

Gene Discovery and Biological Insights into Anxiety Disorders from a Multi-Ancestry Genome-wide Association Study of >1.2 Million Participants

Eleni Friligkou^{1,2}, Solveig Løkhammer^{1,3,4}, Brenda Cabrera-Mendoza^{1,2}, Jie Shen^{1,5}, Jun He^{1,2}, Giovanni Deiana^{1,6}, Mihaela Diana Zanoaga^{1,7}, Zeynep Asgel^{1,8}, Abigail Pilcher¹, Luciana Di Lascio^{1,9}, Ana Makharashvili^{1,2}, Dora Koller^{1,10}, Daniel S. Tylee¹, Gita A. Pathak^{1,2}, Renato Polimanti^{1,2,11,12}

¹Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA

²VA CT Healthcare Center, West Haven, CT, USA

³Department of Clinical Science, University of Bergen, Bergen, Norway

⁴Dr. Einar Martens Research Group for Biological Psychiatry, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway

⁵Department of Cardiology, Children's Hospital of Soochow University, Suzhou, China.

⁶Center for Neuroscience, Pharmacology Unit, School of Pharmacy, University of Camerino

⁷Department of Computational Biology, University of Lausanne, Lausanne, Switzerland

⁸Department of Child and Adolescent Psychiatry, NYU Langone Health, New York Metropolitan Area, USA

⁹IRCCS Istituto Clinico Humanitas, Rozzano, Milan, Italy; Humanitas University, Pieve Emanuele, Milan, Italy.

¹⁰Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Catalonia, Spain

¹¹Wu Tsai Institute, Yale University, New Haven, CT, USA

¹²Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA

Correspondence: Renato Polimanti, PhD. Department of Psychiatry, Yale University School of Medicine. 60 Temple, Suite 7A, New Haven, CT 06511, USA. Phone: +1 (203) 737-6391. E-mail: renato.polimanti@yale.edu

ABSTRACT

We leveraged information from more than 1.2 million participants to investigate the genetics of anxiety disorders across five continental ancestral groups. Ancestry-specific and cross-ancestry genome-wide association studies identified 51 anxiety-associated loci, 39 of which are novel. Additionally, polygenic risk scores derived from individuals of European descent were associated with anxiety in African, Admixed-American, and East Asian groups. The heritability of anxiety was enriched for genes expressed in the limbic system, the cerebral cortex, the cerebellum, the metencephalon, the entorhinal cortex, and the brain stem. Transcriptome- and proteome-wide analyses highlighted 115 genes associated with anxiety through brain-specific and cross-tissue regulation. We also observed global and local genetic correlations with depression, schizophrenia, and bipolar disorder and putative causal relationships with several physical health conditions. Overall, this study expands the knowledge regarding the genetic risk and pathogenesis of anxiety disorders, highlighting the importance of investigating diverse populations and integrating multi-omics information.

MAIN

Anxiety disorders affect many people worldwide, with a lifetime prevalence of approximately 34%, impacting overall health and all-cause mortality^{1,2}. Although, according to the DSM-5³, this diagnostic category contains multiple anxiety disorders with distinct sets of diagnostic criteria, there is a significant phenotypic overlap among them; 48-68% of individuals with an anxiety diagnosis fulfill the criteria for at least a second one⁴. Likewise, twin and genome-wide association studies (GWAS) demonstrated a shared genetic basis for the whole spectrum of anxiety diagnoses⁵. Specifically, anxiety disorders have an estimated twin-based heritability of 20–60% across subtypes and are highly polygenic^{6,7}. GWAS contributed to understanding the genetic architecture of anxiety disorders by identifying up to ten risk loci⁸⁻¹⁰. These were primarily conducted in individuals of European descent, with limited samples available from African American participants enrolled in the Million Veteran Program (MVP)¹⁰. While the lack of diversity is a known issue in human genetic research¹¹, genetic research on anxiety appears to progress slower and be less diverse compared to current studies of other internalizing disorders, such as major depressive disorder (MDD)¹² and posttraumatic stress disorder (PTSD)¹³. To fill this gap, we conducted a multi-ancestry GWAS meta-analysis combining newly generated cohorts with previously reported data, reaching a total sample size of 1,266,780 participants (97,383 anxiety cases; Table 1). We included individuals of European (EUR, N=1,096,458), African (AFR, N=118,071), Admixed-American (AMR, N=36,634), South Asian (SAS, N=10,534), and East Asian (EAS, N=5,083) ancestries available from the All of Us Research Program (AoU)¹⁴, the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH)¹⁵, the FinnGen Project¹⁶, MVP¹⁰, the Psychiatric Genomics Consortium (PGC)¹⁷, and the UK Biobank (UKB)¹⁸. Studying these cohorts, we identified multiple loci associated with anxiety and gained insights into the role of polygenic risk, pleiotropy, tissue-specific regulation, and transcriptomic and proteomic variations in the pathogenesis of anxiety disorders.

RESULTS

Genetic Discovery for Common Anxiety Factor, SNP-Heritability, and Genetic Correlation.

We combined GWAS of anxiety from six cohorts consisting of five ancestries (Table 1). In the EUR cohort, statistically significant single nucleotide polymorphism (SNP)-based heritability was observed in all cohorts, with Z-scores ranging from 2.64 (PGC) to 15.13 (FinnGen), and values proportional to the effective sample size of each cohort. The genetic

correlation (r_g) ranged from 0.46 ($p=2.1 \times 10^{-3}$), between PGC and iPSYCH, to 1.02 ($p=2.0 \times 10^{-5}$), between PGC and AoU, with a median estimate of 0.72 (Figure 1; Supplemental Table 1). To quantify the genetically inferred differences across anxiety phenotypes across EUR datasets (Supplemental Figure 1), we applied PheMED (Phenotypic Measurement of Effective Dilution)¹⁴. FinnGen was considered the reference sample because it had the highest SNP-based heritability z score among the datasets investigated. MVP exhibited the largest phenotypic dilution ($\phi=2.92$, $p=5 \times 10^{-12}$), followed by AoU ($\phi=2.35$, $p=6.0 \times 10^{-13}$), PGC ($\phi=1.81$, $p=2.0 \times 10^{-13}$) and UKB ($\phi=1.73$, $p<1 \times 10^{-300}$), while the iPSYCH anxiety phenotype was not significantly diluted when compared to FinnGen ($\phi=1.09$, $p=0.09$). Leveraging data from cohorts assessed with different approaches, we observed that FinnGen and iPSYCH show more consistent genetic effects than the other samples investigated, with MVP showing the highest phenotypic dilution factor. We did not observe statistically significant SNP-based heritability among non-EUR groups due to their limited sample size (Supplemental Table 2). Therefore, we could not explore the genetic correlation and phenotypic dilution among anxiety datasets in other populations.

We combined the genome-wide information of the 1,096,458 EUR individuals with genomic structural equation modeling (gSEM)¹⁵ to account for the variability in the genetic correlation and the phenotypic heterogeneity among cohorts. The common anxiety factor (ANX; $\chi^2(9)=13.4$, $p=0.15$, comparative fit index=1, standardized root mean square residual=0.08) loaded significantly on all indicators (standardized loadings on AoU=0.85 \pm 0.05, FinnGen=0.82 \pm 0.03, MVP=0.80 \pm 0.04, iPSYCH=0.80 \pm 0.05, PGC=0.80 \pm 0.08, UKB=0.95 \pm 0.05; Figure 2). ANX showed a SNP-based heritability of 0.05 \pm 0.002 z -score=25.1) with no evidence of systematic bias due to population stratification or other confounders (linkage disequilibrium score regression¹⁶ LDSC intercept=1.01 \pm 0.01, ratio=0.01 \pm 0.02). We identified 35 linkage disequilibrium (LD)-independent ($r^2<0.1$) variants with genome-wide significant association ($p<5 \times 10^{-8}$) with ANX (Figure 3; Supplemental Table 3). A conditional analysis¹⁷ yielded five additional genome-wide significant (GWS) variants, leading to a total of 40 genome-wide significant independent SNPs (Supplemental Table 3). Of these, 29 were LD-independent ($r^2<0.1$) from variants reported by previous anxiety GWAS studies.⁸⁻¹⁰ Among the novel associations, the most significant one was rs6689226 (beta=0.01, $p=4.8 \times 10^{-12}$), located in an intergenic region downstream of *LINC01360* gene on chromosome 1.

Due to the limited sample size of other population groups, we could not apply the gSEM approach. These ancestry-specific meta-analyses were conducted using a sample-

size weighted approach¹⁸ to combine genetic effects derived from different assessments (Table 1). In the AFR GWAS meta-analysis of UKB, MVP, and AoU, we observed a GWS association for rs575403075 (beta=-0.11, $p=2.8 \times 10^{-8}$), located within a candidate cis-regulatory element with enhancer-like signature (ENCODE accession number: EH38E2597848) on chromosome 7¹⁹. This AFR-specific locus (non-AFR minor allele frequency<0.01) was previously associated with anxiety in the AFR GWAS in MVP.¹⁰ No GWS associations were observed for the other populations. Because of the limited sample size of the other population groups, we observed cross-ancestry replication ($p<0.01$) only for two EUR GWS associations (Supplemental Table 3): rs12457101 (EUR beta=0.01, $p=2.1 \times 10^{-9}$; SAS beta=0.12, $p=0.008$) and rs10078721 (EUR beta=0.01, $p=4.9 \times 10^{-8}$; AFR beta=0.02, $p=0.009$). EUR-derived polygenic risk scores (PRS) were significantly associated with anxiety phenotypes (fifth quintile vs. first quintile of PRS distribution) in AFR-AoU (OR=1.23, 95%CI=1.12-1.36), AMR-AoU (OR=1.55, 95%CI=1.39-1.74), EAS-AoU (OR=1.69, 95%CI=1.16-2.49), and EUR-AoU (OR=1.79, 95%CI=1.69-1.89) (Figure 4; Supplemental Table 4). The strength of the cross-ancestry PRS associations was proportional to the genetic distance between EUR and the other populations investigated²⁰. Likely due to the limited sample size, null EUR-derived PRS association was observed with respect to SAS-AoU (Figure 4; Supplemental Table 4).

A cross-ancestry meta-analysis revealed 41 GWS LD-independent loci, of which ten were novel (Supplemental Table 5), considering previously reported variants⁸⁻¹⁰ and index variants identified in the present ancestry-specific GWAS. The most significant among the novel cross-ancestry GWS associations was rs2510682 (cross-ancestry $z=5.92$, $p=3.3 \times 10^{-9}$; EUR beta=-0.12, $p=6.2 \times 10^{-8}$; AFR beta=-0.02, $p=0.05$; AMR beta=-0.04, $p=0.18$; EAS beta=-0.04, $p=0.69$; SAS beta=-0.04, $p=0.44$) located in an intronic region of *CNTN5* gene. For this locus and most of the other cross-ancestry associations, the signal was almost exclusively driven by the EUR sample. However, rs11681562 showed a cross-ancestry GWS association ($z=-5.67$, $p=1.4 \times 10^{-8}$) driven by both EUR and AFR samples (EUR beta=-0.01, $p=4.3 \times 10^{-7}$; AFR beta=-0.02, $p=0.005$; EAS beta=-0.05, $p=0.64$; SAS beta=-0.07, $p=0.12$).

