EUROPEAN JOURNAL OF PSYCHOTRAUMATOLOGY
2019, VOL. 10, 1688935
https://doi.org/10.1080/20008198.2019.1688935

EUROPEAN JOURNAL OF PSYCHO-TRAUMATOLOGY
THE OFFICIAL JOURNAL OF THE EUROPEAN SOCIETY FOR TRAUMATIC STRESS STUDIES

Taylor & Francis
Taylor & Francis Group

CLINICAL RESEARCH ARTICLE

A preliminary investigation of rare variants associated with genetic risk for PTSD in a natural disaster-exposed adolescent sample

Christina M. Sheerin[a], Vladimir Vladimirov[a], Vernell Williamson[b], Kaitlin Bountress[a], Carla K. Danielson[c], Kenneth Ruggiero[d]* and Ananda B. Amstadter[a]*

[a]Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA; [b]Molecular Diagnostics Laboratory, Virginia Commonwealth University, Richmond, VA, USA; [c]Departments of Nursing and Psychiatry, Medical University of South Carolina, Charleston, SC, USA; [d]Department of Psychiatry, Medical University of South Carolina, Charleston, SC, USA

ABSTRACT

Background: Posttraumatic stress disorder (PTSD) involves a complex interaction of biological, psychological, and social factors. Numerous studies have demonstrated genetic variation associated with the development of PTSD, primarily in adults. However, the contribution of low frequency and rare genetic variants to PTSD is unknown to date. Moreover, there is limited work on genetic risk for PTSD in child and adolescent populations.
Objective: This preliminary study aimed to identify the low frequency and rare genetic variation that contributes to PTSD using an exome array.
Method: This post-disaster, adolescent sample (n = 707, 51% females, M_{age} = 14.54) was assessed for PTSD diagnosis and symptom count following tornado exposure.
Results: Gene-based models, covarying for ancestry principal components, age, sex, tornado severity, and previous trauma identified variants in four genes associated with diagnosis and 276 genes associated with symptom count (at p_{adj} < .001). Functional class analyses suggested an association with variants in the nonsense class (nonsynonymous variant that results in truncation of, and usually non-functional, protein) with both outcomes. An exploratory gene network pathway analysis showed a great number of significant genes involved in brain and immune function, illustrating the usefulness of downstream examination of gene-based findings that may point to relevant biological processes.
Conclusions: While further investigation in larger samples is warranted, findings align with extant PTSD literature that has identified variants associated with biological conditions such as immune function.

Una investigación preliminar de variantes raras asociadas con riesgo genético para TEPT en una muestra de adolescentes expuestos a un desastre natural

Antecedentes: El Trastorno de estrés postraumático (TEPT) implica una compleja interacción de factores biológicos, psicológicos y sociales. Numerosos estudios han demostrado la asociación de variación genética con el desarrollo de TEPT, principalmente en adultos. Sin embargo, la contribución de variantes genéticas de baja frecuencia y raras para TEPT es desconocida a la fecha. Más aún, hay limitado trabajo en el riesgo genético para TEPT en poblaciones infantiles y adolescentes.
Objetivo: Este estudio preliminar buscó identificar variaciones genéticas de baja frecuencia y raras que contribuyen al TEPT utilizando un arreglo de exomas.
Método: La muestra de adolescentes, post-desastre (n=707, 51% mujeres, M edad = 14.54) fue evaluada para el diagnóstico de TEPT y conteo de síntomas luego de la exposición a un tornado.
Resultados: Los modelos basados en genes, covariando por componentes ancentrales principales, edad, sexo, severidad del tornado, y trauma previo, identificaron variantes en 4 genes que se asociaron al diagnóstico y 276 genes asociados con el conteo de síntomas (en p adj <.001). Los análisis funcionales de clases sugirieron una asociación con variantes en la clase sin sentido (variante no sinónimo que resulta del truncado la proteína, generalmente no funcional) con ambos resultados. Un análisis exploratorio de vías de redes genéticas mostró un gran número de genes significativos involucrados en la función cerebral e inmune, ilustrando la utilidad de la revisión aguas abajo de los hallazgos basados en genes que podrían apuntar a procesos biológicos relevantes.
Conclusiones: A pesar que se requiere mayor investigación en muestras más grandes, los hallazgos se alinean con la literatura en TEPT existente que ha identificado variantes asociadas a condiciones biológicas tales como la función inmune.

ARTICLE HISTORY
Received 4 June 2019
Revised 21 October 2019
Accepted 23 October 2019

KEYWORDS
PTSD; natural disaster; adolescent risk; rare variants; exome array

PALABRAS CLAVE
TEPT; desastre natural; riesgo adolescente; variantes raras; arreglo de exomas

关键词
自然灾害; 青少年风险; 罕见变异; 外显子组

HIGHLIGHTS

- This study aimed to identify low frequency and rare genetic variation that contributes to PTSD using an exome array in a disaster-exposed adolescent sample.
- Gene-based models, identified 4 genes associated with PTSD diagnosis and 276 genes associated with PTSD symptom count.
- Functional class analyses suggested an association with the nonsense class with both outcomes.
- Networks related to immune function appeared to be particularly relevant.

