

Reproducible Risk Loci and Psychiatric Comorbidities in Anxiety:

Results from ~200,000 Million Veteran Program Participants

Daniel F. Levey^{1,2*}, Joel Gelernter^{1,2,*,*x}, Renato Polimanti^{1,2}, Hang Zhou^{1,2}, Zhongshan Cheng^{1,2}, Mihaela Aslan^{3,4}, Rachel Quaden⁵, John Concato^{3,4}, Krishnan Radhakrishnan^{3,6}, Julien Bryois^{7,8}, Patrick F. Sullivan^{7,8}, the Million Veteran Program, Murray B. Stein^{9,10*,x}

Author names, affiliations, contact

¹Division of Human Genetics, Department of Psychiatry, Yale University School of Medicine,

New Haven, Connecticut

²Department of Psychiatry, Veterans Affairs Connecticut Healthcare Center, West Haven,

Connecticut

³ VA Clinical Epidemiology Research Center (CERC), VA Connecticut Healthcare System, West Haven, CT

⁴ Department of Medicine, Yale University School of Medicine, New Haven, CT

⁵Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), VA Boston

Healthcare System, Boston, MA

⁶ College of Medicine, University of Kentucky, Lexington, KY

⁷Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-17177 Stockholm,

Sweden

2

⁸Departments of Genetics, University of North Carolina, Chapel Hill, NC, 27599-7264, USA

⁹Psychiatry Service, VA San Diego Healthcare System, San Diego, CA

¹⁰Departments of Psychiatry and Family Medicine & Public Health, University of California San Diego, La Jolla, CA

Correspondence:

Joel Gelernter, Yale Univ. School of Medicine, VA 116A2, VA CT Healthcare Center, West Haven, CT 06525; OR

Murray B Stein, University of California San Diego, 9500 Gilman Drive, Mailcode 0855, La Jolla, CA 92093-0855

*Equal contribution

^xCorrespondence

Significance

Anxiety disorders are common and often disabling. They are also frequently co-morbid with other mental disorders such as major depressive disorder (MDD); these disorders may share commonalities in their underlying genetic architecture. Using one of the largest homogenously phenotyped cohorts available, the Million Veteran Program sample, we investigated common variants associated with anxiety in genome-wide association studies (GWASes), using survey results from the GAD-2 anxiety scale (as a continuous trait, $n=199,611$), and self-reported anxiety disorder diagnosis (as a binary trait, $n=224,330$). This largest GWAS to date for anxiety and related traits identified numerous novel significant associations, several of which are replicated in other datasets, and allows inference of underlying biology.

Abstract

We used GWAS in the Million Veteran Program sample (nearly 200,000 informative individuals) using a continuous trait for anxiety (GAD-2) to identify 5 genome-wide significant (GWS) signals for European Americans (EA) and 1 for African Americans. The strongest findings were on chromosome 3 ($rs4603973$, $p=7.40 \times 10^{-11}$) near the *SATB1* locus, a global regulator of gene expression and on chromosome 6 ($rs6557168$, $p=1.04 \times 10^{-9}$) near *ESR1* which encodes estrogen receptor α . A locus identified on chromosome 7 near *MAD1L1* ($p=1.62 \times 10^{-8}$) has been previously identified in GWAS of bipolar disorder and of schizophrenia and may represent a risk factor for psychiatric disorders broadly. SNP-based heritability was estimated to be ~6% for GAD-2. We also GWASed for self-reported anxiety disorder diagnoses ($N=224,330$) and identified two GWS loci, one ($rs35546597$, $MAF=0.42$, $p=1.88 \times 10^{-8}$) near the *AURKB* locus, and the other ($rs10534613$, $MAF=0.41$, $p=4.92 \times 10^{-8}$) near the *IQCHE* and *MAD1L1* locus identified in the GAD-2 analysis. We demonstrate reproducibility by replicating our top findings in

the summary statistics from the Anxiety NeuroGenetics Study (ANGST) and a UK Biobank neuroticism GWAS. We also replicated top findings from a large UK Biobank preprint, demonstrating stability of GWAS findings in complex traits once sufficient power is attained. Finally, we found evidence of significant genetic overlap between anxiety and major depression using polygenic risk scores, but also found that the main anxiety signals are independent of those for MDD. This work presents novel insights into the neurobiological risk underpinning anxiety and related psychiatric disorders.

Introduction

Anxiety disorders are common, affecting 1 in 10 Americans each year¹, and are a leading cause of disability worldwide.² A recent analysis of anxiety and depressive disorders showed that cumulatively they cost about \$90 billion in personal health spending for 2013.³ Given their prevalence, associated impairment, and economic costs, anxiety disorders are a major public health concern.⁴

Anxiety is “a future-oriented mood state associated with preparation for possible, upcoming negative events”,⁵ and is usually a normal and adaptive behavioral response to everyday life. In anxiety *disorders*, anxiety is excessive or out of proportion to the actual or anticipated event, and accompanied by clinically significant distress or disability.⁶ Numerous risk factors for anxiety disorders have been studied, including experiential and genetic factors.⁷ For example, neurotic personality traits are predictive of the onset of anxiety disorders.⁸ Twin studies demonstrate a heritable component to anxiety disorders,⁷ but there have been few published GWAS to date investigating anxiety or anxiety related traits. The Anxiety Neuro Genetics Study (ANGST)⁹ meta-analysis has been the most noteworthy published study to date, finding one genome wide significant (GWS) locus each for a categorical case control design for any anxiety disorder diagnosis and a quantitative factor score for anxiety in a cohort of over 18,000 subjects. Understanding of the genetics of anxiety disorders has thus lagged behind other related disorders such as major depression.¹⁰

Only a third of individuals with anxiety disorders receive treatment.¹¹ For those who do enter treatment, psychological approaches such as cognitive behavioral therapy have been shown to be very effective^{12,13}; as have certain pharmacotherapies.^{14,15} Better understanding of genetic risk factors and determinants now inform other aspects of medicine such as oncology and cardiology through identification of causal mutations¹⁶ and variants, and will have important implications for psychiatry.¹⁷ These precision medicine approaches are challenging in complex traits such as anxiety because they are associated with many (perhaps hundreds of thousands) of variants of individually small effect¹⁸. The use of polygenic risk scores will require a suitably large sample size to provide sufficient confidence in these small individual effects that cumulatively account for so much of the heritability. Underlying polygenic risk factors from sufficiently large cohorts may inform an approach to identify those with a predisposition to develop anxiety disorders and improve outcomes.

The Million Veteran Program (MVP) is one of the world's largest databases of genetic, environmental, and medical information, based on data from United States military veterans¹⁹. Using this large genetic dataset and (1) the Generalized Anxiety Disorder 2-item (GAD-2) scale²⁰ as well as (2) self-report of physician diagnosis of anxiety or panic disorder, we discovered novel GWS variants associated with anxiety in European Americans (EA) and African Americans (AA). We examined replication and genetic overlap of these results with those of previous studies examining anxiety and traits with which anxiety disorders are commonly comorbid – major depression, PTSD, and neuroticism. We also examined expression quantitative trait loci (eQTLs) to identify possible gene expression implications of these genetic variants, with eQTL evidence for altered expression in the basal ganglia and cerebellum. These findings, in the largest cohort of individuals GWASed for anxiety and anxiety disorders (199,611 subjects for the quantitative trait, 224,330 for binary diagnosis), indicate shared genetic risk with some other mental disorders, but also point to loci that may be especially important for anxiety and anxiety-related traits.

