720G, New York, NY 10032-3727, USA. Tel: +212 305 7482; E-mail: kck5@cumc.columbia.edu
Received 24 June 2013; revised 9 January 2014; accepted 10 January 2014; accepted article preview online 14 February 2014

(Medicine, 2012), Department of Defense (2012a) and the NIMH (2008). The ability to identify persons at high risk of developing PTSD would enable providers to target evidence-based interventions to high risk groups (Rothbaum *et al*, 2012; Shalev *et al*, 2012). The identification of robust predictive biomarkers may also improve our understanding of the pathophysiology of PTSD and lead to more effective pharmacological interventions. Genetic variants associated with PTSD are promising biomarkers of risk because they remain unchanged throughout life, and DNA can be obtained noninvasively and assayed reliably.

There is abundant evidence that there are genetic determinants of PTSD to be discovered. PTSD risk is transmitted across generations (Roberts *et al*, 2012b; Yehuda *et al*, 2001) and heritability estimates range from 30% (Stein *et al*, 2002; True *et al*, 1993)–70% (Sartor *et al*, 2011). A recent review of the literature identified 45 published candidate gene studies of PTSD covering 22 genes. Several candidate genes, such as *FKBP5* and *PACAP* have been associated with PTSD across several independent samples. However, despite more than two decades of genetic research on PTSD, the authors argued that more work needs to be done to identify robust genetic variants for PTSD (Pitman *et al*, 2012).

Our understanding of the genetic architecture of PTSD and the identification of robust genetic variants associated with the disorder has been hampered by four limitations of extant studies. First, case definition has not been consistent across studies; some studies have examined quantitative outcomes such as current PTSD symptoms (Ressler *et al*, 2011) and others have defined PTSD cases as persons with lifetime PTSD (Chang *et al*, 2012; Koenen *et al*, 2009a). Second, PTSD is conditional on trauma exposure. The optimal design for genetic studies of PTSD, therefore, is to compare PTSD cases to trauma-exposed controls from the same underlying population. The inconsistency in use of trauma-exposed controls makes negative associations difficult to interpret. Third, almost all studies focus on single gene associations and many focus on single SNPs rather than assaying all common variation. Such studies do not advance our understanding of the genetic architecture of PTSD more broadly. They also do not allow for a comparison of the association between multiple candidate genes and PTSD. Fourth, sample sizes are small with only three studies in our recent review including more than 1000 participants. In addition, the majority of studies do not include replication samples.

We sought to examine the genetic architecture of PTSD by conducting a large genetic study of lifetime PTSD diagnosis and severity in 2538 trauma-exposed European-American women selected from the Nurses' Health Study II (NHSII) on whom we tested 3742 SNPs across more than 300 genes. The SNPs in our study were chosen to comprehensively cover candidate genes in the literature as well as top hits from published genome-wide association studies (GWAS) of psychiatric disorders. Since twin studies suggest PTSD has a high genetic overlap with other psychiatric disorders, particularly with major depression (Sartor *et al*, 2012; Wolf *et al*, 2010), we also conducted polygenic analyses using results for bipolar disorder (BP), major depressive depression (MDD), and schizophrenia (SZ) from the Psychiatric Genome-Wide Association

Studies Consortium (PGC) (2012b; Ripke *et al*, 2011; Sklar *et al*, 2011). We found a significant association between PTSD diagnosis and *SLC18A2* after correction for multiple testing. We identified a risk haplotype in *SLC18A2* and found a consistent result in the Detroit Neighborhood Health Study (DNHS) ($N=748$), an epidemiologic sample of primarily African-American adults residing in urban Detroit (Goldmann *et al*, 2011). Our polygenic analyses suggest that there are shared genetic factors between PTSD severity and BP.

MATERIALS AND METHODS

NHSII Sample

We used data from the PTSD diagnostic subsample ($N=3013$) of the NHSII. A subset of 2612 subjects had DNA available for analysis. The ascertainment of the subsample has been described in detail previously (Roberts *et al*, 2012a) and the protocol has been published (Koenen *et al*, 2009b). 2538 Trauma-exposed European-American women from the NHSII were included in this analysis. The mean severity score was 33 (SD:13), higher values indicating more severe symptoms, and 845 (33%) women were diagnosed with PTSD. The mean age at the start of the study was 55.6 (SD:4.5).

A brief description of the recruitment can be found in the Supplementary Material. The Partners Human Research Committee approved this study.