CONTACT Christina M. Sheerin christina.sheerin@vcuhealth.org Virginia Institute of Psychiatric and Behavioral Genetics, 800 E Leigh Street, Biotech One, 980126, Richmond, VA 23298, USA

*Denotes co-final authorship

Supplemental data for this article can be accessed here

一项关于自然灾害暴露青少年样本中与PTSD遗传风险相关罕见变异的初步研究

背景：创伤后应激障碍（PTSD）涉及生物，心理和社会因素间复杂的交互作用。大量研究表明，遗传变异与PTSD的发展相关，这一点主要体现在成年人中。但是，低频和罕见的遗传变异对PTSD的贡献迄今尚不清楚。此外，针对儿童和青少年人群PTSD遗传风险的工作还很有限。
目标：这项初步研究旨在使用外显子组识别促进PTSD的低频和罕见遗传变异。
方法：评估了707名龙卷风灾后青少年（51％为女性，平均年龄为14.54）遭灾后的PTSD诊断及症状计数。
结果：在世系主要成分，年龄，性别，龙卷风严重程度及先前创伤方面共变的基因模型，识别出了与诊断相关的4个基因以及与症状计数相关的276个基因上的变异（调整p值 <.001）。功能类别分析表明，这两个结果均与无义类的变异（会导致蛋白质且通常是非功能性蛋白质被截短的非同义变异）相关。探索性基因网络路径分析显示出大量与大脑及免疫功能有关的重要基因，阐明了针对可能指向相关生物过程的基因发现的下游检查的作用。
结论：尽管有必要在更大的样本中进行进一步研究，研究结果与现有识别出生物环境（例如免疫功能）相关变异的PTSD文献一致。

Natural disasters are a frequently occurring form of trauma exposure with wide-reaching effects on individuals, families, and communities (North & Pfefferbaum, 2013). Given the large numbers of individuals simultaneously affected, and the associated long-term health consequences, disasters are a significant public health concern. Children and adolescents affected by disasters, as compared to non-exposed community controls, are at increased risk for negative post-disaster outcomes, particularly post-traumatic stress symptoms (Furr, Comer, Edmunds, & Kendall, 2010) and diagnosed PTSD, with estimated rates ranging 7-32% (see review in Norris et al., 2002). Youth show greater likelihood of impairment following disasters than adults, as on average they are less well equipped to cope with disasters than are adults (Norris et al., 2002). However, many children and adolescents continue to cope well in the aftermath of a traumatic event suggesting the presence of individual differences that, if understood, might inform prevention and intervention.

The genetic contribution to PTSD has been consistently demonstrated in family (e.g. Sack, Clarke, & Seeley, 1995), twin (e.g. Stein, Jang, Taylor, Vernon, & Livesley, 2002) and, more recently, molecular genetic (Duncan et al., 2018) studies. These studies aim to identify the role that genetic factors play in the range of post-trauma responses and increase understanding of the aetiological underpinnings of PTSD. Molecular genetic studies examine the association between DNA sequence variation and a given phenotype of interest, mostly through the use of single nucleotide polymorphisms (SNPs; a substitution of a single nucleotide at a specific position in the genome, where each allele is present in the population with varying frequency). These SNPs can be used to identify genetic associations in candidate gene studies and large-scale, genome-wide association studies (GWAS) that agnostically examine very large numbers of independent common SNPs (i.e. where the less frequent (i.e. minor allele) occurs at a frequency [MAF] >1% in the population). Over 11 GWAS examining predictors of PTSD have been performed in a number of large samples and have begun to identify variants that account for some of the genetic influences on PTSD (Sheerin, Lind, Bountress, Nugent, & Amstadter, 2017).

PTSD is similar to other complex disorders wherein although independent loci have been identified via GWAS, most of the heritability of PTSD and other complex traits remains unexplained (Manolio et al., 2009; Visscher, Brown, McCarthy, & Yang, 2012). One way of identifying additional potentially causal variants may be through the examination of rare genetic variation (Cirulli & Goldstein, 2010). Less common (MAF 1%) and rare variants (MAF < 1%), are thought to contribute important heritability. However, these variants are often removed from GWAS arrays (Price et al., 2010); as such, their potential contribution to PTSD cannot be determined in standard GWAS designs. The development of exome arrays allows for specific study of these variants, by capturing low frequency and rare coding variants (Page et al., 2015; Perreault et al., 2014). The exome, the collection of known exons in the genome that are translated to proteins, makes up a small portion of the genome but contains the most functionally relevant information (Cirulli et al., 2010). An assumption of exome arrays is that the number of rare variants possessed might be associated with increased risk for a given phenotype. While most exome arrays have examined medical conditions, some have been successful in identifying polygenic rare variant contributions to phenotypes including schizophrenia (Richards et al., 2016) and addiction (Vrieze et al., 2014); however, to date, this has not been extended to PTSD.