Methods

Participants. The MVP cohort has been described previously.¹⁹ Results were analyzed in two separate tranches based on when genotyping results were conducted. Ancestry was assigned using 10 principal components and the 1000 Genomes project phase 3 EUR and AFR as reference within each tranche of data.

Genotyping, Imputation, and Quality Control. Genotyping, imputation, and quality control within MVP has been previously described. Briefly, samples were genotyped using a 723,305-SNP Affymetrix Axiom biobank array, customized for MVP.¹⁹ Imputation was performed with minimac3 using data from the 1000 Genomes project. For post-imputation QC, SNPs with imputation INFO scores of < 0.3 or minor allele frequencies (MAF) below 0.01 were removed from analysis. For the first tranche of data, 22,183 SNPs were selected through linkage disequilibrium (LD) pruning using PLINK,²¹ and then Eigensoft²² was used to conduct principal component analysis on 343,286 MVP samples and 2,504 1000 Genomes samples.²³ The reference population groups (EUR, EAS, AFR, AMR, or SAS) in the 1000 Genomes samples were used to define EUR (n=241,541) and AFR (n=61,796) groups used in this analysis. Similar methods were used in the second tranche of data, which contained 108,416 new MVP samples and the same 2,504 1000 Genomes samples. In Tranche 2, 80,694 participants were defined as EUR and 20,584 were defined as AFR.

Phenotypic Assessment. The GAD-2 scale consists of 2 questions (Table S1) included in a self-report survey, each scored on a 0-3 scale. Participants were asked to respond based on their symptoms during the past 2 weeks. Values for the two responses are summed, resulting in a range of scores between 0-6, which we treated as a continuous trait (Table 1).

Another anxiety phenotype was analyzed based on data collected from the MVP baseline survey. Participants were asked the question: "Please tell us if you have been diagnosed with the

following conditions: Anxiety reaction/Panic disorder.” Answers were recorded as yes/no binary responses and missing responses were excluded from analysis. 224,330 participants (34,189 cases) responded to the diagnosis question and, had available genotype, and had either a European or African ancestry assigned.

Statistical Analysis. GWAS analysis was carried out by linear regression for each ancestry group and tranche using PLINK 2.0 on dosage data, covarying for age, sex, and the first 10 PCs. Meta-analysis was performed using METAL (EA n=175,163, AA n=24,448). Similar analyses were conducted for the self-reported medical history of anxiety diagnosis except that logistic regression was performed.

LD Score Regression (LDSC) and SNP-based Heritability. We used LD score regression (LDSC) through LD Hub²⁴ to estimate SNP-based heritability, and to assess genetic correlation of GAD-2 anxiety with 35 traits selected using the category keywords: psychiatric, personality, cognitive, reproductive, sleeping, smoking behavior, aging, and neurological. The traits from the ANGST GWAS of anxiety case-control and factor scores⁹ were calculated separately in LDSC using summary statistics downloaded from the Psychiatric Genomics Consortium website.

Conditional Analysis for Major Depression. Considering the extensive comorbidity between major depression and anxiety disorders⁷ – we ran a conditional analysis with the mtCOJO method²⁵ using GCTA software with the MVP GAD-2 summary statistics as the primary analysis and the PGC MDD2 (excluding 23andMe)¹⁰ summary statistics to condition the analysis for depression.

Gene-based Tests. Summary statistics from the GWAS were loaded into Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS) to test for gene-level associations using Multi-Marker Analysis of GenoMic Annotation (MAGMA).^{26,27} Input SNPs were mapped to 17,927 protein coding genes. The GWS threshold for the gene-based test was therefore determined to be $p = 0.05/17,927 = 2.79 \times 10^{-6}$.

Results

Primary analysis. GWAS was conducted separately in two tranches of each ancestry in the MVP sample, defined by the time data became available, and meta-analyzed together within population. One genomic locus was GWS in the AA meta-analysis (Figure 1a) while five loci were GWS in the EA meta-analysis (Figure 1b).

The GWS result from the AA analysis was near the *TRPV6* (Transient Receptor Potential Cation Channel Subfamily V Member 6) locus. The top signal in the EA meta-analysis consisted of 64 GWS SNPs in high LD at the *SATB1-AS1* (Special AT-Rich Sequence Binding 1 Antisense RNA 1) locus. The strongest finding (rs4603973, MAF=0.29, $p=7.40 \times 10^{-11}$) was intronic at *SATB1-AS1*. The second strongest independent signal was on chromosome 6 (rs6557168, MAF=0.63, $p=1.04 \times 10^{-9}$) intronic at *ESR1* (Estrogen Receptor 1) with 10 other GWS SNPs in high LD. A third GWS association for EAs was found on chromosome 1 (rs12023347, MAF=0.48, $p=9.73 \times 10^{-9}$) near the long noncoding RNA *LINC01360* and *LRR1Q3*. The fourth GWS association found in EAs was on chromosome 7 (rs56226325, MAF=0.17, $p=1.62 \times 10^{-8}$) in an intron of *MAD1L1* (Mitotic Arrest Deficient 1 Like 1). The fifth association for EAs was in and around the *TCEA2* (Transcription Elongation Factor A2), *RGS19* (Regulator Of G Protein Signaling 19), and *OPRL1* (Opioid Related Nociceptin Receptor 1) genes (rs6090040, MAF=0.48, $p=2.59 \times 10^{-8}$).

We conducted additional analysis in related phenotypes within the same samples using case-control status for self-reported diagnosis of anxiety or panic disorder. For the EA subjects, there were two GWS signals for the self-reported Anxiety/Panic case control GWAS, in a gene-rich region nearest *AURKB* on chromosome 17 (rs35546597, MAF=0.42, $p=1.88 \times 10^{-8}$) and chromosome 7 in an *IQCHE* intron (rs10534613, MAF=0.41, $p=4.92 \times 10^{-8}$) close to the *MAD1L1* locus identified for GAD-2. There were no GWS findings for this phenotype in African Americans.

Replication. For replication we looked up our top 5 SNPs in 3 independent GWAS with anxiety-related phenotypes. We looked up our lead GWS SNPs in GWAS for anxiety case-control,⁹ and neuroticism in the recent large UK Biobank, 23andMe, and Genetics of Personality Consortium²⁸ (Table 2). (The lead SNP on chromosome 3 near the *SATB1* locus, rs4603973, was not available for lookup in the UK Biobank neuroticism GWAS, so we used the strongest LD-proxy available (rs4390955 $R^2=0.91$, $p=8.68E-11$), as a surrogate.) In the UK Biobank neuroticism study, 4 of 5 independent SNPs looked up had the same direction of effect, 3 were significant, and one near *MAD1L1* was nearly GWS (rs56226325, $p=6.59\times 10^{-8}$). A preprint reported results for anxiety from UK Biobank using Case-Control and the GAD-7, scored as a dichotomous trait.²⁹ We find significant replication for 2 of their 4 findings, with suggestive evidence for a third (Table 3).

Genome wide gene based association study (GWGAS) for GAD-2. In the GWGAS, the top gene identified was *OPRL1* ($p=1.15\times 10^{-9}$), which was also significant in the SNPwise analysis as noted above. 31 genes were identified as GWS following Bonferroni correction for multiple comparisons. A more permissive Benjamini-Hochberg correction with a still relatively restrictive 0.05 FDR identified 189 genes (Supplementary Table S3) in total for investigation of biological relevance through the Ingenuity pathway enrichment tool. Ingenuity Pathway Analysis (IPA, Ingenuity Systems, www.ingenuity.com, Redwood City, CA, USA) was used to examine underlying biological process for the 189 genes corrected by Benjamini-Hochberg (Supplementary Table S4).

eQTL. To identify causal implications for genetic variants, eQTLs were assessed for top GWS signals using GTEx v7 brain tissue expression data. Top GWS signals on chromosome 7 and 20 had significant eQTL ($FDR \leq 0.05$) for 4 different genes: *FTSJ2*, *RGS19*, *C20orf201*, and *OPRL1* (Supplementary table S7). The top signals are centered in the basal ganglia and cerebellum.

