ARTICLE

Genome-wide association study of traumatic brain injury in U.S. military veterans enrolled in the VA million veteran program

Victoria C. Merritt [1,2,3 ✉], Adam X. Maihofer[1,2], Marianna Gasperi[1,2,3], Catherine Chanfreau-Coffinier [4], Murray B. Stein [1,2,5], Matthew S. Panizzon[1,2,6], Richard L. Hauger[1,2,3,6], Mark W. Logue [7,8,9,10], Lisa Delano-Wood[1,2,3] and Caroline M. Nievergelt[1,2,3]

Large-scale genetic studies of traumatic brain injury (TBI) are lacking; thus, our understanding of the influence of genetic factors on TBI risk and recovery is incomplete. This study aimed to conduct a genome-wide association study (GWAS) of TBI in VA Million Veteran Program (MVP) enrollees. Participants included a multi-ancestry cohort (European, African, and Hispanic ancestries; N = 304,485; 111,494 TBI cases, 192,991 controls). TBI was assessed using MVP survey data and International Classification of Diseases (ICD) codes from the Veterans Health Administration's electronic health record. GWAS was performed using logistic regression in PLINK, and meta-analyzed in METAL. FUMA was used for post-GWAS analysis. Genomic structural equation modeling (gSEM) was conducted to investigate underlying genetic associations with TBI, and bivariate MiXeR was used to estimate phenotype specific and shared polygenicity. SNP-based heritability was 0.060 (SE = 0.004, $p = 7.83 \times 10^{-66}$). GWAS analysis identified 15 genome-wide significant (GWS) loci at $p < 5 \times 10^{-8}$. Gene-based analyses revealed 14 gene-wide significant genes; top genes included *NCAM1*, *APOE*, *FTO*, and *FOXP2*. Gene tissue expression analysis identified the brain as significantly enriched, particularly in the frontal cortex, anterior cingulate cortex, and nucleus accumbens. Genetic correlations with TBI were significant for risk-taking behaviors and psychiatric disorders, but generally not significant for the neurocognitive variables investigated. gSEM analysis revealed stronger associations with risk-taking traits than with psychiatric traits. Finally, the genetic architecture of TBI was similar to polygenic psychiatric disorders. Neurodegenerative disorders including Alzheimer's and Parkinson's disease showed much less polygenicity, however, the proportion of shared variance with TBI was high. This first well-powered GWAS of TBI identified 15 loci including genes relevant to TBI biology, and showed that TBI is a heritable trait with comparable genetic architecture and high genetic correlation with psychiatric traits. Our findings set the stage for future TBI GWASs that focus on injury severity and diversity and chronicity of symptom sequelae.

INTRODUCTION

Military personnel serving in Iraq and Afghanistan have sustained traumatic brain injuries (TBIs) at a higher frequency than prior conflicts [1], prompting widespread interest in the assessment and management of TBI and its associated sequelae. Within military populations, the overwhelming majority of TBIs are considered "closed" head injuries; among these, approximately 80% are classified as mild in severity [2]. Although designated as "mild," it has become well established that a host of clinically meaningful sequelae (i.e., neurobehavioral symptoms, psychiatric distress, and problems with thinking and cognition) develop acutely post-injury [3–5]. Despite the expectation of full recovery within days to weeks following mild TBI (mTBI) [6], these difficulties can persist for many years for some individuals [7–10]. Moreover, long-term outcome studies, even in mTBI, have demonstrated an association between TBI and increased risk of dementia, including Alzheimer's

disease (AD), as well as chronic symptoms of posttraumatic stress disorder (PTSD) and depression [8, 11].

Accumulating evidence also suggests that Veterans with a history of TBI—especially when combined with psychiatric comorbidities—experience poor health-related quality of life [12–14], increased difficulties with psychosocial functioning [15], and greater global disability [16]. The costs associated with TBI assessment and management are also substantial, with one study estimating that the median annual healthcare cost per patient was approximately four times higher for Veterans with a history of TBI compared to Veterans without a history of TBI [17]. It is thus crucial to deepen our understanding of the risk factors associated with *sustaining* a TBI as well as the underlying factors that contribute to poor outcome and recovery *following* TBI.

It has been hypothesized that there are several mechanisms by which genetics may influence risk for and response to

[1]VA San Diego Healthcare System (VASDHS), San Diego, CA, USA. [2]Department of Psychiatry, University of California San Diego, La Jolla, CA, USA. [3]Center of Excellence for Stress and Mental Health, VASDHS, San Diego, CA, USA. [4]VA Informatics and Computing Infrastructure (VINCI), VA Salt Lake City Health Care System, Salt Lake City, UT, USA. [5]School of Public Health, University of California San Diego, La Jolla, CA, USA. [6]Center for Behavior Genetics of Aging, University of California, San Diego, La Jolla, CA, USA. [7]National Center for PTSD, Behavioral Sciences Division, VA Boston Healthcare System, Boston, MA, USA. [8]Boston University Chobanian & Avedisian School of Medicine, Department of Psychiatry, Boston, MA, USA. [9]Boston University Chobanian & Avedisian School of Medicine, Biomedical Genetics, Boston, MA, USA. [10]Boston University School of Public Health, Department of Biostatistics, Boston, MA, USA. ✉email: victoria.merritt@va.gov

neurotrauma, including pre-injury risk factors for sustaining a TBI (e.g., risk-taking behaviors, impulsivity, substance use); the acute response to neurotrauma (i.e., extent of injury); the repair or recovery processes following injury (i.e., capacity for recovery); and pre- and post-injury cognitive capacity and cognitive reserve [18]. However, large-scale genetic studies of TBI are lacking and thus our understanding of the influence of genetic factors on TBI risk and recovery remains incomplete. Examining the role of genetics in TBI risk and recovery may offer novel mechanistic insights and lead to important discoveries that can inform preventive efforts and innovative treatments for affected individuals.

Among the published studies that have examined genetic associations with TBI risk and recovery, findings are considerably disparate, likely due to varied methodologies as well as inadequate sample sizes and therefore low power to detect meaningful differences in TBI samples [18–23]. Furthermore, existing studies have largely adopted a "candidate gene" approach, focusing on a specific gene of interest—most commonly, the apolipoprotein E (APOE) gene [24–28], though it is now widely believed that genetic predisposition to complex traits or conditions is highly polygenic. For example, it has been well demonstrated through GWAS that all psychiatric disorders as well as other complex phenotypes (e.g., height, body mass index) are polygenic [29, 30]. Thus, not only are adequately powered studies needed to better understand the influence of genetic markers on TBI risk and recovery but, given the likelihood that multiple genes may be involved, critical next steps are to apply the concept of polygenic risk to TBI and capitalize on recent advances in genetics and genomics research —namely, the implementation of genome-wide association studies.

Until recently, it had been very challenging to amass a sufficient number of cases to adequately carry out a genome-wide association study (GWAS) in TBI. However, the VA Million Veteran Program, a nationwide research initiative that examines how genes influence health and behaviors [31], affords a unique opportunity to explore a large cohort of U.S. military Veterans with and without a history of TBI. In this GWAS of TBI using Veterans Health Administration's (VHA) electronic health record data, we assessed TBI history among military Veterans. We examined genome-wide significant (GWS) variants associated with TBI in a multi-ancestry sample, and separately examined Veterans of European ancestry (EA), African ancestry (AA), and Hispanic ancestry (HA). We examined replication of results with data from another biobank, FinnGen (which included individuals of EA only), and assessed SNP-based heritability. We then examined genetic correlations with various health domains, including risk-taking behaviors, psychiatric disorders, neurocognition, and brain morphometrics, and evaluated phenotype-specific and shared polygenicity among these traits.

METHODS
Procedures
This research was conducted as part of project "MVP026" within the Million Veteran Program (MVP). The MVP, sponsored by the Department of Veterans Affairs Office of Research and Development, is a nationwide initiative offering military Veterans the opportunity to participate in research that seeks to better understand how genetic factors influence health. Upon consenting to participate in MVP, participants are asked to complete two comprehensive surveys—the "Baseline Survey" and "Lifestyle Survey"—and provide a blood sample for genetic analysis. Participants also consent to provide MVP investigators access to their VHA electronic health record (EHR). MVP study procedures have been previously described by Gaziano and colleagues [31]. The overarching MVP project was approved by the VA Central Institutional Review Board (IRB) in 2010 and this specific project ("MVP026") received VA Central IRB approval in 2019. All participants provided informed consent prior to MVP enrollment.

Participants
Veterans (N = 304,485; 111,494 TBI cases; 192,991 controls) were included in this study if they had both phenotype and genotyping data available (described below in more detail). The EA cohort included 249,785 Veterans (85,613 cases; 164,172 controls); the AA cohort included 35,470 Veterans (15,714 cases; 19,756 controls); and the HA cohort included 19,230 Veterans (10,167 cases; 9063 controls). See Table 1 for overall (multi-ancestry) sample characteristics and Supplemental Table 1a for EA, AA, and HA sample characteristics. Supplemental Table 1b includes sex-stratified counts for the multi-ancestry, EA, AA, and HA cohorts.

MVP data sources
The following data sources (from MVP v20_1 data release) were used to gather information pertaining to TBI history: MVP Baseline Survey, MVP Lifestyle Survey, and International Classification of Diseases, 9th Revision Clinical Modification (ICD-9) and International Classification of Diseases, 10th Revision Clinical Modification (ICD-10) codes gleaned from Veterans' EHRs.

MVP Baseline Survey. The MVP Baseline Survey gathers information pertaining to sociodemographics, military history, health status and habits, and medical history and health care usage. Within the "Medical History and Health Care Usage" section of the survey, Veterans are provided a list of health conditions and are asked to self-report whether they have ever been diagnosed with any of the health conditions. Two questions pertaining to TBI were extracted from the MVP Baseline Survey and utilized in our TBI phenotype: "Concussion or loss of consciousness" and "Traumatic brain injury" (hereafter labeled "BL-CC" and "BL-TBI," respectively).

MVP Lifestyle Survey. The MVP Lifestyle Survey gathers information pertaining to psychosocial health and activities, men's and women's health, exercise and nutrition, well-being, and military and environmental experiences. Within the "Military and Environmental Experiences" section of the survey, Veterans are asked specifically about deployment-related experiences. If a Veteran indicates that they have been deployed and served in a combat or war zone, they are asked additional questions to assess for history of deployment-related TBI. These questions, based on the Brief Traumatic Brain Injury Screen [32], were designed to assess mechanism of injury, signs or symptoms associated with TBI, and current symptoms. Specific questions were as follows: (1) "Did you have any injury(ies) during your deployment from any of the following?", with response options including "fragment," "bullet," "vehicular," "fall," "blast," "other," and "none"; (2) "Did any injury received while you were deployed result in any of the following?", with response options including "being dazed, confused or seeing stars," "not remembering the injury," "losing consciousness [knocked out] for less than a minute," "losing consciousness for 1-20 minutes," "losing consciousness for longer than 20 minutes," "having any symptoms of concussion afterward (such as headache, dizziness, irritability, etc.)," "head injury," and "none of the above"; and (3) "Are you currently experiencing any of the following problems that you think might be related to a possible head injury or concussion?", with response options including "headache," "dizziness," "memory problems," "balance problems," "ringing in the ears," "irritability," "sleep problems," and "other." Responses to the second question pertaining to signs and symptoms associated with TBI (hereafter labeled "LS-Q39") were extracted from the MVP Lifestyle Survey and utilized in our TBI phenotype.