Transcriptomic and Proteomic Analyses.

Partitioning the EUR-ANX heritability to 205 annotations based on tissue- and cell-specific transcriptomic profiles (Supplemental Table 6), we identified that ANX genetic liability was enriched for genes expressed in multiple brain regions, including the limbic system ($p=2.3 \times 10^{-5}$), the cerebral cortex ($p=4.2 \times 10^{-5}$), the cerebellum ($p=7.4 \times 10^{-5}$), the

metencephalon ($p=1.1 \times 10^{-4}$), the entorhinal cortex ($p=1.5 \times 10^{-4}$), and the brain stem ($p=2.0 \times 10^{-4}$).

We performed a transcriptome-wide association study (TWAS) of the EUR-ANX GWAS using S-PrediXcan²¹. Initially, we conducted a tissue-specific TWAS considering 13 brain tissues available from GTEx v8²². We observed 152 transcriptome-wide significant associations related to 39 genes accounting for the number of genes and brain tissues tested ($N=165,710$, $p<3.0 \times 10^{-7}$) with a Bonferroni correction (Supplemental Table 7). The strongest association was related to the genetically regulated transcriptomic variation of the *DRD2* locus in the cerebellar hemisphere ($z=-6.8$, $p=1.1 \times 10^{-11}$). Interestingly, no associations survived Bonferroni correction for this gene among the other brain tissues tested. Conversely, four genes showed Bonferroni significant associations with 10 or more brain tissues: *CD40* (from hippocampus $z=-6.0$, $p=1.5 \times 10^{-9}$ to nucleus accumbens $z=-5.1$, $p=2.7 \times 10^{-7}$), *GNL3* (from spinal cord $z=5.5$, $p=4.1 \times 10^{-8}$ to cerebellum $z=5.3$, $p=1.0 \times 10^{-7}$), *NEK4* (from cerebellar hemisphere $z=5.6$, $p=5.6 \times 10^{-8}$ to spinal cord $z=5.2$, $p=2.5 \times 10^{-7}$), and *SMIM4* (from anterior cingulate cortex $z=-5.9$, $p=3.4 \times 10^{-9}$ to spinal cord $z=-5.8$, $p=6.3 \times 10^{-9}$). To investigate cross-tissue transcriptomic regulation further, we combined information from 49 tissues available from GTEx V8²² in a multi-tissue TWAS using S-MultiXcan²³. After Bonferroni correction accounting for the number of genes tested ($N=22,045$; $p<2.3 \times 10^{-6}$), we identified 94 loci with genetically regulated transcriptomic associations with ANX (Figure 3, Supplemental Table 8). Fifty-nine of the 94 loci were not observed in the brain-specific TWAS, including the most significant transcriptome-wide association related to *MED19* (S-MultiXcan $p=5.5 \times 10^{-15}$). The second strongest multi-tissue association was *LINC01360* (S-MultiXcan $p=1.5 \times 10^{-11}$) near rs6689226, the most significant novel variant in the present EUR-ANX GWAS.

To understand the potential role of brain proteomic regulation in the pathogenesis of anxiety, we tested the EUR-ANX GWAS against protein quantitative trait loci (pQTL) in the dorsolateral prefrontal cortex (dlPFC) with FUSION²⁴. We identified 24 loci with evidence of genetically regulated proteomic association with ANX after Bonferroni correction accounting for the number of genes tested ($N=1,629$; $p<3.1 \times 10^{-5}$) (Figure 3; Supplemental Table 9). We observed six genes identified by the proteome-wide association study (PWAS) to also be Bonferroni significant in the multi-tissue TWAS: *CTNND1* (PWAS $p=3.5 \times 10^{-16}$; multi-tissue TWAS $p=2.1 \times 10^{-8}$), *CNNM2* (PWAS $p=8.3 \times 10^{-14}$; multi-tissue TWAS $p=2.3 \times 10^{-7}$), *RAB27B* (PWAS $p=9.4 \times 10^{-14}$; multi-tissue TWAS $p=5.9 \times 10^{-8}$), *KHK* (PWAS $p=4.2 \times 10^{-9}$; multi-tissue TWAS $p=8.1 \times 10^{-7}$), *NEK4* (PWAS $p=9.4 \times 10^{-8}$; multi-tissue TWAS $p=7.2 \times 10^{-8}$), and *BTN2A1* (PWAS $p=9.6 \times 10^{-9}$; multi-tissue TWAS $p=1.2 \times 10^{-7}$). Of those, *CTNND1* and *NEK4*

also demonstrated Bonferroni-significant expression in areas of the cortex, the basal ganglia, and the cerebellum in the brain TWAS. *CGREF1* was a PWAS significant hit (PWAS $p=3.3 \times 10^{-7}$), with a significant brain TWAS expression in the frontal cortex (TWAS $p=1.13 \times 10^{-7}$) and the putamen (TWAS $p=2.73 \times 10^{-7}$). The Bonferroni-significant PWAS loci were enriched for multiple synaptic locations and biological processes, including synapse (gene ontology, GO:0045202, $p=4.6 \times 10^{-4}$), synaptic vesicle membrane (GO:0030672, $p=5.9 \times 10^{-4}$), presynapse, synaptic vesicle (GO:0008021, $p=0.001$), process in the synapse (SYNGO:synprocess, $p=0.003$), and process in the presynapse (SYNGO:presynprocess, $p=0.008$).

Pleiotropy with Human Traits and Diseases.

We used MixeR²⁵ to investigate the polygenic architecture of anxiety in the context of other psychiatric disorders. A total of $12,622 \pm 834$ influential variants were estimated for EUR-ANX. A similar estimate was observed for MDD ($N=11,428 \pm 453$, $p_{\text{ANX-difference}}=0.21$). Conversely, for anorexia nervosa (AN, $N=7,869 \pm 387$, $p_{\text{ANX-difference}}=1.9 \times 10^{-6}$), attention-deficit hyperactivity disorder (ADHD, $N=7,978 \pm 387$, $p_{\text{ANX-difference}}=4.4 \times 10^{-7}$), bipolar disorder (BIP, $N=8,772 \pm 387$, $p_{\text{ANX-difference}}=1.4 \times 10^{-5}$), PTSD ($N=7,585 \pm 533$, $p_{\text{ANX-difference}}=3.6 \times 10^{-7}$), and schizophrenia (SCZ, $N=9,636 \pm 262$, $p_{\text{ANX-difference}}=6.4 \times 10^{-4}$), the number of influential variants was significantly lower. No reliable estimate was obtained for autism spectrum disorder (ASD), obsessive-compulsive disorder (OCD), and Tourette syndrome (TS, Supplemental Table 10). Considering statistically meaningful bivariate MixeR models (Akaike information criterion for the best model versus model with minimal possible polygenic overlap >0 ; Supplemental Table 11), ANX shared $91 \pm 4\%$, $75 \pm 9\%$, and $68 \pm 1\%$ of its influential variants with MDD, BIP, and SCZ, respectively (Figure 5A; Supplemental Figure 2).

To estimate the local genetic correlation of ANX with other psychiatric traits, we used Local Analysis of (co-)Variant Association (LAVA)²⁶. After Bonferroni correction for the number of tests performed ($N=12,368$, $p<4.1 \times 10^{-6}$), we observed 47 local genetic correlations of ANX with five psychiatric disorders (i.e., ADHD, BIP, MDD, PTSD, and SCZ) across 35 genomic regions (Figure 5B; Supplemental Table 12). Consistent with the extensive genetic overlap observed in MixeR, we observed local genetic correlations between ANX and MDD in 39 genomic regions. Twelve regions showed evidence of local genetic correlation between ANX and SCZ. Only two regions were genetically correlated between ANX and ADHD, BIP, and PTSD, respectively. ANX demonstrated local genetic correlations with two or more psychiatric disorders in 10 regions. Among them, ANX showed

local genetic correlation with MDD ($\rho=0.97$, $p=2.3 \times 10^{-6}$), SCZ ($\rho=0.84$, $p=7.2 \times 10^{-12}$), and BIP ($\rho=0.73$, $p=8.3 \times 10^{-8}$) within locus 1719 (chr11:112,755,447-113,889,019; Supplemental Figure 3. Within locus 2281 (chr18:52,512,524-53,762,996), ANX local genetic correlation was found with MDD ($\rho=0.88$, $p=1.7 \times 10^{-12}$) and SCZ ($\rho=0.67$, $p=1.9 \times 10^{-7}$). Among the most significant genomic regions, we also observed locus 852 (chr5:91,956,906-93,814,604), where ANX was genetically correlated with ADHD ($\rho=0.97$, $p=4.7 \times 10^{-7}$) and MDD ($\rho=1$, $p=3.4 \times 10^{-8}$).

To decompose the pleiotropic mechanisms linking ANX to multiple psychiatric disorders, we applied LAVA multiple regression models²⁶ to the five genomic regions with the strongest evidence of local genetic correlation with ANX (Supplemental Table 13). ANX was entered in the regression models as an outcome, and the other genetically correlated traits as predictors. We observed that the local genetic correlations of ANX with MDD, SCZ, and BIP in locus 1719 were not independent of each other. Shared pleiotropic mechanisms involving ANX were also observed in locus 62 (chr1:72,513,120-73,992,170) for MDD and SCZ and locus 852 (chr5:91,956,906-93,814,604) for ADHD and MDD. Conversely, for locus 1582 (chr10:106,142,284-107,877,787), we identified independent pleiotropic mechanisms of ANX with MDD ($\rho=0.91$, $p=5.2 \times 10^{-9}$; $\gamma=0.63$, $p=0.04$) and SCZ ($\rho=1$, $p=2.8 \times 10^{-8}$; $\gamma=0.78$, $p=0.004$). For locus 2281 (chr18:52,512,524-53,762,996), the ANX-MDD local genetic correlation ($\rho=0.88$, $p=1.7 \times 10^{-12}$; $\gamma=0.78$, $p=0.003$) was the primary pleiotropic driver that also accounted for ANX-SCZ relationship ($\rho=0.67$, $p=1.9 \times 10^{-7}$; $\gamma=0.16$, $p=0.35$).

To investigate the shared genetic mechanisms between ANX and human traits and diseases, we performed a phenome-wide genetic correlation analysis leveraging 11,175 phenotypes available from UKB, MVP, and FinnGen²⁷⁻²⁹. After Bonferroni correction ($p < 4.47 \times 10^{-6}$), 1,929 showed a statistically significant genetic correlation with ANX (Figure 6, Supplemental Table 14). Among these, the median $|rg|$ was 0.42 (25th percentile=0.29; 75th percentile=0.55), with 77 of the top 100 results (all $|rg| > 0.72$) being mental health outcomes (e.g., FinnGen anxiety disorders $rg=0.91$, $p < 1 \times 10^{-300}$; UKB self-reported depression $rg=0.83$, $p=1.0 \times 10^{-53}$) and related drug prescriptions (e.g., UKB citalopram medication $rg=0.93$, $p=7.3 \times 10^{-8}$; FinnGen depression medications $rg=0.82$, $p < 1 \times 10^{-300}$). Among non-psychiatric outcomes, the top results included myalgia (FinnGen $rg=0.85$, $p=4.0 \times 10^{-12}$), nausea/vomiting (UKB $rg=0.84$, $p=1.5 \times 10^{-11}$), diseases of pulp and periapical tissues (FinnGen $rg=0.84$, $p=4.6 \times 10^{-7}$), dysuria (MVP $rg=0.79$, $p=1.7 \times 10^{-6}$), cystitis (FinnGen $rg=0.77$, $p=8.1 \times 10^{-9}$), constipation (UKB $rg=0.75$, $p=1.2 \times 10^{-20}$), pain related to

temporomandibular disorders (FinnGen $rg=0.75$, $p=1.4 \times 10^{-16}$), unspecified dorsalgia (FinnGen $rg=0.73$, $p=3.1 \times 10^{-45}$), and pain management (UKB $rg=0.73$, $p=6.1 \times 10^{-30}$).