LD score regression (LDSC) analysis. Eighteen of the 35 traits were significantly correlated with GAD-2 following correction for multiple comparisons ($\alpha = 0.05/35 = 0.001429$) (Figure 2 and Supplementary Table S5). The most significantly correlated traits were depressive symptoms ($r_g = 0.8055$, $p = 1.95 \times 10^{-53}$) and neuroticism ($r_g = 0.7174$, $p = 6.53 \times 10^{-53}$).

We also investigated genetic correlation within the MVP cohort using the GWAS for self-reported case-control anxiety. Genetic correlations between GAD-2 and self-reported diagnosis were high, with an r_g of 0.856. The phenotype correlation between GAD-2 and self-reported diagnosis was $r = 0.64$, $p < 2.2 \times 10^{-16}$.

SNP heritability. SNP-based heritability using LDSC for the GAD-2 quantitative trait was estimated to be 0.0558 (SE = 0.004). SNP based inflation was mild considering the sample size and polygenic trait studied ($\lambda = 1.19$); the intercept (1.026) and attenuation ratio (0.1177) estimated by LDSC showed negligible evidence for inflation due to population stratification. SNP based heritability for the phenotype of case-control anxiety was 0.0374 (SE = 0.004).

Polygenic risk score (PRS) analysis. was used to test polygenic risk scores. Summary statistics from the MVP GAD-2 analysis were used as the base data for calculating the PRS (using PRSice v 1.25³⁰). Genetic overlap between anxiety and PTSD or MDD were tested in the Psychiatric Genomics Consortium (PGC) MDD¹⁰ or PTSD³¹ GWAS, respectively, and overlap with case-control anxiety disorder was tested in the largest previously available study.⁹ Significant overlap was identified: the MVP GAD-2 PRS can explain up to 0.24% of the variance in MDD in the PGC GWAS ($p = 2.05 \times 10^{-94}$), 0.23% of the variance of PTSD in the PGC GWAS ($p = 4.23 \times 10^{-12}$), and 0.48% of the variance in ANGST Anxiety Disorder ($p = 3.66 \times 10^{-20}$) (Figure 3).

mtCOJO. mtCOJO GWAS was used to condition the GAD-2 MVP summary statistics for anxiety on the PGC MDD2 summary statistics for MDD¹⁰. There were no new signals and significance of lead findings

dropped by about an order of magnitude, but results on chromosomes 3 (*SATB1*) and 6 (*ESR1*) remained GWS.

Discussion

We present the largest GWAS to date for anxiety traits, employing a quantitative phenotype, the GAD-2, in nearly 200,000 MVP subjects, as well as self-reported anxiety/panic case-control phenotypes in >220,000 MVP subjects. We identified novel genetic variants in and around genes some of which have previously-known functional relationships with anxiety. These genes play roles in the HPA axis, neuronal development, and global regulation of gene expression

There is high comorbidity between anxiety and PTSD.³² We used a PRS derived from the MVP GAD-2 analysis to identify genetic overlap with the independent PGC PTSD GWAS (Figure 3). We found a high degree of genetic overlap between these two traits, providing biological evidence that this known clinical comorbidity is due at least in part to shared genetic etiology. We also used LDSC to identify correlation between traits relevant to brain and behavior. We identified substantial positive correlations with depression and neuroticism as well as a negative correlation with subjective well-being (Figure 2).

The GWS result in AAs is an insertion variant which is rare outside of African ancestry and occurs in a genomic region proposed to be under recent selection in Europeans.³³ The lead SNP is at *TRPV6*, which encodes a Ca²⁺ selective³⁴ membrane cation channel associated with epithelial calcium transport and homeostasis in kidney and intestine. The lead SNP rs575403075 has an MAF range of between 0% and 1% in non-African populations and would often fall below MAF QC thresholds used for common variants in most non-African populations. This highlights the importance of studying genetics in diverse populations, otherwise these signals may be missed entirely.

The top GWS findings for EAs in the GAD-2 analysis were in and around *SATB1* and the antisense gene *SATB1-AS1*. *SATB1* is a global regulator that influences expression of multiple genes³⁵ involved in neuronal development. One gene modulated in expression is Corticotropin Releasing Hormone (*CRH*), encoding the protein product of the same name which plays an essential role in the HPA axis, which has frequently been shown to modulate stress and fear/anxiety response.³⁶ The *CRHR1* (Corticotropin Releasing Hormone Receptor 1) gene was GWS in the gene-based association analysis ($p=3.60\times 10^{-7}$). *CRHR1* is bound by endogenous CRH and *CRHR1* has been a proposed target for treatment, with evidence for anxiolytic-like effects of *CRHR1* antagonists in animal models^{36,37} although not yet in humans.^{38,39} One possible explanation for a lack of anxiolytic effect observed in humans is heterogeneity in the patient population. The present results suggest that this may be a reasonable pathway to address via personalized medicine, as individuals with differing genetic risk that does or does not involve this pathway may differ in their responses to glucocorticoid-targeted therapeutic agents.

The estrogen receptor *ESR1* (also known as estrogen receptor α) has been a focus in animal models of anxiety-like behaviors and these have provided mechanistic validity for the role of *ESR1*. Studies of estradiol administration to ovariectomized rats and *ESR1* null mice have shown consistent evidence that *ESR1* is involved in anxiety-like behavior in non-human animals.⁴⁰ Our finding of an association between *ESR1* and anxiety may have implications for our understanding of sex differences in anxiety disorders, which generally show about a 2:1 female predominance, and greater susceptibility to develop PTSD following traumatic events.⁴¹ Although this increased susceptibility is partially explained by sex-specific exposure to certain kinds of traumatic events, there may also be differential biological context provided in part by the role of the estrogen receptor. Our study in a predominantly male sample identifies *ESR1* as GWS. Estrogen is important in both sexes, with a recent review highlighting the important role for estrogens in men.⁴² Future studies with larger sample sizes of women should continue to investigate sex differences in genetic risk for anxiety-related traits.

The lead SNP from the GAD-2 GWAS near the *LINC01360* and *LRR1Q3* (rs2180945) loci is nominally significant and has the same direction of effect in the PGC MDD 2018 analysis ($p=1.434\times 10^{-6}$). Previous genetic epidemiology studies have shown that common genetic factors can underlie anxiety and depressive traits.⁴³ This variant may be linked to a common risk factor for both disorders.

One GWS signal for GAD-2 was in a gene-rich region on chromosome 20 near *TCEA2*, *C20orf201*, *RGS19*, and *OPRL1*. eQTL data suggest that variants in this region play roles in the regulation of the expression of *RGS19* and *OPRL1* in the cerebellum and in the basal ganglia (Supplementary Table S7). Increased *OPRL1* expression has been found in the amygdala of PTSD-like severe stress exposed mouse model following a cued-fear expression paradigm.⁴⁴

MAD1L1 (GAD-2 lead SNP rs56226325, MAF=0.17, $p=4.68\times 10^{-10}$, case-control lead SNP rs10534613, MAF=0.41, $p=4.92\times 10^{-8}$) has been associated previously with bipolar disorder⁴⁵ and one of the lead SNPs in that study in a *MAD1L1* intron is also nominally associated with anxiety in the current study (rs11764590, $p=2.82\times 10^{-7}$). This locus has also been identified among 108 GWS loci by the PGC schizophrenia study,⁴⁶ and our lead SNP is nominally significant in that study (rs56226325, $p=1.12\times 10^{-3}$). This SNP is also nominally significant (6.71×10^{-4}) in the 2018 PGC depression GWAS¹⁰, so this locus may be a common risk factor for several psychiatric disorders.

