

ARTICLE OPEN



Discovery and replication of blood-based proteomic signature of PTSD in 9/11 responders

Monika A. Waszczuk¹, Pei-Fen Kuan², Xiaohua Yang³, Jiaju Miao⁴, Roman Kotov⁴ and Benjamin J. Luft³✉

© The Author(s) 2023

Proteomics provides an opportunity to develop biomarkers for the early detection and monitoring of post-traumatic stress disorder (PTSD). However, research to date has been limited by small sample sizes and a lack of replication. This study performed Olink Proseek Multiplex Platform profiling of 81 proteins involved in neurological processes in 936 responders to the 9/11 disaster (mean age at blood draw = 55.41 years (SD = 7.93), 94.1% white, all men). Bivariate correlations and elastic net regressions were used in a discovery subsample to identify concurrent associations between PTSD symptom severity and the profiled proteins, and to create a multiprotein composite score. In hold-out subsamples, nine bivariate associations between PTSD symptoms and differentially expressed proteins were replicated: SKR3, NCAN, BCAN, MSR1, PVR, TNFRSF21, DRAXIN, CLM6, and SCARB2 ($|r| = 0.08$ – 0.17 , $p < 0.05$). There were three replicated bivariate associations between lifetime PTSD diagnosis and differentially expressed proteins: SKR3, SIGLEC, and CPM ($OR = 1.38$ – 1.50 , $p < 0.05$). The multiprotein composite score retained 38 proteins, including 10/11 proteins that replicated in bivariate tests. The composite score was significantly associated with PTSD symptom severity ($\beta = 0.27$, $p < 0.001$) and PTSD diagnosis ($OR = 1.60$, 95% CI: 1.17–2.19, $p = 0.003$) in the hold-out subsample. Overall, these findings suggest that PTSD is characterized by altered expression of several proteins implicated in neurological processes. Replicated associations with TNFRSF21, CLM6, and PVR support the neuroinflammatory signature of PTSD. The multiprotein composite score substantially increased associations with PTSD symptom severity over individual proteins. If generalizable to other populations, the current findings may inform the development of PTSD biomarkers.

Translational Psychiatry (2023)13:8; <https://doi.org/10.1038/s41398-022-02302-4>

INTRODUCTION

Post-traumatic stress disorder (PTSD) is a persistent, debilitating psychiatric condition [1, 2] affecting up to 30% of people in high-risk groups, such as veterans and first responders [3–6]. PTSD can have deleterious physical consequences that place substantial strain on both individuals and the healthcare system [7]. Biological alterations associated with PTSD might provide an opportunity to improve detection and intervention for this condition.

Because PTSD has a complex etiology, a downstream biological signature should capture both genetic and environmental contributions to PTSD, and should be functionally proximal to the clinical presentation. Proteins meet this requirement because they perform most functions within cells and facilitate intercellular communication. Protein expression can reflect an organism's biological state, including disease-related alterations that are not detectable at the genetic or transcript level. Thus, proteins might plausibly play roles in pathological processes underpinning PTSD and serve as PTSD biomarkers.

Although a single candidate protein is unlikely to capture the complex pathophysiology of PTSD, a panel of proteins might more accurately reflect PTSD disease status. Recent progress in clinical proteomic technologies has enabled information on protein expression to be collected from many proteins in peripheral

tissues, such as the plasma. Therefore, proteomics is a preferred approach for unbiased biomarker discovery.

To date, proteomics has advanced the understanding of several psychiatric disorders, including schizophrenia, bipolar disorder, and depression [8–10]. For example, proteomics profiles derived from machine learning approaches have been found to discriminate between bipolar disorder and major depressive disorder in several independent studies, with a range from AUC = 0.67 [11] to AUC = 0.81 [12].

Differential expression of peripheral proteins has been reported in PTSD, most consistently for inflammatory cytokines [13, 14]. To our knowledge, the largest proteomics study in PTSD, exploring the expression profiles of 276 proteins, was recently conducted by our team [15]. In that study, machine learning identified 38 differentially expressed proteins in World Trade Center (WTC) responders with a probable PTSD diagnosis. Multiprotein composite scores based on the top differentially expressed proteins achieved a high accuracy in the classification of cases vs controls (AUC = 0.77–0.83). Eighteen of the top differentially expressed proteins were predominantly involved in a range of neurobiological and neuroinflammatory processes. This is in line with the well-established evidence that genes implicated in PTSD are expressed in several cortical and subcortical regions [16–19], as

¹Department of Psychology, Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA. ²Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA. ³Department of Medicine, Stony Brook University, Stony Brook, NY, USA. ⁴Department of Psychiatry, Stony Brook University, Stony Brook, NY, USA. ✉email: Benjamin.Luft@stonybrookmedicine.edu

Received: 19 November 2022 Revised: 28 November 2022 Accepted: 22 December 2022

Published online: 11 January 2023

well as with findings of widespread immune dysregulation in PTSD [20]. These initially promising results must be replicated in independent samples to better validate their biomarker potential.

Overall, a growing body of work has demonstrated that incorporating proteomic profiling strategies might elucidate the pathophysiology of psychiatric conditions and provide a biological basis for detection and intervention. It is important to note that processes such as diurnal variation in protein expression, interactions with other cells in the blood, and patient-related factors such as diet can affect peripheral measurements of proteomics. Nonetheless, plasma remains a more accessible tissue than the brain or specific cells for clinical translation purposes. The existing blood-based proteomics literature is further limited in several aspects [21]. Notably, sample sizes in prior studies have generally been small and have been further subdivided into discovery and test sets for machine learning purposes. Furthermore, replication across studies is lacking. Finally, most studies have used a case-control design, which lowers statistical power and is vulnerable to confounds such as case-control differences in socio-economic status, medical comorbidities, and treatment exposure associated with treatment-seeking [22, 23]. In line with ample evidence that biological vulnerability transcends diagnostic boundaries [24], proteomics might benefit from dimensional psychiatric definitions.