To investigate potential causal relationships underlying those genetic correlations, we employed the latent causal variable approach (LCV)³⁰. We identified 24 traits with statistically significant genetic causal proportion (gcp) after a false discovery rate (FDR) multiple testing correction (FDR $q<0.05$, $p<4.1 \times 10^{-6}$) (Supplemental Table 15). Positive causal effects were observed from other phenotypes to ANX in all cases (LCV $\rho>0$). Only three of these phenotypes were related to mental health: “Number of things worried about during worst period of anxiety” (UKB gcp=0.89, $p=4.1 \times 10^{-63}$; $\rho=0.67 \pm 0.10$), “Ever had period of mania/excitability” (UKB gcp=0.46, $p=1.4 \times 10^{-28}$; $\rho=0.51 \pm 0.08$), and “Number of cigarettes currently smoked daily” (UKB gcp=0.54, $p=2.2 \times 10^{-7}$; $\rho=0.23 \pm 0.07$). Seven of the phenotypes identified were related to physical health concerns, with the strongest relationship implicating Dupuytren's disease (MVP, gcp=0.71, $p=7.1 \times 10^{-93}$; $\rho=0.13 \pm 0.05$). Five were related to the cardiovascular system: “Atherosclerotic cardiovascular disease” (MVP, gcp=0.60, $p=8.9 \times 10^{-29}$; $\rho=0.40 \pm 0.09$), “Other disorders of circulatory system” (UKB, gcp=0.43, $p=5.2 \times 10^{-26}$; $\rho=0.46 \pm 0.07$), “Medication for transient ischemic attack” (MVP, gcp=0.63, $p=6.4 \times 10^{-17}$; $\rho=0.24 \pm 0.08$), “Other specified cardiac dysrhythmias” (UKB, gcp=0.93, $p=7.0 \times 10^{-15}$; $\rho=0.40 \pm 0.08$), and “Heart Surgery” (UKB, gcp=0.96, $p=7.0 \times 10^{-15}$; $\rho=0.40 \pm 0.08$). We also observed two environmental variables with FDR-significant effects: “Workplace had a lot of diesel exhaust” (UKB, gcp=0.95, $p=2.9 \times 10^{-24}$; $\rho=0.45 \pm 0.10$) and “Number of gap periods in employment” (UKB, gcp=0.60, $p=5.4 \times 10^{-7}$; $\rho=0.26 \pm 0.08$).

DISCUSSION

The present study analyzed genome-wide information of more than 1.2 million participants, giving valuable insights into the genetic liability of anxiety disorders. The ANX GWAS in EUR identified 40 GWS associations, 29 of which were novel. Unfortunately, the limited sample size of other ancestry groups permitted us to only confirm a previously reported locus in AFR¹⁰. However, the cross-ancestry GWAS added ten novel anxiety risk loci to our gene discovery. The novel findings of the EUR-specific and cross-ancestry analyses quadrupled the gene discoveries reported by previous studies⁸⁻¹⁰. Among the novel loci identified, the strongest ANX association was with rs6689226, located near *LINC01360*. While the regulatory mechanisms related to this long non-coding RNA (lncRNA) are unclear, this variant was previously associated with several mental health outcomes such as smoking initiation³¹, externalizing behaviors³², and cross-disorder psychopathology³³. Our multi-

tissue transcriptomic analysis identified genetically regulated expression of *LINC01360* in the testis. LncRNAs are more expressed in the testis than other tissues, and some of them play a role in androgen production³⁴, which may be of pathophysiological relevance in certain animal models of anxiety³⁵. In the cross-ancestry GWAS, the strongest ANX association was rs2510682, located in an intron region of the *CNTN5* gene, which encodes a protein mediating cell surface interactions during the development of the nervous system³⁶. While it has not been associated with anxiety by previous GWAS⁸⁻¹⁰, *CNTN5* variants were linked to neuroticism³⁷, suicidal behaviors³⁸, and ASD³⁹. The second most significant cross-ancestry association, rs11681562, was also a novel ANX locus. Variants in high LD with rs11681562 ($r^2 > 0.8$) have been previously associated with multiple phenotypes related to human cortical folding (i.e., vertex-wise sulcal depth, vertex-wise cortical thickness, vertex-wise cortical surface area, cortical thickness, and cortical surface area)⁴⁰. Rs11681562 is an eQTL for multiple genes located in a chromosomal region with high gene density. In this region, our transcriptome-wide and proteome-wide analyses converge on *KHK* as potentially linked to anxiety pathogenesis. This gene is responsible for producing an enzyme involved in fructose metabolism⁴¹. Animal studies highlighted the potential role of early-life high fructose exposure in long-term depression- and anxiety-like behaviors⁴². *KHK* has also been identified as a PTSD risk locus in transcriptomic and proteomic analyses in the brain and blood⁴³. However, we cannot exclude that other genes in this region may contribute to the association observed.

The transcriptome- and proteome-wide analyses converged on five additional loci: *CTNND1*, *CNNM2*, *RAB27B*, *BTN2A1*, and *NEK4*. In both approaches, the strongest statistical evidence was observed for the *CTNND1* gene. This encodes a protein regulating the dendritic spine and synapse development through rho-family GTPases and cadherins⁴⁴. The overexpression of *CTNND1* has been linked to improved memory and reduced anxiety in mice⁴⁵. The *CNNM2* gene, encoding a cyclin protein involved in Mg^{2+} transport, has been associated with impaired neuronal development and epilepsy⁴⁶. *RAB27B* has been identified as a PTSD and depression risk locus in transcriptomic and proteomic analyses in the brain and blood⁴⁷. Reduced *BTN2A1* expression in the placenta was linked to immune system processes in the context of maternal anxiety and depression⁴⁸. *NEK4* has been previously associated with ADHD, BIP, migraines, and cross-disorder psychopathology in genetic and multi-omics analyses⁴⁹⁻⁵².

In line with the cross-disorder overlap observed in the single-locus discoveries described above, our genome-wide analyses indicated pleiotropic mechanisms between ANX and other psychiatric disorders. We observed a similar degree of polygenicity and a

high genetic correlation between ANX and MDD ($r_g=0.9$), consistent with previous studies^{53,54} and their shared symptomatology^{55,56}. However, while 99% of MDD influential variants appear to be shared with ANX, our MixeR analysis suggests that about 10% of the ANX influential variants do not overlap with those of MDD. This may suggest, from a genetic perspective, that MDD is considered an ANX subtype.

LAVA identified several genomic regions with statistical evidence for the local genetic correlation of ANX with multiple psychiatric disorders. In line with the global pleiotropy analysis, ANX shared the most local genetic correlations with MDD. We also observed multiple regions where ANX showed local genetic correlation with two or more mental illnesses. Among them, locus 1719 (chr11:112,755,447-113,889,019) linked ANX to MDD, SCZ, and BIP. Based on data available from the GWAS catalog⁵⁷, over 400 GWS associations involving more than 100 phenotypes have been previously reported in this genomic region (Supplemental Table 16), with the majority related to mental health outcomes, such as anxiety, SCZ, BIP, MDD, PTSD, substance use, and educational attainment. Within this region, we observed the strongest brain-specific TWAS finding, involving the association of ANX with the genetically regulated transcriptomic variation of *DRD2* in the cerebellar hemispheres but not in other brain tissues. This locus is well-known in psychiatric genetics, and considerable research has characterized the function of the encoded dopamine D2 receptor (DR2) protein⁵⁸. In mice, changes in cerebellar D2R levels in Purkinje cells alter sociability and preference for social novelty⁵⁹. The present findings suggest a role for cerebellar D2R in the pathophysiology of human anxiety disorders.

Additionally, we observed a strong genetic overlap between ANX and health-related traits beyond psychiatric disorders. Among the non-psychiatric phenotypes in the 100 highest genetic correlations, we observed several gastrointestinal (43%) and pain-related (17%) phenotypes involving different body areas (e.g., muscle, temporomandibular, and thoracic). Anxiety is highly comorbid with gastrointestinal (GI) disorders. Irritable bowel syndrome has a common genetic basis with anxiety and mood disorders^{60,61}. However, the genomic links between anxiety and the broader spectrum of GI disorders are still unclear. A growing body of literature supports the complex interplay between pain and anxiety disorders^{62,63}. A recent brain-wide analysis integrating imaging and genetic data highlighted the potential role of pain sensitivity in the pathogenesis of anxiety disorders⁶⁴. The results of our LCV analyses also identified genetic evidence of a causal impact of physical health phenotypes on anxiety disorders, particularly for multiple cardiovascular conditions. This is in line with the known comorbidity between these disorders⁶⁵ and the high anxiety prevalence after cardiac events⁶⁶.

Our common ANX factor demonstrated high loadings to all input datasets (standardized loadings ≥ 0.8), suggesting the generalizability of the results across different cohorts. However, FinnGen and iPSYCH showed more consistent genetic effects than the other samples investigated, with MVP showing the highest degree of phenotypic dilution. One potential explanation is that MVP was the only cohort where anxiety was assessed quantitatively. Those results highlight that more homogeneous and detailed assessments may improve the statistical power and interpretability of genetic studies. Unfortunately, the limited sample size and statistical power in non-EUR populations precluded identifying novel loci in these ancestries. However, we demonstrated evidence that the polygenic risk for ANX detected in EUR is translatable to other population groups. Similar to prior observations⁶⁷, the strength of the cross-ancestry PRS association was proportional to the genetic distance between the population group of the training and target samples. To improve inclusion and equity in anxiety genetics, future studies should prioritize recruiting and assessing diverse cohorts.

In conclusion, this large-scale genome-wide and multi-omics investigation yielded novel insights into the biology of anxiety disorders. We significantly expanded the number of known anxiety risk loci, highlighting the importance of analyzing data from ancestrally diverse participants. Our multi-omics analyses identified novel pathways putatively involved in the pathogenesis of anxiety and refined plausible mechanisms related to previously reported loci. We also disentangled the pleiotropic mechanisms linking ANX to MDD and other psychiatric disorders, reinforcing the importance of studying anxiety disorders in the context of physical health.

METHODS

Study Populations.

