Multi-level genomic analysis identifies pleiotropic genetic liability for insomnia, anxiety and depression

Marijn Schipper[1], Alexey A. Shadrin[2], Cato Romero[3], Eleni Friligkou[4,5], Renato Polimanti[4,5,6,7], Danielle Posthuma[1,8], Eus J.W. Van Someren[9*], Elleke Tissink[9*]

[*] Contributed equally to this work

[1] Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam 1081 HV, The Netherlands

[2] Centre for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[3] Department of Child and Adolescent Psychiatry and Psychology, Section Complex Trait Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam UMC, Amsterdam, the Netherlands

[4] Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA

[5] VA CT Healthcare Center, West Haven, CT, USA

[6] Wu Tsai Institute, Yale University, New Haven, CT, USA

[7] Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA

[8] Department of Clinical Genetics, Section Complex Trait Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam University Medical Centre, Amsterdam 1081 HZ, The Netherlands

[9] Department of Sleep and Cognition, Netherlands Institute for Neuroscience, an institute of the Royal Netherlands Academy of Arts and Sciences, The Netherlands

✉ Correspondence: e.tissink@nin.knaw.nl

Abstract

Insomnia (INS), major depression (DEP) and anxiety disorders (ANX) frequently co-occur and share a substantial proportion of their genetic risk. Most genetic studies have focused on shared risk between DEP and ANX, often omitting INS. Other studies that include INS have been limited to genome-wide correlations or overlap. Identifying the specific genetic factors that influence these three conditions could offer a route to identify therapeutic targets with potential transdiagnostic benefits. Realizing this potential requires pinpointing the likely causal variants and genes that are shared, estimating how pleiotropic effects are mediated, as well as identifying the biological processes they affect. Here we conduct a multi-level trivariate genetic analysis of INS, DEP and ANX using genome-wide association studies (GWAS) of more than one million individuals from multiple ancestries. We show that 55% of the genetic signal is shared across all three conditions. Supporting its importance, INS shares a significant fraction of the genetic overlap (75%) and correlation (7%) between DEP and ANX. We identify 195 genomic loci with shared signal for at least two conditions, many of which are likely arising from the same causal variant (at least 50%) or effector gene (60-80%). Pairwise mediation analyses suggest that these shared likely causal variants are more consistent with models of vertical pleiotropy where DEP and INS are risk factors towards ANX, rather than with models of horizontal pleiotropy. We find convergence of shared effector genes on biological processes in inhibitory synaptic transmission, neuronal organization, and axonal development. These results reveal an interconnected but mechanistically diverse basis for shared genetic risk across INS, DEP and ANX, and offer potential candidates for future pathway-tailored therapeutic targets with transdiagnostic benefits.

Introduction

One-third of the general population has difficulties falling or staying asleep, which can severely impact quality of life when manifesting frequently and persistently as insomnia disorder (INS)[1]. INS may develop as a symptom of another condition, such as depression (DEP) or anxiety disorders (ANX), or can alternatively precede and increase the risk of developing these conditions[2,3]. INS, DEP and ANX are the most prevalent mental health conditions with 12-month prevalences of 7%, 6.9%, and 14% respectively[4], and show moderate response rates to existing treatment options[5]. High rates of lifetime multimorbidity[6], transdiagnostic intervention efficacy[7], and a shared neurobiological vulnerability circuit[8] suggest a common aetiology underlying INS, DEP, and ANX.

Shared genetic association between these disorders can be a statistical byproduct of comorbidity[9], phenotype misclassification[10], or can point to real biological shared mechanisms, partially explaining the observed patterns of convergence between these phenotypes. Despite the potential of identifying shared biological mechanisms underlying genetic liability, the molecular genetic signal specifically shared across the triad of INS, DEP, and ANX has not been studied systematically to date. Understanding these shared genetic factors may help find common biological mechanisms and guide the development of transdiagnostic treatments that target these shared vulnerabilities, potentially improving outcomes across multiple conditions simultaneously[11,12].

Based on decades of twin research[13], it is widely recognized that the heritability of INS (39%[14]), DEP (37%[15]), ANX (32%[16]) is largely shared (>50% for DEP, complete for ANX). More recently, the molecular genetic signals of INS[17], ANX[18,19], and DEP[20,21] have been intensively investigated in genome-wide association studies (GWAS). These studies confirm a similar genetic architecture across the three conditions based on common genetic variants, i.e. single nucleotide polymorphisms (SNPs). That is, the phenotypic variance in INS, ANX, or DEP that can be explained by additive effects of SNPs (i.e. observed SNP-heritability, h^2_{SNP}) is modest (6-7%[21–23]), and can be captured by ~12,000 putatively causal SNPs (i.e. polygenicity) with fairly low effect sizes (i.e. discoverability) compared to other psychiatric disorders[24].

The degree to which these SNPs overlap and their effect sizes correlate (i.e. genetic overlap and genetic correlation [r_g] respectively) has been well-documented to be extensive for DEP and ANX (r_g = .87[25], and 85-95% of putative causal SNPs shared, with 82-91% sign concordance[18,21,26]), though recent disorder-specific analyses suggest that this correlation may be inflated by comorbid cases[27]. Comparisons with INS are often omitted[25,28,29]. Few exceptions include two bivariate studies reporting the genetic overlap between DEP and INS to be 65%[30,31], and one cross-disorder study estimating the genetic correlations between INS with DEP and ANX to be r_g = .35 and .45 respectively[24]. Combined analysis of all three phenotypes in a trivariate approach has not been performed previously, but has the potential to

distinguish the portion of overlap shared between all three disorders, and the proportion of overlap unique to DEP and INS, ANX and INS, and ANX and DEP[32].

Pinpointing where in the genome effects between genetically correlated traits co-occur, and onto which biological mechanisms these effects converge, is necessary to interpret the shared genetic aetiology underlying correlated traits. This shared locus discovery has been challenging in the past due to insufficient statistical power[24]. Fortunately, with sample sizes growing exponentially in psychiatric genetics studies[33], we can now leverage the recently published GWAS of INS[22], ANX[18], and DEP[20,21] based on millions of individuals to conduct well-powered local genetic correlation analyses.

Observing non-zero (local) genetic correlations between INS, DEP, and ANX would indicate the presence of pleiotropy[34–36]: the statistical association of one SNP or gene with multiple traits independently (i.e. biological or horizontal pleiotropy) or through a causal pathway (i.e. mediation or vertical pleiotropy). Thorough follow-up analyses are necessary to distinguish biological or mediation pleiotropy from spurious pleiotropy after identifying shared loci[37]. Spurious pleiotropy can arise when distinct causal variants or genes within the same locus are correlated due to linkage disequilibrium (LD), creating the false impression that a single genetic factor influences multiple traits[37]. To address this, colocalization methods can test whether it is likely that the same variant underlies the observed association for both traits[38]. A recently developed framework further extends pleiotropy analyses to the gene level[39], allowing the identification of shared effector genes and enabling us to improve the biological interpretability of genetic overlap between INS, DEP, and ANX.

Mediation analyses can subsequently be used to discriminate between vertical and horizontal pleiotropy and inform us how shared genetic factors are associated to differential diagnoses such as INS, DEP, and ANX[40]. Previous mediation studies have primarily relied on trait-level causal inference using genetic instruments (i.e. Mendelian Randomization), and report bidirectionality between DEP and INS[41], and a shift from a bidirectional to a unidirectional effect from ANX to DEP once comorbid cases are excluded[27]. A recently developed method[42] now extends trait-level inference to the variant level, classifying SNPs as likely confounding, horizontal or vertical pleiotropic. Observing in which direction potential vertical pleiotropic variants act would refine previous reports[27,41], since trait-level bidirectionality is most plausibly understood as a mixture of unidirectional variant-level effects.

Here, we leverage recent GWAS summary statistics of INS, DEP, and ANX together with recently developed statistical genetics tools to hierarchically investigate the genetic overlap across these conditions on the genome-wide, locus, gene and variant level. Specifically, we estimate the trivariate genetic overlap to be 55%, using a new extension of the MiXeR model[32], and identify 195 unique loci of shared genome-wide significance or correlating genetic signal using LAVA[43], for at least two traits. We show with colocalization[38] and FLAMES[39] that at least 50% and 60-80% of these loci also share the same

likely causal variant or effector gene respectively, and observe with PRISM[42] that these variants are most compatible with models of INS and DEP having bidirectional effects and mediating effects on ANX. We use these shared effector genes to search for convergence of implicated biology across ANX, DEP and INS and find their almost exclusive involvement in inhibitory synaptic transmission and organization, and neuronal and axonal development. Altogether, this study shows that many genomic regions associated with ANX, DEP and INS map to the same likely causal variants showing effects consistent with vertical pleiotropy, and map to shared genes and converge, even when acting through distinct genes, on shared biological pathways.

Results

Our analyses are based on previously published GWAS summary statistics for ANX[18], DEP[20,21] and INS[22], (Supplementary Table 1). The original studies had an effective sample size of N_{INS} = 1,114,549, N_{ANX} = 529,764, N_{DEP} = 1,061,269 based on individuals of European (EUR) and African (AFR; N_{ANX} = 80,104; N_{DEP} = 132,002) genetic ancestry, and also performed cross-ancestry meta-analyses (N_{ANX} = 627,041; N_{DEP} = 1,118,664). We obtained access through public and restricted data access protocols (see Methods). With these data, we investigated the genetic relationship between INS, ANX and DEP on a genome-wide, locus, and causal gene/variant level.