Diagnostic Interviews

Participants identified stressful events they had experienced from a list of 25 events used in diagnostic interviews, (Kessler and Ustun, 2004) and PTSD was assessed in relation to the participant's self-selected worst stressful event. Participants were cued to think of the period following the event during which symptoms were most frequent and intense. They were asked whether they had ever been bothered by each of the 17 symptoms and rated each symptom on a Likert-style scale (1: 'not at all' to 5: 'extremely') (Weathers and Ford, 1996). Additional questions assessed the other three DSM-IV criteria: intense fear, horror, or helplessness in response to the event (criterion A2), symptom duration of at least 1 month (criterion E), and clinically significant impairment in functioning owing to symptoms (criterion F) (Kessler and Ustun, 2004). Based on the diagnostic interview, we created two lifetime PTSD phenotypes as follows.

To meet criteria for the lifetime PTSD diagnosis, respondents must have endorsed experiencing one or more of the five reexperiencing symptoms, three or more of the seven avoidance/numbing symptoms, two or more of the five arousal symptoms, and criteria A2, E, and F as defined above. In addition to the diagnostic phenotype, we analyzed lifetime PTSD symptom severity, which was defined as the sum of the symptom ratings across the 17 questions.

We assessed the validity of our identification of PTSD in a separate cohort, the DNHS, via clinical interviews among a random subsample of 51 participants and found excellent concordance (Uddin *et al*, 2010).

NHSII Genotyping and Quality Control

Genotyping was performed utilizing an Illumina Infinium-Select custom 6000 bead chip system which was designed to cover: (1) haplotype-tagging and replication SNPs from candidate genes for PTSD (2) SNPs in a number of important neurotransmitter and neuroendocrine systems associated with PTSD (3) SNPs and rare CNV markers with most significant p -values from GWAS studies conducted in the last 5 years (see Supplementary Text and Supplementary Table 1 for details). Base pair positions for SNPs were annotated using the hg19 reference genome. Genotyping was performed by the University of Michigan DNA Sequencing Core. Duplicates for 10 samples were included in genotyping and the concordance of the genotypes was nearly perfect (mean, 99.97%; min, 99.68%; max, 100%). SNPs with a call rate < 0.95 , minor allele frequency (MAF) < 0.05 , or Hardy-Weinberg equilibrium $p < 10E-06$ were removed. Duplicate samples and subjects with a call rate < 0.95 were dropped. To assess population substructure, principal components analysis (PCA) was conducted using EIGENSTRAT (Price *et al*, 2006) on a set of 1598 SNPs that passed QC and were pruned with respect to LD (window size, 50 SNPs; window shift size, 5; r^2 , 0.20). As expected, NHSII participants were predominantly European-American. Based on the top two principal components, 35 subjects were identified as African-American or Asian and were removed from subsequent analyses (Supplementary Figure 1). In total, 2538 subjects and 3742 SNPs passed the quality control measures and were retained for analysis.

DNHS Sample

Participants were recruited from the DNHS, a longitudinal cohort of predominantly African-American adults living in Detroit, Michigan and screened for lifetime trauma exposure using procedures described elsewhere (Goldmann *et al*, 2011; Uddin *et al*, 2010). DNA samples were collected for 814 subjects and the present study focuses on the subset of subjects who were exposed to at least one traumatic event ($N = 748$). Fifty-six percent were female, 214 subjects had a lifetime PTSD diagnosis, using the same diagnostic algorithm as in the NHSII, and the mean severity score was 38.9 (SD:16.4). The Institutional Review Board of the University of Michigan reviewed and approved the study protocol.

DNHS Genotyping

We genotyped 730 525 SNPs in blood or saliva samples using the Illumina Human Omni Express array (Logue *et al*, 2012). The genotyping was performed at Wayne State University and the Illumina Genome Studio Genotyping Module was used for data normalization and genotype calling (details in (Logue *et al*, 2012)). A total of 778 subjects remained after removing duplicate samples and subjects with a call rate < 0.95 . A total of 688 890 SNPs passed quality control filters (call rate, $> 95\%$; MAF, > 0.01 ; Hardy-Weinberg equilibrium, $p > 1 \times 10^{-6}$). DNHS subjects were imputed to the 1000 genomes African (YRI), African-American (ASW) and European (CEU, FIN, GBR, IBS, TSI) reference panels using BEAGLE (Browning and Browning, 2009). SNPs with an imputation quality score

> 0.8 were retained and a subset of 52 SNPs (50 SNPs with $p < 0.01$ in NHSII and two additional SNPs in top NHSII chr10 region) was selected to replicate top hits in the NHSII. Among the 52 SNPs, 25 were genotyped on the original array and 27 were imputed. A PCA was conducted using EIGENSOFT on 133 542 SNPs that had been pruned with respect to LD in PLINK (window size, 50 SNPs; window shift size, 5; r^2 , 0.20). Since the DNHS samples are predominately African-American, the genotypes for the DNHS were combined with genotypes for the Caucasian (CEU) and African (YRI) Hapmap populations which were used as reference panels in the PCA. The first principal component captured the degree of admixture in the DNHS and was included in the association analysis as a covariate. The first PC captures 5.7% of the variation.