ICD-9/ICD-10 codes. ICD-9/10 codes pertaining to TBI were extracted from the EHR. TBI-specific ICD-9/10 codes were established using surveillance case definitions for TBI set forth by the Armed Forces Health Surveillance Branch of the Military Health System. [33] The list of ICD-9 and ICD-10 codes utilized in this study is provided in Supplemental Tables 2 and 3, respectively.

TBI phenotyping
TBI case definition. To be considered a TBI case, there must have been evidence of TBI as determined by: (1) endorsement of BL-CC or (2) endorsement of BL-TBI or (3) endorsement of at least one sign/symptom associated with TBI on LS-Q39 or (4) at least one inpatient or outpatient ICD-9/10 code for TBI. In other words, Veterans with evidence of any one or more of these criteria were classified as a TBI case.

TBI control definition. To be considered a TBI control, there must have been no evidence of TBI as determined by: (1) no endorsement of BL-CC

and (2) no endorsement of BL-TBI *and* (3) no endorsement of any signs/symptoms associated with TBI on LS-Q39 *and* (4) at least one inpatient or outpatient ICD-9/10 diagnosis in the EHR but no ICD-9/10 codes for TBI at any point in the EHR. In other words, Veterans must have completed the MVP Baseline Survey and MVP Lifestyle Survey and not reported any event consistent with TBI on either survey. Additionally, there must have been evidence of VHA usage as determined by the presence of an inpatient or outpatient ICD-9/10 diagnosis, but no TBI ICD codes in the EHR.

MVP genotyping procedures
The MVP Bioinformatics Core completed the data cleaning and processing for all genetics data using standard methodology. All genetic analyses for this study utilized MVP Release 4 data. MVP genotyping procedures have been described in detail elsewhere [31, 34]; briefly, the "Affymetrix Axiom Biobank Array" was used for genotyping and standard quality control procedures were applied. Phasing occurred using Eagle v2.4 and imputation occurred through Minimac v4 using the Haplotype Reference Consortium panel [35, 36].

Ancestry assignment & principal components. For this study, the harmonized ancestry and race/ethnicity (HARE) method [37] was used to inform genetic ancestry for GWAS; ancestry groups included EA, AA, and HA. KING [38] was used to estimate relatedness (defined as kinship coefficient >0.0884); related individuals were removed from analysis by prioritizing cases and removing controls. When related pairs were both cases (or both controls), one individual was removed at random. FlashPCA2 [39] was used to compute principal components (PCs) within each ancestry group.

GWAS and meta-analysis. Separate GWASs were conducted for each ancestry group (EA, AA, and HA), followed by sex-stratified GWASs. Logistic regression analyses were performed to evaluate associations between genotype dosage data and TBI case-control status using PLINK [40], including the first 10 PCs as covariates. The standard genome-wide significance threshold was used ($p < 5 \times 10^{-8}$). SNPs with MAF < 1% or an imputation information score < 0.6 were excluded from the results. Multi-ancestry meta-analyses were conducted in METAL [41] using SNPs present in all datasets. To produce regional visualizations of GWS loci, LocusZoom 1.4 [42] was used. The 1000 Genomes phase 3 reference data (1KGPp3) [35] was utilized to calculate LD, relying on ancestry-specific reference genotypes.

Replication. GWS SNPs in the EA MVP cohort were compared to the largest available external TBI data repository—FinnGen (https://www.finngen.fi/en). Within FinnGen (Release 7), two TBI phenotypes were

Table 1. Multi-ancestry participant characteristics.

	Overall N = 304,485	TBI Case n = 111,494	Control n = 192,991	p
Age, Mean (SD)	64.36 (13.16)	59.92 (14.55)	66.92 (11.54)	<.001
Sex, N (%)				<.001
Male	281,108 (92.3)	103,378 (92.7)	177,730 (92.1)	
Female	23,370 (7.7)	8109 (7.3)	15,261 (7.9)	
Missing	7 (0.0)	7 (0.0)	0 (0.0)	
Ancestry, N (%)				<.001
European	249,785 (82.0)	85,613 (76.8)	164,172 (85.1)	
African	35,470 (11.6)	15,714 (14.1)	19,756 (10.2)	
Hispanic	19,230 (6.3)	10,167 (9.1)	9063 (4.7)	
Data Source, N (%)			--	--
BL	9039 (3.0)	9039 (8.1)		
LS	4101 (1.3)	4101 (3.7)		
ICD	23,776 (7.8)	23,776 (21.3)		
BL + LS	45,270 (14.8)	44,544 (40.0)		
BL + ICD	11,557 (3.8)	11,557 (10.4)		
LS + ICD	2558 (0.8)	2558 (2.3)		
BL + LS + ICD	208,184 (68.4)	15,919 (14.3)		
TBI Source, N (%)			--	--
BL-CC	6623 (2.2)	6623 (5.9)		
BL-TBI	1251 (0.4)	1251 (1.1)		
LS	4101 (1.3)	4101 (3.7)		
ICD	58,606 (19.2)	57,880 (51.9)		
BL-CC + BL-TBI	1165 (0.4)	1165 (1.0)		
BL-CC + LS	13,982 (4.6)	13,982 (12.5)		
BL-CC + ICD	1399 (0.5)	1399 (1.3)		
BL-TBI + ICD	3138 (1.0)	3138 (2.8)		
LS + ICD	205,154 (67.4)	12,889 (11.6)		
BL-CC + BL-TBI + ICD	3478 (1.1)	3478 (3.1)		
BL-CC + LS + ICD	2106 (0.7)	2106 (1.9)		
BL-TBI + LS + ICD	1419 (0.5)	1419 (1.3)		
All 4 Sources	2063 (0.7)	2063 (1.9)		

BL MVP Baseline Survey, *LS* MVP Lifestyle Survey, *ICD* International Statistical Classification of Diseases.
BL-CC MVP Baseline Survey, Concussion or Loss of Consciousness; *BL-TBI* MVP Baseline Survey, Traumatic Brain Injury.
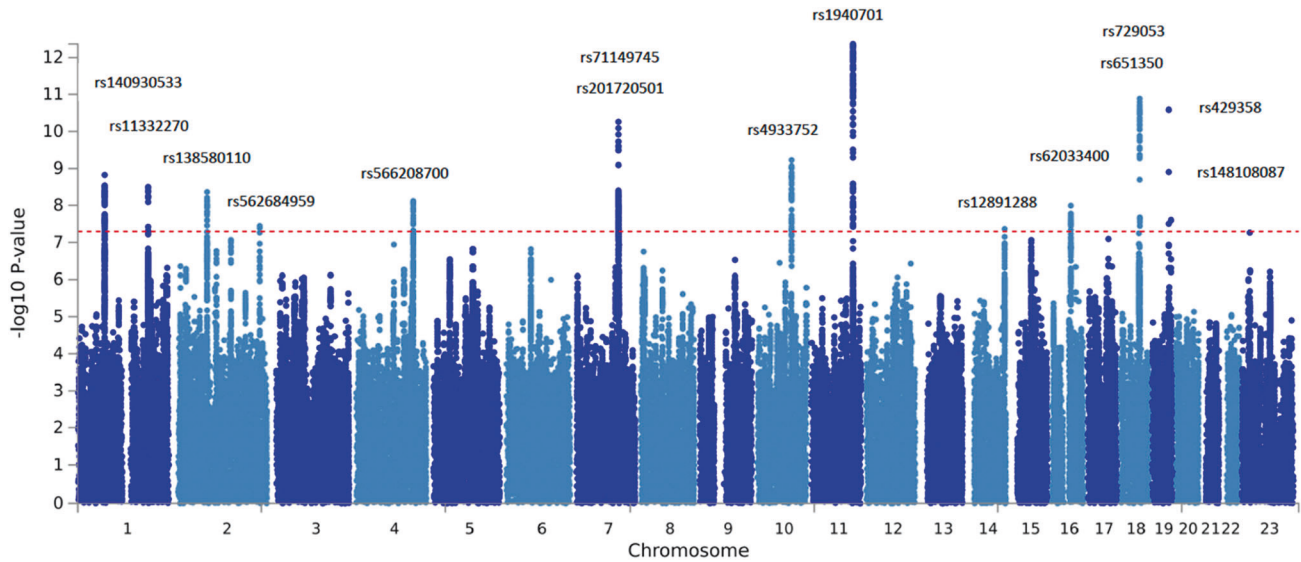
Fig. 1 Manhattan plot of TBI for the multi-ancestry cohort. The GWAS was conducted using logistic regression, including the first 10 principal components of ancestry as covariates. The x-axis displays the chromosome position, and the y-axis shows the GWAS *p*-value on a -log$_{10}$ scale. The red dashed line reflects genome-wide significance at $p < 5 \times 10^{-8}$.

examined, one for "concussion" (15,787 cases, 184,565 controls) and one for "severe TBI" (4927 cases, 304,227 controls).

Functional mapping and annotation. Functional Mapping and Annotation (FUMA) software [43] was used for functional mapping and annotation of GWAS results (default settings were used unless otherwise specified), with annotations derived from human genome assembly GRCh37 (hg19). To delineate independent genomic risk loci (defined by $r^2 > 0.6$) and variants in LD with lead SNPs, the SNP2Gene module was applied using ancestry-appropriate reference genotypes. To map risk loci to protein-coding genes, a 10 kb window size was used. eQTL mapping was completed for significant SNP-gene pairs (FDR < 0.05) using GTEx v8 brain tissue expression data.

Gene-based results and gene-pathway and tissue-enrichment analyses with MAGMA. Multi-Marker Analysis of GenoMic Annotation (MAGMA [44]) was used to conduct analyses pertaining to genes and tissue enrichment (i.e., gene-based, gene-pathway, and tissue-enrichment analyses). Regarding gene-based analyses, SNPs were mapped to 18,873 protein coding genes, and Bonferroni correction was used to establish the GWS threshold for these analyses ($p = 0.05/18,873 = 2.65\times10^{-6}$). To determine whether specific biological pathways were associated with TBI, 15,485 curated gene sets and GO terms (from MsigDB) were used in a pathway analysis; Bonferroni correction was again applied ($p = 0.05/15,485 = 3.23\times10^{-6}$). Finally, tissue-enrichment analyses were carried out using GTEx v8 RNA-seq and BrainSpan RNA-seq expression data.

SNP-based heritability estimation and genetic correlations. SNP-based heritability (SNP-h^2) was computed using linkage disequilibrium score regression (LDSC) [45] in the EA cohort. The 1KGPp3 EA reference sample was used to calculate input LD scores. Both observed scale and liability scale SNP-h^2 are reported; liability scale SNP-h^2 was computed using a population prevalence of 2% and 20%. Genetic correlations (r_g) were also computed using LDSC.