The present study aimed to identify genetic variation that contributes to PTSD outcomes in a trauma-exposed adolescent sample using an exome array of rare variants. Beyond specific variation, functional

classes of these variants associated with PTSD and symptom count were examined to inform upon the potential biological impact of variants. Innovative features of this preliminary study include: the first application of exome arrays to the prediction of PTSD, the first genome-wide epidemiologic studies of PTSD in trauma-exposed adolescents, and the disaster-exposed design offers a powerful method for genetic studies of PTSD (e.g. ensures similar trauma exposure in cases and controls, removes confounding aspects of gene-environment correlation for trauma exposure). The sample is from a larger project, Bounce Back Now (Ruggiero et al., 2015), a web-based intervention developed for disaster-affected adolescents and their parents. This paper has the following aims: 1) to assess the relative contribution of rare genetic variants on lifetime PTSD diagnosis and symptom count, using an allelic burden model (i.e. the number of rare variants possessed) and 2) to assess functional classes of these variants (i.e. synonymous variants that do not effect protein sequence, nonsynonymous variants that do affect protein sequence, or silent variants within unknown impact). Exploratory gene network pathway analyses were also conducted to better understand potential functions of significant findings.

1. Method

1.1. Procedures

The Institutional Review Board at the Medical University of South Carolina approved all procedures. Two thousand unrelated families with adolescent offspring were recruited from areas affected by the spring 2011 Alabama and Missouri tornadoes, the costliest and deadliest tornados in US history, using a targeted address-based sampling frame for geographic areas affected by (i.e. in the paths of) the tornadoes. Recruitment procedures for the parent, longitudinal study are described in detail elsewhere (Ruggiero et al., 2015). Briefly, caregivers were eligible for the study if they were the guardian of an adolescent aged 12–17 years, had internet access in their home and had access to a telephone. One caregiver in each family and one randomly chosen adolescent between ages 12–17 in each family was interviewed via phone following informed consent and assent, respectively. During this phone interview, which occurred between 4 and 13 months following the tornados (Ruggiero et al., 2015), families provided information about demographic information, disaster characteristics, history of life stressors, substance use (e.g. alcohol, marijuana) and symptoms of PTSD and depression. Following this interview, all families were mailed a packet that included reimbursement for their interview, a letter informing them about the optional genetic component of the study for the adolescents, a saliva collection kit with detailed instructions for use, and a self-addressed postage-paid envelope for mailing the saliva specimen to the laboratory. The present study sample is comprised of participants who returned saliva samples.

1.2. Participants

Adolescents between the ages of 12–17 who resided at their household address at the time of the tornado and who provided genotypic data that passed quality control standards were included in the current study. Less than 50% of participants returned samples (n = 780). There were no differences between those who returned samples and those who did not on age or gender. However, those who did return samples were more likely (*ps*<.05) to meet criteria for PTSD, endorsed greater PTSD symptoms, and to be African-American. All of these differences were small effects.

1.3. Measures

1.3.1. Covariates

Primary analyses were adjusted for covariate effects including sex, age, ancestry principal components (see below), a measure of tornado severity, and prior traumatic life events. Tornado severity was assessed via parental report on a questionnaire created for this study, consisting of nine yes/no items, regarding whether or not the adolescent was physically injured by the tornado; was concerned about the safety of others; or experienced damage to their home, furniture, sentimental objects, vehicles, pets, land, or any other item not mentioned (Ruggiero et al., 2015). A sum score of endorsed items was used in the present study (possible range: 0–9, range in the study from 0 to 9). Prior trauma history assessment included a yes/no response to a list of 5 potentially traumatic events prior to, and not including, the tornado (e.g. accident, physical assault, other disaster; range: 0–5).

1.3.2. Lifetime PTSD diagnosis and symptom count

During the phone interview, the PTSD module from the National Survey on Adolescents structured clinical interview (Kilpatrick et al., 2003) was used to assess exposure to potentially traumatic events and assessed PTSD using the DSM-IV criterion with yes/no responses. This yielded a lifetime DSM-IV PTSD diagnosis. A PTSD symptom count was also created by summing the number of PTSD symptoms endorsed (out of 17 total; Cronbach's alpha = .88). Due to the distribution of symptom endorsements, with lower endorsement at the tail of the higher end of the distribution (i.e. 13–17), the symptom count

continuous scale, instead of 0–17, was truncated at 0–13+ symptoms.

1.4. DNA collection, genotyping, and quality control

Oragene kits (DNA Genotek OG-500) were used for the collection of saliva, and they were mailed to Yale University for isolation and genotyping. Samples were processed on the Illumina Human Exome-12 v1.1 BeadChip (Illumina, INC, San Diego, CA), which queries 247,870 variable exonic site, using the Illumina Genome Studio Software (Kermani, 2008) and following best practice approaches (see *Supplemental Methods* for detailed processing procedures). Following standard quality control approaches, SNPs with call rates (<95%) and deviations from Hardy–Weinberg equilibrium ($p < 10^{-6}$) were eliminated. Of the 780 participants who returned samples, 763 were genotyped (remaining samples were not viable); of those genotyped, 56 subjects had a GenCall (Grove et al., 2013) quality score <0.38 or a call rate <.92 and were excluded from further allele frequency calculations, resulting in 707 samples successfully genotyped by the array. Ruggiero et al. (2015) reports additional information on genetic data collection.