Statistical Analysis

We tested the association between each SNP and the two phenotypes (diagnosis, severity score) separately, using logistic and linear regression models in PLINK (Purcell *et al*, 2007). We assumed an additive model for the SNP. In the NHSII, the first five principal components were included as covariates in all analyses. Haplotype analyses were conducted using PLINK and were constructed using the following nine SNPs ranging from chr10, 119021407–119036625; rs10082463; rs2015586; rs363223; rs363256; rs363230; rs363272; rs2244249; rs363276; rs363279. The nine SNPs include the most significant SNP in the NHS (rs363276) and this region begins and ends with SNPs that had a consistent association in the DNHS (rs10082463 and rs363279).

To determine the study-wide significance threshold for testing 3742 correlated SNPs, we performed a permutation analysis with 5000 iterations. For each iteration, the phenotypes were permuted and the association analysis was conducted as described above. We computed the significance threshold for each phenotype by taking the 5th percentile of the minimum p -values for each phenotype (PTSD diagnosis, 2.26E-05; PTSD severity, 2.12E-05). We also used the distribution of minimum p -values to compute a corrected p -value for each SNP. We used simulation to assess the power of detecting a SNP association across various MAFs and odds ratios (ORs) and the details of the simulation are provided in the Supplementary Material.

Polygenic scoring was performed with results from the PGC (Purcell *et al*, 2009). Results for MDD, BP, and SZ were downloaded from the PGC website and included the effect estimate (OR) and p -value for each SNP (Ripke *et al*, 2011; Sklar *et al*, 2011). We selected SNPs common to the NHSII sample and each PGC disorder (2582 MDD SNPs, 3702 BP SNPs, and 2584 SZ SNPs) and removed SNPs in strong linkage disequilibrium using the 'clumping' procedure in PLINK ($p_1 = 0.5$, $p_2 = 1$, $R_2 = 0.10$, $kb = 500$). The clumping procedure preferentially retains significant SNPs in a LD block and yielded 632 MDD SNPs, 849 BP SNPs and 715 SZ SNPs. We used the PGC results to compute polygenic scores in the NHSII at varying p -value thresholds. For each disorder, we selected PGC SNPs passing a set p -value threshold, p_T , and computed a score in the NHSII corresponding to the number of risk alleles carried by each subject weighted by the log of the ORs. We then used linear and

logistic regression to test the association between the polygenic score and PTSD severity and diagnosis, respectively, in the NHSII. Polygenic analyses were adjusted for the top five principal components.

All analyses were conducted in PLINK (Purcell et al, 2007) and the R package (Team, 2013).

RESULTS

The top SNPs for each phenotype are displayed in Table 1. No excess inflation was observed in the test statistics (Supplementary Figure 2) indicating adequate quality control and no population stratification was detected. Rs363276 ($p = 2.1E-05$) was the most significant SNP and had an OR of 1.4 with PTSD cases having an excess of the A allele as compared with the controls. Rs363276 is located in an intron of *SLC18A2* and the association is supported by SNPs in the region that are in strong LD. The significance for rs363276 surpassed the multiple testing correction for the PTSD diagnosis ($p_{corrected} = 0.045$). The next most significant SNP was rs2576573 ($p = 5.5E-05$, $p_{corrected} = 0.131$) with the A allele associated with lower PTSD symptoms. Results for SNPs with a p -value < 0.01 (uncorrected) with either phenotype are provided in Supplementary Table 2. Results for SNPs in genes previously reported in candidate gene studies of PTSD (Pitman et al, 2012) are presented in Table 2. Eight genes (*SLC6A3*, *SLC6A4*, *HTR2A*, *FKBP5*, *APOE*, *BDNF*, *COMT*, and *TPH2*) had one or more SNPs that achieved nominal levels of significance ($p_{uncorrected} < 0.05$) with either PTSD severity or diagnosis although none were significant after correction for multiple testing.