Univariate and bivariate Gaussian mixer model (MiXeR) analysis. MiXeR v1.3 [46] was used to estimate polygenicity and discoverability of causal variants for TBI. Polygenicity estimates reflect the number of loci needed to explain 90% of SNP-h^2. Bivariate MiXeR [47] was utilized to estimate phenotype specific and shared polygenicity. Akaike information criteria (AIC) values were used to evaluate model fit.

Genomic structural equation modeling (gSEM). Genomic structural equation modeling (gSEM) was conducted using GWAS summary statistics in R [48] (using the genomicSEM package [49]). LDSC [45] was carried out in the EA cohort to estimate the genetic covariance matrix and the sampling

covariance matrix using the 1KGPp3 EA reference for the following variables: TBI, general risk tolerance (GRT), major depressive disorder (MDD), attention-deficit/hyperactivity disorder (ADHD), alcohol dependence (ALCH), posttraumatic stress disorder (PTSD), reaction time (RT), verbal numeric reasoning (VNR), neuroticism (NEUR), Parkinson's disease (PD), and Alzheimer's disease (AD). Additionally, we evaluated the following brain morphometric variables: nucleus accumbens (NCA), amygdala (AM), brainstem (BS), caudate nucleus (CN), globus pallidus (GP), putamen (PU), thalamus (TH), hippocampal volume (HV), intracranial volume (ICV), full surface area (FSA), and full average thickness (FT). SNPs within the major histocompatibility region were removed in the calculation of genetic covariance.

Exploratory factor analysis (EFA) was used to assess the number of latent factors among variables pertaining to TBI, risk-taking behaviors, and psychiatric disorders. Using the R *factanal* function, EFA was carried out on the odd chromosomes and then confirmatory factor analyses (CFAs) were carried out on the covariance matrix from the even chromosomes. A factor loading threshold of 0.35 or greater was required from the EFA for variables to be assigned to factors in the CFA. The weighted least squares estimator was used to fit the data. Model fit was evaluated using the chi-square test *p*-value, comparative fit index (CFI), AIC, and standardized root mean square residual (SRMR).

Phenome-wide association study (PheWAS). A PheWAS for GWS SNPs from the TBI GWAS was conducted using GWAS Atlas (https://atlas.ctglab.nl/PheWAS) [50]. SNP rs-ID's were entered and traits were sorted by "Domain and P-value" and the max *p*-value was set to 0.05. Data were downloaded from the GWAS Atlas website and plotted in R. Bonferroni correction was applied ($p = 0.05/3,302 = 1.51\times10^{-5}$).

RESULTS

SNP-based results

The multi-ancestry meta-analysis GWAS of TBI (case-control) was conducted using a sample of 304,485 Veterans (111,494 cases; 192,991 controls). A Manhattan plot of the complete results is shown in Fig. 1. We identified 15 genome-wide significant (GWS) independent loci (Table 2). Regional annotation plots of GWS loci are presented in Supplemental Fig. 1. The leading SNP of the most significant locus (rs1940701, locus 13) was located in *NCAM1* on chromosome 11. The leading SNP of the second most significant locus (rs729053, locus 17) was in *TCF4* on chromosome 18. The leading SNP of the third most significant locus (rs429358, locus 19) was in *APOE* on chromosome 19. Notably, rs429358 is one of the two polymorphisms (alongside rs7412) that determines the *APOE* haplotype [51].

Table 2. Genome-wide significant loci associated with TBI (case-control) for multi-ancestry, EA, AA, and HA cohorts.

Locus	SNP	Chr	Start	Stop	A1	A2	Multi-Ancestry (111,494 cases, 192,991 controls)				European Ancestry (85,613 cases, 164,172 controls)				African Ancestry (15,714 cases, 19,756 controls)				Hispanic Ancestry (10,167 cases, 9,063 controls)				Genes
							A1 Freq	Z-score	p-value	HetPVal	A1 Freq	OR	SE	p-value	A1 Freq	OR	SE	p-value	A1 Freq	OR	SE	p-value	
1	rs140930533[a]	1	73,766,431	73,991,651	CAG	C	0.570	-0.035	1.49E-09	0.680	0.573	0.968	0.006	4.09E-07	0.641	0.956	0.017	0.007	0.406	0.955	0.022	0.037	ZRANB2, LRRC53, LHX8
2	rs623111[b]	1	190,886,930	191,418,368	A	C	0.404	0.034	3.33E-09	0.242	0.373	1.038	0.006	1.78E-09	0.670	1.008	0.017	0.649	0.340	1.027	0.022	0.235	BRINP3, RGS21
2	rs11322270[a]	1	191,154,894	191,418,368	G	GA	0.403	0.034	3.16E-09	0.309	0.372	1.038	0.006	2.28E-09	0.671	1.010	0.017	0.548	0.339	1.026	0.022	0.246	RGS21
3	rs4557006[b]	2	22,430,795	22,493,637	A	G	0.441	0.028	5.96E-07	0.063	0.442	1.034	0.006	4.87E-08	0.342	1.009	0.016	0.568	0.611	0.986	0.023	0.542	NA
4	rs138580110[a]	2	78,623,429	78,838,751	A	ATTAT	0.316	-0.036	4.33E-09	0.751	0.333	0.966	0.006	1.04E-07	0.122	0.948	0.024	0.027	0.277	0.968	0.024	0.170	NA
5	rs34799259[b]	2	144,151,161	144,248,905	A	AT	0.579	-0.030	1.35E-07	0.191	0.565	0.966	0.006	2.95E-08	0.612	0.995	0.017	0.753	0.717	0.988	0.024	0.611	ARHGAP15, GTDC1, ZEB2
6	rs17394072[b]	2	222,869,560	222,872,562	A	G	0.507	-0.031	4.05E-08	0.078	0.476	0.965	0.006	2.95E-09	0.819	1.002	0.020	0.928	0.547	0.997	0.021	0.903	FAM124B
6	rs56284959[a]	2	222,869,560	222,921,656	CT	C	0.535	-0.032	3.54E-08	0.770	0.512	0.970	0.006	8.60E-07	0.668	0.959	0.017	0.017	0.608	0.978	0.023	0.324	PAX3, CCDC140, SGPP2, FAM124B
7	rs4377571[b]	4	154,650,106	154,768,011	A	T	0.454	0.032	1.89E-08	0.073	0.464	1.037	0.006	7.22E-09	0.411	0.996	0.017	0.821	0.403	1.040	0.022	0.069	TMEM154, TIGD4, ARFIP1, FHDC1, MND1, TLR2, RNF175, SFRP2, PLRG1, FGA
7	rs5662700[a]	4	154,650,106	154,768,011	G	GGTGTGTGTGTGT	0.564	-0.033	7.64E-09	0.193	0.539	0.966	0.006	1.58E-08	0.738	0.998	0.019	0.913	0.642	0.952	0.023	0.030	TMEM154, TIGD4, ARFIP1, FHDC1, MND1, TLR2, RNF175, SFRP2, PLRG1, FGA
8	rs147355048[b]	5	107,402,938	107,597,787	A	AGGATGGGATG	0.555	-0.030	1.51E-07	0.144	0.529	0.967	0.006	4.57E-08	0.825	0.975	0.022	0.253	0.620	1.011	0.022	0.617	EFNA5, FBXL17, MAN2A1, TMEM232, SLC25A46
9	rs12705715[b]	7	110,267,021	110,284,071	A	G	0.490	-0.028	7.19E-07	0.057	0.466	0.967	0.006	3.87E-08	0.786	1.004	0.020	0.851	0.430	1.005	0.022	0.826	LRRN3, TMEM168
10	rs71149745[a,b]	7	113,858,363	114,290,415	A	AAATTTCAT	0.597	0.038	5.46E-11	0.848	0.576	1.040	0.006	3.28E-10	0.839	1.028	0.022	0.207	0.613	1.034	0.022	0.134	PPP1R3A, FOXP2, MDFIC
11	rs201720501[a]	7	114,940,147	115,113,279	T	TG	0.561	0.034	1.02E-08	0.812	0.538	1.034	0.006	2.91E-07	0.744	1.031	0.019	0.114	0.575	1.049	0.023	0.036	WNT2
12	rs34350548[b]	10	94,577,629	94,750,745	T	TA	0.362	-0.038	1.96E-09	0.179	0.380	0.958	0.007	4.31E-10	0.204	0.999	0.023	0.978	0.275	0.981	0.028	0.487	KIF11, EXOC6, CYP26C1, CYP26A1, MYOF, CEP55, LGI1
12	rs4933752[a]	10	94,570,602	94,756,363	T	C	0.307	-0.039	6.00E-10	0.937	0.323	0.962	0.007	3.33E-09	0.084	0.954	0.032	0.141	0.191	0.969	0.028	0.252	KIF11, EXOC6, CYP26C1, CYP26A1, MYOF, CEP55, LGI1
13	rs1940701[a,b]	11	112,826,311	113,048,466	T	C	0.595	-0.040	4.38E-13	0.303	0.601	0.961	0.006	1.32E-10	0.609	0.973	0.016	0.078	0.506	0.934	0.021	0.001	IL18, TEX12, AP002884.3, BCO2, NCAM1, TMPRSS5

Table 2. continued

Locus	SNP	Chr	Start	Stop	A1	A2	Multi-Ancestry (111,494 cases, 192,991 controls)				European Ancestry (85,613 cases, 164,172 controls)				African Ancestry (15,714 cases, 19,756 controls)				Hispanic Ancestry (10,167 cases, 9,063 controls)				Genes
							A1 Freq	Z-score	p-value	HetPVal	A1 Freq	OR	SE	p-value	A1 Freq	OR	SE	p-value	A1 Freq	OR	SE	p-value	
14	rs12891288[a]	14	103,849,715	104,053,764	T	C	0.277	0.035	**4.29E-08**	0.758	0.275	1.036	0.007	4.65E-07	0.277	1.039	0.018	0.032	0.306	1.018	0.024	0.452	ANKRD9, TNFAIP2, EIF5, MARK3, CKB, TRMT61A, BAG5, KLC1, RP11-73M18.2, APOPT1, XRCC3, AL049840.1, ZFYVE21, PPP1R13B, C14orf2, TDRD9, KIF26A
15	rs9972653[b]	16	53,797,908	53,845,487	T	G	0.395	-0.030	**4.54E-08**	0.284	0.398	0.966	0.006	**2.31E-08**	0.418	0.991	0.015	0.567	0.289	0.979	0.023	0.350	RPGRIP1L, FTO, RP11-324D17.1, IRX3, IRX5, IRX6, MMP2, CES1, GNAO1, MT1B, MT1F, MT1G, MT1H, MT1X
15	rs62033400[a]	16	53,797,908	53,845,487	A	G	0.628	0.033	**1.02E-08**	0.858	0.604	1.035	0.006	**2.37E-08**	0.897	1.026	0.025	0.317	0.744	1.024	0.024	0.333	RPGRIP1L, FTO, RP11-324D17.1, IRX3, IRX5, IRX6, MMP2, CES1, GNAO1, MT1B, MT1F, MT1G, MT1H, MT1X
16	rs79940062[b]	17	4,708,126	4,879,323	A	G	NA	NA	NA	NA	0.023	1.129	0.020	**9.11E-10**	NA	NA	NA	NA	NA	NA	NA	NA	UBE2G1, MED11, VMO1, GLTPD2, PSMB6, PLD2, MINK1, CHRNE, C17orf107, SLC25A11, RNF167, PFN1, ENO3, SPAG7, CAMTA2, INCA1, KIF1C, SCIMP, NLRP1
17	rs4608411[b]	18	52,765,283	52,827,668	T	C	0.322	0.041	**2.70E-11**	0.210	0.334	1.041	0.007	**8.44E-10**	0.111	1.010	0.026	0.688	0.343	1.073	0.023	0.002	RAB27B, TCF4, TXNL1, WDR7, NARS, NEDD4L
17	rs729053[a]	18	52,765,283	52,827,668	C	G	0.324	0.040	**1.32E-11**	0.302	0.335	1.040	0.006	**1.04E-09**	0.097	1.022	0.027	0.412	0.342	1.073	0.022	0.001	RAB27B, TCF4, TXNL1, WDR7, NARS, NEDD4L
18	rs618869[b]	18	53,198,836	53,456,801	T	C	NA	NA	NA	NA	0.145	0.951	0.009	**3.50E-08**	NA	NA	NA	NA	0.080	0.970	0.043	0.484	STARD6, C18orf54
18	rs651350[a]	18	53,210,302	53,464,917	A	G	0.876	0.049	**2.14E-08**	0.825	0.871	1.051	0.009	5.29E-08	0.975	1.071	0.052	0.186	0.938	1.028	0.044	0.529	STARD6, C18orf54

