

ARTICLE OPEN



Multi-ancestry genome-wide association analyses incorporating SNP-by-psychosocial interactions identify novel loci for serum lipids

Amy R. Bentley^{1,109}, Michael R. Brown^{2,109}, Solomon K. Musani³, Karen L. Schwander⁴, Thomas W. Winkler⁵, Mario Sims⁶, Tuomas O. Kilpeläinen^{7,8}, Hugues Aschard^{9,10}, Traci M. Bartz¹¹, Lawrence F. Bielak¹², Jin-Fang Chai¹³, Kumaraswamy Naidu Chitrala¹⁴, Nora Franceschini¹⁵, Mariaelisa Graff¹⁵, Xiuqing Guo¹⁶, Fernando P. Hartwig¹⁷, Andrea R.V.R. Horimoto^{11,18}, Elise Lim¹⁹, Yongmei Liu²⁰, Alisa K. Manning^{21,22}, Ilja M. Nolte²³, Raymond Noordam²⁴, Melissa A. Richard^{25,26}, Albert V. Smith²⁷, Yun Ju Sung²⁸, Dina Vojinovic^{29,30,31}, Rujia Wang³², Yujie Wang¹⁵, Mary F. Feitosa⁴, Sarah E. Harris³³, Leo-Pekka Lyytikäinen^{34,35}, Giorgio Pistis³⁶, Rainer Rauramaa³⁷, Peter J. van der Most²³, Erin Ware³⁸, Stefan Weiss^{39,40}, Wanqing Wen⁴¹, Lisa R. Yanek⁴², Dan E. Arking⁴³, Donna K. Arnett⁴⁴, Christie Ballantyne^{45,46}, Eric Boerwinkle^{2,47}, Yii-Der Ida Chen¹⁶, Martha L. Daviglus⁴⁸, Lisa de las Fuentes^{49,50}, Paul S. de Vries², Joseph A. C. Delaney^{51,52,53,54}, Amanda M. Fretts^{51,53}, Lynette Ekinwe⁵⁵, Jessica D. Faul³⁸, Linda C. Gallo⁵⁶, Sami Heikkinen⁵⁷, Georg Homuth³⁹, M. Arfan Ikram²⁹, Carmen R. Isasi⁵⁸, Jost Bruno Jonas^{59,60}, Liisa Keltikangas-Järvinen⁶¹, Pirjo Komulainen³⁷, Aldi T. Kraja⁶², Jose E. Krieger¹⁸, Lenore Launer⁶³, Lifelines Cohort Study*, Jianjun Liu⁶⁴, Kurt Lohman⁶⁵, Annemarie I. Luik²⁹, Ani W. Manichaikul^{66,67,68}, Pedro Marques-Vidal⁶⁹, Yuri Milaneschi⁷⁰, Stanford E. Mwasongwe⁷¹, Jeffrey R. O'Connell⁷², Kenneth Rice¹¹, Stephen S. Rich^{66,67,68}, Pamela J. Schreiner⁷³, Lars Schwettmann^{74,75}, James M. Shikany⁷⁶, Xiao-ou Shu⁴¹, Jennifer A. Smith^{12,38}, Harold Snieder⁷⁷, Nona Sotoodehnia⁷⁷, E. Shyong Tai⁷⁸, Kent D. Taylor⁷⁹, Lesley Tinker⁸⁰, Michael Y. Tsai⁸¹, André G. Uitterlinden⁸², Cornelia M. van Duijn^{29,83}, Diana van Heemst²⁴, Melanie Waldenberger⁸⁴, Robert B. Wallace⁸⁵, Hwee-Lin Wee¹³, David R. Weir³⁸, Wen-Bin Wei⁸⁶, Ko Willems van Dijk^{87,88}, Gregory Wilson⁸⁹, Jie Yao¹⁶, Kristin L. Young¹⁵, Xiaoyu Zhang¹⁹, Wei Zhao^{12,38}, Xiaofeng Zhu⁹⁰, Alan B. Zonderman⁶³, Ian J. Deary³³, Christian Gieger^{84,91,92}, Hans Jörgen Grabe^{93,94}, Timo A. Lakka^{37,57,95}, Terho Lehtimäki^{34,35}, Albertine J. Oldehinkel⁹⁶, Martin Preisig³⁶, Ya-Xing Wang⁶⁰, Wei Zheng⁴¹, Michele K. Evans⁶³, Michael Province⁴, James Gauderman⁹⁷, Vilmondur Gudnason^{98,99}, Catharina A. Hartman⁹⁶, Bernardo L. Horta¹⁰⁰, Sharon L. R. Kardia¹², Charles Kooperberg⁸⁰, Ching-Ti Liu¹⁹, Dennis O. Mook-Kanamori^{101,102}, Brenda WJH Penninx⁷⁰, Alexandre C. Pereira¹⁸, Patricia A. Peyser¹², Bruce M. Psaty¹⁰³, Jerome I. Rotter¹⁶, Xueling Sim¹³, Kari E. North¹⁵, Dabeeru C. Rao⁴⁹, Laura Bierut¹⁰⁴, Clint L. Miller^{66,67,68}, Alanna C. Morrison¹⁰², Charles N. Rotimi^{1,109}, Myriam Fornage^{25,109} and Ervin R. Fox^{55,109}