To address these limitations and to build on previous findings from Kuan et al. [15], the current study investigated associations between proteomics and PTSD in a much larger, independent sample of WTC responders. A representative cohort of more than 900 responders with a full range of PTSD symptoms and lifetime PTSD diagnostic data were profiled for 92 proteins involved in neurobiological and neuroinflammatory processes. First, we attempted to replicate findings for the 18 neurology panel proteins reported to be differentially expressed in Kuan et al. [15]. Second, to test for associations between PTSD and all profiled proteins in the current sample, we divided the sample into discovery and hold-out (replication) subsamples. In the discovery subsample, we established bivariate associations and derived a multiprotein composite score for dimensional PTSD symptom severity. The discovered associations were then replicated against PTSD symptoms as well as clinical PTSD diagnosis in the hold-out subsample.

METHODS

Participants

The participants were 1000 male WTC responders enrolled in the Stony Brook University WTC Health Program [25]. This program monitors more than 10,000 WTC responders, primarily from Long Island, NY. Blood samples were drawn routinely during monitoring examinations at the WTC Health Program. Only male participants with available blood samples were randomly selected for participation in the project. We only assayed samples from male responders because women make up <10% of the cohort and show protein expression patterns notably different from those in men [26]. The study was approved by the Institutional Reviewer Board of Stony Brook University, and all participants provided written informed consent to participate.

Sixteen samples that failed QC and 48 duplicate IDs were excluded, thus resulting in an analytic $N = 936$. The mean age at blood draw was 55.41 years ($SD = 7.93$, range = 37–81 years); 94.1% of participants were white, 94.2% were non-Hispanic, 64.4% were employed in law enforcement on 9/11, and 19.3% were exposed to a toxic dust cloud during 9/11. Full demographic and clinical information about the study sample is reported in Supplementary Table 1. The assayed sample is representative of the 9/11 responder population at the Health Program. None of the participants were previously included in the proteomics study by Kuan et al. [15].

PTSD measures

PTSD symptoms were assessed with the PTSD Checklist (PCL)–Specific Version [27]. PCL-17 is a 17-item self-reported questionnaire assessing the severity of WTC-related DSM-IV PTSD symptoms in the past month, on a five-point scale (1 = not at all to 5 = extremely). PCL-17 has been demonstrated to have sound psychometric properties [28] and had

excellent internal consistency in the current sample ($\alpha = 0.96$). The PTSD symptom assessment was concurrent with the blood draws in 70.8% of the sample; otherwise, the most recent available PTSD assessment was used. The mean lag between PTSD symptom assessment and the blood draw in the full sample was 14.59 days ($SD = 26.85$, range = 0–196 days), and among participants without a concurrent blood draw, it was 54.33 days ($SD = 22.96$, range = 1–196 days).

Two sources of diagnostic interview data were available to obtain the lifetime PTSD diagnosis. First, master's level clinical assessors were trained to administer selected modules of the Structured Clinical Interview for DSM-IV [29] for a study of WTC PTSD [6]. The second source of diagnostic data came from the Diagnostic Interview Schedule for DSM-IV [30], which was administered by trained mental health professionals to every responder at the second monitoring visit and at follow-up visits if necessary [25, 31]. Lifetime diagnosis was ascertained if responders met the diagnostic criteria at least once, according to either the Structured Clinical Interview for DSM-IV or the Diagnostic Interview Schedule for DSM-IV. Both interviews were modified to assess PTSD symptoms associated with traumatic WTC exposures (criterion A). Participants reported information on their worst episode since 9/11. The inter-rater agreement for 55 independently rated audio tapes was very good ($\kappa \geq 0.82$) [6]. Overall, diagnostic data were available on $N = 787$, 84.1% of the total sample, resulting in 122 lifetime WTC-related PTSD cases (15.5% prevalence rate). PTSD cases and controls differed significantly on PTSD symptom severity ($t(df) = 16.87$ (768), $p < 0.001$, $d = 11.62$) and prevalence of lifetime depression diagnosis ($\chi^2(1, 766) = 184.92$, $p < 0.001$, $\phi = 0.49$), but not on any demographic and other clinical characteristics, see Supplementary Table 1. No significant differences were observed in demographic characteristics and PCL severity between responders with or without diagnostic interview data.

Proteomics profiling

Fasting plasma samples were collected from the morning to noon during the participant's monitoring visit. Blood was collected in two BD Vacutainer blood collection tubes with K2EDTA and centrifuged (2000 $\times g$, 4 °C for 10 min). Plasma samples were aliquoted into a 1.5 ml polypropylene vial (0.5 ml plasma in each vial) and stored at a –80 °C freezer within 30 min of blood collection. All plasma samples were kept at –80 °C freezer until analysis. Plasma protein expression was profiled with the Olink Proseek Multiplex Platform. The Olink multiplex immunoassay was designed to provide an ultrasensitive, reproducible, highly multiplexed method for measuring protein expression. The measurement was based on state-of-the-art proximity extension assay technology [32]. We used the Olink Neurology panel consisting of 92 proteins, including markers associated with neurobiological processes and neurological diseases (e.g., neural development, axon guidance, synaptic function, or specific conditions such as Alzheimer's disease), as well as with broader roles in processes such as cellular regulation, immunology, development, and metabolism. More details are available from the manufacturer online (<https://www.olink.com>).

Proteomics data preprocessing

Several internal and external controls were added to the plasma samples for quality control to monitor protein–antibody reactions, the DNA extension step, and the detection quality of the qPCR, in order to estimate the background signal and calculate the limit of detection for Olink panels. All values below the limit of detection were coded as missing. A total of 11 proteins were excluded from the panel because of a high missing rate or failed QC, thus resulting in a final panel of 81 proteins. Proteomics data are presented as normalized protein expression values, Olink Proteomics' arbitrary units on a log scale, i.e., a difference of one normalized protein expression unit indicated a doubling of protein concentration.

Analytic approach

First, we tested whether the 18 proteins from the Olink Neurology panel that were differentially expressed in the PTSD groups in Kuan et al. [15] were significantly associated with PTSD symptom severity and PTSD diagnosis, in the full sample. Associations were adjusted for age and the time lag between PTSD assessment and blood draw, and a 5% FDR correction was applied.