Trivariate genetic overlap on the genome-wide level

We modeled genome-wide genetic overlap across three traits simultaneously using Trivariate MiXeR[32], a recently developed extension of the established univariate and bivariate Gaussian mixture models[44,45]. When applied on EUR GWAS summary statistics, we observed a genetic overlap of 55% across all three traits (Figure 1a). The genetic signal of INS overlapping with DEP and with ANX was of comparable size (78% and 79% respectively) and concerned the same SNPs, since 98% of DEP overlapped with ANX (Supplementary Table 2). However, the correlation of effect sizes within the overlap was stronger between INS and ANX ($r_{\text{g-shared}}$ = .85, SD =.09) than between INS and DEP ($r_{\text{g-shared}}$ = .63, SD = .04). As expected based on previous studies[18,21,25,26], the strongest correlation of effect sizes was observed for the overlapping SNPs between DEP and ANX ($r_{\text{g-shared}}$ = .94, SD = .04). Notably, 21% of INS variants were not shared with ANX or DEP. While ANX also contained a unique genetic component, this was accompanied by larger polygenicity (number of non-null SNPs; π_{ANX} = 14,463; π_{DEP} = 12,474, π_{INS} = 12,180; Supplementary Table 2).

These results deepen findings from standard global genetic correlations (r_g) estimated with LD Score regression[46] (LDSC; $r_{g_{\text{INS-DEP}}}$ = .63 [SE = .01], $r_{g_{\text{INS-ANX}}}$ = 0.63 [SE = .02]; $r_{g_{\text{DEP-ANX}}}$ = .91 [SE =

.01]). Given the considerable trivariate genetic overlap that we observed, we aimed to quantify the contribution of each trait in explaining the r_g between the other two conditions. (Figure 1b). Using a novel LDSC extension, partial LDSC[47], we find that accounting for the SNP effects of INS significantly reduces the genetic correlation between DEP and ANX to $r_{g_{DEP\text{-}ANX\,|\,INS}}$ = .87 (SE = .04, T_{diff} = 3.27, p_{diff} = 1.07×10^{-3}). As expected from the strong correlation between DEP and ANX, conditioning the DEP-INS correlation on ANX and the ANX-INS correlation on DEP greatly reduces the found genetic correlation ($r_{g_{INS\text{-}ANX\,|\,DEP}}$ = .17 [SE = .08]; $r_{g_{INS\text{-}DEP\,|\,ANX}}$ = .16 [SE = .08]). Altogether, these genome-wide results show that a significant proportion of the genetic overlap between ANX and DEP also extends to INS, whilst INS also contains a distinct genetic component.

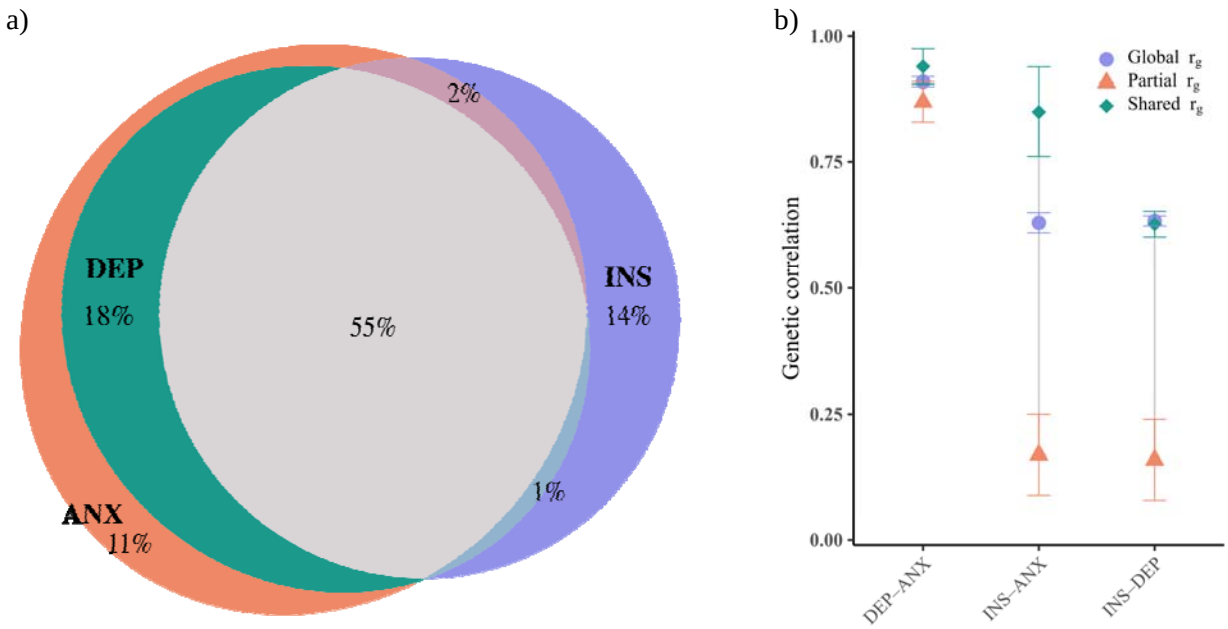

Figure 1. **The genome-wide genetic overlap of INS, ANX and DEP.** a) Trivariate overlap estimated with MiX3R (Supplementary Table 2). Each circle (i.e. trait) has a size proportional to the number of trait-influencing variants (i.e. polygenicity) and overlapping areas proportional to the number of variants shared among respective traits. Percentages in the diagram are given with respect to the combined total area of all three traits (and sum to 101% due to rounding). b) Genetic correlations based on LDSC (global r_g with SE) and Trivariate MiXeR (shared r_g with SD). The partial r_g (with SE) between two traits depends on conditioning on the third trait.

Identifying pleiotropic loci across conditions

We then aimed to move beyond global patterns and identify the genomic loci with shared genetic signal. First, we investigated which genome-wide significant loci ($p < 5 \times10^{-8}$) were physically overlapping across traits. The EUR GWAS summary statistics included 248, 235, and 40 genome-wide significant

(GWS) loci for ANX[18], DEP[21] and INS[22] respectively, based on the FUMA definition[48]. Cross-ancestry GWAS summary statistics of ANX[18] and DEP[20] included 10 and 47 genome-wide significant loci, respectively, that were not identified in the EUR-only analyses. One locus was GWS in the AFR summary statistics for ANX[18], none for DEP[20]. This totals 580 GWS loci in our three traits across these ancestries.

Thirty-two percent of these loci were overlapping with at least one locus of another phenotype in the triad, spanning 87 unique locations in the genome. Of all EUR and cross-ancestry loci (Supplementary Table 3), nine overlapped across all three phenotypes, 49 physically overlapped between INS and DEP, 26 between DEP and ANX, and three between INS and ANX (Figure 2; Supplementary Table 4). The GWS locus of ANX in the AFR sample did not overlap with any locus of DEP or INS.

Second, we explored regions that had significant univariate signal and bivariate local genetic correlations. We applied LAVA[43] to perform local genetic correlation analysis to the EUR summary statistics. The non-overlapping genomic regions (n = 2,495; defined in Methods) were first screened for sufficient univariate genetic signal in each trait (local observed h^2_{SNP}; 582 regions survived $p_{univ} < 1\times10^{-4}$, a threshold used in previous work[49–51]) before bivariate genetic correlations between INS and ANX, INS and DEP, and/or ANX and DEP were estimated. Local r_g estimates were significant after correcting for multiple testing ($p_{bivar} < .05 / 582 = 8.59 \times10^{-5}$) in ten EUR loci for all three trait pairs (Supplementary Table 5-6). Seventeen loci were significant for INS-DEP as well as DEP-ANX, and 3 loci showed significant local r_g between both INS-ANX and DEP-ANX. Most loci were only significant for one trait pair: 46 between INS and DEP, five between INS and ANX, and 27 between ANX and DEP.

Of all 108 unique EUR LAVA-identified loci, n = 39 loci physically overlapped with any shared GWS locus described above. These LAVA regions were almost always identified for the same trait pair or for the same and additional trait pairs as the GWS locus (Supplementary Table 6). In total, 195 loci were identified to be shared across at least one trait pair (n_{GWS} = 87, n_{LAVA} = 108), which could be reduced to 146 unique locations given the overlap across approaches. We continued our downstream analyses with all 195 loci with cross-disorder evidence.