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025

Serum lipid levels, which are influenced by both genetic and environmental factors, are key determinants of cardiometabolic health and are influenced by both genetic and environmental factors. Improving our understanding of their underlying biological mechanisms can have important public health and therapeutic implications. Although psychosocial factors, including depression, anxiety, and perceived social support, are associated with serum lipid levels, it is unknown if they modify the effect of genetic loci that influence lipids. We conducted a genome-wide gene-by-psychosocial factor interaction (GxPsy) study in up to 133,157 individuals to evaluate if GxPsy influences serum lipid levels. We conducted a two-stage meta-analysis of GxPsy using both a one-degree of freedom (1df) interaction test and a joint 2df test of the main and interaction effects. In Stage 1, we performed GxPsy analyses on up to 77,413 individuals and promising associations ($P < 10^{-5}$) were evaluated in up to 55,744 independent samples in Stage 2. Significant findings ($P < 5 \times 10^{-8}$) were identified based on meta-analyses of the two stages. There were 10,230 variants from 120 loci significantly associated with serum lipids. We identified novel associations for variants in four loci using the 1df test of interaction, and five additional loci using the 2df joint test that were independent of known lipid loci. Of these 9 loci, 7 could not have been detected without modeling the interaction as there was no evidence of association in a standard GWAS model. The genetic diversity of included samples was key in identifying these novel loci: four of the lead variants displayed very low frequency in European ancestry populations. Functional annotation highlighted promising loci for further experimental follow-up, particularly rs73597733 (MACROD2), rs59808825 (GRAMD1B), and rs11702544 (RRP1B). Notably, one of the genes in identified loci (RRP1B) was found to be a target of the approved drug Atenolol suggesting potential for drug repurposing. Overall, our findings suggest that taking interaction between genetic variants and psychosocial factors into account and including genetically diverse populations can lead to novel discoveries for serum lipids.

Translational Psychiatry (2025)15:207; <https://doi.org/10.1038/s41398-025-03418-z>

A full list of author affiliations appears at the end of the paper.

Received: 20 June 2024 Revised: 14 May 2025 Accepted: 2 June 2025

Published online: 20 June 2025

INTRODUCTION

The concentrations of key serum lipids, such as high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), and triglycerides (TG) are routinely assessed to determine an individual's cardiometabolic clinical risk profile, and to guide drug therapy (e.g., statins) aiming to reduce the morbidity and mortality associated with diseases such as coronary artery disease, stroke, and type 2 diabetes. Serum lipid levels are known to be influenced both by lifestyle, including diet, physical activity, smoking, and alcohol consumption, as well as genetic factors, with over 700 lipid loci identified using genome-wide association studies (GWAS) [1–3]. Although the importance of both genetic and lifestyle factors is well-established, the interplay between these two factors on serum lipid levels is less well understood. The Gene-Lifestyle Interactions Working Group, under the aegis of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium [4], has developed a framework for studying gene-lifestyle interactions for cardiometabolic traits in large, multi-ancestry meta-analyses [5]. This strategy has facilitated the discovery of novel lipid loci in studies accounting for interactions with smoking [6], physical activity [7], alcohol consumption [8], educational attainment [9], and sleep duration [10], suggesting that these lifestyle factors may indeed modulate genetic effects on serum lipids. The loci identified in these efforts could potentially explain how lifestyle exposures can contribute to disturbances in lipid levels.