Self-reported ancestry was verified using a randomly selected subset of SNPs, and a Principal Components Analysis (PCA) in the larger sample (from Ruggiero et al., 2015) was conducted on 9,827 SNPs known to differentiate members of different population groups. The PCA yielded four components that captured greater than 96% of the variability. The first two PCs with eigenvalues of 1 or higher were retained, which explained 88% of the variance in self-reported ethnicity, respectively, and were used in all analyses to control for population stratification. These two components differentiate European American (EA) and African American (AA; current subsample racial composition: 67.3% EA, 29.5% AA, 3.1% Other). Details of the PCA can be found in the *Supplemental Methods.*

1.5. Statistical analyses

The two phenotypes were considered separately in gene-based analyses (i.e. a sum or aggregate of rare variants in a given gene are used). All analyses were corrected for multiple testing using the conservative false discovery rate (FDR) Benjamini Hochberg method, with an FDR-corrected significance threshold set at $p_{adj} < .001$. Analyses were adjusted to control for ancestry, age, sex, prior trauma history count, and tornado severity. Results from analyses unadjusted for trauma count and tornado severity are included in the supplemental information (*Supplemental Table 1*). It is noted that AA and EA ancestry groups were combined to increase power given the sample size. In addition to controlling for ancestry PCs, we also investigated whether there were significant differences in AA and EA allele frequency in the study. Minor allele frequencies (MAF) were extracted from both ancestry groups (based on self-report) and examined with a t-test in R, using mean MAF binned by chromosome; results showed no differences between AA/EA groups. A chi-square test was conducted as a follow-up ($X^2 = 0.0018$, df = 25, p = 1) and also showed no significant group difference (see *Supplemental Table 2* and *Supplemental Figure 1).*

Rare-variant association tests of gene-based models were conducted for diagnosis and symptom count to address the first study aim. As is common in rare-variant analyses, to increase power, gene-based tests were performed to investigate the aggregate rare variant effect on PTSD diagnostic status and symptom severity. Gene-based tests have also been shown to accommodate complicated linkage disequilibrium (LD) structure among SNPs as well as differences in size (Guo, Liu, Wang, & Zhang, 2013) and are believed to be more powerful than a single-variant-based test. Analyses for diagnostic status were conducted using the sequence kernel association test (SKAT; Wu et al., 2011), a variance component test that is quite powerful when the direction of effect is unknown (i.e. for rare variants with both disease-increasing as well as disease-decreasing impact). As SKAT is not appropriate for symptom count, analyses for symptom count were conducted in PLINK (open-source whole genome association analysis toolset; Purcell et al., 2007).

For the second aim, SKAT was run using functional classes as a parameter for both diagnosis and symptom count and defined each class as either: synonymous, nonsynonymous, missense, nonsense, or silent.

Finally, combining findings from both diagnosis and symptom count, pathway enrichment analyses were performed to explore potential functions of gene variant findings implicated as significant (FDR < 0.1). Analyses were performed with dmGWAS (Jia, Zheng, Long, Zheng, & Zhao, 2011) and Pinbpa (Wang, Matsushita, Madireddy, Mousavi, & Baranzini, 2015). Whereas Pinbpa represents a traditional pathway-based approach to enrichment analyses, dmGWAS's method of dense module searching allows the user greater localized flexibility in gene set definition and more effective utilization of local protein–protein interaction (PPI) data. Consensus-based approaches afford greater precision with respect to gene/network identification and prioritization. Association *p*-values were transformed and grouped into modules to identify cases of

significant enrichment (i.e. groups of highly significant associated variants). These areas were then permuted against disease labels and selected based on an empirical *p*-value. A fisher's exact test against a background of variants from the 1000 genomes project (Phase 3, 2013) was used to assess significance. The findings and *p*-values from both outcomes were examined within the context of six publicly available protein–protein interaction (PPI) databases (MINT, IntACT, DIP, BioGrid, HRPD, and MIPS/MPact) to identify enriched subnetworks within the larger well-known PPI datasets.

2. Results

2.1. Demographics and sample characteristics

Sample characteristics and PTSD group comparisons are presented in *Supplemental Table 1*. As expected, those with PTSD diagnosis had greater symptom severity and also had a higher tornado severity score; there were no demographic differences between those with and without PTSD.

2.2. Gene-based models

The relative contributions of rare variants to PTSD were assessed with gene boundaries defined as all SNPs that fell within (± 20 kb) of known gene boundaries. We chose to expand gene boundaries by such a margin as to include any SNPs that might be found within regulatory regions flanking the protein coding sequences, resulting in a mean of 20.58 (SD = 8.3) SNPs.