*MAD1L1**rs56226325 is also an eQTL with the gene *FTSJ2* (Supplementary Table S7), which exhibits altered expression in the brain. This gene is a mitochondrial RNA methyltransferase important for the proper assembly of the mitochondrial ribosome and cellular respiration.⁴⁷ The protein product of *FTSJ2* is “Mitochondrial rRNA Methyltransferase 2” (*MRM2*), which was implicated in a case-study of a 7 year old Italian boy with a damaging mutation leading to an encephalopathy, lactic acidosis and stoke-like (MELAS) syndrome.⁴⁸ Larger effect mutations at this locus can have devastating effects on the brain;

smaller effect variations may be less deleterious but still cumulatively influence development which may predispose to neurological and psychiatric disorders.

The Brainstorm Consortium has investigated shared heritability between psychiatric and neurological disorders.⁴⁹ Consistent with their findings, we find very strong genetic correlation between anxiety and psychiatric traits such as depression ($r_g = 0.81$, $p=1.95 \times 10^{-53}$) and neuroticism ($r_g = 0.72$, $p=6.53 \times 10^{-53}$) but relatively weaker genetic overlap with neurological disorders. Significant overlap is detected for amyotrophic lateral sclerosis ($r_g = 0.39$, $p=3.00 \times 10^{-4}$) but Parkinson's disease ($r_g = -0.19$, $p=4.80 \times 10^{-3}$) doesn't survive correction for multiple comparisons and no overlap is seen for Alzheimer's disease ($r_g = 0.00$, $p=1.00$). The relatively strong genetic overlap across psychiatric disorders is a common feature in psychiatric genetics and could point to a need for increased diagnostic specificity, perhaps with a focus on more dimensional rather than diagnostic phenotypes. This may also point to opportunities to improve treatment where there is evidence for greater biological and genetic overlap between disorders. The blurred boundaries between some psychiatric disorders in the clinic are increasingly supported by genetic findings.^{10 31 49}

Limitations to this work include the fact that phenotypes were based on self-reported survey data. The GAD-2 asks questions which temporally reference the "past 2 weeks," that is, it assesses state rather than trait anxiety. This measure should still capture a large proportion of those who experience anxiety on a daily or regular basis, but it falls short of the desired trait (lifetime) anxiety measure. Similarly, the question about diagnosis for anxiety or panic relies on self-report. Also, MVP has mostly male participants. While women are included in this analysis, clinically important interactions between sex, phenotype, and genotype could not be addressed. This cohort is growing, and future recruitment will provide additional power to revisit sex stratified analyses of this sample.

In summary, we have identified novel variants for anxiety by performing GWAS in the large MVP cohort. We found replication of many of these findings in independent cohorts assessed with similar or related phenotypes. We replicated results in our GWAS for top findings from other relevant trait-related GWAS. We also identified significant genetic overlap with MDD, PTSD, and Neuroticism using polygenic risk scores and LD score regression. This work provides additional genetic evidence for the overlap between disorders which frequently present co-morbid with anxiety and presents new molecular targets for investigation with a longer view toward the development of new treatments.

Acknowledgements

The authors would like to thank the veterans who participate in the Million Veterans Program, without whom this work could not have been possible. This study was supported by funding from the Veteran Affairs Office of Research and Development Million Veteran Program Grant#MVP011 and CSP575B.

Competing Interests

Dr. Stein reports receiving consulting fees in the past 3 years from Aptinyx, Bionomics, Janssen, Neurocrine, Pfizer, and Oxeia Biopharmaceuticals. All other authors declare that they have no conflict of interest.

References

1. Comer JS, Blanco C, Hasin DS, Liu SM, Grant BF, Turner JB *et al.* Health-related quality of life across the anxiety disorders: results from the national epidemiologic survey on alcohol and related conditions (NESARC). *J Clin Psychiatry* 2011; **72**(1): 43-50.
2. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; **392**(10159): 1789-1858.
3. Roehrig C. Mental Disorders Top The List Of The Most Costly Conditions In The United States: \$201 Billion. *Health Affair* 2016; **35**(6): 1130-1135.
4. Stein MB, Craske MG. Treating Anxiety in 2017: Optimizing Care to Improve Outcomes. *JAMA* 2017; **318**(3): 235-236.
5. Craske MG, Rauch SL, Ursano R, Prenoveau J, Pine DS, Zinbarg RE. What is an anxiety disorder? *Depress Anxiety* 2009; **26**(12): 1066-1085.
6. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th edn: Washington, DC, 2013.
7. Craske MG, Stein MB, Eley TC, Milad MR, Holmes A, Rapee RM *et al.* Anxiety disorders. *Nat Rev Dis Primers* 2017; **3**: 17024.
8. Zinbarg RE, Mineka S, Bobova L, Craske MG, Vrshek-Schallhorn S, Griffith JW *et al.* Testing a Hierarchical Model of Neuroticism and Its Cognitive Facets: Latent Structure and Prospective Prediction of First Onsets of Anxiety and Unipolar Mood Disorders During 3 Years in Late Adolescence. *Clinical Psychological Science* 2016; **4**(5): 805-824.
9. Otowa T, Hek K, Lee M, Byrne EM, Mirza SS, Nivard MG *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry* 2016; **21**(10): 1485.
10. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 2018; **50**(5): 668-681.
11. Baxter AJ, Vos T, Scott KM, Ferrari AJ, Whiteford HA. The global burden of anxiety disorders in 2010. *Psychological Medicine* 2014; **44**(11): 2363-2374.

17

12. Gelernter CS, Uhde TW, Cimbolich P, Arnkoff DB, Vittone BJ, Tancer ME *et al.* Cognitive-behavioral and pharmacological treatments of social phobia. A controlled study. *Archives of general psychiatry* 1991; **48**(10): 938-945.
13. Cuijpers P, Sijbrandij M, Koole S, Huibers M, Berking M, Andersson G. Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clin Psychol Rev* 2014; **34**(2): 130-140.
14. Bandelow B, Reitt M, Rover C, Michaelis S, Gorlich Y, Wedekind D. Efficacy of treatments for anxiety disorders: a meta-analysis. *Int Clin Psychopharm* 2015; **30**(4): 183-192.
15. Zareifopoulos N, Dylja I. Efficacy and tolerability of vilazodone for the acute treatment of generalized anxiety disorder: A meta-analysis. *Asian J Psychiatr* 2017; **26**: 115-122.
16. Lee JK, Wang J, Sa JK, Ladewig E, Lee HO, Lee IH *et al.* Spatiotemporal genomic architecture informs precision oncology in glioblastoma. *Nat Genet* 2017; **49**(4): 594-599.
17. Stein MB, Smoller JW. Precision Psychiatry-Will Genomic Medicine Lead the Way? *JAMA Psychiatry* 2018; **75**(7): 663-664.
18. Boyle EA, Li Yi, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 2017; **169**(7): 1177-1186.
19. 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**: 214-223.
20. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006; **166**(10): 1092-1097.
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**(3): 559-575.
22. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**(8): 904-909.
23. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM *et al.* A global reference for human genetic variation. *Nature* 2015; **526**(7571): 68-74.