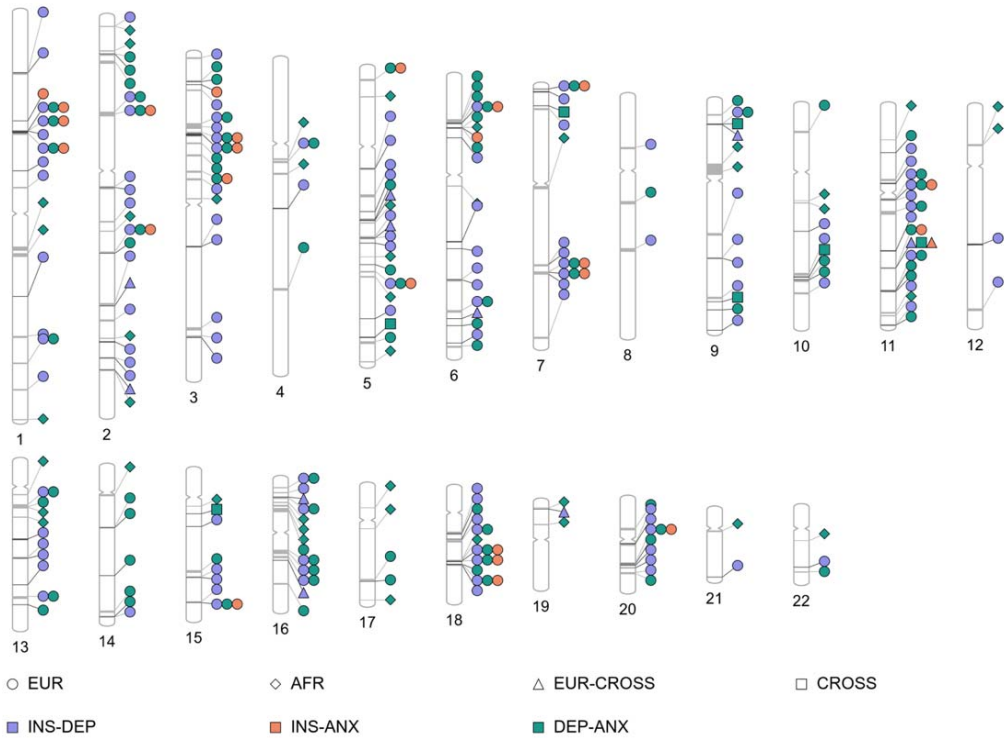

Figure 2. **Loci with shared genetic signal across ANX, DEP and INS.** Ideogram with each chromosome showing the location of genome-wide significant loci (dark grey lines) that were physically overlapping across INS, DEP and ANX GWAS summary statistics (both EUR or cross-ancestry, or EUR for INS and cross-ancestry for DEP/ANX). Markers connected by grey lines had significant correlating GWAS signal across trait pairs (i.e. identified through LAVA). See Supplementary Table 3-6 for the accompanying statistics.

Identifying putative causal variants

Genetic loci with correlated association patterns do not necessarily share a causal signal, but could also be caused by phenotypic correlation of non-genetic origin or emerge from distinct causal signals in linkage disequilibrium[52]. We aimed to disentangle spurious pleiotropy from true pleiotropy by filtering each of the 195 loci identified above for having an observed signal consistent with a shared causal variant. For this purpose, we ran colocalization analysis with COLOC[38] on the GWS and LAVA significant loci for each relevant trait pair. These loci colocalized to the same variants in 50% of locus pairs. Thirteen loci showed colocalization evidence across all three traits (Table 1).

The percentage of significant colocalization is higher (62%) in the GWS loci than in the LAVA loci (42%), which is a significant difference (Fisher's exact test, $p = 1.5\times10^{-3}$). This discrepancy can be explained by (1) LAVA loci generally being larger than GWS loci allowing for different signals residing within the same region, and/or (2) a lack of power. All our LAVA loci were significant in the univariate test of local heritability ($p_{univ} < 1\times10^{-4}$), yet for these loci, the COLOC posterior probability for at least

one trait having no signal for a causal variant in the locus is, on average, 34% compared to 7% for GWS significant loci. Of the remaining posterior probability, on average, 64% of the probability points to colocalization, whereas 36% of the probability points to non-colocalizing signal for both traits. These probabilities are highly similar (65% and 35% respectively) for GWS loci. This might indicate that the COLOC analysis based on LAVA loci is hindered by statistical power, and the colocalization estimate in the LAVA loci might be too conservative.

Across all loci (GWS and LAVA), the average proportion of colocalization probability is 64% vs. 36% non-colocalization probability when signal is detected for both traits. This suggests that the signal observed in about two thirds of the overlapping GWS and LAVA loci is consistent with a shared causal variant.

Table 1. *Overview of loci with pleiotropic evidence on the gene or variant level across all three traits. LAVA regions without significant local h^2 or loci being unresolved by FLAMES are left blank (-). H4 = posterior probability for the hypothesis that both traits are associated and share a single causal variant. Loci that overlap across traits or metric that pass a certain (e.g. significance) threshold are printed in bold. Without colocalization evidence, we did not perform variant-level mediation analyses (-).*

Trait	Ancestry	GWS locus (Chr:Start:End)	Gene	Traitpair	Genetic correlation (EUR)			Colocalization		Variant-level pleiotropy (EUR)
					Region	r_g	*P*	H4	Top SNP	
INS	EUR	**1:72512988:74077588**	*NEGR1 / LRRIQ3*	INS-DEP	Chr 1, BP	.88	**5.17×10^{-12}**	.47	-	-
				INS-ANX	72513120:	.85	**2.20×10^{-7}**	.00	-	-
DEP	EUR	**1:73307570:73991792**	*NEGR1 / LRRIQ3*	DEP-ANX	73992170	.97	**7.31×10^{-10}**	.00	-	-
				INS-DEP	Chr 1, BP	.56	**5.56×10^{-5}**	.41	-	-
ANX	EUR	**1:73766431:73991651**	*LRRIQ3*	INS-ANX	73992171:	.73	**1.04×10^{-6}**	**.77**	rs61767168	*NA*
				DEP-ANX	75132092	.76	**2.72×10^{-6}**	**.94**	rs7540134	DEP → ANX
INS	EUR	-	*ARHGAP15*	INS-DEP	Chr 2, BP	.82	**6.95×10^{-5}**	.03	-	-
DEP	EUR	**2:144146675:144248905**	*ARHGAP15*	INS-ANX	143259781:	.80	**5.74×10^{-7}**	.01	-	-
ANX	EUR	**2:144145478:144272229**	*ARHGAP15*	DEP-ANX	144503688	1.0	**1.06×10^{-6}**	**.96**	rs1320139	DEP → ANX
INS	EUR	3:71433554:71474216	*FOXP1*	INS-DEP	Chr 3, BP	.34	3.42×10^{-3}	-	-	-
DEP	EUR	3:71748419:71820969	*FOXP1*	INS-ANX	71223282:	.75	**6.04×10^{-7}**	.00	-	-
ANX	EUR	-	*FOXP1*	DEP-ANX	72334704	.68	**7.67×10^{-5}**	**.53**	rs1617010	DEP → ANX
INS	EUR	-	*LPHN3*	INS-DEP	Chr 4, BP	.72	**6.67×10^{-5}**	.01	-	-
DEP	EUR	-	*LPHN3*	INS-ANX	61971162:	.41	2.20×10^{-2}	-	-	-
ANX	EUR	-	*LPHN3*	DEP-ANX	63337954	.93	**1.71×10^{-6}**	.01	-	-
INS	EUR	**5:92995013:93494731**	*NR2F1*	INS-DEP	Chr 5, BP	.51	4.96×10^{-4}	**.91**	rs72786643	INS → DEP
DEP	Cross	**5: 92788598:93288598**	*NR2F1*	INS-ANX	91956906:	-	-	-	-	-
ANX	EUR	**5:92377803:92487543**	*NR2F1*	DEP-ANX	93814604	-	-	**.88**	rs10078721	DEP → ANX
INS	EUR	-	*ZNF608*	INS-DEP		.69	**6.47×10^{-6}**	.03	-	-
DEP	EUR	5:124203948:124308088 5:124896039:124955099	*ZNF608*	INS-ANX	Chr 5, BP 124109486:	.91	**3.02×10^{-7}**	.00	-	-
ANX	EUR	-	*ZNF608*	DEP-ANX	125131720	1.0	**3.03×10^{-9}**	.02	-	-
INS	EUR	**7:1928823:2003043**	-	INS-DEP	Chr 7, BP	.38	3.13×10^{-3}	**.73**	rs868754	DEP → INS
DEP	EUR	**7:1860733:2110850**	-	INS-ANX	1366973:	-	-	**.70**	rs11773627	INS → ANX
ANX	EUR	**7:1855531:2110850**	-	DEP-ANX	2473749	-	-	**.96**	rs868754	DEP → ANX
INS	EUR	**7: 113772805:114348256**	*FOXP2*	INS-DEP	Chr 7, BP	.70	**2.32×10^{-8}**	**.99**	rs8180817	*NA*
DEP	EUR	**7:114015707:114194615**	*FOXP2*	INS-ANX	113339387:	.79	**5.74×10^{-7}**	**.99**	rs8180817	*NA*
ANX	EUR	**7:114015707:114194615**	*FOXP2*	DEP-ANX	115321301	.98	**4.19×10^{-8}**	**.99**	rs8180817	*NA*
INS	EUR	**9:14534422:14534505**	*NFIB*	INS-DEP	Chr 9, BP	.63	**2.32×10^{-5}**	.06	-	-
DEP	Cross	**9:14401283:14901283**	*NFIB / ZDHHC21*	INS-ANX	14244748:	.47	3.97×10^{-3}	-	-	-
ANX	Cross	**9:14439816:14466354**	*NFIB*	DEP-ANX	14902924	.90	**2.07×10^{-7}**	.05	-	-