Second, the total sample was randomly split in a 7:3 ratio into discovery ($N = 657$) and hold-out ($N = 279$) subsamples. Given that diagnostic data were used only for test purposes, the random split was conditional on PTSD diagnostic information being available for participants in the hold-

Table 1. Bivariate associations between proteins discovered in Kuan et al. (2020) and PTSD symptoms and diagnosis in the current sample.

Kuan et al., 2020		Total sample			
Significant proteins	Direction	PTSD symptoms (<i>N</i> = 936)		PTSD Diagnosis (<i>N</i> = 787)	
		<i>r</i>	<i>p</i> value	OR (95% CI)	<i>p</i> value
SKR3	Up	0.14	0.000	1.38 (1.14–1.68)	0.001
NCAN	Down	−0.12	0.001	0.80 (0.66–0.97)	0.022
BCAN	Down	−0.10	0.001	0.84 (0.68–1.02)	0.075
MSR1	Up	0.11	0.001	1.28 (1.04–1.58)	0.018
PVR	Up	0.09	0.006	1.24 (1.02–1.52)	0.035
TNFRSF21	Up	0.10	0.002	1.27 (1.03–1.56)	0.023
DRAXIN	Up	0.08	0.018	1.14 (0.94–1.38)	0.193
CTSS	Up	0.05	0.139	1.12 (0.92–1.36)	0.256
MDGA1	Up	0.04	0.201	1.23 (1.01–1.51)	0.041
CPA2	Up	0.01	0.713	1.12 (0.91–1.36)	0.283
EDA2R	Up	0.03	0.447	1.27 (1.00–1.62)	0.052
FLRT2	Up	0.04	0.237	1.30 (1.06–1.60)	0.011
ADAM22	Down	−0.05	0.130	1.03 (0.84–1.26)	0.775
CTSC	Up	0.00	0.915	1.12 (0.93–1.36)	0.227
CD200	Up	0.03	0.357	1.29 (1.06–1.58)	0.013
DDR1	Up	0.05	0.160	1.11 (0.91–1.35)	0.325
sFRP3	Up	−0.01	0.666	1.06 (0.87–1.29)	0.537
VWC2	Up	0.05	0.125	1.28 (1.04–1.58)	0.022

Replicated results after 5% FDR correction are bolded. Associations were adjusted for age and the time lag between PTSD assessment and blood draw.

out subsample. The discovery and hold-out subsamples did not significantly differ in PTSD diagnosis prevalence ($\chi^2(1, N = 787) = 1.93$, $p = 0.17$, $\phi = -0.05$) and PCL severity ($t(482.45) = 1.49$, $p = 0.14$, $d = 0.11$). All analyses were adjusted for age and the time lag between PTSD assessment and blood draw.

Partial correlations between PCL symptom severity and the expression levels of the remaining 63 proteins (i.e., excluding significant proteins from Kuan et al. analyzed above) were calculated in the discovery subsample. Proteins with associations at nominal $p < 0.05$ in the discovery subsample were taken forward to the hold-out subsample. Partial correlations between proteins and PCL symptom severity, and logistic regressions between proteins and lifetime PTSD diagnosis, in the hold-out subsample, were considered statistically significant replications at nominal $p < 0.05$.

To jointly test the association between PTSD and all relevant proteins, we created a multiprotein composite score by using elastic net regression. All 81 proteins and covariates were used as a candidate feature set in the discovery subsample. The optimal tuning parameters were determined via fivefold cross-validation. Next, in the hold-out subsample, linear and logistic regressions were used to estimate the associations between the multiprotein composite scores and the PCL symptom severity and PTSD diagnosis, respectively.

RESULTS

Of the 18 neurology panel proteins that emerged in Kuan et al. [15], seven showed significant bivariate associations with PTSD symptom severity at the 5% FDR level: BCAN, NCAN, MSR1, PVR, TNFRSF21, SKR3, and DRAXIN ($|r| = 0.08$ – 0.14 ; Table 1). Only SKR3 was associated with PTSD diagnosis at the 5% FDR level ($OR = 1.38$, 95% CI: 1.14–1.68). Nearly all (16/18) proteins showed a sign consistent with that in Kuan et al. [15] across PTSD symptoms and PTSD diagnosis comparisons (P for a sign test for consistency = 7.0×10^{-4}).

Across the remaining 63 profiled proteins in the current study, nine showed significant bivariate associations with PTSD symptom severity in the discovery subsample (Table 2). In the hold-out subsample, we replicated two associations with PTSD symptom

severity: SCARB2 ($r = 0.17$) and CLM6 ($r = 0.13$), as well as replicated two associations with PTSD diagnosis: CPM ($OR = 1.50$, 95% CI: 1.08–2.09) and SIGLEC1 ($OR = 1.45$ CI: 1.04–2.03). Across the discovery and hold-out subsamples, all nine proteins were associated in a consistent direction with PTSD symptom severity, and 7/9 were associated in a consistent direction with PTSD diagnosis. Overall, 11 unique protein markers were associated with PTSD; their functional pathways included cell adhesion, cellular metabolic processes, neurogenesis, and immune response (Table 3).

The multiprotein composite score created in the discovery subsample retained 38 differentially expressed proteins (Fig. 1). The proteins that loaded most strongly on the composite score were NCAN and SKR3. The composite score retained 10/11 differentially expressed proteins that replicated in bivariate tests, all with a consistent direction of effect. Finally, in the hold-out subsample, the multiprotein score was significantly associated with PCL severity ($\beta = 0.27$, $p < 0.001$) and with PTSD diagnosis status ($OR = 1.60$, 95% CI: 1.17–2.19, $p = 0.003$) (Fig. 2). Sensitivity analyses demonstrated that additional demographic covariates—race, Hispanic ethnicity, employment in law enforcement during 9/11, and 9/11 dust cloud exposure—did not change the findings.

DISCUSSION

To our knowledge, the current study reports the largest discovery and replication of blood-based proteomics findings in PTSD to date. By analyzing dimensional PTSD symptom severity in over 900 responders to the 9/11 disaster, we replicated findings for seven proteins reported previously in an independent study, identified and replicated two new protein markers associated with PTSD symptom severity, and two associated with PTSD diagnosis. Overall, we report a total of 11 unique replicated proteins associated with PTSD. Finally, we derived the multiprotein composite score and demonstrated that it significantly predicted

Table 2. Bivariate associations between significant proteins in the discovery subsample, and PTSD symptoms and diagnoses in the hold-out subsample.