The present study leveraged genome-wide data from six cohorts including AoU⁶⁸, FinnGen⁶⁹, iPSYCH⁷⁰, MVP¹⁰, PGC⁷¹, and UKB²⁸ (Table 1). This permitted us to investigate a total of 1,266,780 participants with diverse ancestral backgrounds. AoU is a research program aiming to create a representative cohort of the US population to accelerate biomedical research and improve healthcare through precision medicine⁶⁸. Among the samples investigated in the present study, the AoU cohort was the most ancestrally diverse, with information on five genetically inferred population groups (i.e., AFR, AMR, EAS, EUR, and SAS). Details regarding AoU recruitment, assessment, and whole-genome sequencing

have been previously described⁶⁸, and the genomic quality control is described at <https://support.researchallofus.org/hc/en-us/articles/4617899955092-All-of-Us-Genomic-Quality-Report>. In the present study, we analyzed genetic and phenotypic information for 201,334 individuals. The definition of anxiety disorder cases and controls was derived from AoU electronic health records (EHR) using the SNOMED (Systematized Nomenclature of Medicine) classification system⁷². Cases were defined as AoU participants with a lifetime diagnosis of generalized anxiety disorder (SNOMED ID: 434613), panic disorder (SNOMED ID: 4138454), phobic disorder (SNOMED ID: 4304010), chronic stress disorder (SNOMED ID: 436074), acute stress disorder (SNOMED ID: 440083), posttraumatic stress disorder (SNOMED ID: 436676), obsessive-compulsive disorder (SNOMED ID: 440374), or mixed anxiety and depressive disorder (SNOMED ID: 4338031). Any participant without a documented EHR diagnosis falling into the above categories until the maximum follow-up age was considered a control. Ancestry-stratified GWAS was performed using the following quality control criteria: biallelic variants with minor allele frequency >1%, Hardy-Weinberg equilibrium $p < 10^{-6}$, call rate >95%, and per-individual genotyping rate >95%. A PLINK 2⁷³ logistic regression model estimated genetic associations, including covariates for sex, age, and the first ten within-ancestry principal components.

FinnGen is a project focused on developing a large Finnish cohort combining genotype data from biobanks and digital health record data from health registries²⁷. In the present study, we used genome-wide association statistics generated from Release 9 (May 11, 2023). A detailed description of FinnGen data is available at <https://finngen.gitbook.io/documentation/v/r9/>. The FinnGen GWAS was performed using REGENIE⁷⁴ including covariates for age, sex, the top-ten within-ancestry PCs, and genotyping batch. In our analysis, we used GWAS data generated from 40,191 cases and 277,526 controls of EUR descent for the phenotype "anxiety disorders" (KRA_PSY_ANXIETY), defined based on International Classification of Diseases (ICD-10 F40-F48; ICD-9 300.0-300.3, 300.6-300.9, 3078A, 309; ICD-8 300.0-300.2, 30030, 300.5-300.9, 305, 30680, 30799). Details regarding the KRA_PSY_ANXIETY definition are available at https://risteys.finregistry.fi/endpoints/KRA_PSY_ANXIETY.

iPSYCH is a large Danish population-based cohort aimed at unraveling the genetic and environmental architecture of severe mental disorders⁷⁵. In the present study, we used genome-wide association statistics generated from the analysis of 12,655 cases and 19,225 controls of EUR descent⁷⁰. Cases were defined as individuals with an anxiety and stress-related diagnosis according to the ICD-10 F40.0-F41.9 and F43.0-F43.9. codes. Controls included individuals without any ICD-10 diagnoses of anxiety, stress-related disorders, or

mood disorders. Genetic associations were estimated using logistic regression models with the imputed marker dosages using the first four principal components as ancestry covariates. The analysis was stratified by genotyping batch, and the results were meta-analyzed using an inverse variance–weighted fixed-effect models.

MVP is a biobank funded by the US Department of Veterans Affairs to understand how genes, lifestyle, military experiences, and exposures affect health and wellness⁷⁶. The GWA statistics used in the present study were generated from an analysis of 61,796 and 241,541 participants of AFR and EUR descent, respectively¹⁰. The anxiety phenotype was the total score of the generalized anxiety disorder 2-item scale, which ranged from 0 to 6. The genetic association analysis was performed using a PLINK 2⁷³ logistic regression model with covariates for age, sex, and the first ten within-ancestry principal components.

PGC genome-wide association statistics were derived from a meta-analysis conducted by the Anxiety NeuroGenetics STudy (ANGST)⁷¹. Cases were defined as individuals with diagnoses of generalized anxiety disorder, panic disorder, social phobia, agoraphobia, or specific phobias⁷¹. A total of 5,761 cases and 11,765 controls of EUR descent were included. The GWAS was performed using a logistic regression model with covariates for sex, age, and the first ten within-ancestry principal components.

UKB is a large population-based study that collected information regarding over 500,000 participants²⁸. GWA statistics used in the present study were obtained from the Pan-UKB analysis. Details regarding the quality control, ancestry assignment, and statistical analyses are available at <https://pan.ukbb.broadinstitute.org/>. Briefly, ancestry-stratified GWASs were performed using the Scalable and Accurate Implementation of Generalized (SAIGE)⁷⁷ mixed models, including a kinship matrix as a random effect and the remaining covariates (i.e., top-10 within-ancestry PC, sex, age, age², sex×age, and sex×age²) as fixed effects. We analyzed UKB a total of 10,751 cases and 383,235 controls of AFR, EUR, and SAS descents for the phecode 300 “Anxiety, dissociative and somatoform disorders”.

Phenotype Definition Comparisons.

Because of the different anxiety definitions, we compared their genetic architecture using two methods. LD Score regression⁷⁸ was used to estimate their SNP-based heritability and their pairwise genetic correlations. LD scores were calculated using HapMap 3 variants and the 1000 Genomes Project reference populations corresponding to each ancestry group⁷⁹. Phenotype dilution among anxiety definitions was quantified using PheMED¹⁴,

considering the FinnGen cohort as the reference sample because it showed the highest SNP-based heritability z-score. Because of the lack of significant SNP-based heritability in non-EUR samples, genetic correlation, and phenotype dilution analyses were performed only in EUR datasets.

Ancestry-Specific and Cross-Ancestry Meta-Analyses and Conditional Analysis.

Due to moderate genetic correlation (median $rg=0.72$) and statistically significant phenotype dilution among anxiety phenotypes, we combined the EUR GWASs using the gSEM approach¹⁵. As recommended⁸⁰, we used the effective sample size for each EUR anxiety GWAS. For binary definitions, this was calculated as $4/(1/N_{cases} + 1/N_{controls})$. For quantitative phenotypes, we used the total sample size. The effective sample size of the common ANX factor obtained from the gSEM analysis was estimated as $mean(1/(2 \times MAF \times (1 - MAF) \times SE^2))$, where MAF is the minor allele frequency and SE is the standard error. Due to the limited sample sizes, we could not use gSEM to combine genome-wide information available for AFR (AoU, UKB, and MVP) and SAS (AoU and UKB). Thus, we performed ancestry-specific meta-analyses using the sample-size weighted approach available in METAL¹⁸. To present ancestry-specific effects with the same statistics, AFR and SAS z-scores obtained from the sample-size weighted approach were converted to betas using the following formula: $beta = z/\sqrt{2 \times p \times (1 - p) \times (n + z^2)}$, where p is the effect allele frequency and n is the sample size. The sample-size weighted approach, using effective sample sizes, was also used to perform the cross-ancestry meta-analysis. Independent associations were identified by clumping GWS associations considering LD $r^2 < 0.1$. To identify secondary associations accounting for the primary GWS signals, we applied the stepwise model selection procedure available in the COJO approach implemented in the GCTA package^{17,81}. Among the GWS associations identified by the conditional analysis, we considered only those with LD $r^2 < 0.8$ with respect to the corresponding index GWS variant.

Cross-Ancestry Polygenic Risk Scoring.

To investigate how anxiety polygenic risk translates across ancestry groups, we derived PRS from the EUR ANX factor and tested them in other populations available from the AoU cohort (i.e., AFR, AMR, EAS, and SAS). To compare within-ancestry to cross-ancestry PRS associations, we derived a PRS from the EUR ANX factor excluding AoU-EUR and used the latter as the EUR target sample. Posterior variant-level effect sizes of the

EUR ANX factor were computed using PRS-CS⁸² and the 1000 Genomes European population as LD reference. ANX PRS were computed using PLINK⁷³. We tested the quintiles of PRS distribution using the first quintile as reference with logistic regression models available in the *stats* R package in AoU population groups.

Heritability Analyses.

LDSC⁷⁸ was used to estimate the SNP-based heritability of anxiety leveraging ancestry-specific genome-wide association statistics using the methods described above. For the EUR ANX factor, we also performed a partitioned heritability analysis⁸³, testing the enrichment of 205 tissue- and cell-type specific transcriptomic profiles. These annotations were derived from GTEx V8²² (53 tissues) and the Franke lab^{84,85} (152 tissues/cell types).

Transcriptome- and Proteome-Wide Association Studies.

To understand the role of genetically regulated transcriptomic and proteomic variation in anxiety pathogenesis, we performed transcriptome-wide and proteome-wide association analyses. A brain-specific transcriptome-wide investigation was conducted using S-PrediXcan²¹ to integrate EUR-ANX genome-wide association statistics with GTEx eQTL data available for 13 brain tissues. A Bonferroni correction accounting for the number of tests performed ($N=165,710$, $p < 3 \times 10^{-7}$) was applied to define the statistically significant associations. To explore cross-tissue transcriptomic regulation, we performed a multi-tissue transcriptome-wide association analysis using S-MultiXcan²³. This permitted us to boost the statistical power of the association analysis via a joint multi-tissue analysis accounting for transcriptomic correlation across the 49 GTEx tissues tested. Multi-tissue genetically regulated transcriptomic associations were defined after Bonferroni correction accounting for the number of the genes tested ($N=22,045$; $p < 2.27 \times 10^{-6}$).

Proteome-wide associations were estimated using the FUSION approach²⁴ to integrate EUR-ANX GWAS with dIPFC pQTLs available from Religious Orders Study and Memory and Aging Project (ROSMAP)⁸⁶ and the Banner Sun Health Research Institute (BSHRI)⁸⁷. Analyses were performed separately in two pQTL datasets and the results were meta-analyzed using an inverse-variance weighted approach available in METAL¹⁸. Bonferroni correction accounting for the number of genes tested ($N=1,629$; $p < 3.07 \times 10^{-5}$) was applied to define proteome-wide significant genes. To understand their implications on synaptic localization and function, these loci were analyzed using SynGO⁸⁸, an online knowledge base that organizes research on synaptic proteins using GO annotations. SynGO

enrichments were estimated for each subset of cellular components and biological processes by a one-sided Fisher exact test.

Pleiotropy Analyses.

We used multiple methods and data sources to estimate the pleiotropy between EUR ANX and human traits and diseases. Initially, we focused on understanding the shared genetic mechanisms of ANX with other psychiatric disorders previously investigated using GWAS, including ADHD⁸⁹, AN⁹⁰, ASD⁹¹, BIP⁹², MDD⁹³, OCD⁹⁴, PTSD⁹⁵, SCZ, and TS⁹⁶. After excluding the MHC region (chr6:26,000,000-34,000,000), we applied mixture models available from MixeR²⁵ to estimate the number of influential variants ANX shares with other psychiatric disorders. We then applied LAVA²⁶ to estimate local genetic correlations of ANX with mental illnesses. Because we observed local genetic correlation with multiple psychiatric disorders in the same genomic regions, we used LAVA multivariate models²⁶ to dissect whether the observed local genetic correlations were common or distinct among the implicated disorders. For MixeR and LAVA analyses, European populations available from 1000 Genomes Project Phase 3 were used as LD reference.