To test whether an underlying haplotype was driving this pattern of associated SNPs, given the differences in LD pattern between the NHSII European-American sample and the DNHS African-American sample in the *SLC18A2* SNPs, we attempted to replicate the association using a haplotype analysis on the SNPs spanning the region starting with rs10082463 (chr10, 119021407) and ending with rs363279 (chr10,119036625). The top SNP, rs363276, which was imputed, did not replicate in the DNHS. However, three other SNPs in the region were associated with PTSD diagnosis with the same direction of effect as the NHSII at $p < 0.05$: rs10082463, rs363256, and rs363279 (Supplementary Table 3). Note that rs10082463 and rs363256 are in high LD in both the DNHS ($r2 = 1$) and the NHSII ($r2 = 0.99$). The haplotype consisted of nine SNPs spanning 15kb and included the top SNP, rs363276. The most significant

haplotype in the NHSII (CGGCGGAAG, case frequency = 0.10, control frequency = 0.08, OR = 1.35, $p = 0.0046$) replicated in the DNHS (CGGCGGAAG, case frequency = 0.10, control frequency = 0.07, OR = 1.48, $p = 0.049$) with the same direction of effect and similar frequencies in cases and controls. The risk allele (A) of the top SNP (rs363276) is present in the risk haplotype (CGGCGGAAG; 8th position). The significance of the risk haplotype is lower than the peak SNP, partially owing to a lower allele frequency, but the

Table 2 Results for Candidate Genes Reported in the Literature

Gene	# SNPs	Severity		Diagnosis	
		SNP	P	SNP	P
<i>DRD2</i>	37	rs4274224	0.237	rs1124492	0.266
<i>DRD4</i>	18	rs752306	0.292	rs7126715	0.210
<i>SLC6A3</i>	44	rs11564751	0.009	rs11564751	0.056
<i>DBH</i>	29	rs1076153	0.154	rs2073837	0.071
<i>SLC6A4</i>	43	rs12939390	0.003	rs12939390	0.006
<i>HTR2A</i>	58	rs9534507	0.027	rs10507545	0.054
<i>FKBP5</i>	90	rs2766551	0.089	rs2766581	0.041
<i>NR3C1</i>	35	rs7701443	0.209	rs7701443	0.059
<i>CRHR1</i>	28	rs242941	0.144	rs242941	0.084
<i>RGS2</i>	17	rs2746073	0.355	rs6650088	0.079
<i>CNR1</i>	11	rs7766029	0.156	rs806377 ^a	0.125
<i>APOE</i>	7	rs584007	0.020	rs584007	0.066
<i>BDNF</i>	38	rs1464896	0.062	rs2140887	0.012
<i>NPY</i>	16	rs16472	0.362	rs16472	0.358
<i>GABRA2</i>	22	rs17537141	0.303	rs693547	0.083
<i>COMT</i>	35	rs165774	0.034	rs165774	0.015
<i>ADCYAP1R1</i>	3	rs2284225	0.304	rs10226318	0.161
<i>DTNBP1</i>	2	rs3213207	0.434	rs1011313	0.096
<i>TPH1</i>	12	rs11606304	0.120	rs11606304	0.077
<i>TPH2</i>	74	rs7300641	0.048	rs17110747	0.163

^aindicates that the SNP was reported in at least one previous PTSD candidate gene study.

Table lists results for PTSD candidate genes reported in the literature and genotyped in the PTSD genetic subsample of the NHSII.¹³ For each gene, we report the number of SNPs genotyped, the genic p -value (assesses evidence of nominally significant associations in gene), the top SNP in the gene and the SNP p -value (not corrected for multiple testing). The p -value cut off after adjusting for multiple testing is 2.26E-05 for PTSD diagnosis and 2.12E-05 for PTSD severity.

Table 1 SNPs with $p < 0.001$ (uncorrected) in at Least One PTSD Phenotype

Chromosome number	SNP	BP	A1	A2	MAF	Diagnosis		Severity		Gene
						OR	P	β	P	
8	rs2576573	57357975	A	G	0.46	0.83	2.6E-03	-0.047	5.5E-05	<i>PENK</i>
10	rs363276	119033809	A	G	0.15	1.42	2.1E-05	0.032	5.0E-02	<i>SLC18A2</i>

The best SNP in each LD region is presented. The effects sizes (OR, β) correspond to the risk of the A1 allele. The effect size for PTSD severity (β) is on the cube root scale. The p -value cut off after adjusting for multiple testing is 2.26E-05 for PTSD diagnosis and 2.12E-05 for PTSD severity. Only rs363276 is significant after multiple testing adjustment.

effect size of the haplotype ($OR = 1.35$) is quite similar to the effect size of rs363276 ($OR = 1.42$). We reran this analysis, adjusting for sex, as the NHSII sample is exclusively female. The results were slightly more significant ($p = 0.03$ for the risk haplotype) and are available upon request.