Table 2. continued

Locus	SNP	Chr	Start	Stop	A1	A2	Multi-Ancestry (111,494 cases, 192,991 controls)				European Ancestry (85,613 cases, 164,172 controls)				African Ancestry (15,714 cases, 19,756 controls)				Hispanic Ancestry (10,167 cases, 9,063 controls)				Genes
							A1 Freq	Z-score	p-value	HetPVal	A1 Freq	OR	SE	p-value	A1 Freq	OR	SE	p-value	A1 Freq	OR	SE	p-value	
19	rs429358[a,b]	19	45,392,254	45,424,351	T	C	0.851	-0.051	**2.60E-11**	0.335	0.862	0.949	0.009	**1.08E-09**	0.787	0.945	0.018	0.002	0.890	0.996	0.033	0.905	PVRL2, TOMM40, APOE, APOC1, APOC4, APOC2, APOC4, APOC2, CTB-129P6.11, CLPTM1, CLASRP, TRAPPC6A, BLOC1S3, AC005779.2
20	rs148108087[a]	19	51,781,455	51,815,081	A	G	0.053	-0.079	**2.50E-08**	0.857	0.055	0.925	0.015	**1.13E-07**	0.010	0.907	0.086	0.259	0.019	0.884	0.088	0.161	NAPSA, C19orf48, SIGLEC11, IGLON5, VSIG10L, ETFB, CTD-2616J11.11, CLDND2, NKG7, ZNF577

Bolded p-values reflect statistically significant loci.

For multi-ancestry analysis, a SNP had to be present in all 3 ancestries (EA, AA, and HA).

There were no independently significant hits in the African or Hispanic ancestries.

[a]Top SNP (within each locus) in the multi-ancestry cohort (in total, 15 SNPs were GWS in the multi-ancestry cohort).

[b]Top SNP (within each locus) in the European ancestry cohort (in total, 15 SNPs were GWS in the EA cohort).

When examining Veterans of EA only (N = 249,785; 85,613 cases, 164,172 controls), we similarly identified 15 GWS independent loci. As in the multi-ancestry analysis, the most significant locus was in *NCAM1*, with the same leading variant (rs1940701) identified. Of the 15 loci, 4 were distinct from the multi-ancestry analysis (Table 2, Supplemental Fig. 2), but all had been suggestive of significance ($p < 1 \times 10^{-6}$) in the multi-ancestry meta-analysis. The leading SNP of the most significant of these 4 distinct loci was in *ARHGAP15* on chromosome 2 (rs34799259, locus 5). Supplemental Table 4 shows functional mapping and annotation of the TBI GWAS results in the EA cohort. No GWS loci were identified in Veterans of AA (N = 35,470; 15,714 cases, 19,756 controls) or HA (N = 19,230; 10,167 cases, 9063 controls; Table 2, Supplemental Figs. 3 and 4).

Sex-stratified results
Within the multi-ancestry cohort, males accounted for 93% of the sample. The results of the male-stratified multi-ancestry GWAS were similar to the overall (non-sex stratified) multi-ancestry GWAS, though fewer loci were identified in the smaller males-only GWAS compared to the overall GWAS. Specifically, the males-only GWAS of TBI (103,146 cases, 177,507 controls) identified 13 GWS independent loci (Supplemental Table 5; Supplemental Fig. 5). Of these loci, all but one had already been identified in the overall GWAS. The single locus that was novel in the males-only multi-ancestry GWAS compared to the overall multi-ancestry GWAS (leading SNP rs70948955, located on chromosome 2, $p = 3.85 \times 10^{-8}$) was suggestive of significance in the overall GWAS ($p = 1.7 \times 10^{-7}$). No significant loci were identified in the females-only GWAS (8104 cases, 15,257 controls).

Replication
We performed replication analysis of the 15 EA GWS loci in the FinnGen concussion phenotype GWAS (15,787 cases, 184,565 controls; Supplemental Table 6); rs71149745 (locus 10, *FOXP2*) remained significant after adjustment for multiple comparisons. Three additional loci were at least nominally significant (at $p < .05$): (1) rs1940701, nearest to *NCAM1* (locus 13); (2) rs12705715 (locus 9); and (3) rs1320139 (a proxy variant for rs34799259, locus 5), nearest to *ARHGAP15*. A sign test was performed to determine whether the direction of effects corresponded across studies at a rate beyond chance (50%). The direction of effects matched for 13 of the 15 variants, a significantly greater rate of correspondence than chance ($p = .001$). When comparing the MVP EA case-control TBI phenotype to the FinnGen severe TBI phenotype (4927 cases, 304,227 controls; Supplemental Table 4), two of the 15 possible SNPs nominally replicated at $p < .05$: (1) rs71149745, nearest to *FOXP2* (locus 10) and (2) rs4933752 (a proxy variant for rs34350548, locus 12), nearest to *EXOC6*.

Gene-based analyses
In the genome-wide gene-based association study in the multi-ancestry cohort, 14 genes were determined to be gene-wide significant following Bonferroni correction for multiple comparisons (Table 3, Supplemental Table 7, Supplemental Fig. 6a). The top gene identified was *NCAM1* ($p = 4.10 \times 10^{-12}$), which was already named in positional mapping of the SNP-based analysis results, followed by *APOE* ($p = 1.30 \times 10^{-9}$), *FTO* ($p = 6.94 \times 10^{-9}$), and *FOXP2* ($p = 2.68 \times 10^{-8}$), also all positionally mapped to significant loci from the SNP-based analysis. In contrast, five of the significant genes from this analysis (*MSRA, FAM120A, NAV3, RBFOX1*, and *DCC*) were not identified based on positional mapping of loci from the SNP-based analysis. *MSRA* is located on the chromosome 8 inversion region [52] where LD patterns are such that the association may be tagging the inversion itself rather than *MSRA* specifically.

We also examined GWS genes in Veterans of EA, AA, and HA separately. In the EA cohort, 18 genes were determined to be GWS

Table 3. Gene-based loci analysis (Bonferroni significant) in the multi-ancestry and European ancestry cohorts.

SYMBOL	CHR	START	STOP	Multi-Ancestry				European Ancestry			
				NSNPS	NPARAM	ZSTAT	P	NSNPS	NPARAM	ZSTAT	P
SFRP2[a,b]	4	154701744	154710272	16	4	4.761	9.65E-07	18	5	4.579	2.34E-06
RNF175[a,b]	4	154631277	154681387	208	12	4.951	3.70E-07	216	13	5.040	2.33E-07
TSSK1B[b]	5	112768251	112770728	-	-	-	-	15	4	4.605	2.06E-06
FBXL17[b]	5	107194736	107717799	-	-	-	-	1390	38	4.703	1.28E-06
FOXP2[a,b]	7	113726382	114333827	889	34	5.439	2.68E-08	1122	49	5.794	3.43E-09
MSRA[a]	8	9911778	10286401	1471	59	4.707	1.25E-06	-	-	-	-
FAM120A[a,b]	9	96214004	96328397	300	15	4.704	1.28E-06	354	19	5.054	2.17E-07
PHF2[b]	9	96338689	96441869	-	-	-	-	491	26	5.270	6.84E-08
EXOC6[a,b]	10	94590935	94819250	525	19	4.775	8.97E-07	652	26	4.663	1.56E-06
NCAM1[a,b]	11	112831997	113149158	848	38	6.835	4.10E-12	999	57	6.195	2.92E-10
NAV3[a,b]	12	78224685	78606790	1251	63	5.152	1.29E-07	1463	78	5.097	1.73E-07
TRMT61A[a]	14	103995521	104003410	18	3	4.488	3.59E-06	-	-	-	-
RCN2[b]	15	77223960	77242601	-	-	-	-	34	12	4.499	3.42E-06
TCF12[b]	15	57210821	57591479	-	-	-	-	1009	18	4.515	3.16E-06
FTO[a,b]	16	53737875	54155853	1303	93	5.675	6.94E-09	1503	116	5.617	9.72E-09
RBFOX1[a]	16	6069095	7763340	9543	385	4.937	3.98E-07	-	-	-	-
AKAP10[b]	17	19807615	19881656	-	-	-	-	123	8	4.674	1.48E-06
DCC[a,b]	18	49866542	51057784	4415	85	4.993	2.97E-07	5169	101	4.926	4.19E-07
APOE[a,b]	19	45409011	45412650	5	2	5.955	1.30E-09	6	3	5.682	6.66E-09
APOC1[a]	19	45417504	45422606	11	3	4.621	1.91E-06	-	-	-	-
FAM46D[b]	X	79591003	79700810	-	-	-	-	58	3	5.055	2.15E-07
BRWD3[b]	X	79926353	80065187	-	-	-	-	180	10	4.540	2.81E-06

N=304,485 for the multi-ancestry cohort; N=249,785 for the European ancestry cohort.
[a]Top gene in the multi-ancestry cohort.
[b]Top gene in the European ancestry cohort.

following Bonferroni correction (Table 3, Supplemental Table 8, Supplemental Fig. 6b). Eight of these genes (*AKAP10, BRWD3, FAM46D, FBXL17, PHF2, RCN2, TCF12,* and *TSSK1B*) were distinct from the multi-ancestry gene-based analysis. Among these, the leading association was with *PHF2*. In the AA and HA cohorts, there were no gene-wide significant genes identified (data not shown).