To assess ANX pleiotropic mechanisms across the human phenotypic spectrum, we performed an LDSC genetic correlation analysis⁷⁸ for 11,175 phenotypes available from UKB (details available at <https://pan.ukbb.broadinstitute.org/>), FinnGen (Release 9, details available at <https://finngen.gitbook.io/documentation/v/r9/>), and MVP²⁹. Bonferroni correction accounting for the number of tests performed was applied to define statistically significant genetic correlations ($p < 4.47 \times 10^{-6}$). To investigate the potential causality underlying the observed genetic correlations, we conducted a phenome-wide LCV³⁰ analysis using FinnGen, UKB, and MVP data. This was performed using LD scores calculated from European populations available from the 1000 Genomes Project Phase 3. LCV permitted us to calculate gcp estimates. Positive and negative gcp values reflect the direction of the putative causal effect (i.e., phenotype #1 → phenotype #2 and phenotype #2 → phenotype #1, respectively). In the present study, positive gcp values correspond to the effect of a phenotype on ANX. Information regarding the sign of the LCV effect is determined by the LCV rho statistics: rho>0 corresponds to a positive effect, while rho<0 corresponds to a negative effect.

ACKNOWLEDGMENTS

This study was supported by grants from the National Institutes of Health (RF1 MH132337, R33 DA047527, and K99 AG078503), One Mind, the Alzheimer's Association (Research Fellowship AARF-22-967171), the American Foundation for Suicide Prevention (PDF-1-022-21), Horizon 2020 (Marie Skłodowska-Curie Individual Fellowship 101028810), University of Bergen (International Training Grant), and the Yale Franke Program in Science and Humanities. We also acknowledge the contribution of the participants and the investigators involved in the UK Biobank, the FinnGen Project, the Million Veteran Program, the All of Us Research Program, the iPSYCH study, and the Psychiatric Genomics Consortium. The All of Us Research Program is supported by the National Institutes of Health, Office of the Director: Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 OD026556; 1 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 OD026551; 1 OT2 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The Participant Center: U24 OD023176; Participant Technology Systems Center: 1 U24 OD023163; Communications and Engagement: 3 OT2 OD023205; 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276.

COMPETING INTERESTS

Dr. Polimanti is paid for their editorial work on the journal Complex Psychiatry and reports a research grant from Alkermes. The other authors declare no competing interests.

REFERENCES

1. Szuhany, K.L. & Simon, N.M. Anxiety Disorders: A Review. *JAMA* **328**, 2431-2445 (2022).
2. Mutz, J., Choudhury, U., Zhao, J. & Dregan, A. Frailty in individuals with depression, bipolar disorder and anxiety disorders: longitudinal analyses of all-cause mortality. *BMC Med* **20**, 274 (2022).
3. Association, A.P. *Diagnostic and Statistical Manual of Mental Disorders*.
4. Penninx, B.W.J.H., Pine, D.S., Holmes, E.A. & Reif, A. Anxiety disorders. *The Lancet* **397**, 914-927 (2021).
5. Meier, S.M. & Deckert, J. Genetics of Anxiety Disorders. *Curr Psychiatry Rep* **21**, 16 (2019).
6. Hettema, J.M., Neale, M.C. & Kendler, K.S. A review and meta-analysis of the genetic epidemiology of anxiety disorders. *Am J Psychiatry* **158**, 1568-78 (2001).
7. Purves, K.L. *et al.* A major role for common genetic variation in anxiety disorders. *Mol Psychiatry* **25**, 3292-3303 (2020).
8. Hettema, J.M. *et al.* Genome-wide association study of shared liability to anxiety disorders in Army STARRS. *Am J Med Genet B Neuropsychiatr Genet* **183**, 197-207 (2020).
9. Li, W. *et al.* Genome-wide meta-analysis, functional genomics and integrative analyses implicate new risk genes and therapeutic targets for anxiety disorders. *Nat Hum Behav* (2023).
10. Levey, D.F. *et al.* Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *Am J Psychiatry* **177**, 223-232 (2020).
11. Sirugo, G., Williams, S.M. & Tishkoff, S.A. The Missing Diversity in Human Genetic Studies. *Cell* **177**, 26-31 (2019).
12. Meng, X. *et al.* Multi-ancestry genome-wide association study of major depression aids locus discovery, fine mapping, gene prioritization and causal inference. *Nat Genet* (2024).
13. Nievergelt, C.M. *et al.* Discovery of 95 PTSD loci provides insight into genetic architecture and neurobiology of trauma and stress-related disorders. *medRxiv* (2023).
14. Burstein, D. *et al.* Detecting and Adjusting for Hidden Biases due to Phenotype Misclassification in Genome-Wide Association Studies. *medRxiv* (2023).
15. Grotzinger, A.D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513-525 (2019).
16. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291-295 (2015).
17. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369-75, S1-3 (2012).
18. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
19. Moore, J.E. *et al.* Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**, 699-710 (2020).
20. Peterson, R.E. *et al.* Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell* **179**, 589-603 (2019).
21. Barbeira, A.N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* **9**, 1825 (2018).

22. Consortium, G. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318-1330 (2020).
23. Barbeira, A.N. *et al.* Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet* **15**, e1007889 (2019).
24. Gusev, A. *et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet* **48**, 245-52 (2016).
25. Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat Commun* **10**, 2417 (2019).
26. Werme, J., van der Sluis, S., Posthuma, D. & de Leeuw, C.A. An integrated framework for local genetic correlation analysis. *Nat Genet* **54**, 274-282 (2022).
27. Kurki, M.I. *et al.* FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**, 508-518 (2023).
28. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).
29. Verma, A. *et al.* Diversity and Scale: Genetic Architecture of 2,068 Traits in the VA Million Veteran Program. *medRxiv* (2023).
30. O'Connor, L.J. & Price, A.L. Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet* **50**, 1728-1734 (2018).
31. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* **51**, 237-244 (2019).
32. Karlsson Linnér, R. *et al.* Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat Neurosci* **24**, 1367-1376 (2021).
33. plee0@mgh.harvard.edu, C.-D.G.o.t.P.G.C.E.a. & Consortium, C.-D.G.o.t.P.G. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e11 (2019).
34. Otsuka, K. *et al.* A Testis-Specific Long Noncoding RNA,. *Front Endocrinol (Lausanne)* **12**, 665874 (2021).
35. Domonkos, E., Hodosy, J., Ostatníková, D. & Celec, P. On the Role of Testosterone in Anxiety-Like Behavior Across Life in Experimental Rodents. *Front Endocrinol (Lausanne)* **9**, 441 (2018).
36. Zuloaga, D.G., Heck, A.L., De Guzman, R.M. & Handa, R.J. Roles for androgens in mediating the sex differences of neuroendocrine and behavioral stress responses. *Biol Sex Differ* **11**, 44 (2020).
37. Wendt, F.R. *et al.* Sex-Specific Genetic and Transcriptomic Liability to Neuroticism. *Biol Psychiatry* **93**, 243-252 (2023).
38. Wendt, F.R. *et al.* Sex-stratified gene-by-environment genome-wide interaction study of trauma, posttraumatic-stress, and suicidality. *Neurobiol Stress* **14**, 100309 (2021).
39. Narita, A. *et al.* Clustering by phenotype and genome-wide association study in autism. *Transl Psychiatry* **10**, 290 (2020).
40. Shadrin, A.A. *et al.* Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. *Neuroimage* **244**, 118603 (2021).
41. Brouwers, M.C.G.J. Fructose 1-phosphate, an evolutionary signaling molecule of abundancy. *Trends Endocrinol Metab* **33**, 680-689 (2022).
42. Hyldgaard Andersen, S., Black, T., Grassi-Oliveira, R. & Wegener, G. Can early-life high fructose exposure induce long-term depression and anxiety-like behaviours? - A preclinical systematic review. *Brain Res* **1814**, 148427 (2023).
43. Pathak, G.A. *et al.* Genetically regulated multi-omics study for symptom clusters of posttraumatic stress disorder highlights pleiotropy with hematologic and cardio-metabolic traits. *Mol Psychiatry* **27**, 1394-1404 (2022).
44. Elia, L.P., Yamamoto, M., Zang, K. & Reichardt, L.F. p120 catenin regulates dendritic spine and synapse development through Rho-family GTPases and cadherins. *Neuron* **51**, 43-56 (2006).

45. Ryu, T. *et al.* Improved memory and reduced anxiety in δ -catenin transgenic mice. *Exp Neurol* **318**, 22-31 (2019).
46. Li, X. *et al.* Case Report: CNNM2 Mutations Cause Damaged Brain Development and Intractable Epilepsy in a Patient Without Hypomagnesemia. *Front Genet* **12**, 705734 (2021).
47. Deng, Y.T. *et al.* Identifying causal genes for depression via integration of the proteome and transcriptome from brain and blood. *Mol Psychiatry* **27**, 2849-2857 (2022).
48. Martinez, C.A. *et al.* Prenatal stress, anxiety and depression alter transcripts, proteins and pathways associated with immune responses at the maternal-fetal interface†. *Biol Reprod* **106**, 449-462 (2022).
49. Al-Mubarak, B.R. *et al.* Whole exome sequencing in ADHD trios from single and multi-incident families implicates new candidate genes and highlights polygenic transmission. *Eur J Hum Genet* **28**, 1098-1110 (2020).
50. Shang, M.Y. *et al.* Genetic associations between bipolar disorder and brain structural phenotypes. *Cereb Cortex* **33**, 6990-7000 (2023).
51. Gedik, H. *et al.* Identifying potential risk genes and pathways for neuropsychiatric and substance use disorders using intermediate molecular mediator information. *Front Genet* **14**, 1191264 (2023).
52. Meyers, T.J. *et al.* Transcriptome-wide association study identifies novel candidate susceptibility genes for migraine. *HGG Adv* **4**, 100211 (2023).
53. Wendt, F.R. *et al.* Characterizing the effect of background selection on the polygenicity of brain-related traits. *Genomics* **113**, 111-119 (2021).
54. Wendt, F.R., Pathak, G.A., Tylee, D.S., Goswami, A. & Polimanti, R. Heterogeneity and Polygenicity in Psychiatric Disorders: A Genome-Wide Perspective. *Chronic Stress (Thousand Oaks)* **4**, 2470547020924844 (2020).
55. Marshall, M. The hidden links between mental disorders. *Nature* **581**, 19-21 (2020).
56. Huang, Y.C. *et al.* Anxiety comorbidities in patients with major depressive disorder: the role of attachment. *Int J Psychiatry Clin Pract* **23**, 286-292 (2019).
57. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res* **51**, D977-D985 (2023).
58. Ike, K.G.O. *et al.* The human neuropsychiatric risk gene *Drd2* is necessary for social functioning across evolutionary distant species. *Mol Psychiatry* (2023).
59. Cutando, L. *et al.* Cerebellar dopamine D2 receptors regulate social behaviors. *Nat Neurosci* **25**, 900-911 (2022).
60. Avramidou, M. *et al.* Epidemiology of gastrointestinal symptoms in young and middle-aged Swiss adults: prevalences and comorbidities in a longitudinal population cohort over 28 years. *BMC Gastroenterol* **18**, 21 (2018).
61. Eijsbouts, C. *et al.* Genome-wide analysis of 53,400 people with irritable bowel syndrome highlights shared genetic pathways with mood and anxiety disorders. *Nat Genet* **53**, 1543-1552 (2021).
62. Kremer, M., Becker, L.J., Barrot, M. & Yalcin, I. How to study anxiety and depression in rodent models of chronic pain? *Eur J Neurosci* **53**, 236-270 (2021).
63. Gómez Penedo, J.M. *et al.* The Complex Interplay of Pain, Depression, and Anxiety Symptoms in Patients With Chronic Pain: A Network Approach. *Clin J Pain* **36**, 249-259 (2020).
64. Zanoaga, M.D. *et al.* Brainwide Mendelian Randomization Study of Anxiety Disorders and Symptoms. *Biol Psychiatry* (2023).
65. Celano, C.M., Daunis, D.J., Lokko, H.N., Campbell, K.A. & Huffman, J.C. Anxiety Disorders and Cardiovascular Disease. *Curr Psychiatry Rep* **18**, 101 (2016).
66. Murphy, B., Le Grande, M., Alvarenga, M., Worcester, M. & Jackson, A. Anxiety and Depression After a Cardiac Event: Prevalence and Predictors. *Front Psychol* **10**, 3010 (2019).
67. (!!! INVALID CITATION !!! 25).