We performed a candidate gene version of polygenic scoring using results from the PGC for MDD, BP, and SZ. As the p -value threshold increased in the PGC, the BP-derived scores showed a stronger association with PTSD severity but not PTSD diagnosis possibly owing to the lack of power (Table 3). The results suggest that there are shared genetic factors captured by SNPs in this array between PTSD severity and BP. The overlap exists among SNPs that exert a weak effect on BP. We note that the evidence for the MDD-derived score only appears for the most liberal p -value threshold (Table 3) and given that multiple thresholds were tested, future studies with larger sample sizes and a large panel of SNPs are needed to verify this finding.

DISCUSSION

We present the results of 3742 SNPs in more than 300 genes in a large study of PTSD with trauma-exposed European-American women. We found an association between PTSD and SNPs in *SLC18A2* and identified a risk haplotype in *SLC18A2* that had a consistent association in an independent cohort of trauma-exposed African Americans. The LD structure in *SLC18A2* differs between European and African Americans, which is one possible explanation for the lack of replication in DNHS for rs363276, the top SNP in *SLC18A2* found in the NHSII. Despite differences in LD, identifying the same risk haplotype in two different populations provides evidence that there is a causal variant (unobserved) tagged by this haplotype in both populations.

An association between PTSD and *SLC18A2* has not been previously reported. *SLC18A2*, also known as *VMAT2*, is responsible for transporting monoamine neurotransmitters (serotonin, noradrenaline, dopamine) into synaptic vesicles (Erickson *et al*, 1996). It has been implicated in several complex disorders including Parkinson's disease (Sala *et al*,

2010), depression (Christiansen *et al*, 2007), alcohol, and nicotine dependence (Schwab *et al*, 2005), prostate cancer (Sorensen *et al*, 2009), and Tourette's syndrome (Ben-Dor *et al*, 2007). *SLC18A2* is also the target of reserpine and tetrabenazine (Peter *et al*, 1993), drugs used to treat high blood pressure, agitation, tardive dyskinesia, and involuntary movements of Huntington's disease. Further research is needed to identify the mechanisms via which variation in this gene may influence risk for PTSD.

We also present the first evidence from molecular genetic data for shared genetic factors between PTSD severity and BP using the same method by Purcell *et al*. (Purcell *et al*, 2009) used to show genetic overlap between BP and SZ. Although the evidence for overlap between MDD scores and PTSD severity is weaker than for BP, the p -values for the MDD scores and PTSD severity steadily decrease with an increase in the number of SNPs suggesting that a relationship may exist; however, studies with larger samples sizes and/or a larger number of SNPs are needed to confirm this relationship. PTSD is highly comorbid with other psychiatric disorders (Breslau, 2002) and twin studies suggest that this comorbidity is at least partially explained by common genetic influences. In particular, genetic influences common to MDD account for the majority of the genetic variance in PTSD (Sartor *et al*, 2012; Wolf *et al*, 2010). Analyses of the structure of common mental disorders suggest PTSD loads with MDD on a common internalizing 'distress' or 'anxious-misery' factor (Cox *et al*, 2002; Slade and Watson, 2006). We speculate that the genetic overlap between PTSD and MDD, in particular, may reflect shared genetic influences on internalizing disorders (Wolf *et al*, 2010). The association between PTSD and BP has been less well studied. BDNF function has been associated with BP in humans and fear extinction in mice and has been proposed as a potential neurobiological factor for BP and PTSD comorbidity (Rakofsky *et al*, 2012). The evidence for overlap between PTSD and MDD and BP may be stronger using a genome-wide panel of SNPs and a larger sample.

Our study includes the following limitations. Telephone interviews were used to collect phenotypic data, which could result in PTSD misclassification, biasing our results towards the null. Although the sample size in this study is

Table 3 Polygenic Scoring

P-value threshold	Major depressive disorder			Bipolar disorder			Schizophrenia		
	N_{SNPs}	$P_{Diagnosis}$	$P_{Severity}$	N_{SNPs}	$P_{Diagnosis}$	$P_{Severity}$	N_{SNPs}	$P_{Diagnosis}$	$P_{Severity}$
0.05	103	0.908	0.907	176	0.226	0.077	140	0.981	0.631
0.1	172	0.339	0.559	282	0.429	0.197	244	0.831	0.623
0.2	309	0.067	0.127	446	0.360	0.119	381	0.555	0.285
0.3	423	0.093	0.109	611	0.479	0.091	514	0.675	0.485
0.4	530	0.086	0.073	736	0.401	0.041	621	0.663	0.383
0.5	632	0.067	0.044	849	0.397	0.045	715	0.433	0.257