Gene tissue expression

We used FUMA software [43] to evaluate gene-tissue expression in the multi-ancestry cohort. Results showed that among 30 general tissues, the brain was the only significantly enriched tissue, particularly in the frontal cortex, anterior cingulate cortex, and nucleus accumbens (tissues aggregated: Fig. 2A, Supplemental Table 9; stratified by tissue subtype: Fig. 2B, Supplemental Table 10). Additionally, gene-set analyses were conducted to understand potential gene pathways of importance; however, results did not identify any significant gene sets in the multi-ancestry cohort (Supplemental Table 11). When evaluating Veterans of EA separately, there were no significant gene-tissue expression results and no significant gene sets (Supplemental Tables 12-13; Supplemental Fig. 7a, b).

SNP-based heritability

Linkage disequilibrium score regression (LDSC) [45] was used to estimate SNP-based heritability (SNP-h^2) of the case-control TBI phenotype in the EA cohort. Observed SNP-h^2 was estimated to be 0.060 (SE = 0.004, $p = 7.83 \times 10^{-66}$). On the liability scale, at 2% population prevalence, SNP-h^2 was 0.044 (SE = 0.003); at 20% population prevalence, SNP-h^2 was 0.087 (SE = 0.005). To obtain further insights into the genetic architecture of TBI, we used univariate MiXeR analysis to decompose heritability into discoverability and polygenicity (number of influential variants) subcomponents. An estimated 10,470 (SE = 848) influential variants explained 90% of the SNP-h^2 and the discoverability was 9.2×10^{-6} (SE = 6.7×10^{-7}). The 15 variants identified in the EA GWAS explained an estimated 0.2% of the heritability of TBI.

We also evaluated sex differences in SNP-h^2. There was no significant difference in SNP-h^2 of TBI between men and women (z-test for different means; z-score = 1.49, $p = 0.135$), with SNP-$h^2 = 0.06$ in men (SE = 0.0034, $p = 2.24 \times 10^{-70}$) and SNP-$h^2 = 0.11$ in women (SE = 0.036, $p = 0.0015$). The genetic correlation between men and women was 0.74 (SE = 0.11, $p = 5.6 \times 10^{-11}$). However, given that the female SNP-h^2 z-score was <4 (attributable to the low sample size of the female cohort [53]), the analyses evaluating sex differences in heritability and genetic correlation were not well powered and should be interpreted with caution.

Genetic overlap

We examined genetic correlations (r_g) between the MVP TBI variable and a variety of a priori selected health domains including risk-taking behaviors, psychiatric disorders, neurocognitive variables, and brain morphometrics in the EA cohort (Fig. 3, Supplemental Tables 14–17, Supplemental Figs. 8 and 9). The strongest genetic correlation was with PTSD ($r_g = 0.69$, SE = 0.04, $p = 2.5 \times 10^{-75}$). Broadly, genetic correlations with TBI were moderate-to-large for risk-taking behaviors and psychiatric disorders. TBI was also negatively correlated with verbal numeric reasoning ($r_g = -0.35$, SE = 0.03, $p = 4.2 \times 10^{-42}$). Genetic correlations with reaction time and neurogenerative disorders (i.e., AD and PD) were not significantly different from 0 after adjustment for multiple comparisons. With regard to brain morphometrics, TBI was weakly genetically correlated to ICV ($r_g = -0.12$, SE = 0.04, $p = 0.002$), but all other brain morphometric correlations were not significantly different from 0.

We used univariate and bivariate MiXeR to perform a more detailed evaluation of the genetic overlap between TBI and these phenotypes. Univariate analyses suggest that the genetic architecture of the TBI phenotype (i.e., the polygenicity and discoverability of risk variants) was similar to polygenic psychiatric disorders (Supplemental Tables 18-19). In contrast, neurodegenerative disorders such as AD were characterized by relatively low polygenicity but high discoverability. In bivariate analyses, for five phenotypes investigated (AD, neuroticism, risky behaviors, reaction time (RT), and verbal numeric reasoning), there was evidence of polygenic overlap with TBI beyond what was explained by a genetic correlation model (best vs min AIC > 0). Notably, while the genetic correlations of TBI with AD and RT were not significantly different from 0, MiXeR reported that 60% (number of shared influential variants = 57, SE = 23) of the influential variation in AD and 94% (number of shared influential variants = 8902, SE = 502) of the influential variation in RT were shared with TBI (Supplemental Figs. 10 and 11). MiXeR also reported the existence of genetic variation specific to PD (best vs max AIC = 0.57), where the r_g estimated among only shared influential variants was -0.83 (SE = 0.14).

Genomic structural equation modeling (gSEM)

Given the strong genetic correlations between TBI and risk-taking behaviors and psychiatric disorders, genomic structural equation models [49] were conducted within the EA cohort to better understand whether TBI is more genetically associated with a latent factor representing risk-taking behaviors versus a latent factor representing psychiatric disorders. Pairwise genetic correlations were calculated between all phenotypes to produce a genetic correlation matrix (Supplemental Figs. 8 and 9) for use in factor analysis. In genomic confirmatory factor analysis, the TBI variable loaded onto its own factor while the psychiatric disorders loaded onto one factor and the risk-taking behaviors loaded onto another factor (CFI = 0.908) (Fig. 4). The correlation between the TBI and psychiatric disorders factors was 0.55 (SE = 0.04), and the correlation between the TBI and risk-taking behaviors factors was 0.73 (SE = 0.06), indicating that TBI has a stronger genetic correlation with risk-taking behaviors than with the psychiatric disorders. When examining genetic associations between TBI and individual psychiatric disorders, PTSD, alcohol dependence, and ADHD showed the strongest associations with TBI (Supplemental Fig. 8). Given this observation, a final set of analyses was run to examine the genetic correlation between these individual disorders and TBI, controlling for the genetic association with risk-taking. These multiple regression models are presented in Fig. 5; the model with PTSD continued to show the strongest relationship with TBI even after accounting for risk-taking.

PheWAS

Our PheWAS examined associations of leading variants in TBI risk loci with phenotypes across 28 domains (Supplemental Fig. 12, Supplemental Table 20). The most common associations were with phenotypes in psychiatric (11 loci), activities (9 loci), metabolic (8 loci), and environment (7 loci) domains (Supplemental Fig. 13). Considering the degree of pleiotropy of the loci, the (median) average locus was associated with phenotypes across three different domains. Loci 15 (*FTO*) and 19 (*APOE*) had significant associations with the largest number of domains (16 and 15 domains, respectively). In contrast, six loci were only associated with phenotypes within a single domain (e.g., locus 7 was only associated with a phenotype in the skeletal domain [osteoarthritis]).

DISCUSSION

Examining the role of genetics in TBI risk and recovery may offer novel mechanistic insights and lead to important discoveries that can inform preventive efforts and innovative treatments for patients who have sustained TBIs. In this study, we present the
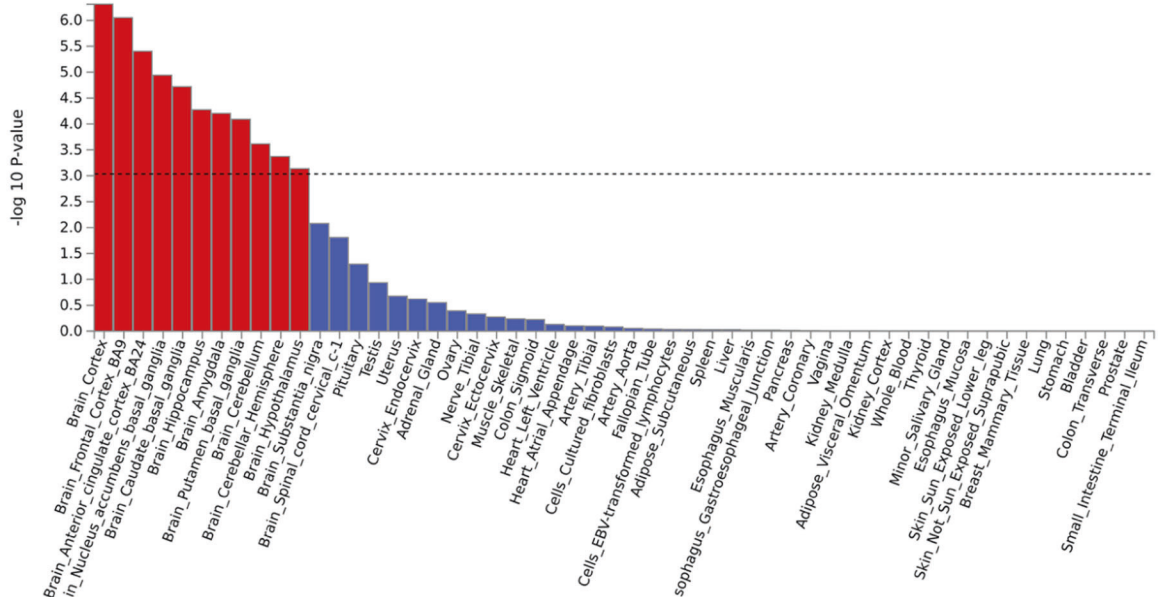
A)
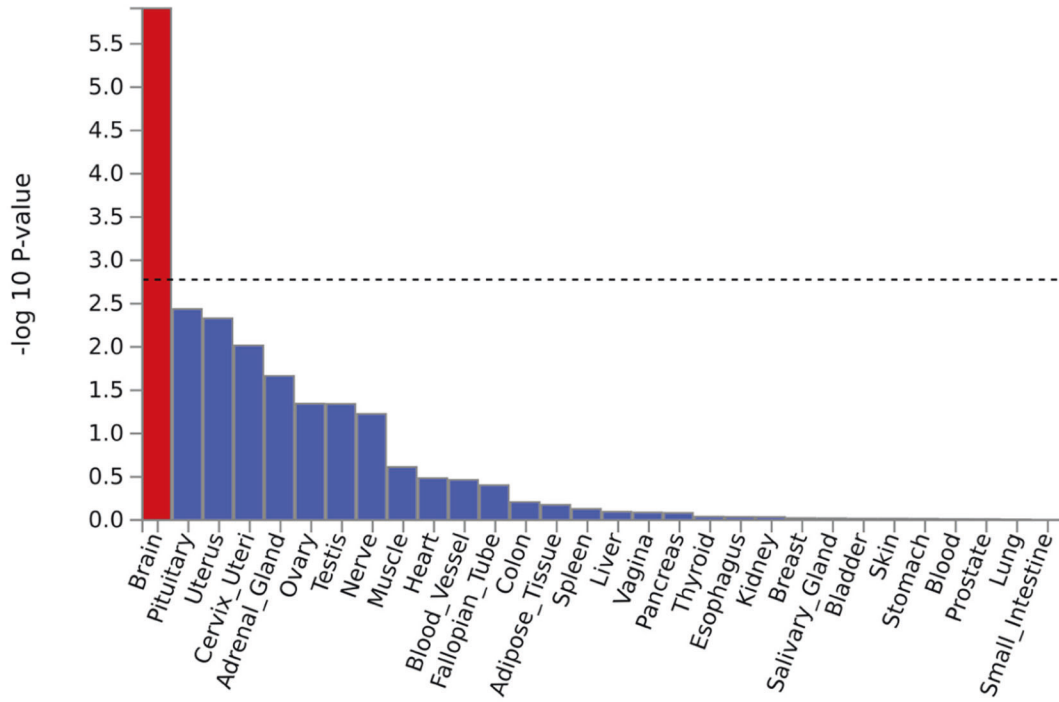
B)

Fig. 2 Gene tissue expression in the multi-ancestry cohort. A Gene tissue expression: tissues aggregated. **B** Gene tissue expression: stratified by tissue subtype.

largest GWAS to date for TBI using data from the VA Million Veteran Program. Our results showed GWS variants associated with a TBI case-control phenotype in a multi-ancestry cohort of Veterans. These associations were primarily driven by the EA cohort, which comprised the majority of the overall sample (82%), and findings replicated in the FinnGen cohort (notably, the direction of effects matched for 13 of the 15 variants). We also

independently evaluated Veterans of African and Hispanic ancestry, but found no independent GWS loci, likely due to the smaller sample size and therefore lower power in these groups to detect any significant hits. When conducting sex-stratified analyses, the males-only GWAS was similar to the overall GWAS, whereas the females-only GWAS found no significant loci, which we again attribute to limited statistical power in the females-only cohort.