INS	EUR	**11:99064188:99200206**	*CNTN5*	INS-DEP	Chr 11, BP	.66	4.23×10^{-4}	.01	-	-
DEP	Cross	**11:98891095:99391095**	*CNTN5*	INS-ANX	98929381:	.85	$\mathbf{7.21\times10^{-5}}$	.03	-	-
ANX	Cross	**11:99083930:99141095**	*CNTN5*	DEP-ANX	99619347	.94	$\mathbf{1.28\times10^{-5}}$	**.98**	rs2510682	DEP → ANX
INS	EUR	**11:57369008:57681828**	-	INS-DEP	Chr 11, BP	.79	$\mathbf{1.97\times10^{-6}}$	**.96**	rs2509805	*NA*
DEP	EUR	**11:57404779:57681828**	*CTNND1*	INS-ANX	56623495:	.69	2.98×10^{-4}	**.95**	rs12790660	INS → ANX
ANX	EUR	**11:57385856:57681828**	-	DEP-ANX	57675986	1.0	$\mathbf{1.28\times10^{-6}}$	**.97**	rs673344	*NA*
INS	EUR	-	*NCAM1*	INS-DEP	Chr 11, BP	.87	$\mathbf{1.66\times10^{-7}}$	.01	-	-
DEP	EUR	**11:112826867:112922254**	*NCAM1*	INS-ANX	112755447:	.71	1.27×10^{-4}	-	-	-
ANX	EUR	**11:112826867:112922254**	*NCAM1*	DEP-ANX	113889019	1.0	$\mathbf{4.33\times10^{-15}}$	**.95**	rs2298527	DEP → ANX
INS	EUR	-	*HMBG1*	INS-DEP	Chr 13, BP	.72	$\mathbf{6.65\times10^{-5}}$	.00	-	-
DEP	EUR	-	*HMBG1*	INS-ANX	30068526:	.68	3.09×10^{-4}	-	-	-
ANX	EUR	-	*HMBG1*	DEP-ANX	31433218	.84	$\mathbf{4.46\times10^{-6}}$	.00	-	-
INS	EUR	**15:91403674:91437388**	*FURIN*	INS-DEP	Chr 15, BP	.72	$\mathbf{9.62\times10^{-6}}$	**.99**	rs4702	INS → DEP
DEP	EUR	**15:91416550:91429042**	*FURIN*	INS-ANX	90632719:	-	-	**.99**	rs4702	INS → ANX
ANX	EUR	**15:91416550:91429042**	*FURIN*	DEP-ANX	91593623	-	-	**.99**	rs4702	DEP → ANX
INS	EUR	16:51163120:51190740	*SALL1*	INS-DEP	Chr 16, BP	.49	$\mathbf{9.78\times10^{-6}}$	**.98**	rs11645288	INS → DEP
DEP	EUR	16:51460871:51490422	*SALL1*	INS-ANX	50467034:	.35	5.36×10^{-3}	-	-	-
ANX	EUR	-	*SALL1*	DEP-ANX	52041336	.70	$\mathbf{1.24\times10^{-5}}$	**.99**	rs728538	*NA*
INS	EUR	-	*RBFOX1*	INS-DEP	Chr 16, BP	.64	$\mathbf{8.55\times10^{-6}}$	**.57**	rs10852673	*NA*
DEP	EUR	16:6303301:6343315	*RBFOX1*	INS-ANX	5782969:	.60	1.92×10^{-4}	-	-	-
ANX	EUR	-	*RBFOX1*	DEP-ANX	6446081	.96	$\mathbf{3.40\times10^{-7}}$	**.97**	rs10852673	*NA*
INS	EUR	**16:12259730:12464821**	*SHISA9*	INS-DEP	Chr 16, BP	.62	$\mathbf{2.02\times10^{-5}}$	**.99**	rs830733	*NA*
DEP	Cross	**16:12101254:12601254**	*SHISA9*	INS-ANX	12793150:	.15	3.44×10^{-1}	-	-	-
ANX	EUR	-	*SHISA9*	DEP-ANX	13893407	.68	$\mathbf{5.47\times10^{-5}}$	.32	-	-
INS	EUR	**18: 50358109:50908403**	*DCC*	INS-DEP	Chr 18, BP	.67	$\mathbf{4.29\times10^{-5}}$	**.61**	rs7231178	INS → DEP
DEP	EUR	**18:50620087:50907127**	*DCC*	INS-ANX	50372824:	.86	$\mathbf{5.66\times10^{-6}}$	**.90**	rs6508210	INS → ANX
ANX	EUR	**18:50555931:50907127**	*DCC*	DEP-ANX	51060986	1.0	$\mathbf{2.81\times10^{-9}}$	**.95**	rs8084351	DEP → ANX
INS	EUR	-	-	INS-DEP	Chr 20, BP	.73	$\mathbf{1.18\times10^{-5}}$	**.79**	rs6131010	INS → DEP
DEP	EUR	**20:44680412:44747947**	-	INS-ANX	44072211:	.85	$\mathbf{1.87\times10^{-6}}$	**.98**	rs6032660	ANX
ANX	EUR	**20:44680412:44747947**	-	DEP-ANX	45673603	.78	$\mathbf{4.33\times10^{-6}}$	**.98**	rs6131010	DEP → ANX
INS	EUR	20:50912495:51236973	*TSHZ2*	INS-DEP	Chr 20, BP	.95	$\mathbf{6.29\times10^{-6}}$	**.76**	rs13043475	*NA*
				INS-ANX	50653621:	-	-	-	-	-
				DEP-ANX	51533750	-	-	-	-	-
DEP	Cross	20:51338847:51838847	*TSHZ2*	INS-DEP	Chr 20, BP	.62	9.47×10^{-5}	-	-	-
				INS-ANX	51533751:	.39	3.88×10^{-3}	-	-	-
ANX	EUR	-	*TSHZ2*	DEP-ANX	52411532	.80	$\mathbf{1.52\times10^{-6}}$	.00	-	-

Distinguishing horizontal from vertical pleiotropy

We performed an exploratory analysis to investigate whether different forms of pleiotropy (i.e. vertical, horizontal, or confounding) could be distinguished among the 95 colocalizing variants. For this purpose we used variant-level PRISM's[42] pairwise analyses (see Methods), which is based on the trait-level LHC-MR[53] framework. LHC-MR assumes classical Mendelian Randomization assumptions but allows for the modelling of a genetic confounder. All of the 86 variants that successfully completed the PRISM pipeline were assigned to models consistent with pairwise vertical pleiotropy (i.e., a direct effect on one trait and a mediated effect on the other) rather than horizontal pleiotropy or shared confounding (see Methods; Supplementary Table 7).

Out of the 25 and 7 unique variants colocalizing between only DEP-ANX or INS-ANX respectively, all but one (25 and 6 respectively) were assigned by PRISM to have a direct effect on INS or DEP. Given strong LHC-MR evidence for INS → ANX (β = 0.65, SE=0.04, $p = 3.39\times10^{-11}$) and DEP → ANX (β = 0.85, SE = 0.02, $p < 1.00\times10^{-300}$), these variants were subsequently classified as vertical pleiotropic of which genetic effects on ANX are mediated via DEP or INS. For the only PRISM direct effect on ANX (an INS-ANX colocalizing variant), we found no evidence for vertical pleiotropy with respect to INS, as LHC-MR did not support ANX→INS (β = 0.07, SE = 0.08, p = 0.37).

The 28 unique variants colocalizing between DEP and INS were more often attributed to have a direct effect on INS (61%) than on DEP (39%). Given bidirectional evidence from LHC-MR (INS → DEP β = 0.56, SE = 0.08, $p = 9.43\times10^{-12}$; DEP → INS β = 0.30, SE=0.08, $p = 1.16\times10^{-}$), all of these variants were classified as vertical pleiotropic. Overall, this suggests that the data for colocalizing variants are more compatible with INS and DEP being risk factors for ANX, while the relationship between INS and DEP shows a directional mixture across colocalizing variants.

Of the thirteen loci with colocalization evidence across all three traits, three loci were also assigned the same top SNP with the highest posterior probability (PP) of being the most likely shared causal variant across two traits (Table 1). rs868754 ($PP_{INS\text{-}DEP}$ = .47; $PP_{DEP\text{-}ANX}$ = .54) was assigned to have a direct effect on DEP in both pairwise PRISM analyses. Combined with the LHC-MR results described above, this led to classification as vertical pleiotropic (DEP → INS and DEP → ANX). Two pairwise PRISM analyses assigned rs6131010 ($PP_{INS\text{-}DEP}$ = .69; $PP_{DEP\text{-}ANX}$ = .70) to have a direct effect on INS and a direct effect on DEP, both of which could again be classified as pairwise vertical pleiotropic effects (INS → DEP and DEP → ANX respectively).

Two other loci with colocalization evidence across all three traits were assigned the same top SNP across all three traits (rs4702 all PP = 1.0; rs8180817 $PP_{INS\text{-}DEP}$ = .50, $PP_{INS\text{-}ANX}$ = .55, $PP_{DEP\text{-}ANX}$ = .40). Following the same pairwise methodology, rs4702 was classified vertical pleiotropic as INS → DEP, INS → ANX and DEP → ANX. Although one might be inclined to interpret these findings as a

trivariate cascade (i.e. INS → DEP → ANX), we caution that we did not perform PRISM's trait wise analysis to integrate evidence across all three trait pairs simultaneously. This would require >30 traits for reliable inference[42] of pleiotropy, necessitating adding traits beyond the scope of this study. Additionally, it is debated whether in practice all the assumptions underlying Mendelian Randomization hold[54]. Violations of assumptions could have impacted the findings presented here.

Shared INS-DEP-ANX loci converge on the gene level

To further explore the potential pleiotropy in these loci, we also assessed the level of pleiotropy on the gene-level. We predicted the effector genes of our shared loci using the FLAMES model, which integrates multimodal evidence (e.g. fine-mapping, annotations, eQTLs, co-expression) to identify the most likely causal gene[39]. Our aim was to examine how often shared loci also shared the same effector gene. We ran FLAMES with default settings (see Methods), on all overlapping GWS and LAVA loci. In total, 17 shared genes are predicted to be effector genes shared between INS, DEP and ANX (Table 1), 38 genes are uniquely shared between DEP and INS, 2 genes are uniquely shared between ANX and INS, and 32 genes are uniquely shared between ANX and DEP (Figure 3a; Supplementary Table 8). Out of all predicted effector genes, 60% of predicted genes are shared with at least one other trait (81% ANX, 80% INS, 71% DEP). However, this includes loci where FLAMES resolved only one trait and cannot with certainty predict an effector gene for the second trait and dilutes estimates of concordance. When considering overlapping loci where at least 2 genes are prioritized, these genes are concordant in 80% of overlapping loci. These results show that in the majority of resolved loci of ANX, DEP and INS, the GWAS signal converges into the same genes.