Polygenic scores were derived for NHS participants using weights from PGC for major depressive disorder, bipolar disorder, and schizophrenia at varying p -value thresholds in the Psychiatric GWAS Consortium. For each disorder and p -value threshold, we report the number of SNPs used to derive the polygenic score (N_{SNPs}) and the p -value testing the association between the polygenic score and the PTSD phenotypes ($P_{Diagnosis}$, $P_{Severity}$) in NHS. The association between PTSD severity and the major depressive disorder and bipolar disorder polygenic scores becomes stronger at more liberal p -value thresholds. The most significant model for bipolar occurs at a p -value threshold of 0.4 ($p = 0.041$, $R^2 = 0.3\%$).

large in comparison to other PTSD studies, we had low power to detect weak genetic effects, as shown in the power analysis. Power analyses (Supplementary Material) found that we had adequate power (>80%) to detect SNPs with ORs 1.5–1.7 when the MAF >0.20 and OR of 1.8 with MAF of 10% (Supplementary Figure 2). These effect sizes are comparable with those observed for association between candidate genes such as *FKBP5* and *PACAP* and PTSD. Thus, we had adequate power to detect effect sizes reported by previous candidate gene studies. However, we did not have power to detect weak effects. Our polygenic analysis suggests that there may be many variants that exert weak effects on the disorder. Although our study included ~300 genes, large GWAS of PTSD are needed to identify novel regions in the genome. At this writing, three GWAS of PTSD have been published. The first GWAS of PTSD has been published using a sample of 295 cases and 196 controls and implicated retinoid-related orphan receptor alpha gene (*RORA*) (Logue *et al*, 2012). The *RORA* gene was assayed separately and was not associated with PTSD in our sample (Guffanti *et al*, under review). The second GWAS of PTSD was conducted in a European-American sample of men and women and found a genome-wide significant hit at rs406001 with the second strongest association in the Toll-like 1 gene (*TLL1*). The *TLL1* gene association was replicated in an independent sample of European Americans. There were no genome-wide significant effects in their African-American samples (Xie *et al*, 2013). The *TLL1* gene was assayed separately and was not associated with PTSD in our sample (Koenen, unpublished data). The third GWAS of PTSD was published by our group and found genome-wide significant association for one marker mapping to a novel RNA gene, lincRNA *AC068718.1*, for which we found suggestive evidence of replication in NHSII.

Our SNP array captured the majority of common variation in 20 genes previously examined in association studies of PTSD. *APOE*, *BDNF*, *COMT*, *FKBP5*, *HTR2A*, *SLC6A3*, *SLC6A4*, and *TPH2* each had a nominally significant SNP association ($p_{\text{uncorrected}} < 0.05$) with the severity score and/or diagnosis (Table 2); however, none remained significant after adjusting for multiple testing. Only one of these SNPs, rs806377 was previously reported in the literature in relation to PTSD and it was not significant in that report (Lu *et al*, 2008). Our relatively large sample size gave us power to detect effect sizes previously reported in the literature for these genes. We note that for several of the genes, such as *FKBP5*, significant effects were reported only for genotype-environment interactions. We did not test such interactions here, but plan in future work.

In addition to *SLC18A2*, the one other gene (*PENK*) with a SNP associated with PTSD at $p < 0.001$ has also not been reported in relation to PTSD. These findings suggest that testing a wide range of genes in relation to PTSD may help identify new risk variants. Very large mega analyses of GWAS for SZ and BP have shown that very large sample sizes produce many, over 100 in the case of SZ (S Ripke, WCPG oral presentation, October 2013), robust genetic associations. Growing evidence suggests that psychiatric disorders are highly polygenic and that very large samples are required to detect weak effects on disease (Sullivan *et al*, 2012). The genetic architecture of PTSD may be similar, but the architecture will be determined with large

scale GWAS of PTSD that will require pooling data across studies such as the one currently being undertaken by the newly formed PTSD working group in the PGC (Koenen *et al*, 2013).

FUNDING AND DISCLOSURE

Dr Wildman receives funds from Elsevier that are unrelated to the present work. The other authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was funded by National Institute on Mental Health awards RO1 MH093612, MH078928 to Karestan C Koenen. The Detroit Neighborhood Health Study (PI: Allison Aiello) is funded by DA022720, DA 022720-S1, and MH088283. Karestan C Koenen is also funded by grants, P51RR000165 and RC4MH092707.