2.2.1. Lifetime PTSD diagnosis

In the analysis with lifetime diagnosis as the phenotype of interest, gene-based sum scores in four genes were significant at the $p_{adj} \leq 0.001$ and many more showed a suggestive trend towards significance (see *Supplemental Table 2*). The four genes were: M-phase phosphoprotein 9 (*MPHOSPH9*, $p_{adj} = .0002$), lectin, galactoside binding soluble 13 (*LGALS13*, $p_{adj} = .0002$), chromosome 12 open reading frame 50 (*C12orf50*, $p_{adj} = .0002$), and solute carrier family 2, member 2 (*SCL2A2*, $p_{adj} = .001$).

2.2.2. Lifetime PTSD symptom count

In the analysis with the PTSD symptom count, controlling for covariates, gene-based sum scores of 276 genes were significant at $p_{adj} \leq 0.001$, and the top 50 are presented in *Supplemental Table 2*. Specific genes of interest will be further addressed in the discussion. *Supplemental Table 3* presents additional results data from these analyses.

2.3. Functional class models

The functional component of the array was selected from multiple exome sequencing studies and was comprised of synonymous (variants without an observable or known impact on outcome), nonsynonymous (variants that do impact protein sequence), missense (results in a change in an amino acid protein), nonsense class (results in truncation of, and usually non-functional, protein), and silent (unknown impact) variants. The predicted functionality of each SNP on this platform was verified by comparing annotation files to predictions generated by the variant effect predictor (VEP; McLaren et al., 2010). The assigned functions from the original annotation file provided by Illumina were strongly correlated with that reported by VEP ($r^2 = 0.762$) and SNPs not showing complete agreement were eliminated from analyses. The total number of remaining variants was 218,813 with an average of 804.2 SNPs (SE = 283.8) per class. *Supplemental Table 3* presents the number of SNPs, breakdown of rare and common variants, and results of functional class analyses for both outcomes. In analyses predicting diagnostic status, the silent functional class and the nonsense were significantly associated with diagnostic status (corrected sum scores $p = .006$ and $p = .018$, respectively), with nominally significant findings in the missense class (combined corrected $p = .054$). Neither the synonymous nor the nonsynonymous classes were shown to be significant ($p = .379$, $p = .318$, respectively). When symptom count was the predictor, only the silent class remained significant (combined adjust $p = 7.09 \times 10^{-5}$).

2.4. Pathway analysis models

Pathway enrichment analyses of genes implicated in our primary analyses from both PTSD diagnosis and symptom count were combined to determine if there were any significant genes and/or gene pairs, using dmGWAS (Jia et al., 2011) and Pinbpa (Wang et al., 2015) programs. Genes significant in either outcome category and gene pairs that have been experimentally shown to interact in human PPI networks were evaluated in networks using a Fisher's exact test ($p < .05$), with results presented in Supplemental Table 4. Among the networks of potential interest, one of note is the *DOCK1* gene, which was included in the same network as the gene *FYN*, a protein tyrosine kinase which has been shown to have a functional role in the regulation of axon and dendrite outgrowth (Liu, Nakazawa, Tezuka, & Yamamoto, 2006). This cluster of genes, suggestive of significant interactions of those genes, is presented in Supplemental Figure 2.

3. Discussion

Although the molecular genetics literature is growing rapidly for PTSD, to date, no study has examined rare variation with regard to the disorder. Thus, the first aim of the present study was to identify genetic variation, using an exome array of rare and less common variants, that contributes to PTSD diagnosis and symptom count. The second aim was to assess functional classes of variants across both outcomes. An exploratory aim was to conduct network analyses to better understand potential functions of significant findings. We couch the following discussion of study findings within the context of the small sample size; thus, results and interpretation of findings should be considered preliminary.

3.1. Gene-based: PTSD diagnosis

Findings from the first aim, which used a gene-based, allelic burden model for rare variants and adjusted for factors known to be associated with PTSD status (e.g. exposure severity), identified a large number of nominally significant rare variants found to be associated with lifetime PTSD status, with rare variants in four genes meeting the *p*-value cut-off. First, previous work has shown that expression changes in solute carrier family 2 (*SCL2A2*), also known at *GLUT2*, are associated with physiological stress (e.g. oxidative stress; Uetake et al., 2014). Existing GWAS have linked *GLUT2* variants with increased risk for transition to type 2 diabetes and cardiovascular diseases (see review by Thorens, 2015). Second, *LGALS13* is a type of galectin, or glycan-binding protein that has been shown to regulate innate and adaptive immune responses (Johnson, Jones, Ryan, & Cobb, 2013). This finding aligns with an established link between PTSD and the immune response. Third, a large body of GWAS literature has identified variants in the *MPHOSPH9* gene associated with type 2 diabetes, in varied ancestry groups (Matsuba et al., 2016). Finally, although the variant *C12orf50 (CL050)* showed an association with the outcome, there exists no apparent evidence from the literature that seems to implicate it with any psychiatric of physical conditions. Further work is needed to determine if this is a novel finding that can be replicated.