In the multi-ancestry cohort, our SNP-based results revealed 15 GWS loci, and our gene-based results revealed 14 gene-wide significant genes, many of which have presumed roles in psychiatric and neurologic functioning. The top GWS SNPs within the multi-ancestry sample were in and around *NCAM1*, followed by *APOE*, *FTO*, and *FOXP2*. Notably, many of these genes have emerged as significant genes in prior studies of neuropsychiatric and neurodegenerative conditions. For example, NCAM1 (neural cell adhesion molecule 1) is a glycoprotein expressed on the surface of neurons, glia, and skeletal muscle, and has been shown to have a role in several brain-related biological processes such as neuronal development and maintenance as well as synaptic plasticity [54–56]. The *NCAM1* gene has been associated with cognitive functioning, particularly learning and memory, and has been implicated in a range of neuropsychiatric and neurodegenerative disorders including schizophrenia, mood and anxiety disorders, and AD [55].

Another notable gene to emerge as significant was *APOE*. APOE is a protein expressed in peripheral tissues and the central nervous system and its primary function is facilitating the transportation and distribution of lipids throughout the body. APOE plays a role in neuronal maintenance, neural transmission, neuroinflammation, and synaptic plasticity within the brain and has been associated with response and repair processes following neurotrauma [57, 58]. The *APOE* ε4 allele is a major risk factor for AD [59–61], and other
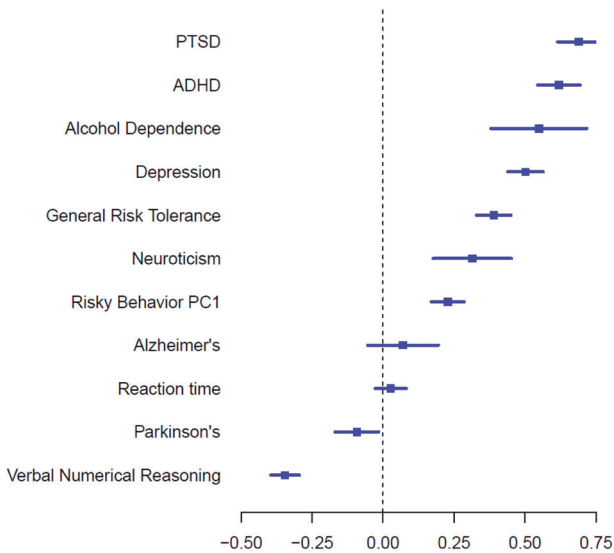

Fig. 3 Genetic correlation between TBI and other phenotypes. The x-axis denotes the magnitude of genetic correlation. The blue squares represent the genetic correlation between TBI and each given phenotype, with lines representing 95% confidence intervals.
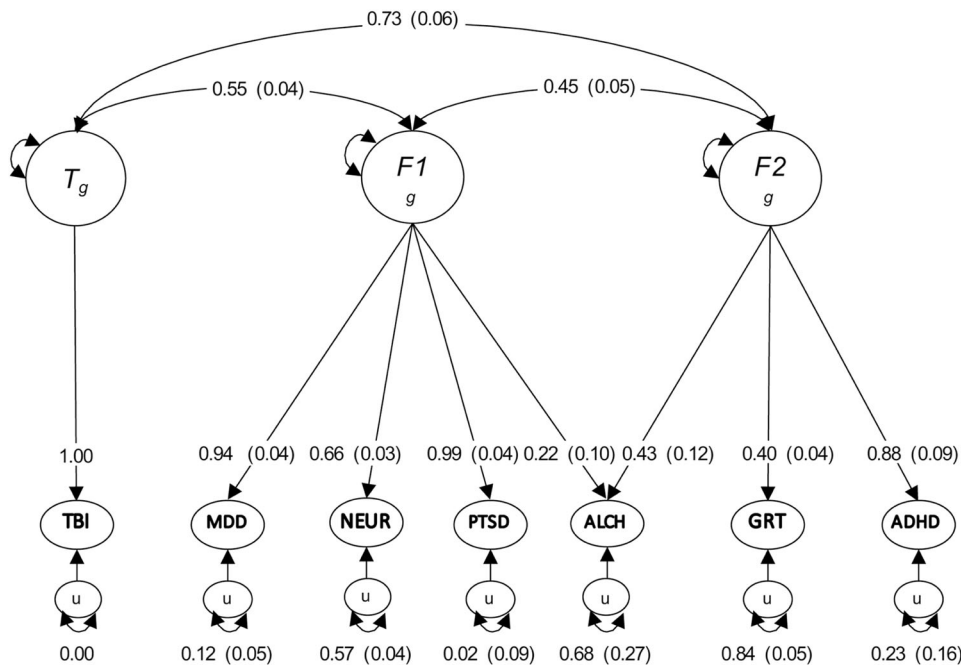

Fig. 4 Genomic structural equation modeling (gSEM) results. Correlated factors model with standardized estimates and standard error estimates in parentheses. Genetic components of disorders are in circles and latent factors are represented by ovals. Single-headed arrows represent regression relationships and parameter estimates that can be squared to estimate the proportion of variance accounted for by components. Double-headed arrows represent correlations. Residual variance represented by double-headed arrow connecting the variable to itself. Each phenotype loads onto at least one factor. The TBI risk variable is loaded onto its own factor (T_g) while the psychiatric disorders loaded onto one factor ($F1_g$) and the risk-taking behaviors loaded onto another factor ($F2_g$); CFI = 0.908. TBI traumatic brain injury; MDD major depressive disorder; NEUR neuroticism; PTSD posttraumatic stress disorder; ALCH alcohol dependence; GRT general risk tolerance; ADHD attention-deficit/hyperactivity disorder.
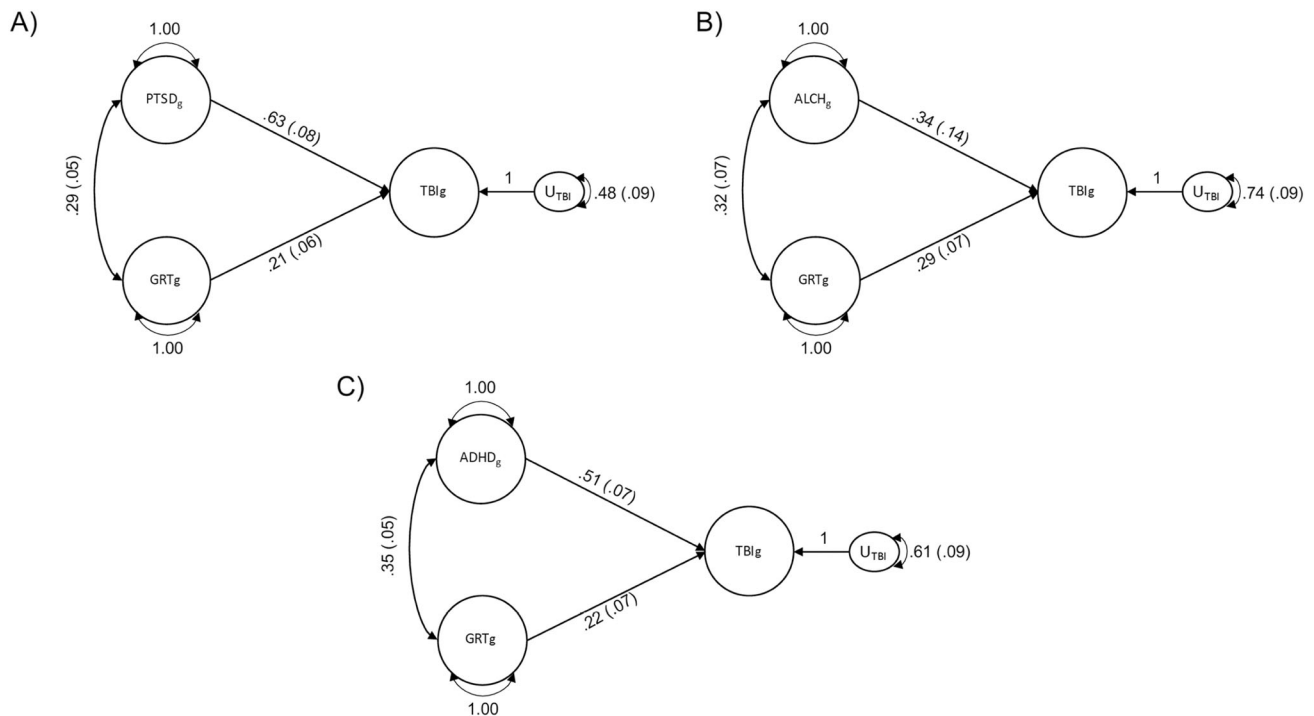
Fig. 5 Multiple regression results. Genetic components of disorders are in circles. Single-headed arrows represent regression relationships and parameter estimates that can be squared to estimate the proportion of variance accounted for. Double-headed arrows represent correlations. Residual variance represented by double-headed arrow connecting the variable to itself. The models reflect the genetic association between TBI and (**A**) PTSD, (**B**) alcohol dependence, and (**C**) ADHD while controlling for general risk taking. PTSD posttraumatic stress disorder; GRT general risk tolerance; TBI traumatic brain injury; ALCH alcohol dependence; ADHD attention-deficit/hyperactivity disorder.

research has shown associations between *APOE* and medical conditions such as vascular dementia, diabetes mellitus, and cardiovascular disease [62, 63], as well as with cognitive functioning [64]. Furthermore, our findings of an association between *APOE* and TBI align with previous candidate gene studies in TBI [18, 19, 65–68] and may help to explain the heterogeneous neurocognitive outcomes observed chronically following TBI. However, it is important to note that our association was observed with a single *APOE* SNP (and not the *APOE* haplotype which defines the *APOE* ε4 allele, which was not examined), and the effect of the haplotype may be different. Follow-up studies by our group are planned to more comprehensively address the role of the *APOE* haplotype in the context of TBI.