Discovery subsample (<i>N</i> = 657)		Hold-out subsample (<i>N</i> = 279)			
Significant proteins	Direction	PTSD symptoms		PTSD diagnosis	
		<i>r</i>	<i>p</i> value	OR (95% CI)	<i>p</i> value
CLM6	Up	0.13	0.031	1.28 (0.94–1.75)	0.116
SCARB2	Up	0.17	0.006	1.36 (0.99–1.86)	0.057
SIGLEC1	Up	0.10	0.091	1.45 (1.04–2.03)	0.030
CPM	Up	0.10	0.094	1.50 (1.08–2.09)	0.016
RGMA	Down	−0.07	0.263	0.99 (0.73–1.36)	0.969
SMPD1	Up	0.10	0.120	0.96 (0.70–1.31)	0.786
CNTN5	Down	−0.10	0.088	0.77 (0.55–1.07)	0.114
CLM1	Up	0.09	0.132	1.20 (0.85–1.70)	0.307
TNR	Down	−0.03	0.592	1.10 (0.80–1.50)	0.560

Replicated results at nominal $p < 0.050$ are bolded. Associations were adjusted for age and the time lag between PTSD assessment and blood draw.

Table 3. Functional pathways of 11 replicated proteins associated with PTSD.

Replicated protein	PTSD definition	Functional pathways
SKR3	Symptom severity and diagnosis	Cell adhesion, cell differentiation, cell growth, cellular metabolic process, signal transduction
NCAN	Symptom severity	Cell adhesion, Cellular metabolic process
BCAN	Symptom severity	Cell adhesion, Cellular metabolic process
MSR1	Symptom severity	Cell differentiation
PVR	Symptom severity	Cell adhesion, immune response
TNFRSF21	Symptom severity	Cell adhesion, cell death, cell differentiation, immune response, neurogenesis, signal transduction
DRAXIN	Symptom severity	Axon development, axon guidance, cell death, cell differentiation, cell growth, neurogenesis, signal transduction
CLM6	Symptom severity	Immune response
SCARB2	Symptom severity	Membrane transportation
CPM	Diagnosis	Cellular metabolic process, Proteolysis
SIGLEC1	Diagnosis	Cell adhesion

PTSD symptom severity, as well as diagnostic status, in a hold-out sample. Together, the current findings make an important contribution to the understanding of the pattern of differential protein expression characterizing PTSD. If generalizable to other populations, they may aid in the development of biomarkers for detection and monitoring PTSD.

The current study replicated findings for seven differentially expressed proteins that were originally discovered by Kuan et al. (2020). Most notably, serine/threonine-protein kinase receptor R3 (SKR3) was associated with both dimensional PTSD symptom severity and PTSD diagnosis in the current sample. SKR3 is expressed in neurons and regulates normal blood vessel development.

The remaining six proteins replicated only in association with dimensional PTSD symptom severity. Neurocan (NCAN) and brevican (BCAN) were found to be downregulated in patients with higher PTSD symptom severity. NCAN is thought to be involved in modulating cell adhesion and migration, whereas BCAN plays a role in the formation of the brain's extracellular matrix, and has been implicated in brain development and synaptic plasticity. The NCAN gene variants independently emerged as genome-wide-significant risk loci for bipolar disorder and schizophrenia [33], as well as a major depressive disorder [34]. Similarly, both the NCAN and BCAN genes have been found to be

differentially expressed in postmortem brains of patients with schizophrenia [35], and differential NCAN gene expression has been reported in PTSD in a blood-based sample [36], but NCAN and BCAN expression was non-significant in postmortem brain samples of PTSD patients [17, 19].

The third replicated protein, macrophage scavenger receptor types I and II (MSR1, also known as CD204), has been implicated in many macrophage-associated physiological and pathological processes, including atherosclerosis, Alzheimer's disease, prostate cancer, and host defense [37]. The MSR1 gene has also been found to be differentially expressed in PTSD [38]. The fourth replicated protein, poliovirus receptor (PVR), has multiple roles in the immune response, including mediating NK cell adhesion and triggering their effector functions. PVR has been found to be associated with bipolar disorder, schizophrenia [39], and Alzheimer's disease [40, 41].

The fifth replicated protein, tumor necrosis factor receptor superfamily member 21 (TNFRSF21), plays a role in neuronal apoptosis and the negative regulation of oligodendrocyte maturation [42]. The TNFRSF21 gene has been found to be differentially expressed in PTSD [43, 44], including in postmortem brain samples in orbitofrontal cortex and dorsal anterior cingulate cortex regions [17], and genetic variants within this gene have emerged in genome-wide association studies of panic disorder