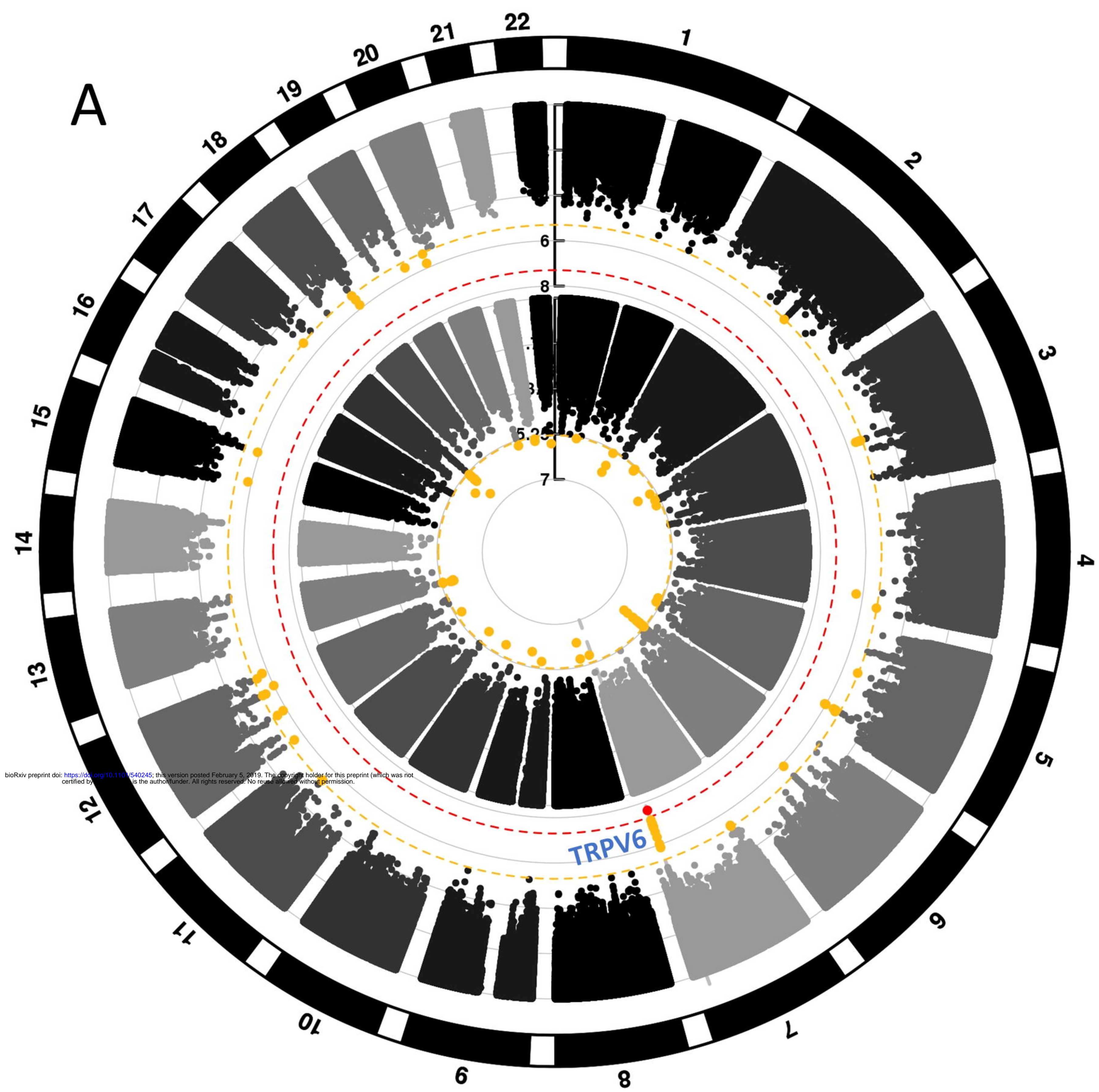
18

24. 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**(2): 272-279.
25. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun* 2018; **9**(1): 224.
26. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017; **8**(1): 1826.
27. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; **11**(4): e1004219.
28. Nagel M, Jansen PR, Stringer S, Watanabe K, de Leeuw CA, Bryois J *et al.* Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* 2018; **50**(7): 920-927.
29. Purves KL, Coleman JR, Rayner C, Hetttema JM, Deckert J, McIntosh AM *et al.* The Common Genetic Architecture of Anxiety Disorders. *bioRxiv* 2017: 203844.
30. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics* 2015; **31**(9): 1466-1468.
31. Duncan LE, Ratanatharathorn A, Aiello AE, Almli LM, Amstadter AB, Ashley-Koch AE *et al.* Largest GWAS of PTSD (N=20 070) yields genetic overlap with schizophrenia and sex differences in heritability. *Mol Psychiatry* 2018; **23**(3): 666-673.
32. Ginzburg K, Ein-Dor T, Solomon Z. Comorbidity of posttraumatic stress disorder, anxiety and depression: a 20-year longitudinal study of war veterans. *J Affect Disord* 2010; **123**(1-3): 249-257.
33. Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA *et al.* Population history and natural selection shape patterns of genetic variation in 132 genes. *Plos Biol* 2004; **2**(10): 1591-1599.
34. Sobolevsky A. Structure and Function of the Epithelial Calcium Channel TRPV6. *Biophys J* 2018; **114**(3): 10a-10a.