Psychosocial factors, especially depression, contribute to the pathogenesis of cardiovascular diseases (e.g., myocardial infarction) and increased mortality in patients with coronary heart disease [11–13], and Mendelian Randomization (MR) analysis suggests that this association is causal [14]. Depression and depressive symptoms are associated with serum lipid concentrations [13, 15], the plasma lipidome [16, 17], and lipid metabolism, with distinct metabolic signatures associated with various symptom dimensions [18]. *APOE* alleles associated with serum lipids have also been associated with anxiety and depression [19, 20]. Some evidence from MR analyses suggests that depression increases TG and decreases HDL [21]. Serum lipids may also mediate the association between depression and cardiovascular disease. The association between depression and coronary artery disease was attenuated when an MR analysis was adjusted for serum lipids [14]. Similarly, in one study, nearly a third of the association of depression with arterial stiffness, a key intermediary of major cardiovascular events, was found to be mediated by metabolic syndrome, particularly hypertriglyceridemia among men [22]. Low social support has been associated with high cholesterol in a nationally representative US nonelderly population [23] and with high cholesterol, LDL, and non-HDL cholesterol among type 2 diabetics and their families [24]. Both anxiety and depression have been associated with elevated triglycerides [25]. Proposed mechanisms through which psychosocial factors and serum lipids may influence each other include high dietary intake of saturated fat and cholesterol, gut dysbiosis, the hypothalamic-pituitary-adrenal axis, and neuroinflammation [26–28]. Both direct and indirect mechanisms, such as psychosocial factor-associated changes in lifestyle or medication use, are plausible. This may confound interaction effects. Importantly, there is evidence for a genetic contribution to some of these psychosocial factors, particularly depression [29–33].

In this study, we assess how incorporating interaction between genetic variants and psychosocial factors (depressive symptoms, anxiety symptoms, and low social support) helps identify lipid loci missed by standard marginal genetic effect GWAS. To maximize the transferability of our results and to address known gaps in the field, we prioritized the inclusion of diverse population groups, as ancestry can influence both genetic (e.g. frequency of variants, linkage disequilibrium around associated signals) and psychosocial factors (e.g. presence of stigma, availability of healthcare

access). We conducted multi-ancestry meta-analyses of genome-wide variant $\times$ psychosocial factor interaction (G $\times$ Psy) studies on serum lipids in up to 133,157 individuals.

RESULTS

In this study of psychosocial factors and serum lipids, we meta-analyzed data on up to 133,157 individuals from 50 genome-wide interaction studies using a two-stage study design (Fig. 1; Study Details Supplementary Tables S1–2; Supplementary Note). Sample sizes and descriptive statistics of the studies participating in Stages 1 and 2 analyses are summarized in Supplementary Table 2. Study participants included European ancestry (EUR; 67.5%; $n = 89,939$), African ancestry (AFR; 14.4%; $n = 19,133$), Hispanic/Latino (HIS; 12.0%; $n = 15,949$), Asian ancestry (ASN; 3.5%; $n = 4672$), and Brazilian individuals (BRZ; 2.6%; $n = 3464$; see Methods for further details regarding selection of groups and population descriptors). All psychosocial factors were coded as binary variables. On average, 17.5% of Stage 1 participants were reported to have elevated depressive symptoms (DEPR) and 24.2% had elevated anxiety symptoms [ANXT], based on standard cutpoints, with a similar distribution among Stage 2 participants (15.8% [DEPR] and 24.3% [ANXT]). Low social support (SOCS) was defined as the lowest quartile of the distribution (Methods). Fewer studies had data on SOCS ($n = 23$) and ANXT ($n = 21$) than DEPR ($n = 50$; harmonization of psychosocial factors is described in Supplementary Table 3).