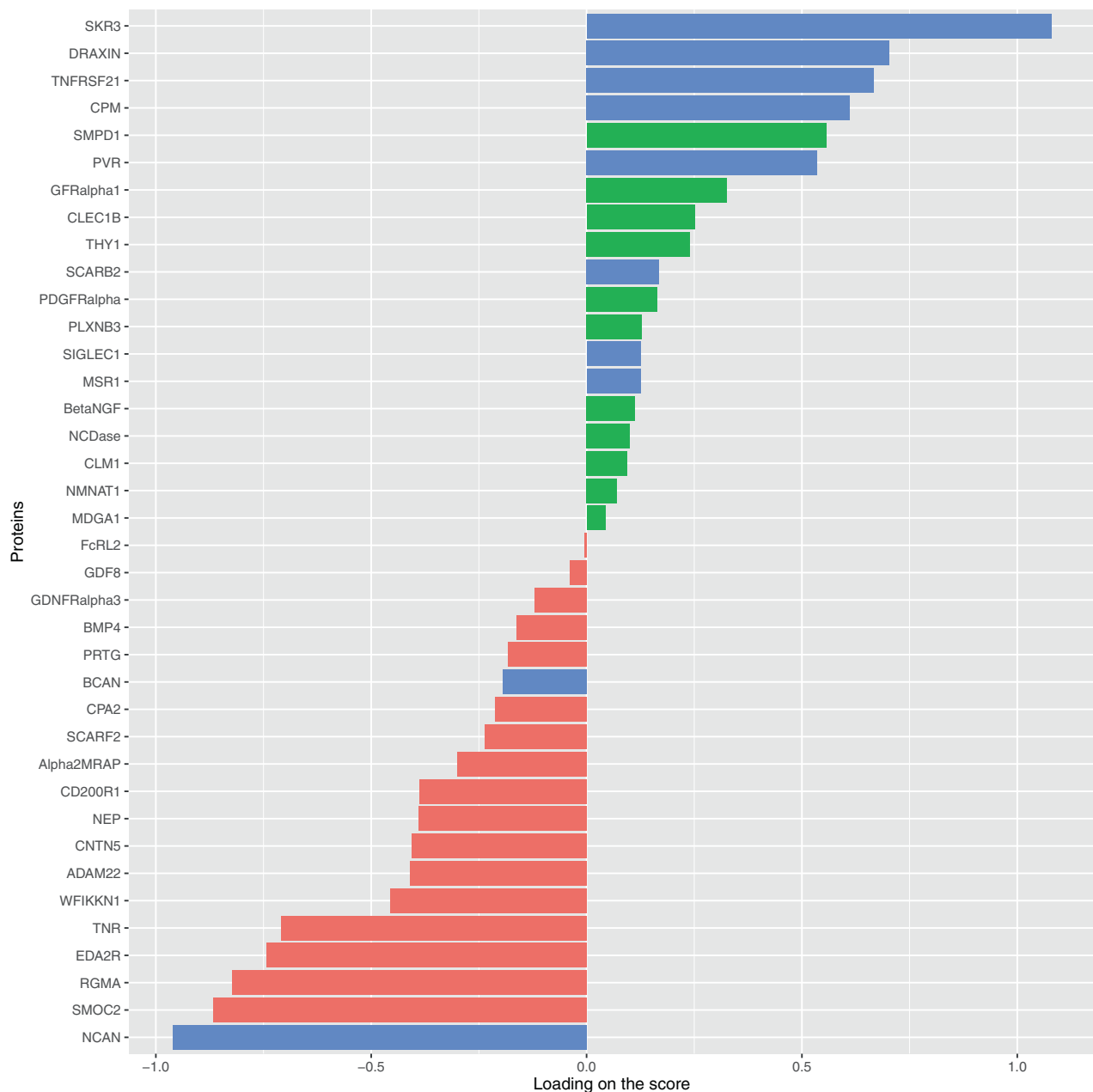


Fig. 1 Proteins retained in the multiprotein composite score. Replicated proteins from bivariate analyses of PTSD symptom severity and PTSD diagnosis are denoted by the blue bars.

[45] and depression [46, 47]. Finally, *DRAXIN* plays a role in neural development and has been implicated in autism spectrum disorder and obsessive-compulsive disorder [48].

The current study discovered and replicated four additional proteins associated with PTSD. Two proteins were upregulated in patients with higher PTSD symptom severity. The first, CMRF35-like molecule 6 (*CLM6*, also known as CD300c), is an activating receptor expressed on monocytes that play a role in immune system processes [49]. The second, lysosome membrane protein 2 (*SCARB2*, also known as CD36), is a glycoprotein that is located in the membranes of lysosomes and endosomes, and may participate in membrane transportation [50]. Differential gene expression of *SCARB2* has been found to be significantly associated with PTSD [36], and genetic variants within the *SCARB2* gene have been associated with Parkinson's disease in several genetic studies [51, 52]. Of note, although *CLM6* and *SCARB2* are intracellular

proteins, they can be released into plasma during natural cell processes as well as sample storage [53].

Two proteins found in the discovery subsample were significantly associated with PTSD diagnosis in the hold-out subsample. The first, sialoadhesin (*SIGLEC1*), is a cell adhesion molecule found on the surfaces of macrophages. Differential gene expression of *SIGLEC1* has been found to be significantly associated with PTSD [36]. The second, carboxypeptidase M (*CPM*), is an enzyme that is associated with monocyte-to-macrophage differentiation and has been implicated in metabolizing bioactive peptides, hormones, and cytokines [54].

To investigate the clinical utility of our proteomic findings, we constructed a PTSD composite multiprotein score by aggregating 38 proteins selected in the discovery sample using machine learning. All replicated proteins discussed above, with the exception of *CLM6*, contributed to the score, together with