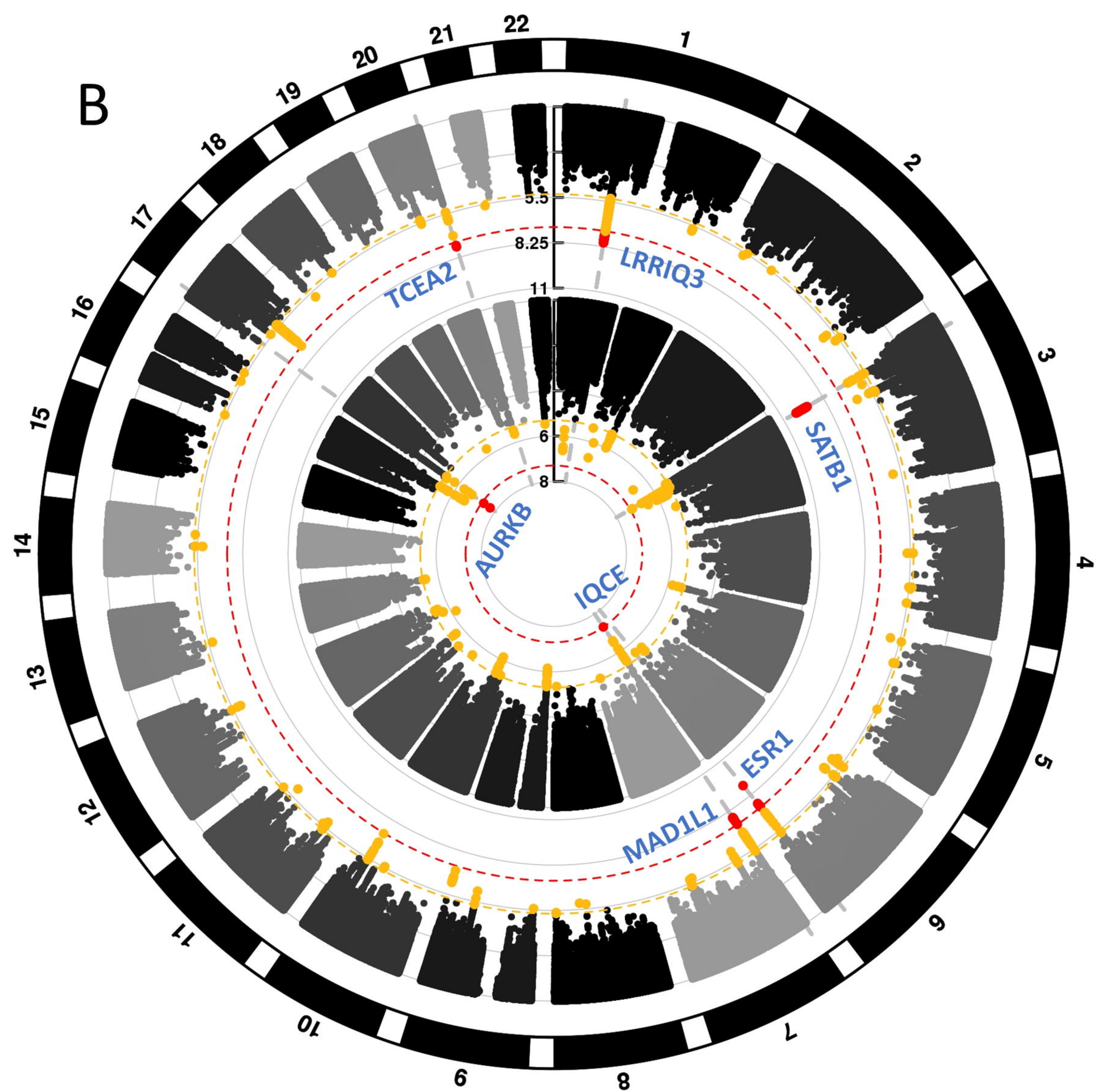
19

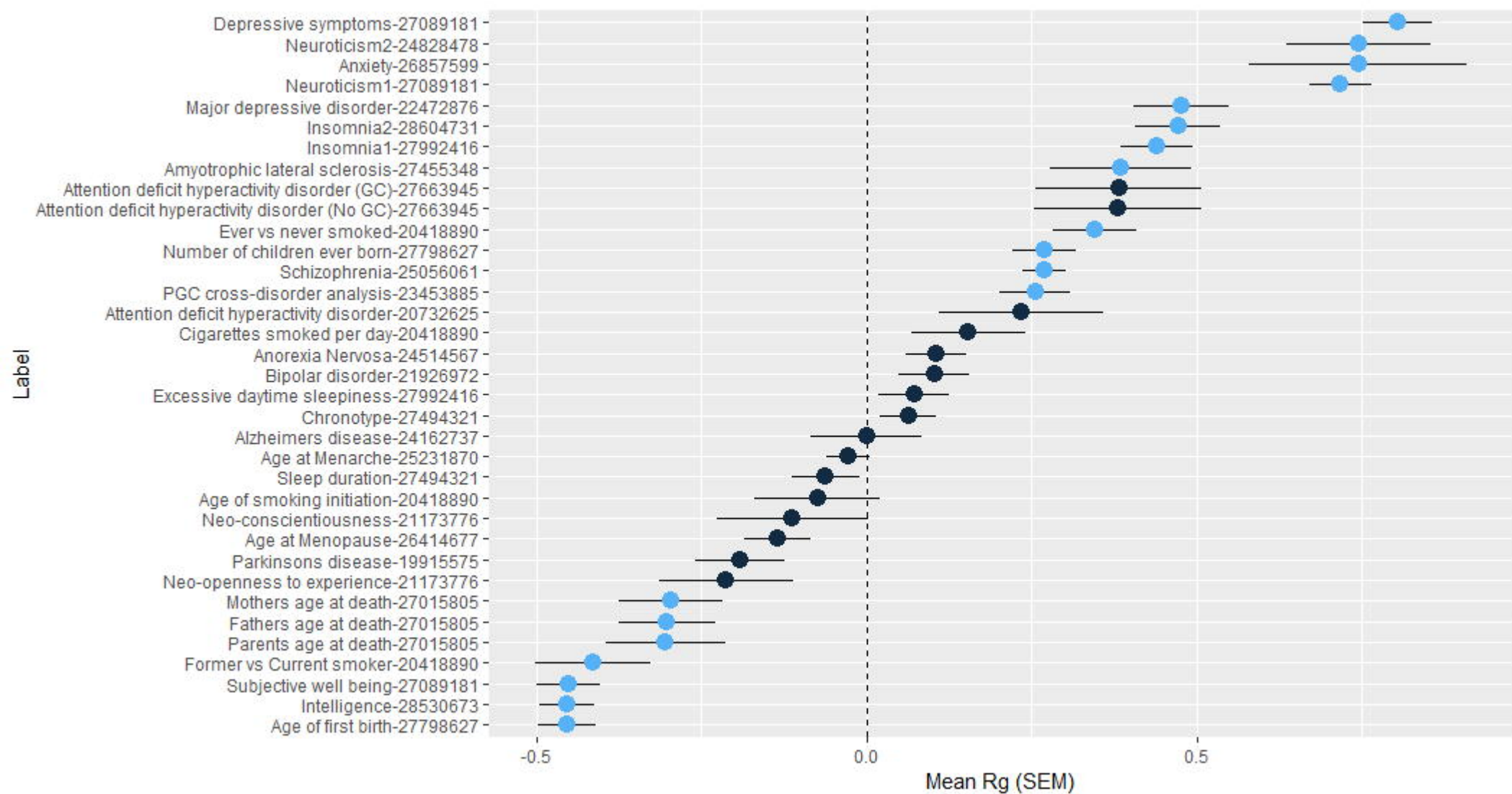
35. Balamotis MA, Tamberg N, Woo YJ, Li J, Davy B, Kohwi-Shigematsu T *et al.* Satb1 ablation alters temporal expression of immediate early genes and reduces dendritic spine density during postnatal brain development. *Mol Cell Biol* 2012; **32**(2): 333-347.
36. Zorrilla EP, Valdez GR, Nozulak J, Koob GF, Markou A. Effects of antalarmin, a CRF type 1 receptor antagonist, on anxiety-like behavior and motor activation in the rat. *Brain Res* 2002; **952**(2): 188-199.
37. Keck ME, Welt T, Wigger A, Renner U, Engelmann M, Holsboer F *et al.* The anxiolytic effect of the CRH1 receptor antagonist R121919 depends on innate emotionality in rats. *Eur J Neurosci* 2001; **13**(2): 373-380.
38. Coric V, Feldman HH, Oren DA, Shekhar A, Pultz J, Dockens RC *et al.* Multicenter, randomized, double-blind, active comparator and placebo-controlled trial of a corticotropin-releasing factor receptor-1 antagonist in generalized anxiety disorder. *Depress Anxiety* 2010; **27**(5): 417-425.
39. Kunzel HE, Ising M, Zobel AW, Nickel T, Ackl N, Sonntag A *et al.* Treatment with a CRH-1-receptor antagonist (R121919) does not affect weight or plasma leptin concentration in patients with major depression. *J Psychiat Res* 2005; **39**(2): 173-177.
40. Borrow AP, Handa RJ. Estrogen Receptors Modulation of Anxiety-Like Behavior. *Vitam Horm* 2017; **103**: 27-52.
41. Haskell SG, Gordon KS, Mattocks K, Duggal M, Erdos J, Justice A *et al.* Gender Differences in Rates of Depression, PTSD, Pain, Obesity, and Military Sexual Trauma Among Connecticut War Veterans of Iraq and Afghanistan. *J Womens Health* 2010; **19**(2): 267-271.
42. Cooke PS, Nanjappa MK, Ko C, Prins GS, Hess RA. Estrogens in Male Physiology. *Physiol Rev* 2017; **97**(3): 995-1043.
43. Mosing MA, Gordon SD, Medland SE, Statham DJ, Nelson EC, Heath AC *et al.* Genetic and Environmental Influences on the Co-Morbidity between Depression, Panic Disorder, Agoraphobia, and Social Phobia: A Twin Study. *Depression and Anxiety* 2009; **26**(11): 1004-1011.
44. Andero R, Brothers SP, Jovanovic T, Chen YT, Salah-Uddin H, Cameron M *et al.* Amygdala-dependent fear is regulated by Oprl1 in mice and humans with PTSD. *Sci Transl Med* 2013; **5**(188): 188ra173.
45. Cichon S, Muhleisen TW, Degenhardt FA, Mattheisen M, Miro X, Strohmaier J *et al.* Genome-wide Association Study Identifies Genetic Variation in Neurocan as a Susceptibility Factor for Bipolar Disorder. *Am J Hum Genet* 2011; **88**(3): 372-381.

46. Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**(7510): 421-+.
47. Rorbach J, Boesch P, Gammage PA, Nicholls TJ, Pearce SF, Patel D *et al.* MRM2 and MRM3 are involved in biogenesis of the large subunit of the mitochondrial ribosome. *Mol Biol Cell* 2014; **25**(17): 2542-2555.
48. Garone C, D'Souza AR, Dallabona C, Lodi T, Rebelo-Guiomar P, Rorbach J *et al.* Defective mitochondrial rRNA methyltransferase MRM2 causes MELAS-like clinical syndrome. *Hum Mol Genet* 2017; **26**(21): 4257-4266.
49. Brainstorm C, Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J *et al.* Analysis of shared heritability in common disorders of the brain. *Science* 2018; **360**(6395).

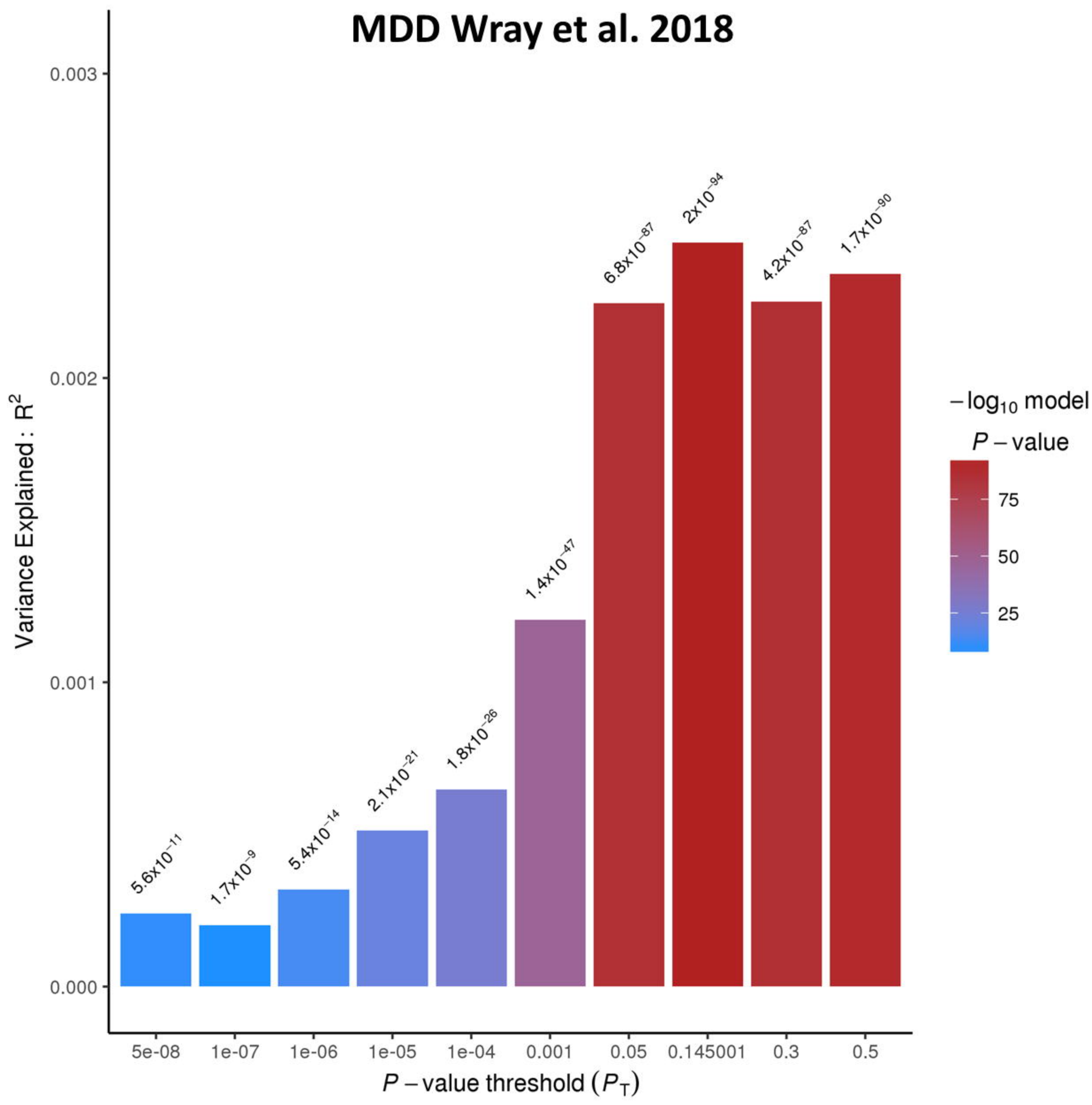


bioRxiv preprint doi: <https://doi.org/10.1101/20240245>; this version posted February 5, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