68. Denny, J.C. *et al.* The "All of Us" Research Program. *N Engl J Med* **381**, 668-676 (2019).
69. Kurki, M.I. *et al.* FinnGen: Unique genetic insights from combining isolated population and national health register data. *medRxiv*, 2022.03.03.22271360 (2022).
70. Meier, S.M. *et al.* Genetic Variants Associated With Anxiety and Stress-Related Disorders: A Genome-Wide Association Study and Mouse-Model Study. *JAMA Psychiatry* **76**, 924-932 (2019).
71. Otowa, T. *et al.* Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry* **21**, 1391-9 (2016).
72. Chang, E. & Mostafa, J. The use of SNOMED CT, 2013-2020: a literature review. *J Am Med Inform Assoc* **28**, 2017-2026 (2021).
73. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
74. Mbatchou, J. *et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet* **53**, 1097-1103 (2021).
75. Pedersen, C.B. *et al.* The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol Psychiatry* **23**, 6-14 (2018).
76. Gaziano, J.M. *et al.* Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol* **70**, 214-23 (2016).
77. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet* **50**, 1335-1341 (2018).
78. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
79. Altshuler, D.M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52-8 (2010).
80. Grotzinger, A.D., Fuente, J., Privé, F., Nivard, M.G. & Tucker-Drob, E.M. Pervasive Downward Bias in Estimates of Liability-Scale Heritability in Genome-wide Association Study Meta-analysis: A Simple Solution. *Biol Psychiatry* **93**, 29-36 (2023).
81. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
82. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C.A. & Smoller, J.W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications* **10**, 1776 (2019).
83. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
84. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
85. Fehrmann, R.S. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet* **47**, 115-25 (2015).
86. Beach, T.G. *et al.* Arizona Study of Aging and Neurodegenerative Disorders and Brain and Body Donation Program. *Neuropathology* **35**, 354-89 (2015).
87. Beach, T.G. *et al.* The Sun Health Research Institute Brain Donation Program: description and experience, 1987-2007. *Cell Tissue Bank* **9**, 229-45 (2008).
88. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217-234.e4 (2019).
89. Demontis, D. *et al.* Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nat Genet* **55**, 198-208 (2023).
90. Watson, H.J. *et al.* Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet* **51**, 1207-1214 (2019).

91. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* **51**, 431-444 (2019).
92. Mullins, N. *et al.* Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet* **53**, 817-829 (2021).
93. Levey, D.F. *et al.* Bi-ancestral depression GWAS in the Million Veteran Program and meta-analysis in >1.2 million individuals highlight new therapeutic directions. *Nat Neurosci* **24**, 954-963 (2021).
94. Yu, D. *et al.* Interrogating the Genetic Determinants of Tourette's Syndrome and Other Tic Disorders Through Genome-Wide Association Studies. *Am J Psychiatry* **176**, 217-227 (2019).
95. Stein, M.B. *et al.* Genome-wide association analyses of post-traumatic stress disorder and its symptom subdomains in the Million Veteran Program. *Nat Genet* **53**, 174-184 (2021).
96. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).

Table 1. Cohorts and Anxiety Definitions investigated. Population groups included African (AFR), Admixed-American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS) descent. Case and control sample sizes are reported for binary anxiety definitions, while the total sample size is reported for quantitative anxiety phenotypes. Other abbreviations include International Classification of Diseases (ICD) and Systemized Nomenclature of Medicine (SNOMED).

Cohort	Population	Sample Size	Anxiety Definition
All of Us Research Program (AoU)	AFR	4,937 - 45,092	Generalized Anxiety Disorder (SNOMED ID: 434613), Panic Disorder (SNOMED ID: 4138454), Phobic Disorder (SNOMED ID: 4304010), Chronic Stress Disorder (SNOMED ID: 436074), Acute Stress Disorder (SNOMED ID: 440083), Posttraumatic Stress Disorder (SNOMED ID: 436676), Obsessive-compulsive Disorder (SNOMED ID: 440374), or Mixed Anxiety and Depressive Disorder (SNOMED ID: 4338031)
	AMR	4,157 - 32,477	
	EAS	337 - 4,746	
	EUR	18,461 - 88,954	
	SAS	133 - 2,040	
FinnGen Study	EUR	40,191 - 277,526	KRA_PSY_ANXIETY endpoint
Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH)	EUR	12,655 - 19,225	ICD-10 F40.0-F41.9; F43.0-F43.9
Million Veteran Program (MVP)	AFR	61,796	Generalized Anxiety Disorder 2-item scale
	EUR	241,541	
Psychiatric Genomics Consortium (PGC)	EUR	5,761 - 11,765	Generalized anxiety disorder, panic disorder, social phobia, agoraphobia, or specific phobias
UK Biobank (UKB)	AFR	130 - 6,116	Phecode 300, "Anxiety, dissociative and somatoform disorders"
	EUR	10,449 - 369,930	
	SAS	172 - 8,189	

FIGURES

Figure 1. Genetic correlations among anxiety phenotypes assessed in participants of European descent (EUR). Full statistics are reported in Supplemental Table 1.

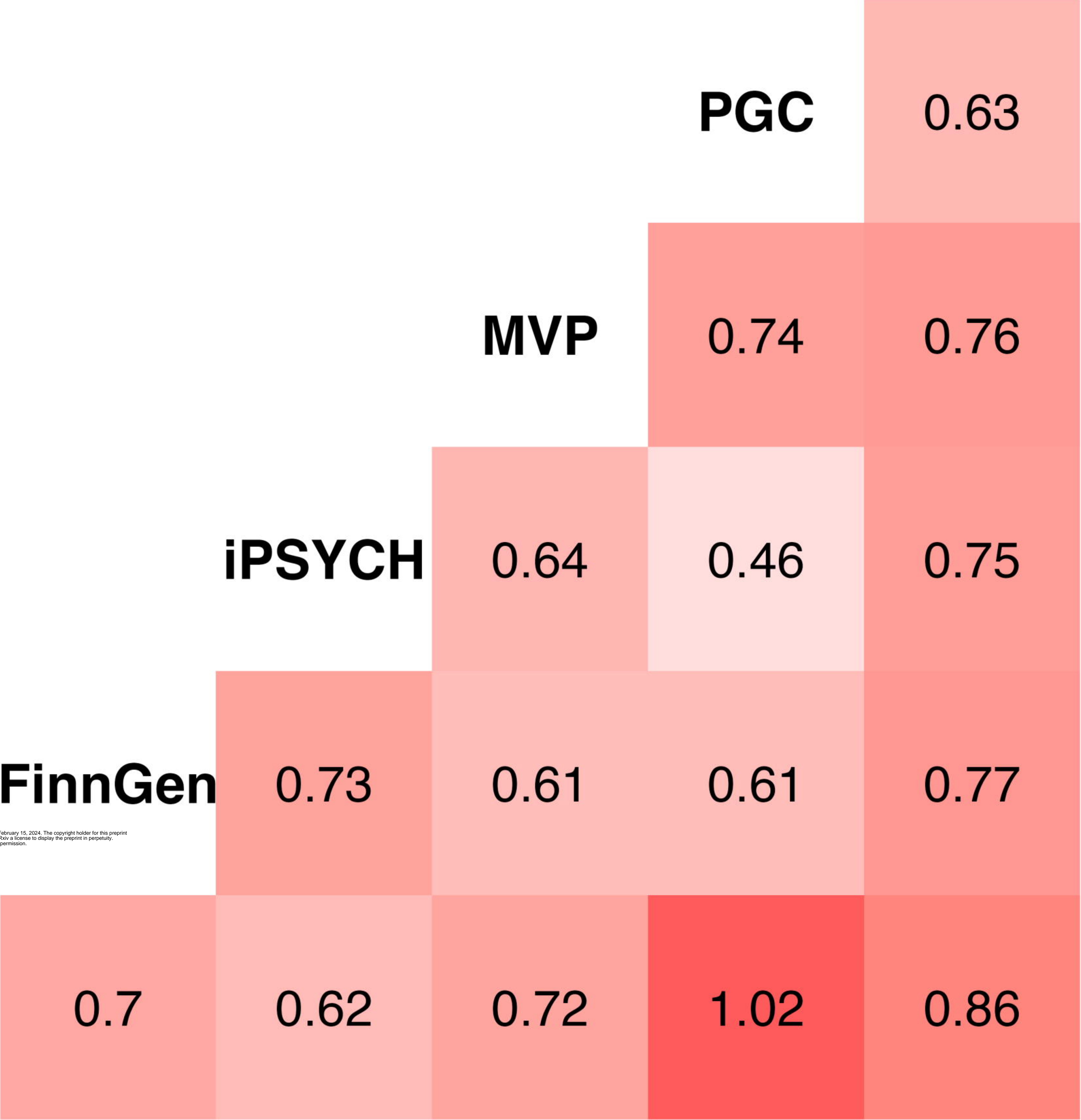
Figure 2. Factor structure of anxiety phenotypes assessed in individuals of European descent. Factor loadings and model fit for the confirmatory factor analysis model of the common anxiety factor (ANX) are reported.

Figure 3. Manhattan plots of genome-wide, transcriptome-wide, and proteome-wide association statistics (bottom, center, and top, respectively) related to the common anxiety factor (ANX). Transcriptome-wide data are those obtained from the multiple-tissue analysis (Supplemental Table 8). Dashed lines represent Bonferroni multiple testing correction applied in each analysis. The labeled genes are those with convergent evidence across analyses.

Figure 4. Within-ancestry and cross-ancestry polygenic risk score (PRS) associations of the common anxiety factor (ANX). Full statistics are available in Supplemental Table 4. In the within-ancestry, the EUR-ANX PRS (excluding EUR AoU from the training sample) is tested against the EUR AoU sample. In the cross-ancestry analysis, EUR-ANX PRSs (with and without EUR AoU in the training sample) are tested against AoU samples of African (AoU), Admixed-American (AMR), East Asian (EAS), and South Asian (SAS) descent.