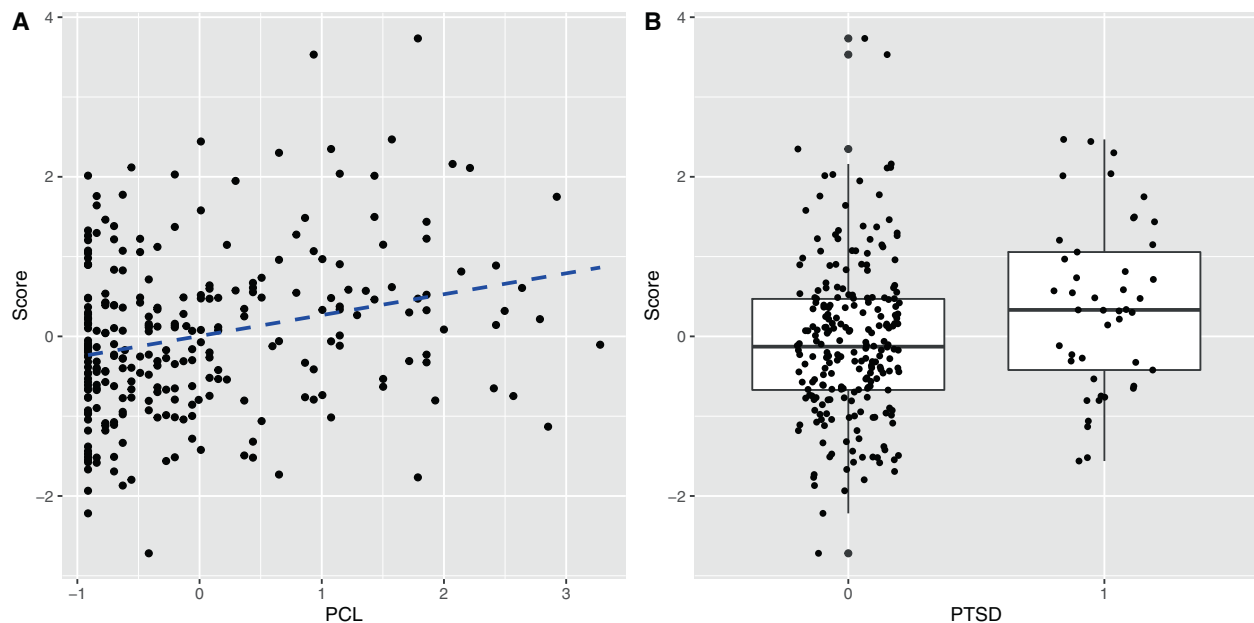


Fig. 2 The associations between the multiprotein score and **A** PTSD symptoms and **B** PTSD diagnosis, in the hold-out subsample ($N = 279$). Associations between the multiprotein score and PTSD were adjusted for age and the time lag between PTSD assessment and blood draw.

proteins that did not reach significance in bivariate comparisons. The composite multiprotein score significantly predicted PTSD symptom severity and diagnostic status in the independent hold-out sample, suggesting that the score indicates illness activity and severity, and may potentially aid in treatment monitoring. Although the current findings are promising, future longitudinal studies are needed to establish whether the multiprotein score might predict the onset and/or chronicity of future PTSD after trauma exposure.

Limitations

The strengths of the current study include its use of a state-of-the-art multiplex proteomics approach to profile a comprehensive and validated panel of neurology proteins; the replication of results in independent subsamples; and the inclusion of many participants, all of whom were exposed to a common trauma. Nonetheless, this study has several notable limitations. First, because the reported associations are cross-sectional, we were unable to determine whether the observed alterations in protein expression were a consequence of PTSD or part of its etiology. The inclusion of trauma-exposed participants without PTSD symptoms suggested that the proteomic signature is not just a consequence of trauma but is linked to PTSD; however, a longitudinal study design is needed to determine the direction of the association of protein expression with PTSD. Second, the current results must be replicated in women to better understand sex differences in protein expression, and in more diverse cohorts with other trauma exposures to clarify the generalizability of the findings.

Third, although blood testing is a relatively feasible, scalable, and non-invasive approach for obtaining biomarkers and could serve as a first step of the multistage diagnostic process, there are notable considerations for relying on blood-based proteomics in neuro-psychiatric conditions [55–58]. Protein expression in the blood reveals the proteins secreted into the blood from multiple organs, tissues, and cells. All proteins reported here are involved in neurobiology and the brain is their most likely source, but this needs to be verified in postmortem studies. Moreover, proteins originating from the brain and the central nervous system might be present at low concentrations in blood once they have crossed

the blood–brain barrier, with the dilution confounded by interactions with other plasma proteins and blood cells, as well as by degradation processes in the blood and in the liver. Diurnal variations in protein concentrations can further confound the findings, given that blood was drawn at various times throughout the morning in the current study. Numerous other factors can influence the measurement of protein expression, including demographic characteristics, medication use, diet, and other lifestyle factors. Although fasting plasma samples were collected, in the current large population-based study, we were not able to control all of these factors. However, the confounds likely have contrasting effects on protein expression, and while potentially lowering the overall observed associations, they reflect the clinical reality of translating proteomic findings to clinical practice. Studies of cognitive performance, Alzheimer's disease, and others demonstrate the utility and validity of studying neurology-related proteins in blood [58–60]. Finally, whole blood is a complex tissue and protein expression could differ if the study used serum instead of plasma, or utilized a different blood processing or storage pipeline. The current study also was not designed to test cell-specific protein expression.

CONCLUSIONS

The current study reports replication of findings for 11 differentially expressed proteins associated with PTSD, many of which play active roles in neuroinflammatory mechanisms and have been implicated in psychiatric and neurologic conditions in prior studies. The multiprotein composite score enabled significant prediction of dimensional PTSD symptom severity as well as PTSD diagnosis in a hold-out sample. Therefore, proteomics might be applied to develop useful biomarkers of PTSD for research and clinical practice. The current results will require replication in other traumatized populations to better understand their generalizability.

DATA AVAILABILITY

The proteomics data will be available at synapse.org upon publication (<https://www.synapse.org/#!Synapse:syn32140679>, <https://doi.org/10.7303/syn32140679>).