3.2. Gene-based: PTSD symptom count

A much larger number of significant rare variants were found to be associated with lifetime PTSD symptom count. Findings of note for rare variants in two genes in particular will be discussed in more detail. First, the *DOCK2* gene, which was also nominally associated with PTSD diagnosis, is involved in several inflammatory diseases, and a recent genome-wide DNA methylation analysis identified methylation changes in sites within this gene associated with PTSD symptom severity (Mehta et al., 2017). Second, it is noteworthy that our rare-variant gene-based analysis found an association for the gene *ADCYAP1R1*. Given its relevance to the stress response system, it has been associated with PTSD (particularly in females) across several candidate gene studies (see meta-analysis by Lind et al., 2017), and represents one of the only cases within the PTSD literature wherein a candidate gene of interest was found in an agnostic study.

3.3. Functional class: PTSD diagnosis and symptom count

When analysing exons, and thus dealing with a broad class of protein coding, nonsynonymous, SNPs, there are a number of different functions that would speak to the potential biological impact on an individual. Establishing the functional classes within the present data associated with our outcomes of interest is one approach to prioritize further examination of potential biological significance of associated variants.

The functional class analysis for PTSD status, identifying association of the nonsense class is notable, as this class of variants gives rise to the most impactful type of variant in the genome (i.e. stop codons and truncation of variants). This suggests the variants that may be associated with increased risk for PTSD may be also ones that result in a non-functional protein, which likely has notable downstream effects. Continued examination of specific functional variants is useful for drug development, which generally targets proteins. The silent class was also significant with regard to PTSD status, and was the only significant class associated with symptom count. However, the biological impact of the silent class findings is difficult to interpret as they do not have known effects. It may be that these silent classes are tagging other nearby variants, which themselves could be missense or nonsense variants. Future research should examine nearby SNPs that may be a nonsynonymous variant, and therefore have an impact on protein coding.

3.4. Network analyses: combining PTSD diagnosis and symptom count

The goal of network analyses is to identify potential functional relatedness of genes to each other and to the trait of interest, looking for enrichment in the network to identify potentially interesting modules made up of relationships or interactions between the genes in the module. The network of particular interest (Supplemental Figure 2) contains the greatest number of significant genes with high confidence

interactions; these genes have roles in both brain and immune function. Descriptions of the 10 genes in this network (*e.g.* ROCK2, which plays a role in regulation of synaptic properties in the hippocampus) are presented in Supplemental Table 4. Gene ontology (GO) and KEGG pathways (see Supplemental Table 5) include a number of pathways associated with extracellular matrix organization and receptor interaction as well as signalling pathways. The number of genes from the present study examined in these analyses is small and the findings are to be considered preliminary. However, this illustrates useful downstream examinations of gene-based findings that can point to relevant processes with the ultimate goal of targeting future therapies.

3.5. Implications and future directions

Present findings should also be considered in the context of the developing genetics literature as a whole. Examination of rare variation, and use of exome arrays have increased in recent years and have proven useful for psychiatric conditions (McCarthy et al., 2016). Some of these studies have identified rare variant association signals enriched among genes that map to loci that have reached genome-wide significance in common variant GWAS (Lescai et al., 2017).

To our knowledge, this study reports on the first systematic research of rare variants using exome arrays in PTSD. While the small sample size of the present study warrants caution for over-interpreting significant gene findings, it is interesting that these genes appear to be associated with cardiovascular and immune responses, which aligns with some of the novel variants identified in the more recent PTSD GWAS (e.g. Nievergelt et al., 2015). As the literature develops in this area, post-replication agnostic approaches will be useful for identification of pathways to be examined in functional studies (e.g. Binder, 2017).

Further, potential low power due to small sample size is attenuated by the larger effect sizes that are typical for rare variants. It is also worth noting that many more significant genes were found in the examination of symptom count as compared to diagnostic status. Beyond the potential power implications, this may be due to the fact that experiencing some symptoms is normative and expected; while these do not meet the threshold of diagnosis, such symptoms are nonetheless relevant for understanding the biological processes underlying outcome. We also note that future work will be needed to gather larger powered datasets given the now known polygenicity (i.e. many genes of small effect) of complex phenotypes such as PTSD as well as their now established genetic overlap with other disorders such as bipolar disorder (Duncan, Cooper, & Shen, 2018).

3.6. Strengths and limitations

A number of study limitations are important to note. Most importantly, we may have been under-powered to detect genetic effects given the small number of individuals who carry these rare risk variants. Given the small sample, findings are considered preliminary and replication is needed; particularly in larger samples, which is particularly important when examining rare variants (Grove et al., 2013) as well as non-clinical samples. However, by selecting a relatively homogenous population, the impact of confounders that may dilute a signal was likely reduced. Second, while it was noted that AA and EA ancestries were combined, analyses suggested that these groups did not differ in minor allele frequency for these rare variants. Third, although PTSD interviews occurred following tornado exposure, we cannot be certain that these symptoms are a direct result of the tornado (versus another trauma) as lifetime PTSD was assessed. Further, while in this sample, 16% of adolescents met criteria for lifetime PTSD based on the diagnostic interview, symptom presentation as a whole in the sample was quite low. Given sample size limitations, we did not examine sex differences, but it is recognized that this is an important future direction, especially in adolescents given developmental changes. Finally, these findings only pertain to disaster-exposed youth. Future research is needed to clarify predictors of PTSD among adolescents with different types of trauma exposure (e.g. interpersonal trauma exposure). The population-based sampling approach used in the present study results in a tornado-affected sample that is likely to be more generalizable with regard to greater variability in demographics as well as in mental health and service utilization histories as compared to selectively recruited, clinical samples (Ruggiero et al., 2015).