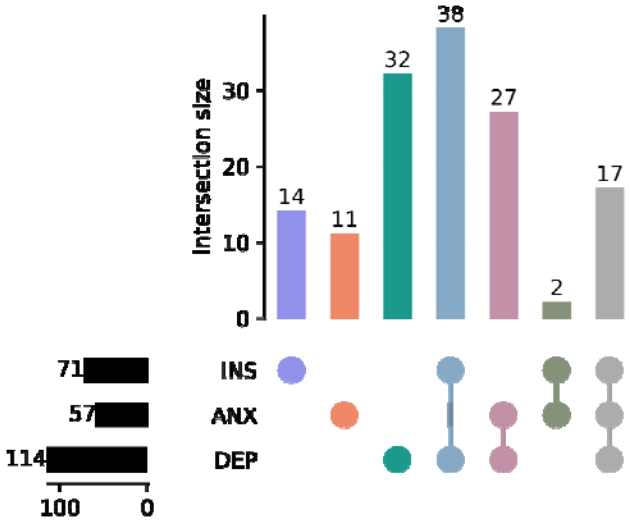

Figure 3. **Number of overlapping effector genes across conditions.** We aggregated the identified effector genes (by FLAMES) across ancestries and across traits to give insight into the number of shared genes within each of these domains.

Biological mechanisms of shared effector genes

We performed Gene Ontology (GO)[55] enrichment analysis on three distinct set of genes; the 17 effector genes shared between all three phenotypes, 27 effector genes uniquely shared between ANX and DEP, and the 38 effector genes uniquely shared between INS and DEP. GO terms from the MSigDB database are evaluated in a logistic regression framework (see Methods). We find significantly associated GO terms with all three sets of genes (Figure 4, Supplementary Table 9). Interestingly, despite having no genes in common, the effector genes uniquely shared by ANX-DEP and by INS-DEP were enriched in the same gene-set on two occasions (Cellular Component GABAergic synapse; Biological Processes synaptic membrane adhesion), providing evidence for pathway-level convergence.

Effector genes uniquely shared between INS and DEP were furthermore enriched in processes related to cell adhesion and synaptic structuring, neuronal development and differentiation, synaptic signaling and plasticity. The effector genes uniquely shared between ANX and DEP were uniquely enriched in the (post)synaptic membrane, postsynaptic modulation of synaptic transmission and the regulation of GABAergic neurotransmission. The two shared effector genes between ANX and INS (*PDE4B, FTO*) were not tested for enrichment, but their functions align with the significantly enriched gene-sets. In rodents, acute stress upregulates Pde4b within a defined GABAergic projection, reducing postsynaptic cyclic adenosine monophosphate (cAMP) signalling and producing anxiety-like behaviour. This effect can be reversed by projection-specific *Pde4b* knockdown in the same circuit[56]. Pharmacological *PDE4D* inhibition also mitigates sleep deprivation-induced memory impairments[57]. *FTO* is known to be involved in the regulation of dopaminergic midbrain circuitry[58].

Many enriched gene-sets are strongly related to neurological development and synapse function despite testing for enrichment in the broad array of gene-sets included in MSigDB. It is clear that with processes such as the generation of neurons, neuron axonogenesis, neuron nucleus development and axon guidance, as well as broader forebrain (i.e. striatum, subpallium and telencephalon) developmental processes, general processes related to the development of the brain are implicated in the trivariate genetic risk for INS-DEP-ANX (Figure 4). Additionally we find that processes related to viral infection and inflammation are implicated, which converges with prior observations[59]. Note that it is possible that genes now uniquely shared between INS-DEP or DEP-ANX will become shared across INS-DEP-ANX when GWAS power further increases.

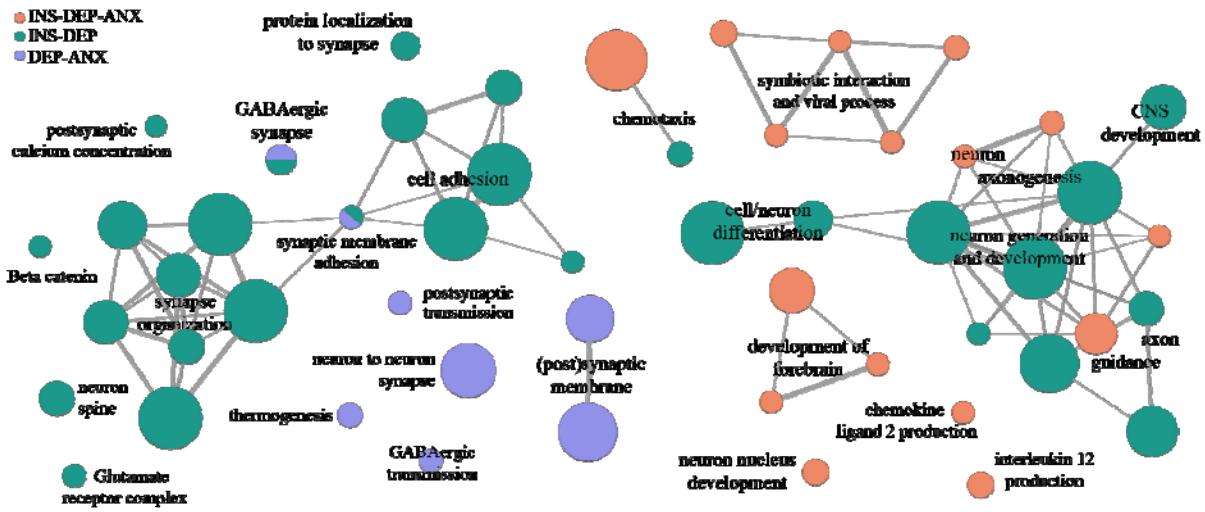

Figure 4. **Biological processes and cellular component enriched for true pleiotropic genes**. Effector genes that were shared among 2 or 3 traits were significantly enriched ($p < 1.90\times10^{-6}$) in Gene Ontology (GO) biological processes and cellular components. Node size reflects gene-set size, edges reflect pathway similarity scores (see Methods). Note that only one gene was uniquely shared between INS and ANX, and was therefore not tested for enrichment.

Discussion

Insomnia, depression and anxiety disorders frequently co-occur and share a substantial proportion of their genetic risk[13]. Most previous genetic studies have concentrated on shared risk between DEP and ANX[26], often omitting INS or restricting analyses to correlations of (overlapping) genetic effects summarized across the genome[30,31]. This study includes the first multi-level cross-ancestry investigation into the shared molecular genetic signal across insomnia, depression and anxiety. We find 55% genetic overlap across all three traits with high correlations of effect sizes of overlapping variants ($r_{g\text{-shared}}$ = .63 - .94). We are able to identify 195 unique genomic loci with statistical evidence of shared genetic signal across (pairs of) traits. In our attempts to rule out instances of spurious pleiotropy by colocalization analyses and effector gene prediction, we find strong evidence that the majority of these loci contain a causal variant and/or shared effector gene shared between INS, ANX, and DEP. Exploratory analyses suggest that these shared causal variants are more consistent with models of vertical than horizontal pleiotropy. Enrichment analyses on shared effector genes also indicate the presence of pathway-level convergence, implying that the genes regulating processes in the GABAergic synapse and synaptic membrane adhesion contain genetic liability factors shared across insomnia, depression and anxiety. Altogether this shows how large-scale datasets and state of the art methodology can be used to shed light on the shared genetic risk across insomnia, depression and anxiety.

Previous work has suggested that the same subset of 10,000-15,000 variants seem to explain most of the SNP heritability of major psychiatric disorders, with genetic differences between conditions arising mainly from the directions and magnitudes of effects rather than from distinct loci[21]. In contrast, our Trivariate MiXeR analysis shows that insomnia does not fully conform to this pattern. Although insomnia shares a substantial proportion of its genetic architecture with depression and anxiety, about one fifth of its genetic signal appears distinct. This observation aligns with previous results, showing that loci from other traits that colocalize with insomnia loci cluster into a psychiatric and metabolic component[17]. Future work could investigate whether the genetic signal that is attributed to be unique to insomnia in this study, overlaps with the genetics of metabolic traits. This may help to disentangle the genetic heterogeneity of insomnia and inform stratified therapeutic approaches.

The current study applied a hierarchical multi-level framework and could thereby demonstrate that the majority of the shared genetic signal found on the locus level consistently points to the same likely causal shared signal on the variant and gene level. Specifically, we find that approximately two thirds of loci, when containing enough genetic signal, colocalize between these three traits. Effector genes appear to be shared in approximately three quarters of the investigated loci. This underscores that, generally, overlapping signal between these traits can be considered shared pleiotropic signal. The ability to prune loci for spurious pleiotropy prior to biological interpretation highlights the importance of incorporating colocalization and effector gene prediction in cross-disorder genetic studies.

In some instances, variant-level and gene-level results diverged. The prioritization of different effector genes in the presence of evidence for a shared causal variant could be a result of difference in overall fine-mapped credible sets despite significant colocalization, which could be a consequence of the single variant assumption used. Alternatively, these differences could be due to different weighting of eQTL evidence based on tissue relevance weighting step in FLAMES based on genome-wide MAGMA scores. The differences could also be due to different PoPS scores for genes in the locus between these traits, which could capture tissue/cell-type relevance of the GWAS signal based on the large quantity of tissue-specific co-expression networks included in the features. This complexity underscores that, as emphasized in a recent review[40], the loci we highlight should be considered a prioritized set of hypotheses for laboratory validation, rather than definitive evidence of shared biology.