REFERENCES

- Kessler RC, Aguilar-Gaxiola S, Alonso J, Benjet C, Bromet EJ, Cardoso G, et al. Trauma and PTSD in the WHO world mental health surveys. *Eur J Psychotraumatology*. 2017;8:1353383. sup5
- Koenen K, Ratanatharathorn A, Ng L, McLaughlin K, Bromet E, Stein D, et al. Posttraumatic stress disorder in the world mental health surveys. *Psychological Med*. 2017;47:2260–74.
- Klimley KE, Van Hasselt VB, Stripling AM. Posttraumatic stress disorder in police, firefighters, and emergency dispatchers. *Aggression Violent Behav*. 2018;43:33–44.
- Petrie K, Milligan-Saville J, Gayed A, Deady M, Phelps A, Dell L, et al. Prevalence of PTSD and common mental disorders amongst ambulance personnel: a systematic review and meta-analysis. *Soc Psychiatry Psychiatr Epidemiol*. 2018;53:897–909.
- Berger W, Coutinho ESF, Figueira I, Marques-Portella C, Luz MP, Neylan TC, et al. Rescuers at risk: a systematic review and meta-regression analysis of the worldwide current prevalence and correlates of PTSD in rescue workers. *Soc Psychiatry Psychiatr Epidemiol*. 2012;47:1001–11.
- Bromet E, Hobbs M, Clouston S, Gonzalez A, Kotov R, Luft B. DSM-IV post-traumatic stress disorder among World Trade Center responders 11–13 years after the disaster of 11 September 2001 (9/11). *Psychological Med*. 2016;46:771–83.
- Ryder AL, Azcarate PM, Cohen BE. PTSD and physical health. *Curr Psychiatry Rep*. 2018;20:1–8.
- Sabherwal S, English JA, Föcking M, Cagney G, Cotter DR. Blood biomarker discovery in drug-free schizophrenia: the contribution of proteomics and multiplex immunoassays. *Expert Rev Proteom*. 2016;13:1141–55.
- Comes AL, Papiol S, Mueller T, Geyer PE, Mann M, Schulze TG. Proteomics for blood biomarker exploration of severe mental illness: pitfalls of the past and potential for the future. *Transl psychiatry*. 2018;8:1–15.
- Preece RL, Han SYS, Bahn S. Proteomic approaches to identify blood-based biomarkers for depression and bipolar disorders. *Expert Rev Proteom*. 2018;15:325–40.
- Kittel-Schneider S, Hahn T, Haenisch F, McNeill R, Reif A, Bahn S. Proteomic profiling as a diagnostic biomarker for discriminating between bipolar and unipolar depression. *Front psychiatry*. 2020;11:189.
- Shin D, Rhee SJ, Lee J, Yeo I, Do M, Joo E-J, et al. Quantitative proteomic approach for discriminating major depressive disorder and bipolar disorder by multiple reaction monitoring-mass spectrometry. *J Proteome Res*. 2021;20:3188–203.
- Rosen RL, Levy-Carrick N, Reibman J, Xu N, Shao Y, Liu M, et al. Elevated C-reactive protein and posttraumatic stress pathology among survivors of the 9/11 World Trade Center attacks. *J Psychiatr Res*. 2017;89:14–21.
- Michopoulos V, Powers A, Gillespie CF, Ressler KJ, Jovanovic T. Inflammation in fear- and anxiety-based disorders: PTSD, GAD, and beyond. *Neuropsychopharmacology*. 2017;42:254–70.
- Kuan P-F, Clouston S, Yang X, Kotov R, Bromet E, Luft BJ. Molecular linkage between post-traumatic stress disorder and cognitive impairment: a targeted proteomics study of World Trade Center responders. *Transl Psychiatry*. 2020;10:1–15.
- Stein MB, Levey DF, Cheng Z, Wendt FR, Harrington K, Pathak GA, et al. Genome-wide association analyses of post-traumatic stress disorder and its symptom subdomains in the Million Veteran Program. *Nat Genet*. 2021;53:174–84.
- Girgenti MJ, Wang J, Ji D, Cruz DA, Stein MB, Gelernter J, et al. Transcriptomic organization of the human brain in post-traumatic stress disorder. *Nat Neurosci*. 2021;24:24–33.
- Logue MW, Zhou Z, Morrison FG, Wolf EJ, Daskalakis NP, Chatzinakos C, et al. Gene expression in the dorsolateral and ventromedial prefrontal cortices implicates immune-related gene networks in PTSD. *Neurobiol Stress*. 2021;15:100398.
- Jaffe AE, Tao R, Page SC, Maynard KR, Pattie EA, Nguyen CV, et al. Decoding shared versus divergent transcriptomic signatures across cortico-amygdala circuitry in PTSD and depressive disorders. *Am J Psychiatry*. 2022;179:673–86.
- Passos IC, Vasconcelos-Moreno MP, Costa LG, Kunz M, Brietzke E, Quevedo J, et al. Inflammatory markers in post-traumatic stress disorder: a systematic review, meta-analysis, and meta-regression. *Lancet Psychiatry*. 2015;2:1002–12.
- Vinberg M. Searching for the needles in a haystack; is it needless? The search for peripheral biomarkers in psychiatry. *Front Psychiatry*. 2020;11:689.
- Preacher KJ, Rucker DD, MacCallum RC, Nicewander WA. Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychological Methods*. 2005;10:178.
- Uher R, Rutter M. Basing psychiatric classification on scientific foundation: problems and prospects. *Int Rev Psychiatry*. 2012;24:591–605.
- Waszczuk MA. The utility of hierarchical models of psychopathology in genetics and biomarker research. *World Psychiatry*. 2021;20:65.
- Dasaro CR, Holden WL, Berman KD, Crane MA, Kaplan JR, Lucchini RG, et al. Cohort profile: World Trade Center Health Program General Responder Cohort. *Int J Epidemiol*. 2017;46:e9.
- Miike K, Aoki M, Yamashita R, Takegawa Y, Saya H, Miike T, et al. Proteome profiling reveals gender differences in the composition of human serum. *Proteomics*. 2010;10:2678–91.
- Weathers FW, Litz BT, Herman DS, Huska JA, Keane T M. The PTSD checklist (PCL): reliability, validity, and diagnostic utility. *Proceedings of the Annual Convention of the International Society for Traumatic Stress Studies*;1993. San Antonio: International Society for Traumatic Stress Studies.
- Wilkins KC, Lang AJ, Norman SB. Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. *Depression Anxiety*. 2011;28:596–606.
- First MB, Spitzer RL, Gibbon M, Williams JB. Structured clinical interview for DSM-IV axis I disorders: patient edition (February 1996 Final), SCID-I/P. New York: Biometrics Research Department, New York State Psychiatric Institute; 1998.
- Robins L, Cottler L, Bucholz K, Compton W, North C, Rourke K. Diagnostic interview schedule for DSM-IV. St Louis, MO: Washington University Press; 1995.
- North CS, Pollio DE, Smith RP, King RV, Pandya A, Suris AM, et al. Trauma exposure and posttraumatic stress disorder among employees of New York City companies affected by the September 11, 2001 attacks on the World Trade Center. *Disaster Med Public Health Pre*. 2011;5:5205–13.
- Assarsson E, Lundberg M, Holmquist G, Björkstén J, Bucht Thorsén S, Ekman D, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE*. 2014;9:e95192.
- Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet*. 2019;51:793–803.
- Peyrot WJ, Price AL. Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. *Nat Genet*. 2021;53:445–54.
- Pantazopoulos H, Katsel P, Haroutunian V, Chelini G, Klengel T, Berretta S. Molecular signature of extracellular matrix pathology in schizophrenia. *Eur J Neurosci*. 2021;53:3960–87.
- Breen MS, Tylee DS, Maihofer AX, Neylan TC, Mehta D, Binder EB, et al. PTSD blood transcriptome mega-analysis: shared inflammatory pathways across biological sex and modes of trauma. *Neuropsychopharmacology*. 2018;43:469–81.
- Kelley JL, Ozment TR, Li C, Schweitzer JB, Williams DL. Scavenger receptor-A (CD204): a two-edged sword in health and disease. *Crit Rev Immunol*. 2014;34:241–61.
- Segman R, Shefi N, Goltser-Dubner T, Friedman N, Kaminski N, Shalev A. Peripheral blood mononuclear cell gene expression profiles identify emergent post-traumatic stress disorder among trauma survivors. *Mol Psychiatry*. 2005;10:500–13.
- Shao L, Vawter MP. Shared gene expression alterations in schizophrenia and bipolar disorder. *Biol Psychiatry*. 2008;64:89–97.
- Harwood JC, Leonenko G, Sims R, Escott-Price V, Williams J, Holmans P. Defining functional variants associated with Alzheimer's disease in the induced immune response. *Brain Commun*. 2021;3:fcab083.
- Zhu Z, Lin Y, Li X, Driver JA, Liang L. Shared genetic architecture between metabolic traits and Alzheimer's disease: A large-scale genome-wide cross-trait analysis. *Hum Genet*. 2019;138:271–85.
- Mi S, Lee X, Hu Y, Ji B, Shao Z, Yang W, et al. Death receptor 6 negatively regulates oligodendrocyte survival, maturation and myelination. *Nat Med*. 2011;17:816–21.
- Tylee DS, Chandler SD, Nievergelt CM, Liu X, Pazol J, Woelk CH, et al. Blood-based gene-expression biomarkers of post-traumatic stress disorder among deployed marines: a pilot study. *Psychoneuroendocrinology*. 2015;51:472–94.
- Neylan TC, Sun B, Rempel H, Ross J, Lenoci M, O'Donovan A, et al. Suppressed monocyte gene expression profile in men versus women with PTSD. *Brain Behav Immun*. 2011;25:524–31.
- Otowa T, Yoshida E, Sugaya N, Yasuda S, Nishimura Y, Inoue K, et al. Genome-wide association study of panic disorder in the Japanese population. *J Hum Genet*. 2009;54:122–6.
- Thalamuthu A, Mills NT, Berger K, Minnerup H, Grotegerd D, Dannlowski U, et al. Genome-wide interaction study with major depression identifies novel variants associated with cognitive function. *Mol Psychiatry*. 2022;27:1111–9.
- Luciano M, Huffman JE, Arias-Vásquez A, Vinkhuyzen AA, Middeldorp CM, Giegling I, et al. Genome-wide association uncovers shared genetic effects among personality traits and mood states. *Am J Med Genet Part B Neuropsychiatr Genet*. 2012;159:684–95.
- Guo W, Samuels J, Wang Y, Cao H, Ritter M, Nestadt P, et al. Polygenic risk score and heritability estimates reveals a genetic relationship between ASD and OCD. *Eur Neuropsychopharmacol*. 2017;27:657–66.
- Simhadri VR, Mariano JL, Gil-Krzewska A, Zhou Q, Borrego F. CD300c is an activating receptor expressed on human monocytes. *J Innate Immun*. 2013;5:389–400.
- Heybrock S, Kanerva K, Meng Y, Ing C, Liang A, Xiong Z-J, et al. Lysosomal integral membrane protein-2 (LIMP-2/SCARB2) is involved in lysosomal cholesterol export. *Nat Commun*. 2019;10:1–12.