4. Conclusions

Given the modest sample size, this paper represents preliminary genetic findings and further replication is needed. Continued molecular genetic efforts to identify both common and rare genetic variation (including use of exome arrays on functional variants) associated with risk for PTSD holds much promise for future clinical implications. This includes aiding in targeted primary and secondary prevention efforts, assisting in diagnosis and treatment planning, and beginning to target genes that may serve as useful candidates for drug development. Overall, such findings can impact upon our broad understanding of the

neurobiologic aetiology of PTSD (Lanius & Olff, 2017).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The main study was supported by NIMH under Grant R01 MH081056 (PI: Ruggiero). The genetic component was supported by supplement grant from NIMH, MH081056-01S1 (PI: Amstadter). Dr Sheerin's time is funded by NIAAA grant K01 AA025692; Dr Amstadter's time is partially funded by NIAAA grant K02 AA023239; Dr Bountress's time was funded by NIMH grant T32 MH018869. Funding sources had no involvement in the design, analysis of data, writing of the report, or decision to submit this manuscript

ORCID

Carla K. Danielson http://orcid.org/0000-0001-5069-5072

References

Binder, E. B. (2017). Dissecting the molecular mechanisms of gene x environment interactions: Implications for diagnosis and treatment of stress-related psychiatric disorders. *European Journal of Psychotraumatology*, *8*, 1412745.

Cirulli, E. T., & Goldstein, D. B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics*, *11*(6), 415–425.

Cirulli, E. T., Singh, A., Shianna, K. V., Ge, D., Smith, J. P., Maia, J. M., … Goldstein, D. B. (2010). Screening the human exome: A comparison of whole genome and whole transcriptome sequencing. *Genome Biology*, *11*(5), R57.

Duncan, L. E., Cooper, B. N., & Shen, H. (2018). Robust findings from 25 years of PTSD genetics research. *Current Psychiatry Reports*, *20*(12), 115.

Duncan, L. E., Ratanatharathorn, A., Aiello, A. E., Almli, L. M., Amstadter, A. B., Ashley-Koch, A. E., … Koenen, K. C. (2018). Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Molecular Psychiatry*, *23*(3), 666–673.

Furr, J. M., Comer, J. S., Edmunds, J. M., & Kendall, P. C. (2010). Disasters and youth: A meta-analytic examination of posttraumatic stress. *Journal of Consulting and Clinical Psychology*, *78*(6), 765–780.

Grove, M. L., Yu, B., Cochran, B. J., Haritunians, T., Bis, J. C., Taylor, K. D., … Boerwinkle, E. (2013). Best practices and joint calling of the HumanExome BeadChip: The CHARGE consortium. *PloS One*, *8*(7), e68095.

Guo, X., Liu, Z., Wang, X., & Zhang, H. (2013). Genetic association test for multiple traits at gene level. *Genetic Epidemiology*, *37*(1), 122–129.

Jia, P., Zheng, S., Long, J., Zheng, W., & Zhao, Z. (2011). dmGWAS: Dense module searching for genome-wide association studies in protein-protein interaction networks. *Bioinformatics*, *27*(1), 95–102.

Johnson, J. L., Jones, M. B., Ryan, S. O., & Cobb, B. A. (2013). The regulatory power of glycans and their binding partners in immunity. *Trends in Immunology*, *34*(6), 290–298.

Kermani, B. (2008). *Artificial intelligence and global normalization methods for genotyping*. United States, Patent no. US7035740B2. Retrieved from Artificial intelligence and global normalization methods for genotyping.

Kilpatrick, D. G., Ruggiero, K. J., Acierno, R., Saunders, B. E., Resnick, H. S., & Best, C. L. (2003). Violence and risk of PTSD, major depression, substance abuse/dependence, and comorbidity: Results from the National Survey of Adolescents. *Journal of Consulting and Clinical Psychology*, *71*(4), 692–700.

Lanius, R. A., & Olff, M. (2017). The neurobiology of PTSD. *European Journal of Psychotraumatology*, *8*(1), 1314165.

Lescai, F., Als, T. D., Li, Q., Nyegaard, M., Andorsdottir, G., Biskopsto, M., … Borglum, A. D. (2017). Whole-exome sequencing of individuals from an isolated population implicates rare risk variants in bipolar disorder. *Translational Psychiatry*, *7*(2), e1034.

Lind, M. J., Marraccini, M. E., Sheerin, C. M., Bountress, K., Bacanu, S. A., Amstadter, A. B., & Nugent, N. R. (2017). Association of posttraumatic stress disorder with rs2267735 in the ADCYAP1R1 gene: A meta-analysis. *Journal of Traumatic Stress*, *30*(4), 389–398.