As for *FTO* (Fat and Obesity-Associated), less is known about the physiological function of this gene in the brain; however, *FTO* has been found to have important actions in metabolic and cardiovascular systems and may be associated with body mass index, obesity risk, and type 2 diabetes [69, 70]. *FTO* gene variants have also been causally associated with risk of AD [71] and have been reported to promote phosphorylation of tau by activating the mTor pathway [72], which has been implicated in AD pathophysiology. Finally, FOXP2 (forkhead box P2) is a protein that is active in several tissues including the brain, and was initially implicated in speech and language development [73, 74]. *FOXP2* has also been shown to play a role in other biological processes including brain development, cell differentiation and proliferation, and neurodegeneration [73, 74], and has previously been implicated as a genetic variant associated with ADHD, risk-taking behaviors, and psychiatric disorders such as PTSD [74, 75]. Taken together, these observations offer critical insights into the genes that are associated with TBI. Although the functional significance of these genes within the context of TBI will still need to be determined, these initial findings are promising and offer targets for future research.

To better understand the role of TBI-associated genes in higher-order systems, we examined gene tissue expression. As expected, our results showed that the brain was the only significantly enriched region, with TBI-associated genes specifically implicated in the cortex, frontal cortex, anterior cingulate cortex (ACC), nucleus accumbens, caudate, hippocampus, amygdala, putamen, cerebellum, cerebellar hemisphere, and hypothalamus. These findings converge with a recent GWAS examining risk tolerance that similarly showed significant gene expression in brain regions such as the frontal cortex, ACC, and cortex [76]. Of note, the ACC is well-known to have connections with both the limbic system and the prefrontal cortex and has been implicated in decision-making and impulse control [77, 78], underscoring the biological link between TBI and risk-taking.

With regard to SNP-based heritability of the TBI phenotype, we found an observed heritability of 6%, which is generally consistent with other heritability statistics reported for risk-taking behaviors (e.g., general risk tolerance: SNP-$h^2 = 0.05$ [76]), psychopathology (e.g., PTSD [MVP]: SNP-$h^2 = 0.05$-0.07 [75]), and neurocognition (e.g., AD, SNP-$h^2 = 0.07$ [79]). Notably, there are no adequately powered studies reporting SNP-h^2 for TBI for comparison. Future studies are thus needed to further explore the heritability of TBI and to specifically tease apart heritability associated with TBI acquisition versus TBI sequelae.

Finally, we examined genetic correlations and our LDSC results suggest that there is significant genetic overlap between the TBI case-control phenotype and risk-taking behaviors and psychopathology. While these findings are certainly consistent with expectations, this is the first empirical data to establish these important *genetic* relationships. In contrast, there was weak genetic correlation between the TBI case-control phenotype and cognition and brain morphometrics, with the exception of verbal numeric reasoning. gSEM analyses further confirmed underlying genetic associations between TBI and both risk-taking behaviors

and psychiatric disorders, though the association was stronger between TBI and risk-taking. Still, when evaluating individual traits/conditions, PTSD had the strongest genetic overlap with TBI, followed by ADHD and alcohol use when controlling for general risk tolerance, providing evidence of shared genetic etiology. Finally, MiXeR results showed that the genetic architecture of TBI was similar to polygenic psychiatric disorders and risk-taking traits. In contrast, neurodegenerative disorders such as AD showed much less polygenicity but had a high degree of shared variance with TBI. The discrepancies between the architectures of TBI and neurodegenerative disorders suggest that results need to be interpreted with caution; however, they are at least indicative of shared genetics.

This is the first large-scale GWAS of TBI in the VHA and findings significantly advance understanding of the genetic basis of TBI. Nevertheless, there are caveats to consider when interpreting our findings. To maximize sample size and ensure sufficient power, we defined a TBI case as any positive indication of TBI on either the MVP Baseline Survey or MVP Lifestyle Survey, or at least one inpatient or outpatient ICD code for TBI. In other words, our objective was to curate as large of a sample of TBI cases as possible to maximize power to detect significant hits. In theory, this approach allowed us to capture *lifetime* history of TBI. However, the diverse methods of data collection, as well as the lack of precise information on the timing between injury event and reporting or documentation of the event, may have impacted results and we acknowledge this limitation. Of note, the MVP Baseline Survey specifically assessed for lifetime history of concussion/TBI whereas the MVP Lifestyle Survey evaluated deployment-related TBI. Furthermore, the ICD codes were limited to the data available within the Veterans' VA EHRs, which may omit historical TBIs (i.e., TBI sustained prior to military service/ enrollment in VA). Given the data, our TBI phenotype focused on the presence vs. absence of TBI (i.e., TBI history), though it is possible that our phenotype also captured aspects of TBI recovery. Future MVP studies are planned using more fine-grained phenotyping approaches to further understand the genetics associated with TBI-associated *sequelae* (e.g., "persistent post-concussive symptoms" and TBI and mental health comorbidities) and other aspects of TBI *recovery*, as it is likely that different genes may be associated with TBI acquisition/susceptibility vs. TBI outcome. Moreover, given that we did not evaluate or consider TBI severity in our analyses, future research may also consider evaluating TBI or TBI-associated variables on a continuum—for example, examining injury severity or injury characteristics such as LOC or PTA.

Other study caveats relate to sample characteristics. Notably, the MVP cohort is predominately male (~ 90%); given our findings and the strong associations observed with risk-taking behaviors and psychiatric traits, it will be important for future research to further examine whether genetic associations differ as a function of biological sex. While we did examine sex-stratified analyses in this study, it is difficult to draw any conclusions regarding possible sex differences given the low statistical power we had when examining the female-only cohort. Nevertheless, prior research has shown that risk for TBI and recovery following TBI differs by sex [80–83]; thus, it is certainly possible that there are important genetic differences between males and females that would emerge when examining a larger cohort of females. Future research with better female representation is needed to answer this question. Finally, the MVP cohort may be unique in other ways (e.g., high frequency of medical comorbidities in this population) and possibly distinct from other major biobanks; thus, it will be necessary for other biobanks to also examine the genetics associated with TBI and to determine whether there are unique genetic risk factors for Veterans with a history of TBI compared to civilian samples with TBI history.

CONCLUSIONS

Learning more about the genetics associated with TBI risk and recovery is necessary for developing a better understanding of the pathophysiology of this injury and could aid in developing novel therapeutics. The results of this first large-scale GWAS examining TBI in military Veterans identified 15 loci in the multi-ancestry cohort, including genes previously known to be relevant to TBI biology that will be important to investigate in future studies. Findings also showed that TBI is a heritable trait with comparable genetic architecture and high genetic correlation with psychiatric and risk-taking traits. Results set the stage for future TBI GWAS studies within MVP that focus on diversity and chronicity of symptom sequelae as well as severity of injury.

DATA AVAILABILITY

REFERENCES

1. Helmick KM, Spells CA, Malik SZ, Davies CA, Marion DW, Hinds SR. Traumatic brain injury in the US military: Epidemiology and key clinical and research programs. Brain Imaging Behav. 2015;9:358–66.
2. DoD Numbers for Traumatic Brain Injury Worldwide - Totals (Defense Health Agency) (2021).
3. Karr JE, Areshenkoff CN, Garcia-Barrera MA. The neuropsychological outcomes of concussion: a systematic review of meta-analyses on the cognitive sequelae of mild traumatic brain injury. Neuropsychology 2014;28:321–36.
4. McCrea M, Guskiewicz K, Doncevic S, Helmick K, Kennedy J, Boyd C, et al. Day of injury cognitive performance on the Military Acute Concussion Evaluation (MACE) by US military service members in OEF/OIF. Mil Med. 2014;179:990–7.
5. Nelson LD, Temkin NR, Dikmen S, Barber J, Giacino JT, Yuh E, et al. Recovery after mild traumatic brain injury in patients presenting to US level I trauma centers: A transforming research and clinical knowledge in traumatic brain injury (TRACK-TBI) study. JAMA Neurol. 2019;76:1049–59.
6. Chapman JC, Diaz-Arrastia R. Military traumatic brain injury: A review. Alzheimer's Dement. 2014;10:S97–104.
7. Schwab K, Terrio HP, Brenner LA, Pazdan RM, McMillan HP, MacDonald M, et al. Epidemiology and prognosis of mild traumatic brain injury in returning soldiers: A cohort study. Neurology 2017;88:1571–9.
8. Perry DC, Sturm VE, Peterson MJ, Pieper CF, Bullock T, Boeve BF, et al. Traumatic brain injury is associated with subsequent neurologic and psychiatric disease: A meta-analysis. J Neurosurg. 2016;124:511–26.
9. Stein MB, Ursano RJ, Campbell-Sills L, Colpe LJ, Fullerton CS, Heeringa SG, et al. Prognostic indicators of persistent post-concussive symptoms after deployment-related mild traumatic brain injury: A prospective longitudinal study in US Army soldiers. J Neurotrauma. 2016;33:2125–32.
10. Stein MB, Jain S, Giacino JT, Levin H, Dikmen S, Nelson LD, et al. Risk of post-traumatic stress disorder and major depression in civilian patients after mild traumatic brain injury: A TRACK-TBI study. JAMA Psychiatry. 2019;76:249–58.
11. Barnes DE, Byers AL, Gardner RC, Seal KH, Boscardin WJ, Yaffe K. Association of mild traumatic brain injury with and without loss of consciousness with dementia in US military veterans. JAMA Neurol. 2018;75:1055–61.
12. Brickell TA, Lange RT, French LM. Health-related quality of life within the first 5 years following military-related concurrent mild traumatic brain injury and polytrauma. Mil Med. 2014;179:827–38.
13. Sakamoto MS, Delano-Wood L, Schiehser DM, Merritt VC Predicting veteran health-related quality of life following mild traumatic brain injury. *Rehabilitation Psychology*. 2021; https://doi.org/10.1037/rep0000392.
14. Agtarap SD, Campbell-Sills L, Jain S, Sun X, Dikmen S, Levin H, et al. Satisfaction with life after mild traumatic brain injury: A TRACK-TBI study. J Neurotrauma. 2021;38:546–54.
15. Pietrzak RH, Johnson DC, Goldstein MB, Malley JC, Southwick SM. Posttraumatic stress disorder mediates the relationship between mild traumatic brain injury and health and psychosocial functioning in veterans of Operations Enduring Freedom and Iraqi Freedom. J Nerv Ment Dis. 2009;197:748–53.

16. MacDonald CL, Barber J, Jordan M, Johnson AM, Dikmen S, Fann JR, et al. Early clinical predictors of 5-year outcome after concussive blast traumatic brain injury. JAMA Neurol. 2017;74:821–9.

17. Taylor BC, Hagel EM, Carlson KF, Cifu DX, Cutting A, Bidelspach DE, et al. Prevalence and costs of co-occurring traumatic brain injury with and without psychiatric disturbance and pain among Afghanistan and Iraq War Veteran VA users. Med Care. 2012;50:342–6.

18. McAllister TW. Genetic factors modulating outcome after neurotrauma. PMR 2010;2:S241–52.

19. Jordan BD. Genetic influences on outcome following traumatic brain injury. Neurochem Res. 2007;32:905–15.

20. Weaver SM, Chau A, Portelli JN, Grafman J. Genetic polymorphisms influence recovery from traumatic brain injury. Neuroscientist. 2012;18:631–44.

21. Dardiotis E, Fountas KN, Dardioti M, Xiromerisiou G, Kapsalaki E, Tasiou A, et al. Genetic association studies in patients with traumatic brain injury. Neurosurg Focus. 2010;28:E9.