51. Alcalay RN, Levy OA, Wolf P, Oliva P, Zhang XK, Waters CH, et al. SCARB2 variants and glucocerebrosidase activity in Parkinson's disease. *npj Parkinson's Dis.* 2016;2:1–4.
52. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet.* 2014;46:989–93.
53. Halvey P, Farutin V, Koppes L, Gunay NS, Pappas DA, Manning AM, et al. Variable blood processing procedures contribute to plasma proteomic variability. *Clin Proteom.* 2021;18:1–14.
54. Krause SW, Rehli M, Andreesen R. Carboxypeptidase M as a marker of macrophage maturation. *Immunol Rev.* 1998;161:119–27.
55. Smith BJ, Silva-Costa LC, Martins-de-Souza D. Human disease biomarker panels through systems biology. *Biophys Rev.* 2021;13:1179–90.
56. Martins-de-Souza D. Proteomics, metabolomics, and protein interactomics in the characterization of the molecular features of major depressive disorder. *Dialogues Clin Neurosci.* 2022;16:63–73.
57. Deutsch EW, Omenn GS, Sun Z, Maes M, Pernemalm M, Palaniappan KK, et al. Advances and utility of the human plasma proteome. *J Proteome Res.* 2021;20:5241–63.
58. Hampel H, O'Bryant SE, Molinuevo JL, Zetterberg H, Masters CL, Lista S, et al. Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. *Nat Rev Neurol.* 2018;14:639–52.
59. Dencker M, Björgell O, Hlebowicz J. Effect of food intake on 92 neurological biomarkers in plasma. *Brain Behav.* 2017;7:e00747.
60. Harris SE, Cox SR, Bell S, Marioni RE, Prins BP, Pattie A, et al. Neurology-related protein biomarkers are associated with cognitive ability and brain volume in older age. *Nat Commun.* 2020;11:1–12.

ACKNOWLEDGEMENTS

The study was funded by the NIOSH/CDC award U01OH011864 to Monika Waszczuk and Roman Kotov. We thank participants for their commitment to the project. The findings and conclusions in this article are those of the authors and do not represent the official position of NIOSH or CDC.

AUTHOR CONTRIBUTIONS

MAW, RK, and BJL conceptualized the study and secured funding; XY processed biological samples and conducted laboratory procedures; JM cleaned and described

quantitative datasets; MAW and P-FK performed statistical analyses and interpreted the results; MAW wrote the initial draft; BJL provided scientific oversight, all authors provided feedback on analyses and edited the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-022-02302-4>.

Correspondence and requests for materials should be addressed to Benjamin J. Luft.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023