Our genetic overlap, genetic correlation and vertical pleiotropy directionality results are in line with previous studies[24,26,30,60], though a recent non-comorbid cross-disorder analysis between depression and anxiety challenges these previous conclusions[27]. Disorder-specific GWAS reduce genetic correlation by design, since excluding comorbid cases removes a large proportion of the overlapping liability. In contrast, our multi-level framework tests whether shared loci reflect true colocalization and gene-level pleiotropy. Colocalization strengthens the evidence that overlap is not merely a product of linkage

disequilibrium between nearby but distinct causal variants, yet it cannot eliminate the possibility that comorbidity in case definitions contributes to the magnitude of overlap. Whereas non-comorbid GWAS clarify disorder-specific liability at the cost of sample size and thus statistical power, cross-disorder pleiotropy modeling identifies the common genetic backbone that underpins multimorbidity in the general population. Together these approaches provide complementary perspectives on specific and shared genetic risk.

The findings of this study should be considered within the context of a small total explained variance of the genetic component investigated. Despite focusing on the shared likely causa of genetic signal between traits that all have very modest SNP-based heritability. Nevertheless, the biological processes we identify to be shared have been implicated in insomnia, depression, and anxiety based on pharmacological[61–63] and neurobiological[64–67] evidence. Our findings extend this literature by showing that genes in these pathways are impacted by cross-disorder genetic risk variants. Gene-sets such as the GABAergic synapse, regulation of (GABAergic) synaptic transmission, postsynaptic signalling, and AMPA receptor complexes map plausibly onto the circuitry that normally silences locus coeruleus noradrenergic neurons during rapid eye-movement (REM) sleep[68]. Incomplete locus coeruleus silencing is hypothesized to relate to restless REM sleep and impaired overnight relief of emotional distress[69], and is central in the leading models hypothesizing about the role of insomnia in anxiety- and stress-related conditions[70]. At the same time, we note that genes with pleiotropic evidence across all three conditions were enriched in processes that are very typically found across psychiatric genetic literature[71], such as axon guidance, and forebrain development. This may reflect more general neurodevelopmental liability rather than mechanisms specifically shared between insomnia, depression and anxiety. Recent evidence also highlights that copy number variations shared across psychiatric conditions converge on pathways, but diverge by cell type, brain region, and gene dosage[72]. This highlights another future direction when focussing on divergence, rather than convergence, of shared genetic risk across INS, DEP, ANX.

Important limitations of this work include a source of potential inflation in the observed genetic overlap cannot be accounted for in this study, that is phenotype misclassification. As mentioned above, the initial INS, ANX, and DEP GWAS analyses could have been biased by a level of misdiagnosis, unrecognized comorbidity resulting from ascertainment bias[9], or the use of both unscreened and supernatural controls[73]. Future studies may address these potential sources of spurious pleiotropy by careful longitudinal assessment, and the selection of control subjects screened to only exclude the target conditions[74]. Similarly, underestimation of true genetic overlap can occur when the original GWAS summary statistics lack sufficient statistical power to detect genome-wide significant loci or significant local heritability. Anxiety was the least powered phenotype in our study, which likely contributed to e.g. the smaller number of overlapping effector genes observed, which should not be interpreted as evidence

of no overlap. Lastly, our exploratory findings that pairwise colocalized variants were attributable primarily to vertical pleiotropic effects should be interpretated with caution. Since we did not consider a large set (i.e. >30) of traits, we had limited statistical power to robustly evaluate the significance of the variant-trait associations, and lack a broader biological context that might reveal additional pleiotropic mechanisms[42]. Furthermore, there is valid criticism in the field that Mendelian Randomization is likely to violate its modelling assumptions which can bias the results[75]. Although simulations show that LHC-MR is robust to some violations of its modelling assumptions[53], we cannot prove that all assumptions are met. Definite conclusions on causal relationships require the triangulation of methods[76].

In conclusion, this study highlights a genetically interconnected basis for insomnia, depression and anxiety. We provide a shortlist of pleiotropic variants, genes and pathways, offering potential candidates for future pathway-tailored therapeutic targets with transdiagnostic benefits.

Methods

GWAS summary statistics

We obtained access to the largest available GWAS summary statistics of INS[17], ANX[18], and DEP[20,21] originally including participants from UK Biobank, 23andMe Research Institute, Million Veteran Program (MVP), Psychiatric Genetics Consortium (PGC), iPSYCH, FinnGen, and All of Us. Cross-ancestry summary statistics were available for ANX and DEP, including individuals from European (EUR), South Asian (SAS), East Asian (EAS), African (AFR) genetic ancestries. Only EUR individuals were included in the original GWAS of INS and compared to the EUR subset of ANX and DEP where necessary.

In accordance with recent methodological recommendations[77], effective sample size (N_{eff}) of each trait was first calculated for each individual sample included in the GWAS meta-analysis and subsequently summed. Essential columns (SNP, CHR, BP, A1, A2, Z, Neff) were extracted using https://github.com/precimed/python_convert (v.0.9.3). We defined genome-wide significant loci according to the FUMA definition[48].

Genome-wide genetic overlap

We applied Trivariate MiXeR[32] to disentangle the pattern of polygenic overlap among the EUR GWAS summary statistics of INS[17], ANX[18], and DEP[21]. This recently released tool (https://github.com/precimed/mix3r) allows for sample overlap and uses the 489 unrelated subjects of European ancestry from 1000 Genomes phase 3 as LD reference panel[78] including SNPs with a minor allele frequency (MAF) > .01 (n SNPs = 9,997,231). Following recommendations from the authors, we ignored all inter-variant correlations with $r^2 < 0.05$, considering them to be very noisy in a sample of 489

subjects. The MHC region (CHR6:BP25000000-34000000, GRCh37/hg19 build) was also excluded as recommended by the authors[44]. The analysis was restricted to overlapping SNPs across summary statistics (N_{SNPs} = 3,122,674). Output parameters such as the proportion of non-null SNPs (polygenicity; π) where multiplied by the total number of SNPs in the LD reference panel for interpretability.

Genome-wide (partial) genetic correlations

We computed genetic correlations (r_g) for INS, ANX and DEP summary statistics (EUR) using LD Score Regression (LDSC)[46]. 1000 Genomes EUR[78] served as reference panel. LDSC analysis is restricted to HapMap3 SNPs[79], ensuring a comparable set of SNPs available across traits with global genome coverage. We additionally applied partial LDSC[47] with the same input and settings, which is an R-package extending standard r_g from LDSC (https://github.com/GEMINI-multimorbidity/partialLDSC). We used this tool to condition the r_g between DEP and ANX on INS, between INS and DEP on ANX, and between INS and DEP on ANX.

Local genetic correlations

LAVA (Local Analysis of [co]Variant Association)[43] was used for local r_g analysis, i.e. detecting genomic regions of shared association between our three phenotypes. Sample overlap was estimated with LDSC as recommended by the authors. The analysis was restricted to overlapping SNPs across summary statistics. We used 2,495 predefined loci that are roughly independent and the 1000 Genomes reference data for analysis with EUR GWAS summary statistics as provided on the GitHub repository (https://github.com/josefin-werme/LAVA). Using the AFR GWAS summary statistics was not feasible due to the absence of suitable reference panels, SAS and EAS summary statistics did not have sufficient statistical power to run this analysis. For each region, traits devoid of local heritability ($p < 1\times10^{-4}$) were filtered out before estimating bivariate r_g. This resulted in 580 EUR bivariate tests performed. We corrected the alpha (α) level for statistical significance of local bivariate r_g to $\alpha = (.5 / 582) = 8.59 \times10^{-5}$.

Shared most likely causal SNPs

We ran COLOC (v5.2.3[80]) under a single variant assumption, in all overlapping LAVA locus pairs and GWS locus pairs. When only Z-scores were available, effect sizes (β) and standard errors (SE) were estimated from Z-scores (Z), effective sample size (N) and the minor allele frequency (p) as:

$$\beta = \frac{Z}{\sqrt{2p(1-p)(N+Z^2)}} \quad (1)$$

$$SE = \frac{1}{\sqrt{2p(1-p)(N+Z^2)}}. \quad (2)$$

We deemed every locus-pair with a posterior probability > .5 for the case that both traits are associated and share a single causal variant to be significantly colocalizing (H4).

Pleiotropy analysis

We ran the PRISM (Pleiotropic Relationships to Infer the SNP Model)[42] pairwise analysis pipeline to classify colocalizing variants into models of horizontal, vertical or confounding pleiotropy. PRISM uses results from Latent Heritable Confounder Mendelian Randomization (LHC-MR; https://github.com/LizaDarrous/lhcMR)[53] as input, a bidirectional causal estimation model that accounts for the presence of a heritable confounder (U). We ran LHC-MR with default settings, using EUR GWAS summary statistics and the provided LD scores and HapMap3[79] SNPs as input. PRISM pairwise analysis then uses LHC-MR's causal path estimates to parameterize the expected cross-trait covariance and assigns each variant posterior probabilities (PP) across eight Gaussian components for each trait pair (X,Y): variant has a direct effect 1) on neither X,Y or U, 2) on X, 3) on U, 4) on Y, 5) on X and U, 6) on X and Y, 7) on Y and U, or 8) on X, Y and U. Variants with the highest PP for model 6 are classified as horizontal pleiotropic, for model 3, 5, 7 or 8 as confounded. Variants with the highest PP for model 2 or 4 get classified as vertical pleiotropic when there was significant causal inference ($p < (.05/3 = .0167)$) from LHC-MR for mediation in that direction.