In Stage 1 population-specific and cross-population meta-analyses (CPMA), we identified 15,774 variants that met our selection criteria of $P < 10^{-5}$ for either the 1 degree of freedom interaction test (1df) or the 2df joint test of the main and interaction effects (2df; Fig. 1). These variants were carried forward for further analysis in Stage 2 samples. In the meta-analysis of stages 1 and 2, there were 10,230 variants from 120 loci that were significantly associated in at least one model using either statistical test (P_{1df} or $P_{2df} < 5 \times 10^{-8}$; Supplementary Table 4). We found seven variants in four loci that were associated with serum lipids with a genome-wide significant p-value for the 1df test of interaction (Table 1). For instance, among those reporting low social support, the A allele of rs11949029 was associated with a much lower LDL concentration than among those who did not ($\beta_{\text{Interaction}} = -19.2$, SE 3.5 mg/dL, $P_{1df} = 4.1 \times 10^{-8}$; Fig. 2A). Thus, among those with low social support, this allele was associated with 12.6 mg/dL lower LDL, but 6.6 mg/dL higher LDL among those not reporting low social support. In meta-analyses that did not include a multiplicative term (i.e. a standard GWAS model; available only for stage 1 studies), no association of the variant with LDL was observed ($P = 0.69$), even after adjustment for SOCS ($P = 0.75$).

A significant association using the 2df joint test of the main effect and interaction can represent the main effect of a variant, its interaction with the exposure, or both. To exclude associations driven primarily by the main effect, we considered as previously unidentified only those variants with $P_{2df} < 5 \times 10^{-8}$ that were independent of known loci (defined as ± 500 kb from the 95% credible sets reported in Graham *et al* [1] or variants reported in other major publications [34–39]). There were 14 variants and 8 loci that were significantly associated with serum lipids using the 2df test and independent from known loci (Table 2). Six of these loci displayed nominal significance for an interaction effect ($P_{1df} < 0.05$). This includes three of the four loci that were genome-wide significant using the 1df test, with the remaining two near this threshold. Among these is rs59808825 (*GRAMD1B* [nearest gene]), for which the main effect of the C allele on LDL was positive ($\beta = 5.0$, SE = 2.1 mg/dL) but an inverse interaction effect with ANXT ($\beta_{\text{Interaction}} = -22.9$, SE = 4.2 mg/dL), so that the total effect of the C allele among those reporting anxiety symptoms was negative ($P_{2df} = 8.8 \times 10^{-9}$; Fig. 2B). In the main