REFERENCES

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders*. 4th ed. American Psychiatric Association: Washington, DC.
- Ben-Dor DH, Zimmerman S, Sever J, Roz N, Apter A, Rehavi M *et al* (2007). Reduced platelet vesicular monoamine transporter density in Tourette's syndrome pediatric male patients. *Eur Neuropsychopharmacol* 17: 523–526.
- Breslau N (2002). Epidemiologic studies of trauma, posttraumatic stress disorder, and other psychiatric disorders. *Can J Psychiatry* 47: 923–929.
- Breslau N, Chilcoat HD, Kessler RC, Peterson EL, Lucia VC (1999). Vulnerability to assaultive violence: further specification of the sex difference in post-traumatic stress disorder. *Psychol Med* 29: 813–821.
- Browning BL, Browning SR (2009). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet* 84: 210–223.
- Chang SC, Koenen KC, Galea S, Aiello AE, Soliven R, Wildman DE *et al* (2012). Molecular variation at the *SLC6A3* locus predicts lifetime risk of PTSD in the Detroit Neighborhood Health Study. *PLoS One* 7: e39184.
- Christiansen L, Tan Q, Iachina M, Bathum L, Kruse TA, McGue M *et al* (2007). Candidate gene polymorphisms in the serotonergic pathway: influence on depression symptomatology in an elderly population. *Biol Psychiatry* 61: 223–230.
- Cox BJ, Clara IP, Enns MW (2002). Posttraumatic stress disorder and the structure of common mental disorders. *Depress Anxiety* 15: 168–171.
- Department of Defense (2012a): Departments of Defense and Veterans Affairs to fund Consortia Studying Combat Related PTSD and TBI.
- Erickson JD, Schafer MK, Bonner TI, Eiden LE, Weihe E (1996). Distinct pharmacological properties and distribution in neurons and endocrine cells of two isoforms of the human vesicular monoamine transporter. *Proc Natl Acad Sci USA* 93: 5166–5171.
- Goldmann E, Aiello A, Uddin M, Delva J, Koenen K, Gant LM *et al* (2011). Pervasive exposure to violence and posttraumatic stress disorder in a predominantly African American Urban Community: the Detroit Neighborhood Health Study. *J Trauma Stress* 24: 747–751.
- Institute of Medicine of the National Academies (2012). *Treatment of posttraumatic stress disorder in military and veteran*