Shared most likely causal genes

We applied the FLAMES (Fine-mapped Locus Assessment Model of Effector geneS)[39] to identify the most likely causal gene in each shared locus identified in our study. The framework consists of five steps, applied on each condition (INS, ANX, DEP) with the available EUR and cross-ancestry summary statistics (https://github.com/Marijn-Schipper/FLAMES). In short, FLAMES integrates SNP-to-gene evidence and convergence-based evidence into a single prediction for each fine-mapped GWAS signal.

First, each locus was fine-mapped to obtain 95% credible sets. We ran PolyFun (POLYgenic FUNctionally-informed fine-mapping)[81] on each locus with default settings except for keeping SNPs in the HLA region. We extracted the estimated prior causal probabilities for the SNPs, which were then used by SuSiE (Sum of Single Effects)[82] to perform non-functionally informed fine-mapping assuming one causal variant per locus. Second, we ran MAGMA[83] as implemented in the CoMorMent Containers v1.8.1[84] to obtain gene-level *z*-scores. We then ran MAGMA tissue type analysis using these *z*-scores on the preformatted GTEx[85] tissue expression file provided by the FLAMES authors. Third, Polygenic Priority Score (PoPS; v0.2)[86] was used to compute convergence-based gene prioritization scores. The FLAMES authors provided PoPS features including an extensive set of public bulk and single-cell expression datasets and predicted protein-protein interactions (https://zenodo.org/records/12635505). We

intentionally excluded biological pathways in the PoPs features, aiming to prevent circularity in downstream pathway analyses on FLAMES prioritized genes[39]. Fourth, we annotated credible sets with FLAMES annotate, which creates gene-level evidence scores for genes within 750kb of the centroids of credible sets from SNP-to-gene annotations, MAGMA Z-scores and PoPS scores. We used the FLAMES predictive model to predict the effector gene of each credible set from the generated annotations.

Note that a few loci will be left unresolved by FLAMES in the case of (1) unsuccessful fine-mapping due to complex LD (e.g. MHC region), (2) no proximal genes (<750 kilobases) from the credible set, or (3) an estimated cumulative precision < 0.75.

Pathway analysis

We performed pathway analyses to find patterns of enrichment of effector genes shared across INS, DEP and ANX, and of effector genes uniquely shared across two traits. As previously described in the FLAMES paper[39], pathway membership of genes was encoded in a binary vector, and FLAMES prioritized genes were also encoded in a binary vector. We subsequently performed logistic regression, regressing the FLAMES genes of a phenotype on the pathways. Pathways were defined as Gene Ontology (GO)[55] Biological Processes and Cellular Components (n = 8,760) from MSigDB (v.2023.1). All significant (p < (.05/8,760)/3 = 1.90×10^{-6}) non-zero regression coefficients where ≥2 genes overlapped between the gene-set and the FLAMES effector genes were visualized as a graph using Cytoscape[87] and EnrichmentMap[65]. Pathway similarity scores (Jaccard and overlap combined coefficient = .5) were used as edges between the gene-sets (see the Nature Protocol by Reimand et al.[89] for further details).

Data availability

This study relies on previously published GWAS summary statistics as cited in the main text. Access to 23andMe summary statistics requires application via https://research.23andme.com/dataset-access/. The data generated in the current manuscript is available in the Supplementary Tables.

Code availability

The code to obtain the results presented in this manuscript will be available via https://github.com/EPTissink/ upon publication.

Acknowledgements

This work has received funding from ZonMw, The Hague, The Netherlands, project 09120011910032 REMOVE, and from the European Union: the European Research Council (ERC), Brussels, Belgium, Advanced Grant 101055383 OVERNIGHT (E.T. and E.J.W.V.S.). M.S. was funded by NWO

Gravitation: BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology (Grant No. 024.004.012). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The funders had no role in the design, data collection, data analysis, and reporting of this study. We would like to thank Marie Verbanck and Martin Tournaire for insightful discussions. We would like to thank the research participants and employees of 23andMe Research Institute for making this work possible.

Author Contributions

M.S.: Formal Analysis, Visualization, Writing – Original. A.A.S.: Software, Writing – Review & Editing. C.R.: Software, Writing – Review & Editing. E.F.: Resources, Writing – Review & Editing. R.P.: Resources; Writing – Review & Editing. D.P.: Conceptualization, Writing – Review & Editing. E.J.W.v.S.: Conceptualization, Funding Acquisition, Writing – Review & Editing. E.T.: Conceptualization, Data Curation, Formal Analysis, Supervision, Visualization, Writing – Original.

Conflicts of interest

M.S., A.A.S., C.R., E.F., D.P., E.J.W.v.S., and E.T. declare no conflict of interest. R.P. received an honorarium from Karger Publishers for his work on Complex Psychiatry and a research grant from Alkermes outside the scope of this manuscript.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. (DSM-5). https://cir.nii.ac.jp/crid/1573950399819987840 (2013).
2. Soehner, A. M. & Harvey, A. G. Prevalence and Functional Consequences of Severe Insomnia Symptoms in Mood and Anxiety Disorders: Results from a Nationally Representative Sample. *Sleep* **35**, 1367–1375 (2012).
3. Ohayon, M. M. & Roth, T. Place of chronic insomnia in the course of depressive and anxiety disorders. *J. Psychiatr. Res.* **37**, 9–15 (2003).
4. Wittchen, H. U. *et al.* The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur. Neuropsychopharmacol.* **21**, 655–679 (2011).
5. Cuijpers, P. The Challenges of Improving Treatments for Depression. *JAMA* **320**, 2529–2530 (2018).
6. Roth, T. *et al.* Sleep Problems, Comorbid Mental Disorders, and Role Functioning in the National Comorbidity Survey Replication. *Biol. Psychiatry* **60**, 1364–1371 (2006).

7. Leerssen, J. *et al.* Treating Insomnia with High Risk of Depression Using Therapist-Guided Digital Cognitive, Behavioral, and Circadian Rhythm Support Interventions to Prevent Worsening of Depressive Symptoms: A Randomized Controlled Trial. *Psychother. Psychosom.* **91**, 168–179 (2022).
8. de Lange, S. C. *et al.* Multimodal brain imaging of insomnia, depression and anxiety symptoms indicates transdiagnostic commonalities and differences. *Nat. Ment. Health* **3**, 517–529 (2025).
9. Smoller, J. W., Lunetta, K. L. & Robins, J. Implications of comorbidity and ascertainment bias for identifying disease genes. *Am. J. Med. Genet.* **96**, 817–822 (2000).
10. Wray, N. R., Lee, S. H. & Kendler, K. S. Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. *Eur. J. Hum. Genet.* **20**, 668–674 (2012).
11. Ruberto, V. L., Jha, M. K. & Murrough, J. W. Pharmacological Treatments for Patients with Treatment-Resistant Depression. *Pharmaceuticals* **13**, 116 (2020).
12. Ansara, E. D. Management of treatment-resistant generalized anxiety disorder. *Ment. Health Clin.* **10**, 326–334 (2020).
13. Lind, M. J. *et al.* An examination of the etiologic overlap between the genetic and environmental influences on insomnia and common psychopathology. *Depress. Anxiety* **34**, 453–462 (2017).
14. Madrid-Valero, J. J., Rubio-Aparicio, M., Gregory, A. M., Sánchez-Meca, J. & Ordoñana, J. R. The heritability of insomnia: Systematic review and meta-analysis of twin studies. *Sleep Med. Rev.* **58**, 101437 (2021).
15. Sullivan, P. F., Michael Neale, F. C. & Kendler, K. S. *Reviews and Overviews Genetic Epidemiology of Major Depression: Review and Meta-Analysis*. *Am J Psychiatry* vol. 157 10 http://views.vcu.edu/pub/mx/examples/mdreview. (2000).
16. 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**, 1568–1578 (2001).
17. Watanabe, K. *et al.* Genome-wide meta-analysis of insomnia prioritizes genes associated with metabolic and psychiatric pathways. *Nat. Genet.* **54**, 1125–1132 (2022).
18. Friligkou, E. *et al.* Gene discovery and biological insights into anxiety disorders from a large-scale multi-ancestry genome-wide association study. *Nat. Genet.* **56**, 2036–2045 (2024).
19. Strom, N. I. *et al.* Genome-wide association study of major anxiety disorders in 122,341 European-ancestry cases identifies 58 loci and highlights GABAergic signaling. 2024.07.03.24309466 Preprint at https://doi.org/10.1101/2024.07.03.24309466 (2024).
20. 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.* **56**, 222–233 (2024).