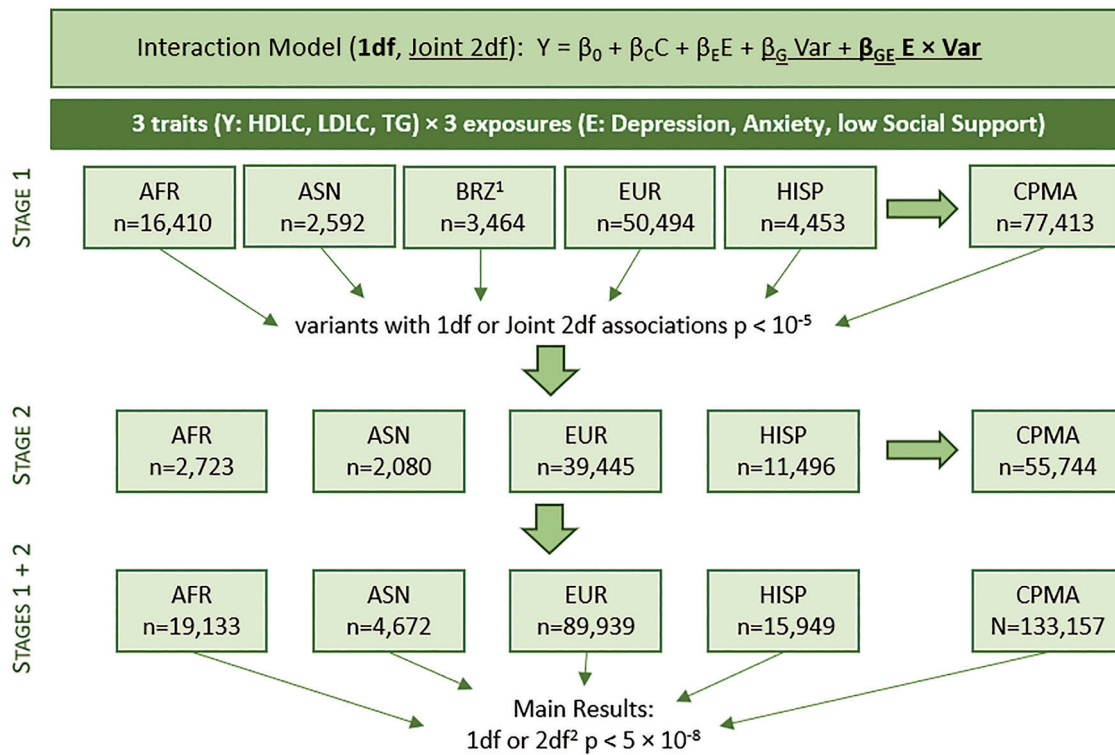


Fig. 1 Study design. African ancestry (AFR), Asian ancestry (ASN), Brazilian (BRZ), European ancestry (EUR), Hispanic (HISP), cross-population meta-analysis (CPMA) ¹ Brazilian samples were only available in Stage 1, so a Stage 1 + 2 meta-analysis of BRZ was not possible. These samples are include in the cross-population Stage 1 + 2 meta-analysis; ² As the 2df results jointly measure the variant's main and interaction effects, our main results only include those 2df findings that are also more than 500 kb from known lipids loci.

effect only meta-analysis, no association between rs59808825 and LDLC was observed ($P = 0.13$). Of the 8 loci identified through the 2df test, two associations had $P_{1df} \geq 0.1$, suggesting that the 2df test may reflect novel main effect associations (though we cannot exclude that G×Psy interactions may contribute to these findings but are undetectable at the current sample sizes). For instance, the joint 2df test was significant for rs34636484 (*CD96*) × SOCS on LDLC ($P_{2df} = 3.1E-8$), while the 1df test was not ($P_{1df} = 0.27$; Supplementary Figure 4C). rs11702544 (*RRP1B*) × DEPR on HDLC was also plausibly driven by a main effect (Supplementary Figure 4H). Importantly, both of these potential main effect associations were identified in the CPMA results, highlighting the importance of including diverse populations for novel discoveries.

The inclusion of underrepresented population groups in this study also provided an advantage in identifying novel interaction associations, with associations observed at four loci at which no data from EUR studies were available because of a minor allele frequency < 0.01 . For instance, an interaction of rs11949029 (*CTC-207P7.1*) and SOCS on LDLC was statistically significant for both the 1df test of interaction (Fig. 2A) and the 2df test of interaction and main effect (Supplementary Figure 4D) and was driven by data from the AFR and HISP populations. In this case, there was consistency of both the main and interaction associations across backgrounds. Such consistency was common among these lead findings; however, there were a few associations that were driven predominantly by one population, despite the availability of data for other groups. In the interaction between rs61248562 (*UNC13C*) and DEPR on HDLC, the observed association among EUR reached statistical significance ($\beta = 0.14$; $SE = 0.025$; $P_{1df} = 5.2 \times 10^{-9}$; Supplementary Figure 3C), yet this association was not seen in other populations despite a comparable number of samples with available data (EUR 15,052 vs. AFR 13,069 and HISP 15,977) and larger effect allele frequencies (EUR 0.02 vs. AFR 0.16 and HISP 0.07). As expected, the CPMA for this association was greatly

reduced in statistical significance ($P_{1df} = 1.5 \times 10^{-3}$). Similarly, the rs73597733 (*MACROD2*) × DEPR interaction on HDLC in AFR ($P_{1df} = 8.4 \times 10^{-9}$) was not seen in HISP ($P_{1df} = 0.43$; Supplementary Figure 3D).