- populations: *Initial assessment*. The National Academies Press: Washington, DC, p1–4.
- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 62: 593–602.
- Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry* 52: 1048–1060.
- Kessler RC, Ustun TB (2004). The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res* 13: 93–121.
- Koenen KC, Aiello AE, Bakshis E, Amstadter AB, Ruggiero KJ, Acierno R *et al* (2009a). Modification of the association between serotonin transporter genotype and risk of posttraumatic stress disorder in adults by county-level social environment. *Am J Epidemiol* 169: 704–711.
- Koenen KC, De Vivo I, Rich-Edwards J, Smoller JW, Wright RJ, Purcell SM (2009b). Protocol for investigating genetic determinants of posttraumatic stress disorder in women from the Nurses' Health Study II. *BMC Psychiatry* 9: 29.
- Koenen KC, Duncan LE, Liberzon I, Ressler KJ (2013). From Candidate Genes to Genome-wide Association: The Challenges and Promise of Posttraumatic Stress Disorder Genetic Studies. *Biological psychiatry* 74: 634.
- Logue MW, Baldwin C, Guffanti G, Melista E, Wolf EJ, Reardon AF *et al* (2012). 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* 18: 937–942.
- Lu AT, Ogdie MN, Järvelin MR, Moilanen IK, Loo SK, McCracken JT *et al* (2008). Association of the cannabinoid receptor gene (CNR1) with ADHD and post-traumatic stress disorder. *Am J Med Genet B Neuropsychiatr Genet* 147: 1488–1494.
- Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium ((2012b)). A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*.
- NIMH (2008): Network(s) for Developing PTSD Risk Assessment Tools (R21).
- Peter D, Finn JP, Klisak I, Liu Y, Kojis T, Heinzmann C *et al* (1993). Chromosomal localization of the human vesicular amine transporter genes. *Genomics* 18: 720–723.
- Pitman RK, Rasmusson AM, Koenen KC, Shin LM, Orr SP, Gilbertson MW *et al* (2012). Biological studies of post-traumatic stress disorder. *Nat Rev Neurosci* 13: 769–787.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al* (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460: 748–752.
- Rakofsky JJ, Ressler KJ, Dunlop BW (2012). BDNF function as a potential mediator of bipolar disorder and post-traumatic stress disorder comorbidity. *Mol Psychiatry* 17: 22–35.
- Ressler KJ, Mercer KB, Bradley B, Jovanovic T, Mahan A, Kerley K *et al* (2011). Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor. *Nature* 470: 492–497.
- Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA *et al* (2011). Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 43: 969–976.
- Roberts AL, Dohrenwend BP, Aiello AE, Wright RJ, Maercker A, Galea S *et al* (2012a). The stressor criterion for posttraumatic stress disorder: does it matter? *J Clin Psychiatry* 73: e264–e270.
- Roberts AL, Galea S, Austin SB, Cerda M, Wright RJ, Rich-Edwards JW *et al* (2012b). Posttraumatic stress disorder across two generations: concordance and mechanisms in a population-based sample. *Biol Psychiatry* 72: 505–511.
- Roberts AL, Gilman SE, Breslau J, Breslau N, Koenen KC (2011). Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. *Psychol Med* 41: 71–83.
- Rothbaum BO, Kearns MC, Price M, Malcoun E, Davis M, Ressler KJ *et al* (2012). Early intervention may prevent the development of posttraumatic stress disorder: a randomized pilot civilian study with modified prolonged exposure. *Biol Psychiatry* 72: 957–963.
- Sala G, Brighina L, Saracchi E, Fermi S, Riva C, Carrozza V *et al* (2010). Vesicular monoamine transporter 2 mRNA levels are reduced in platelets from patients with Parkinson's disease. *J Neural Transm* 117: 1093–1098.
- Sartor CE, Grant JD, Lynskey MT, McCutcheon VV, Waldron M, Statham DJ *et al* (2012). Common heritable contributions to low-risk trauma, high-risk trauma, posttraumatic stress disorder, and major depression. *Arch Gen Psychiatry* 69: 293–299.
- Sartor CE, McCutcheon VV, Pommer NE, Nelson EC, Grant JD, Duncan AE *et al* (2011). Common genetic and environmental contributions to post-traumatic stress disorder and alcohol dependence in young women. *Psychol Med* 41: 1497–1505.
- Schwab SG, Franke PE, Hoefgen B, Guttenthaler V, Lichtermann D, Trixler M *et al* (2005). Association of DNA polymorphisms in the synaptic vesicular amine transporter gene (SLC18A2) with alcohol and nicotine dependence. *Neuropsychopharmacology* 30: 2263–2268.
- Shalev AY, Ankri Y, Israeli-Shalev Y, Peleg T, Adesky R, Freedman S (2012). Prevention of posttraumatic stress disorder by early treatment: results from the Jerusalem Trauma Outreach And Prevention study. *Arch Gen Psychiatry* 69: 166–176.
- Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N *et al* (2011). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 43: 977–983.
- Slade T, Watson D (2006). The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychol Med* 36: 1593–1600.
- Sorensen KD, Wild PJ, Mortezaei A, Adolf K, Topping N, Heeboll S *et al* (2009). Genetic and epigenetic SLC18A2 silencing in prostate cancer is an independent adverse predictor of biochemical recurrence after radical prostatectomy. *Clin Cancer Res* 15: 1400–1410.
- Stein MB, Jang KJ, Taylor S, Vernon PA, Livesley WJ (2002). Genetic and environmental influences on trauma exposure and posttraumatic stress disorder: A twin study. *Am J Psychiatry* 159: 1675–1681.
- Sullivan PF, Daly MJ, O'Donovan M (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet* 13: 537–551.
- Team RDC (2013). *R: A language and environment for statistical computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing: Vienna, Austria, <http://www.R-project.org>.
- True WJ, Rice J, Eisen SA, Heath AC, Goldberg J, Lyons MJ *et al* (1993). A twin study of genetic and environmental contributions to liability for posttraumatic stress symptoms. *Arch Gen Psychiatry* 50: 257–264.
- Uddin M, Aiello AE, Wildman DE, Koenen KC, Pawelec G, de los Santos R *et al* (2010). Epigenetic and immune function profiles associated with posttraumatic stress disorder. *Proc Natl Acad Sci USA* 107: 9470–9475.

Weathers FW, Ford J (1996). Psychometric review of PTSD Checklist (PCL-C, PCL-S, PCL-M, PCL-PR). In: Stamm BH (eds) *Measurement of Stress, Trauma, and Adaptation*. Sidran Press: Lutherville, MD, pp250–251.

Wolf EJ, Miller MW, Krueger RF, Lyons MJ, Tsuang MT, Koenen KC (2010). Posttraumatic stress disorder and the genetic structure of comorbidity. *J Abnorm Psychol* 119: 320–330.

Xie P, Kranzler HR, Yang C, Zhao H, Farrer LA, Gelernter J (2013). Genome-wide association study identifies new susceptibility loci for posttraumatic stress disorder. *Biol Psychiatry* 74: 656–663.

Yehuda R, Halligan SL, Bierer LM (2001). Relationship of parental trauma exposure and PTSD to PTSD, depressive and anxiety disorders in offspring. *J Psychiatr Res* 35: 261–270.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (<http://www.nature.com/npp>)