21. Als, T. D. *et al.* Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. *Nat. Med.* **29**, 1832–1844 (2023).
22. Jansen, P. R. *et al.* Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* **51**, 394–403 (2019).
23. Friligkou, E. *et al.* Gene discovery and biological insights into anxiety disorders from a large-scale multi-ancestry genome-wide association study. *Nat. Genet.* 1–10 (2024) doi:10.1038/s41588-024-01908-2.
24. Romero, C. *et al.* Exploring the genetic overlap between twelve psychiatric disorders. *Nat. Genet.* **54**, 1795–1802 (2022).
25. Grotzinger, A. D. *et al.* Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nat. Genet.* **54**, 548–559 (2022).
26. Tesfaye, M. *et al.* Identification of Novel Genomic Loci for Anxiety and Extensive Genetic Overlap with Psychiatric Disorders. *medRxiv* 2023.09.01.23294920 (2024) doi:10.1101/2023.09.01.23294920.
27. Tesfaye, M. *et al.* Comorbidity alters the genetic relationship between anxiety disorders and major depression. 2024.11.19.24317523 Preprint at https://doi.org/10.1101/2024.11.19.24317523 (2024).
28. The Brainstorm Consortium *et al.* Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).
29. Lee, P. H. *et al.* Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e11 (2019).
30. O'Connell, K. S. *et al.* Characterizing the Genetic Overlap Between Psychiatric Disorders and Sleep-Related Phenotypes. *Biol. Psychiatry* **90**, 621–631 (2021).
31. Baranova, A., Cao, H. & Zhang, F. Shared genetic liability and causal effects between major depressive disorder and insomnia. *Hum. Mol. Genet.* **31**, 1336–1345 (2022).
32. Shadrin, A. A. *et al.* Distinct patterns of genetic overlap among multimorbidities revealed with trivariate MiXeR. *Genome Med.* **17**, 106 (2025).
33. Cai, J., Zhan, J., Arking, D. E. & Bader, J. S. Priors, population sizes, and power in genome-wide hypothesis tests. *BMC Bioinformatics* **24**, 170 (2023).
34. Fisher, R. A. The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans. R. Soc. Edinb.* **52**, 399–433 (1919).
35. Falconer, D. S. *Introduction to Quantitative Genetics.* (Pearson Education India, 1996).
36. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits*. (Sinauer, 1998).
37. Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in complex traits: challenges and strategies. *Nat. Rev. Genet.* **14**, 483–495 (2013).

38. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLOS Genet.* **10**, e1004383 (2014).
39. Schipper, M. *et al.* Prioritizing effector genes at trait-associated loci using multimodal evidence. *Nat. Genet.* **57**, 323–333 (2025).
40. Mackay, T. F. C. & Anholt, R. R. H. Pleiotropy, epistasis and the genetic architecture of quantitative traits. *Nat. Rev. Genet.* 1–19 (2024) doi:10.1038/s41576-024-00711-3.
41. Pasman, J. A. *et al.* An encompassing Mendelian randomization study of the causes and consequences of major depressive disorder. *Nat. Ment. Health* 1–10 (2025) doi:10.1038/s44220-025-00471-x.
42. Tournaire, M., Nouira, A., Rozenholc, Y. & Verbanck, M. Inferring genetic variant causal network by leveraging pleiotropy. 2024.06.01.24308193 Preprint at https://doi.org/10.1101/2024.06.01.24308193 (2024).
43. 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).
44. Holland, D. *et al.* Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genet.* **16**, (2020).
45. Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, (2019).
46. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
47. Mounier, N. *et al.* Genetics identifies obesity as a shared risk factor for co-occurring multiple long-term conditions. 2024.07.10.24309772 Preprint at https://doi.org/10.1101/2024.07.10.24309772 (2024).
48. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, (2017).
49. Uffelmann, E., Leeuw, C. de & Posthuma, D. Local Genetic Sex Differences in Quantitative Traits. 2023.05.04.539410 Preprint at https://doi.org/10.1101/2023.05.04.539410 (2023).
50. Werme, J. *et al.* Local genetic correlation analysis links depression with molecular and brain imaging endophenotypes. *medRxiv* https://doi.org/10.1101/2023.03.01.23286613 (2023) doi:10.1101/2023.03.01.23286613.
51. Tissink, E. *et al.* The Genetic Architectures of Functional and Structural Connectivity Properties within Cerebral Resting-State Networks. *eNeuro* **10**, (2023).
52. Grotzinger, A. D. Shared genetic architecture across psychiatric disorders. *Psychol. Med.* **51**, 2210–2216 (2021).

53. Darrous, L., Mounier, N. & Kutalik, Z. Simultaneous estimation of bi-directional causal effects and heritable confounding from GWAS summary statistics. *Nat. Commun.* **12**, 7274 (2021).
54. de Leeuw, C., Savage, J., Bucur, I. G., Heskes, T. & Posthuma, D. Understanding the assumptions underlying Mendelian randomization. *Eur. J. Hum. Genet.* **30**, 653–660 (2022).
55. Carbon, S. *et al.* The Gene Ontology resource: Enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).
56. Xiao, Z.-X. *et al.* Pde4b-regulated cAMP signaling pathway in the AUDGABA-S1TrSst circuit underlies acute-stress-induced anxiety-like behavior. *Cell Rep.* **44**, (2025).
57. Zhao, H., Blokland, A., Prickaerts, J., Havekes, R. & Heckman, P. R. A. Treatment with the selective PDE4B inhibitor A-33 or PDE4D inhibitor zatolmilast prevents sleep deprivation-induced deficits in spatial pattern separation. *Behav. Brain Res.* **459**, 114798 (2024).
58. Hess, M. E. *et al.* The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. *Nat. Neurosci.* **16**, 1042–1048 (2013).
59. Prather, A. A., Vogelzangs, N. & Penninx, B. W. J. H. Sleep duration, insomnia, and markers of systemic inflammation: Results from the Netherlands Study of Depression and Anxiety (NESDA). *J. Psychiatr. Res.* **60**, 95–102 (2015).
60. Pasman, J. A. *et al.* Causes and consequences of major depressive disorder: An encompassing Mendelian randomization study. 2024.05.21.24307678 Preprint at https://doi.org/10.1101/2024.05.21.24307678 (2024).
61. Cutler, A. J., Mattingly, G. W. & Maletic, V. Understanding the mechanism of action and clinical effects of neuroactive steroids and GABAergic compounds in major depressive disorder. *Transl. Psychiatry* **13**, 228 (2023).
62. Krystal, J. H., Kavalali, E. T. & Monteggia, L. M. Ketamine and rapid antidepressant action: new treatments and novel synaptic signaling mechanisms. *Neuropsychopharmacology* **49**, 41–50 (2024).
63. Lucassen, P. J. *et al.* Regulation of adult neurogenesis by stress, sleep disruption, exercise and inflammation: Implications for depression and antidepressant action. *Eur. Neuropsychopharmacol.* **20**, 1–17 (2010).
64. Duman, R. S., Sanacora, G. & Krystal, J. H. Altered Connectivity in Depression: GABA and Glutamate Neurotransmitter Deficits and Reversal by Novel Treatments. *Neuron* **102**, 75–90 (2019).
65. Duman, R. S., Aghajanian, G. K., Sanacora, G. & Krystal, J. H. Synaptic plasticity and depression: new insights from stress and rapid-acting antidepressants. *Nat. Med.* **22**, 238–249 (2016).
66. Bisaz, R., Conboy, L. & Sandi, C. Learning under stress: A role for the neural cell adhesion molecule NCAM. *Neurobiol. Learn. Mem.* **91**, 333–342 (2009).

67. Manitt, C. *et al.* dcc orchestrates the development of the prefrontal cortex during adolescence and is altered in psychiatric patients. *Transl. Psychiatry* **3**, e338–e338 (2013).
68. Breton-Provencher, V. & Sur, M. Active control of arousal by a locus coeruleus GABAergic circuit. *Nat. Neurosci.* **22**, 218–228 (2019).
69. Wassing, R., Benjamins, J. S., Talamini, L. M., Schalkwijk, F. & Van Someren, E. J. W. Overnight worsening of emotional distress indicates maladaptive sleep in insomnia. *Sleep* **42**, (2019).
70. Van Someren, E. J. W. Brain mechanisms of insomnia: new perspectives on causes and consequences. *Physiol. Rev.* **101**, 995–1046 (2021).
71. Lee, P. H. *et al.* Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e11 (2019).
72. Engchuan, W. *et al.* Psychiatric disorders converge on common pathways but diverge in cellular context, spatial distribution, and directionality of genetic effects. 2025.07.11.25331381 Preprint at https://doi.org/10.1101/2025.07.11.25331381 (2025).
73. Kendler, K. S., Chatzinakos, C. & Bacanu, S.-A. The impact on estimations of genetic correlations by the use of super-normal, unscreened, and family-history screened controls in genome wide case–control studies. *Genet. Epidemiol.* **44**, 283–289 (2020).
74. Lee, P. H., Feng, Y.-C. A. & Smoller, J. W. Pleiotropy and Cross-Disorder Genetics Among Psychiatric Disorders. *Biol. Psychiatry* **89**, 20–31 (2021).
75. de Leeuw, C., Savage, J., Bucur, I. G., Heskes, T. & Posthuma, D. Understanding the assumptions underlying Mendelian randomization. *Eur. J. Hum. Genet.* **30**, 653–660 (2022).
76. Treur, J. L., Lukas, E., Sallis, H. M. & Wootton, R. E. A guide for planning triangulation studies to investigate complex causal questions in behavioural and psychiatric research. *Epidemiol. Psychiatr. Sci.* **33**, e61 (2024).
77. Grotzinger, A. D., Fuente, J. de la, 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).
78. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
79. Altshuler, D. M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
80. Pullin, J. M. & Wallace, C. Variant-specific priors in colocalisation analysis. 2024.08.21.608957 Preprint at https://doi.org/10.1101/2024.08.21.608957 (2024).
81. Weissbrod, O. *et al.* Functionally-informed fine-mapping and polygenic localization of complex trait heritability. https://doi.org/10.1101/807792 (2020) doi:10.1101/807792.

82. Wang, G., Sarkar, A., Carbonetto, P., Stephens, M. & Matthew Stephens, C. A simple new approach to variable selection in regression, with application to genetic fine-mapping. https://doi.org/10.1101/501114 doi:10.1101/501114.
83. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **11**, (2015).
84. Akdeniz, B. C. *et al.* COSGAP: COntainerized Statistical Genetics Analysis Pipelines. *Bioinforma. Adv.* **4**, vbae067 (2024).
85. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
86. Weeks, E. M. *et al.* Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases.
87. Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
88. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, (2010).
89. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* **14**, 482–517 (2019).