Liu, H., Nakazawa, T., Tezuka, T., & Yamamoto, T. (2006). Physical and functional interaction of Fyn tyrosine kinase with a brain-enriched Rho GTPase-activating protein TCGAP. *The Journal of Biological Chemistry*, *281*(33), 23611–23619.

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., … Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747–753.

Matsuba, R., Imamura, M., Tanaka, Y., Iwata, M., Hirose, H., Kaku, K., … Maeda, S. (2016). Replication study in a Japanese population of six susceptibility loci for type 2 diabetes originally identified by a transethnic meta-analysis of genome-wide association studies. *PloS One*, *11*(4), e0154093.

McCarthy, N. S., Melton, P. E., Ward, S. V., Allan, S. M., Dragovic, M., Clark, M. L., … Jablensky, A. (2016). Exome array analysis suggests an increased variant burden in families with schizophrenia. *Schizophrenia Research*. doi:10.1016/j.schres.2016.12.007

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., & Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP effect predictor. *Bioinformatics*, *26*(16), 2069–2070.

Mehta, D., Bruenig, D., Carrillo-Roa, T., Lawford, B., Harvey, W., Morris, C. P., … Voisey, J. (2017). Genomewide DNA methylation analysis in combat veterans reveals a novel locus for PTSD. *Acta Psychiatrica Scandinavica*, *136*(5), 493–505.

Nievergelt, C. M., Maihofer, A. X., Mustapic, M., Yurgil, K. A., Schork, N. J., Miller, M. W., … Baker, D. G. (2015). 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*, *51*, 459–471.

Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims

speak: Part I. An empirical review of the empirical literature, 1981-2001. *Psychiatry*, *65*(3), 207–239.

North, C. S., & Pfefferbaum, B. (2013). Mental health response to community disasters: A systematic review. *JAMA*, *310*, 507–518.

Page, C. M., Baranzini, S. E., Mevik, B. H., Bos, S. D., Harbo, H. F., & Andreassen, B. K. (2015). Assessing the power of exome chips. *PloS One*, *10*(10), e0139642.

Perreault, L. P., Legault, M. A., Barhdadi, A., Provost, S., Normand, V., Tardif, J. C., & Dube, M. P. (2014). Comparison of genotype clustering tools with rare variants. *BMC Bioinformatics*, *15*, 52.

Price, A. L., Kryukov, G. V., de Bakker, P. I., Purcell, S. M., Staples, J., Wei, L. J., & Sunyaev, S. R. (2010). Pooled association tests for rare variants in exon-resequencing studies. *The American Journal of Human Genetics*, *86*(6), 832–838.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575.

Richards, A. L., Leonenko, G., Walters, J. T., Kavanagh, D. H., Rees, E. G., Evans, A., … O'Donovan, M. C. (2016). Exome arrays capture polygenic rare variant contributions to schizophrenia. *Human Molecular Genetics*, *25*(5), 1001–1007.

Ruggiero, K. J., Davidson, T. M., McCauley, J., Gros, K. S., Welsh, K., Price, M., … Amstadter, A. B. (2015). Bounce Back Now! Protocol of a population-based randomized controlled trial to examine the efficacy of a Web-based intervention with disaster-affected families. *Contemporary Clinical Trials*, *40*, 138–149.

Sack, W. H., Clarke, G. N., & Seeley, J. (1995). Posttraumatic stress disorder across two generations of Cambodian refugees. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*(9), 1160–1166.

Sheerin, C. M., Lind, M. J., Bountress, K., Nugent, N. R., & Amstadter, A. B. (2017). The genetics and epigenetics of PTSD: Overview, recent advances, and future directions. *Current Opinion in Psychology*, *14*, 5–11.

Stein, M. B., Jang, K. L., Taylor, S., Vernon, P. A., & Livesley, W. J. (2002). Genetic and environmental influences on trauma exposure and posttraumatic stress disorder symptoms: A twin study. *American Journal of Psychiatry*, *159*(10), 1675–1681.

Thorens, B. (2015). GLUT2, glucose sensing and glucose homeostasis. *Diabetologia*, *58*(2), 221–232.

Uetake, R., Sakurai, T., Kamiyoshi, A., Ichikawa-Shindo, Y., Kawate, H., Iesato, Y., … Shindo, T. (2014). Adrenomedullin-RAMP2 system suppresses ER stress-induced tubule cell death and is involved in kidney protection. *PloS One*, *9*(2), e87667.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *The American Journal of Human Genetics*, *90*(1), 7–24.

Vrieze, S. I., Feng, S., Miller, M. B., Hicks, B. M., Pankratz, N., Abecasis, G. R., … McGue, M. (2014). Rare nonsynonymous exonic variants in addiction and behavioral disinhibition. *Biological Psychiatry*, *75*(10), 783–789.

Wang, L., Matsushita, T., Madireddy, L., Mousavi, P., & Baranzini, S. E. (2015). PINBPA: Cytoscape app for network analysis of GWAS data. *Bioinformatics*, *31*(2), 262–264.

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., & Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *The American Journal of Human Genetics*, *89*(1), 82–93.