22. Zeiler FA, McFadyen C, Newcombe VF, Synnot A, Donoghue EL, Ripatti S, et al. Genetic influences on patient-oriented outcomes in traumatic brain injury: A living systematic review of non-apolipoprotein E single-nucleotide polymorphisms. J Neurotrauma. 2021;38:1107–23.

23. Panenka WJ, Gardner AJ, Dretsch MN, Crynen GC, Crawford FC, Iverson GL. Systematic review of genetic risk factors for sustaining a mild traumatic brain injury. J Neurotrauma. 2017;34:2093–9.

24. Lawrence DW, Comper P, Hutchison MG, Sharma B. The role of apolipoprotein E episilon (ε)-4 allele on outcome following traumatic brain injury: A systematic review. Brain Inj. 2015;29:1018–31.

25. Sun X-C, Jiang Y. Genetic susceptibility to traumatic brain injury and apolipoprotein E gene. Chin J Traumatol. 2008;11:247–52.

26. Zhou W, Xu D, Peng X, Zhang Q, Jia J, Crutcher KA. Meta-analysis of APOE 4 allele and outcome after traumatic brain injury. J Neurotrauma. 2008;25:279–90.

27. Zeng S, Jiang J-X, Xu M-H, Xu L-S, Shen G-J, Zhang A-Q, et al. Prognostic value of apolipoprotein E epsilon4 allele in patients with traumatic brain injury: A meta-analysis and meta-regression. Genet Test Mol Biomark. 2014;18:202–10.

28. McFadyen CA, Zeiler FA, Newcombe V, Synnot A, Steyerberg E, Gruen RL, et al. Apolipoprotein E4 polymorphism and outcomes from traumatic brain injury: A living systematic review and meta-analysis. J Neurotrauma. 2021;38:1124–36.

29. Duncan LE, Ostacher M, Ballon J. How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. Neuropsychopharmacology 2019;44:1518–23.

30. Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM. Research review: Polygenic methods and their application to psychiatric traits. J Child Psychol Psychiatry. 2014;55:1068–87.

31. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, et al. Million Veteran Program: A mega-biobank to study genetic influences on health and disease. J Clin Epidemiol. 2016;70:214–23.

32. Schwab KA, Ivins B, Cramer G, Johnson W, Sluss-Tiller M, Kiley K, et al. Screening for traumatic brain injury in troops returning from deployment in Afghanistan and Iraq: initial investigation of the usefulness of a short screening tool for traumatic brain injury. J Head Trauma Rehabilitation. 2007;22:377–89.

33. U.S. Department of Defense. Traumatic Brain Injury (TBI): DoD Standard Surveillance Case Definition for TBI Adapted for AFHSD Use. U.S. Department of Defense.

34. Hunter-Zinck H, Shi Y, Li M, Gorman BR, Ji S-G, Sun N, et al. Genotyping array design and data quality control in the Million Veteran Program. Am J Hum Genet. 2020;106:535–48.

35. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526:68–74.

36. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48:1443–8.

37. Fang H, Hui Q, Lynch J, Honerlaw J, Assimes TL, Huang J, et al. Harmonizing genetic ancestry and self-identified race/ethnicity in genome-wide association studies. Am J Hum Genet. 2019;105:763–72.

38. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26:2867–73.

39. Abraham G, Qiu Y, Inouye M. FlashPCA2: Principal component analysis of Biobank-scale genotype datasets. Bioinformatics 2017;33:2776–8.

40. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JL. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015;4:7.

41. Willer CJ, Li Y, Abecasis GR. METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–1.

42. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: Regional visualization of genome-wide association scan results. Bioinformatics 2010;26:2336–7.

43. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.

44. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. Am J Hum Genet. 2011;88:294–305.

45. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizofrenia Working Group of the Psychiatric Genomics Consortium. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47:291–5.

46. Holland D, Frei O, Desikan R, Fan C-C, Shadrin AA, Smeland OB, et al. Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. PLoS Genet. 2020;16:e1008612.

47. Frei O, Holland D, Smeland OB, Shadrin AA, Fan C-C, Maeland S, et al. Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. Nat Commun. 2019;10:2417.

48. R Core Team. R: A language and environment for statistical computing. https://www.R-project.org/.

49. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav. 2019;3:513–25.

50. Watanabe Kea. A global view of genetic architecture in human complex traits. [under preparation].

51. Seripa D, D'Onofrio G, Panza F, Cascavilla L, Masullo C, Pilotto A. The genetics of the human APOE polymorphism. Rejuvenation Res. 2011;14:491–500.

52. Broman KW, Matsumoto N, Giglio S, Martin CL, Roseberry JA, Zuffardi O, et al. Common long human inversion polymorphism on chromosome 8p. Lecture Notes-Monograph Series. 2003;40:237–45.

53. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. LD Hub: A centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics 2017;33:272–9.

54. Schmid RS, Maness PF Cell recognition molecules and disorders of neurodevelopment. International Handbook on Brain and Behaviour in Human Development. Kluwer Academic Publishers; 2001:199-218.

55. Brennaman LH, Maness PF NCAM in neuropsychiatric and neurodegenerative disorders. vol 663. Structure and Function of the Neural Cell Adhesion Molecule NCAM: Advances in Experimental Medicine and Biology Springer; 2010:299-317.

56. Vukojevic V, Mastrandreas P, Arnold A, Peter F, Kolassa I-T, Wilker S, et al. Evolutionary conserved role of neural cell adhesion molecule-1 in memory. Transl Psychiatry. 2020;10:217.

57. Mahley RW, Huang Y. Apolipoprotein e sets the stage: Response to injury triggers neuropathology. Neuron 2012;76:871–85.

58. Mahley RW, Weisgraber KH, Huang Y. Apolipoprotein E4: A causative factor and therapeutic target in neuropathology, including Alzheimer's disease. Proc Natl Acad Sci. 2006;103:5644–51.

59. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: A meta-analysis. JAMA 1997;278:1349–56.

60. Horsburgh K, McCarron MO, White F, Nicoll JA. The role of apolipoprotein E in Alzheimer's disease, acute brain injury and cerebrovascular disease: Evidence of common mechanisms and utility of animal models. Neurobiol Aging. 2000;21:245–55.

61. Verghese PB, Castellano JM, Holtzman DM. Apolipoprotein E in Alzheimer's disease and other neurological disorders. Lancet Neurol. 2011;10:241–52.

62. Reinvang I, Espeseth T, Westlye LT. APOE-related biomarker profiles in non-pathological aging and early phases of Alzheimer's disease. Neurosci Biobehav Rev. 2013;37:1322–35.

63. Van Giau V, Bagyinszky E, An SSA, Kim SY. Role of apolipoprotein E in neurodegenerative diseases. Neuropsychiatr Dis Treat. 2015;11:1723–37.

64. Wisdom NM, Callahan JL, Hawkins KA. The effects of apolipoprotein E on nonimpaired cognitive functioning: A meta-analysis. Neurobiol Aging. 2011;32:63–74.

65. McAllister TW, Flashman LA, Harker Rhodes C, Tyler AL, Moore JH, Saykin AJ, et al. Single nucleotide polymorphisms in ANKK1 and the dopamine D2 receptor gene affect cognitive outcome shortly after traumatic brain injury: A replication and extension study. Brain Inj. 2008;22:705–14.

66. Merritt VC, Clark AL, Sorg SF, Evangelista ND, Werhane M, Bondi MW, et al. Apolipoprotein E ε4 genotype is associated with elevated psychiatric distress in veterans with a history of mild to moderate traumatic brain injury. J Neurotrauma. 2018;35:2272–82.

67. Merritt VC, Clark AL, Sorg SF, Evangelista ND, Werhane ML, Bondi MW, et al. Apolipoprotein E (APOE) ε4 genotype is associated with reduced neuropsychological performance in military veterans with a history of mild traumatic brain injury. J Clin Exp Neuropsychol. 2018;40:1050–61.

68. Merritt VC, Lange RT, Lippa SM, Brickell TA, Soltis AR, Dalgard CL, et al. Apolipoprotein e (APOE) ε4 genotype influences memory performance following remote traumatic brain injury in US military service members and veterans. Brain cognition. 2021;154:105790.

69. Zhao X, Yang Y, Sun B-F, Zhao Y-L, Yang Y-G. FTO and obesity: Mechanisms of association. Curr Diabetes Rep. 2014;14:1–9.

70. McCarthy MI. Genomics, type 2 diabetes, and obesity. N. Engl J Med. 2010;363:2339–50.

71. Reitz C, Tosto G, Mayeux R, Luchsinger JA, Group N-LNFS, Initiative tAsDN. Genetic variants in the Fat and Obesity Associated (FTO) gene and risk of Alzheimer's disease. PloS One. 2012;7:e50354.

72. Li H, Ren Y, Mao K, Hua F, Yang Y, Wei N, et al. FTO is involved in Alzheimer's disease by targeting TSC1-mTOR-Tau signaling. Biochemical Biophysical Res Commun. 2018;498:234–9.

73. Co M, Anderson AG, Konopka G. FOXP transcription factors in vertebrate brain development, function, and disorders. Wiley Interdiscip Rev: Developmental Biol. 2020;9:e375.

74. Den Hoed J, Devaraju K, Fisher SE. Molecular networks of the FOXP2 transcription factor in the brain. EMBO Rep. 2021;22:e52803.

75. 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.

76. Linnér RK, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet. 2019;51:245–57.

77. Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci. 2000;4:215–22.

78. Stevens FL, Hurley RA, Taber KH. Anterior cingulate cortex: unique role in cognition and emotion. J Neuropsychiatry Clin Neurosci. 2011;23:121–5.

79. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. Nat Genet. 2019;51:414–30.

80. Gupte R, Brooks W, Vukas R, Pierce J, Harris J. Sex differences in traumatic brain injury: what we know and what we should know. J Neurotrauma. 2019;36:3063–91.

81. Kim LH, Quon JL, Sun FW, Wortman KM, Adamson MM, Harris OA. Traumatic brain injury among female veterans: A review of sex differences in military neurosurgery. Neurosurgical Focus. 2018;45:E16.

82. Merritt VC, Padgett CR, Jak AJ. A systematic review of sex differences in concussion outcome: What do we know? Clin Neuropsychologist. 2019;33:1016–43.

83. Mollayeva T, Mollayeva S, Colantonio A. Traumatic brain injury: sex, gender and intersecting vulnerabilities. Nat Rev Neurol. 2018;14:711–22.

AUTHOR CONTRIBUTIONS
Overall study coordination: VCM, CMN. Study concept and design: VCM, CMN, AXM, LDW. Data curation: CCC. Phenotype analysis: CCC (lead), VCM, LDW. Statistical analysis: AXM (lead), MG. Data interpretation: AXM, MG, CMN, MBS, MSP, RLH, MWL, VCM. Writing: VCM (lead), AXM, CMN. All authors provided edits, feedback, and approved the final version of the paper for submission.

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/s41380-023-02304-8.

Correspondence and requests for materials should be addressed to Victoria C. Merritt.

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.