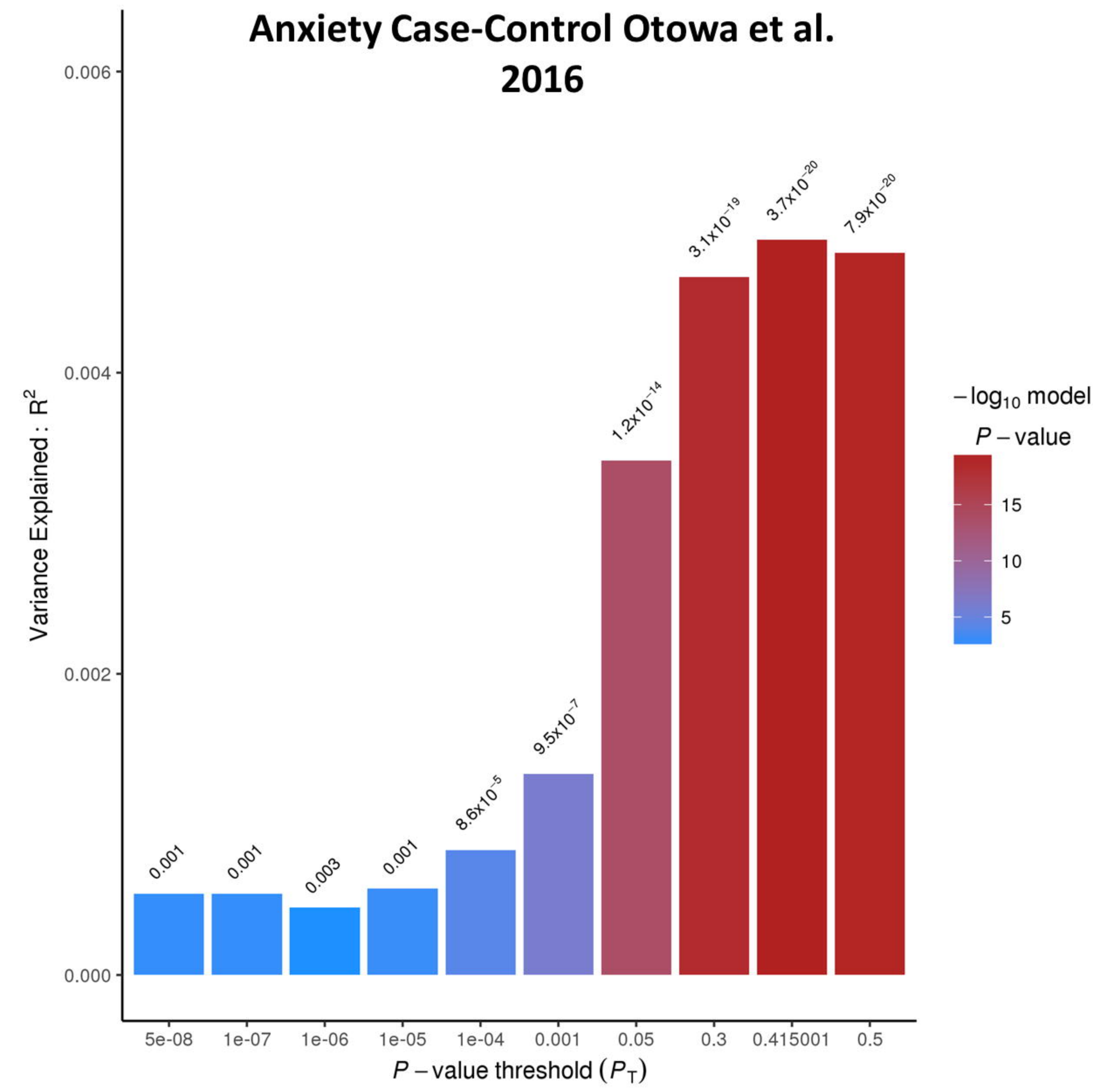




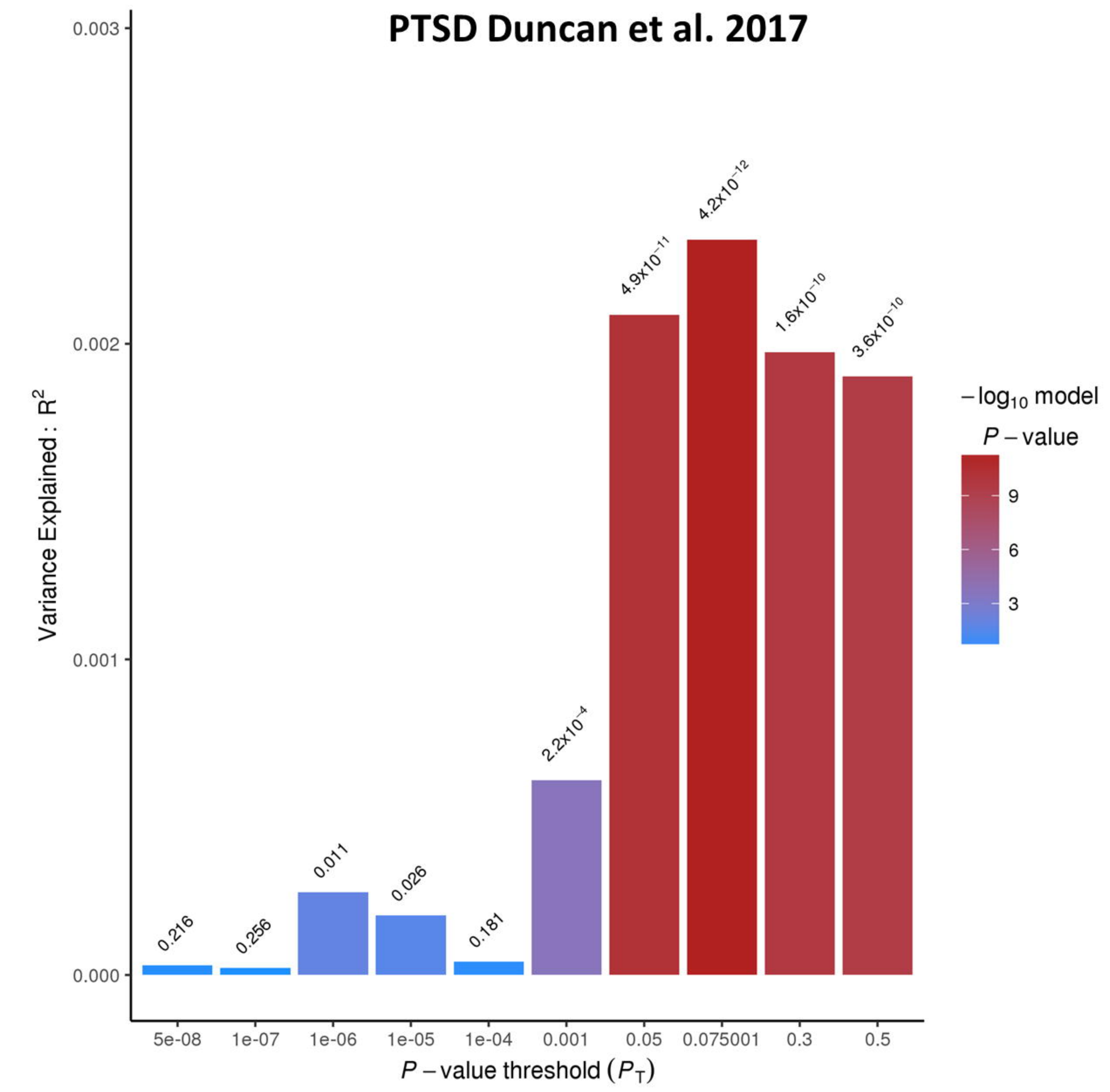
MDD Wray et al. 2018



Anxiety Case-Control Otowa et al. 2016



PTSD Duncan et al. 2017



thresh	nsnps	r2	pval
5.00E-08	4	0.000241	5.55E-11
1.00E-07	6	0.000202	1.75E-09
1.00E-06	12	0.000319	5.39E-14
1.00E-05	57	0.000514	2.11E-21
1.00E-04	371	0.000647	1.81E-26
0.001	2209	0.001205	1.41E-47
0.05	42817	0.002246	6.83E-87
0.145001	89152	0.002445	2.05E-94
0.3	140386	0.002252	4.15E-87
0.5	186003	0.002341	1.73E-90

thresh	nsnps	r2	pval
5.00E-08	4	0.000541	0.001213
1.00E-07	7	0.000541	0.001212
1.00E-06	12	0.000448	0.002894
1.00E-05	45	0.000573	0.000896
1.00E-04	257	0.000831	8.55E-05
0.001	1410	0.001334	9.48E-07
0.05	23025	0.003417	1.19E-14
0.3	69997	0.004637	3.08E-19
0.415001	83033	0.004884	3.66E-20
0.5	90913	0.004795	7.89E-20

thresh	nsnps	r2	pval
5.00E-08	4	3.08E-05	0.216368
1.00E-07	6	2.15E-05	0.256134
1.00E-06	13	0.000262	0.011035
1.00E-05	60	0.000189	0.025942
1.00E-04	384	4.16E-05	0.180868
0.001	2321	0.000618	0.00022
0.05	44293	0.002091	4.89E-11
0.075001	58854	0.00233	4.23E-12
0.3	144415	0.001974	1.63E-10
0.5	190873	0.001897	3.58E-10