Figure 5 A. Venn diagrams of unique and shared causal variants showing polygenic overlap (gray) of the common anxiety factor (ANX) with major depressive disorder (MDD), bipolar disorder (BIP), and schizophrenia (SCZ). The numbers indicate the estimates of causal variants and their standard errors (in thousands), explaining 90% of the SNP heritability in each phenotype. The size of the circles reflects the degree of polygenicity. **B.** Chord diagram representing ANX local genetic correlations with respect to attention deficit and hyperactivity disorder (ADHD), BIP, MDD, post-traumatic stress disorder (PTSD), and SCZ.

Figure 6. Phenome-wide genetic correlation of the common anxiety factor (ANX). The x-axis reports the genetic correlation of ANX with traits available from UK Biobank, FinnGen, and Million Veteran Program. The y-axis corresponds to two-tailed $-\log_{10}(\text{p-value})$. Bonferroni-significant results are color-coded based on the corresponding categories. Full statistics of the Bonferroni-significant results are available in Supplemental Table 14. Non-significant results are reported in grey.



rG

1.0

0.9

0.8

0.7

0.6

0.5

FinnGen

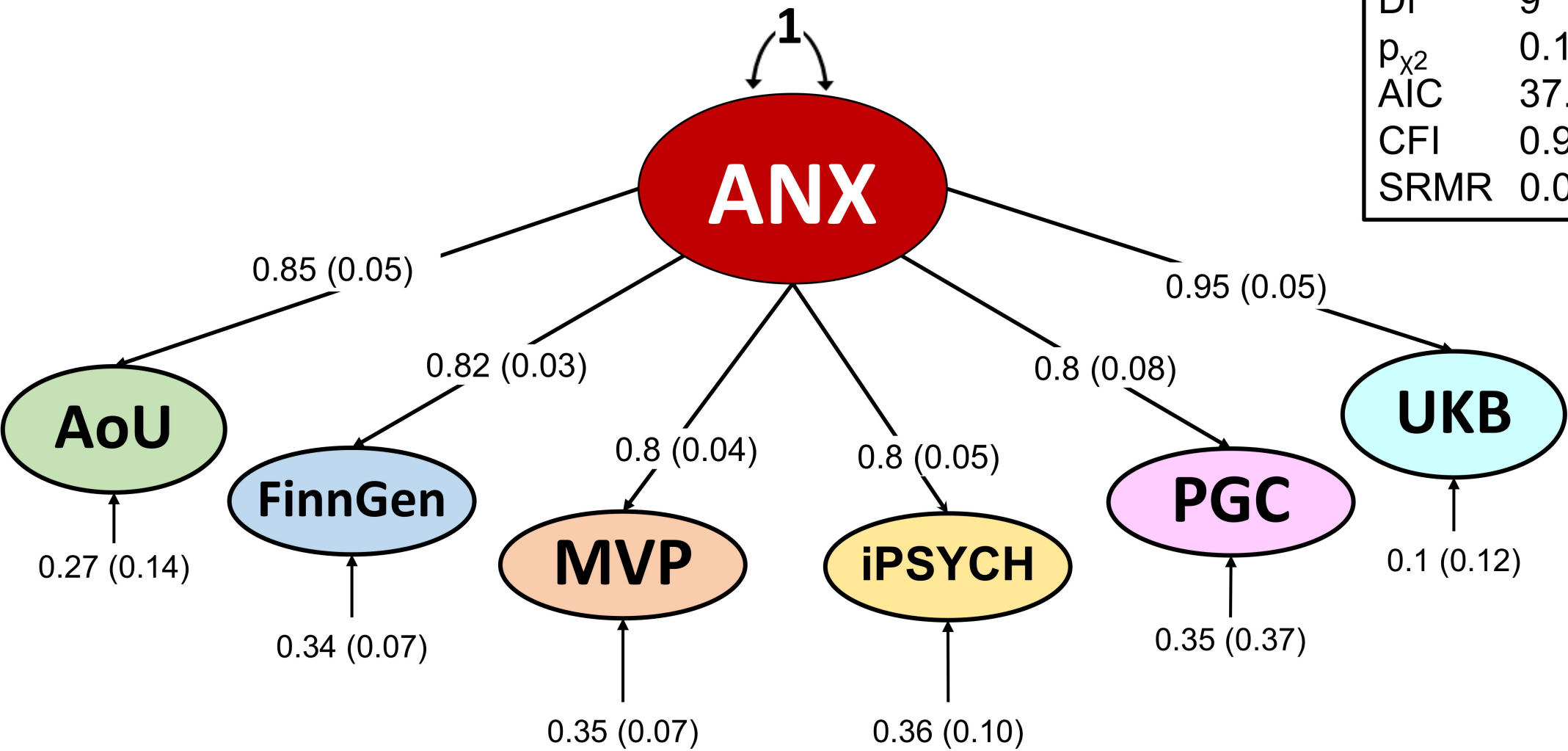
iPSYCH

MVP

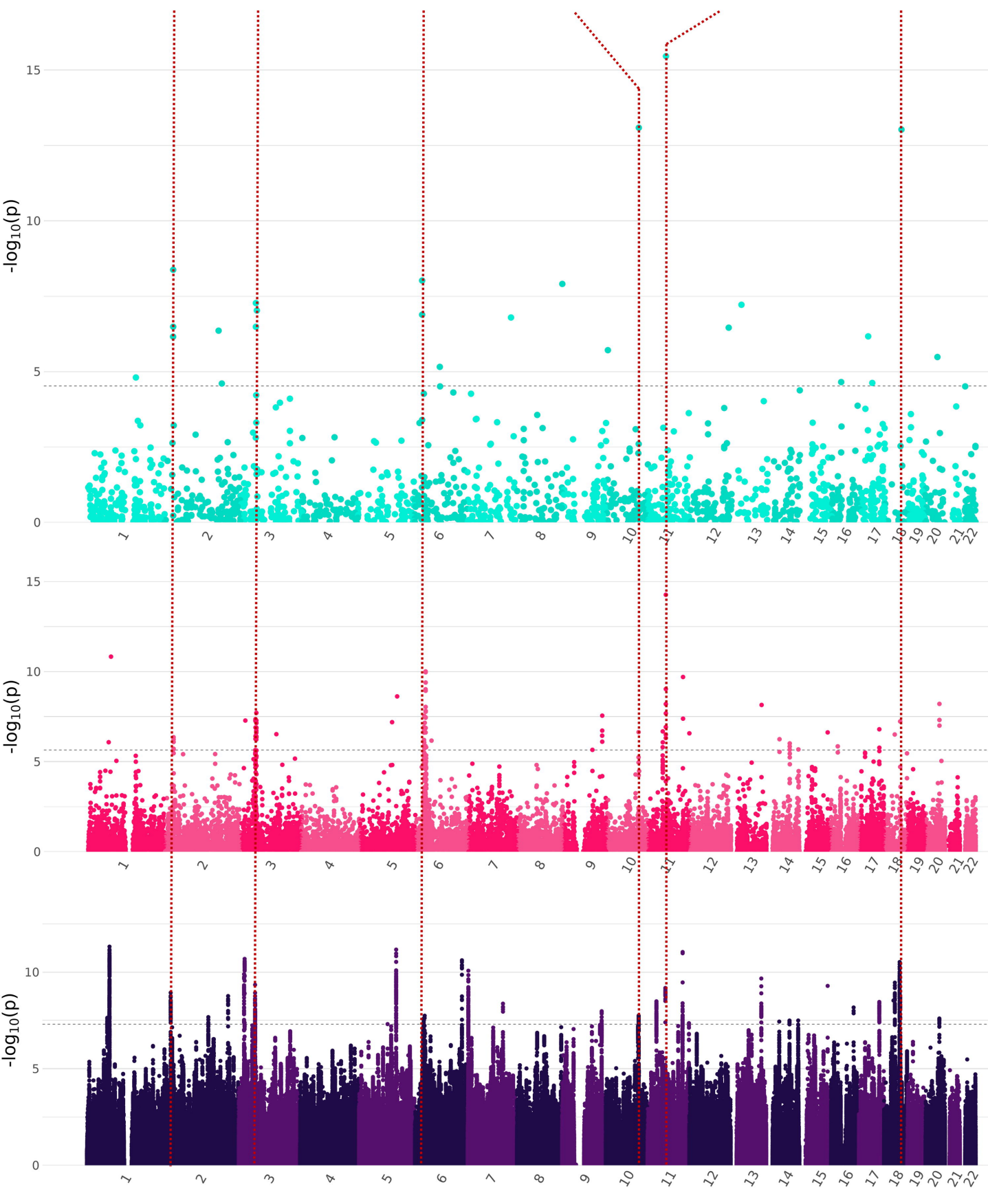
PGC

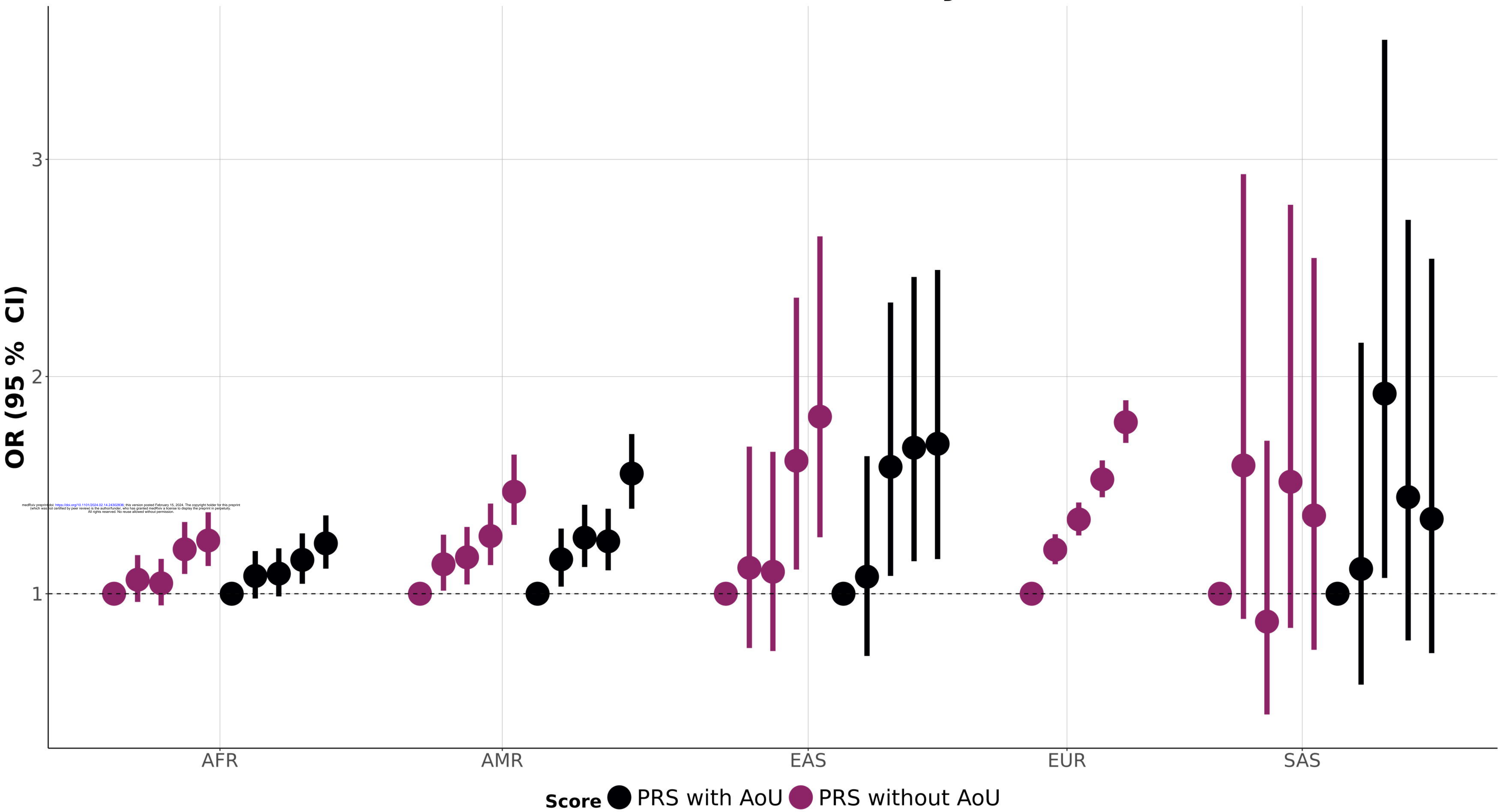
UKB

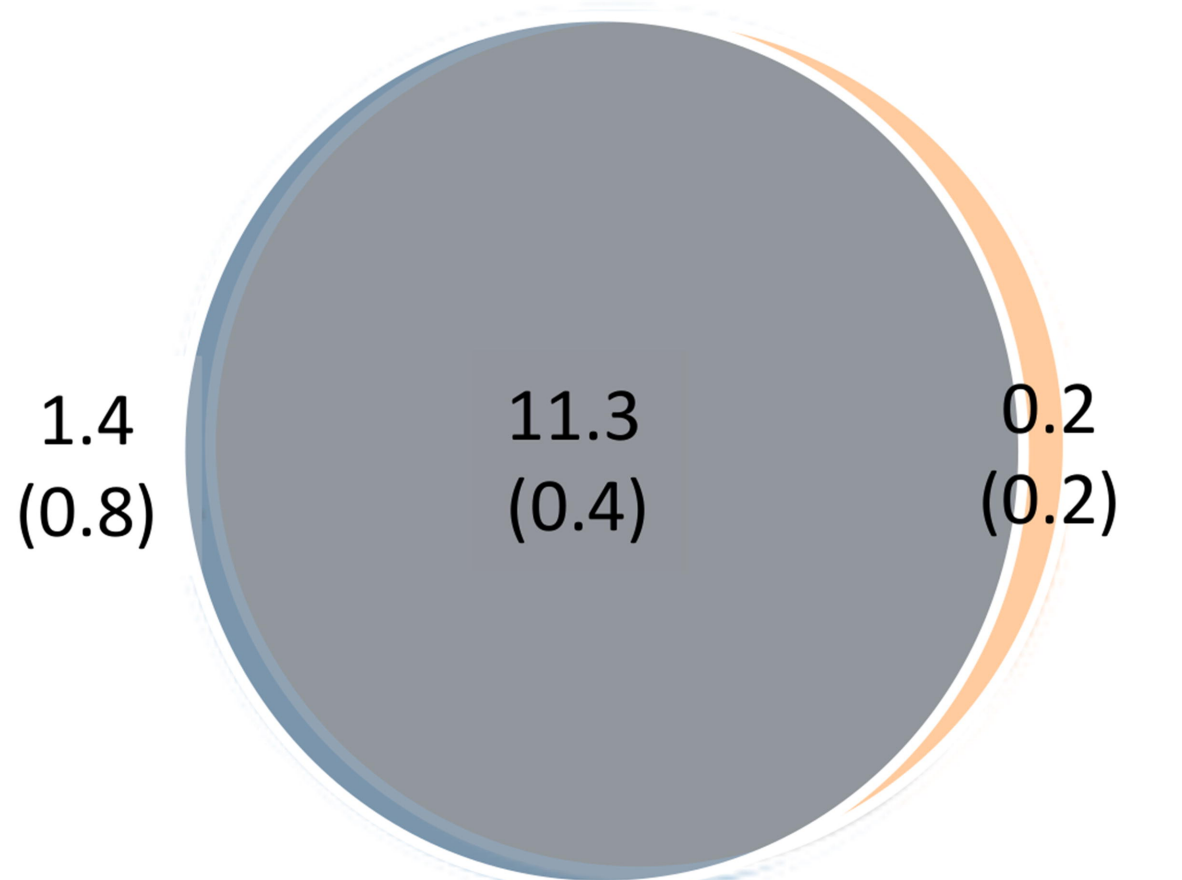
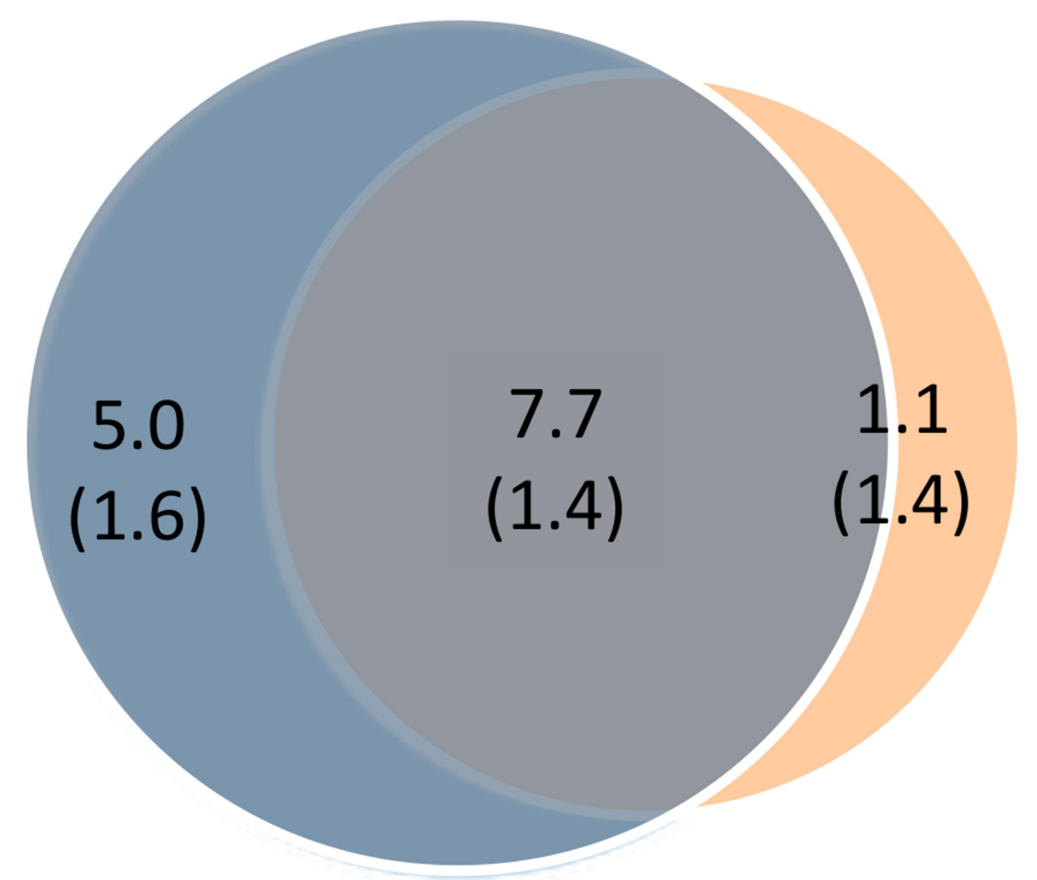
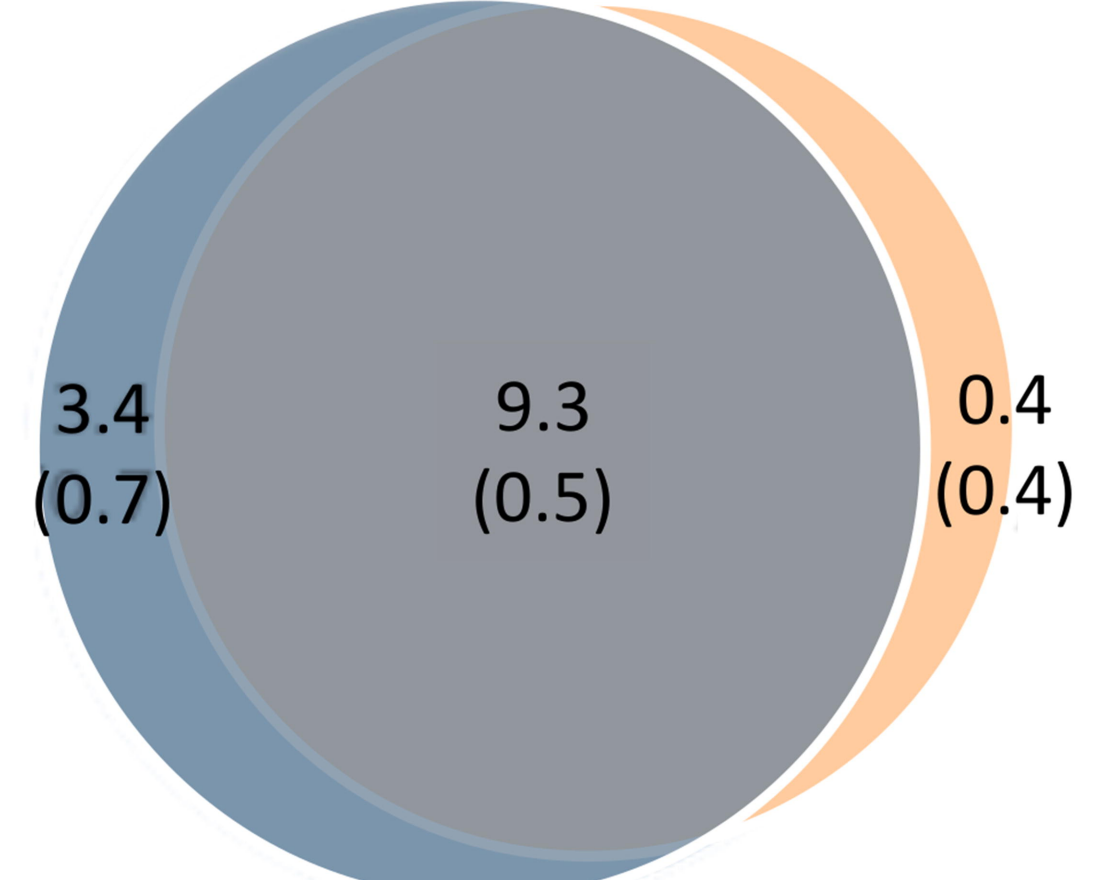
χ^2	13.35
Df	9
p_{χ^2}	0.15
AIC	37.35
CFI	0.996
SRMR	0.080



KHK **NEK4** **BTN2A1** **CNNM2** **CTNND1** **RAB27B**





A**ANX-MDD****ANX-BIP****ANX-SCZ****B**