Of the 10 lead associations in 9 loci that reached genome-wide statistical significance in the meta-analyses of Stages 1 and 2 (for one locus there were significant associations in two population groups), 4 were significant in both stages ($P < 0.05$) while 6 were only significant in stage 1 (Supplementary Table 5). There were 16 variants in 9 loci that were considered as the novel associated variant set for annotation and follow-up: those associated with either $P_{1df} < 5 \times 10^{-8}$ (seven variants in four loci) or $P_{2df} < 5 \times 10^{-8}$ and independent of known lipids loci (14 variants in 8 loci; five variants in three loci overlapping in 1df and 2df findings). The novel associated variants were characterized using FUMA. As expected, most of the variants were annotated to be intronic ($n = 10$) or intergenic ($n = 3$; Supplementary Table 6). While a single signal was detected for most of the described loci in Tables 1 and 2, the associated region on chromosome 21 (CPMA-HDL-DEPR) had three independent genomic signals at variants rs11702544, rs6518309, and rs9977076. Each of these variants is an eQTL for three genes in a variety of tissues, including whole blood: *PDXK*, *RRP1B*, and *HSF2BP* [40] (Supplementary Table 7).

We also evaluated 257 variants in LD with our lead variants ($R^2 \geq 0.6$ in 1000 Genomes, Phase 3 ALL; Supplementary Table 8). Evaluation of these variants in RegulomeDB identified 75 variants (29.2%) with functional prediction scores ≤ 3 , indicating moderate to high potential for regulatory effects. Variants within the locus on chromosome 21 characterized by rs11702544, rs6518309, and rs9977076 (*RRP1B*) had the lowest RegulomeDB scores in this set: 1a ($n = 1$) and 1b ($n = 6$), which indicates that they are likely to affect transcription factor binding to the gene targets, in this case *HSF2BP*, *RRP1B*, or *LINC00313*. These variants were also tested in our data and nearly reached statistical significance for the 2df

Table 1. Genome-wide variants significantly associated with serum lipids using a 1 df test of interaction in Meta-Analysis of Stages 1 & 2.

rsid Index Variant Nearest Gene ^a	Chr:BP (GRCh37)	Tested Allele: Freq	Main Effect	SE	Interaction Effect	SE	P _{1df} ^b	P _{2df} ^b	Population, Sample Size	Lipid ^c	Psychosocial Factor
rs11949029 RNU4-73P	5:163470425	A:0.02	6.6	2.2	-19.2	3.5	4.1E-08	2.4E-08	CPMA, n = 16,927	LDLC	SOCS
rs59808825 GRAMD1B	11:123152496	C:0.03	5.0	2.1	-22.9	4.2	4.3E-08	8.8E-09	CPMA, n = 5,973	LDLC	ANXT
rs61248562 RP11-643A5.2	15:54249794	T ^d :0.02	-0.024	0.015	0.14	0.025	5.2E-09	6.9E-08	EUR, n = 15,052	HDLC	DEPR
rs73597733 MACROD2	20:15941137	A:0.04	-0.015	0.010	0.14	0.025	8.4E-09	1.0E-08	AFR, n = 11,234	HDLC	DEPR

^aNearest gene determined based on ANNOVAR annotations implemented in FUMA. While the “Nearest Gene” is listed for simplicity, we acknowledge that the nearest gene may not be of functional relevance for the given variant.

^bBolded values indicate $P < 5 \times 10^{-8}$.

^cAnalyses were conducted on natural log-transformed HDLC and TG values, while LDLC was untransformed (all original values in mg/dL).

^dInsertion.

interaction with DEPR on HDLC (P_{2df} range 2.4×10^{-6} to 4.2×10^{-7}), with similar effect sizes in all.

Next, we assessed the predicted chromatin state around our 16 novel associated variants using the minimum 15-core chromatin state models calculated across 127 tissue/cell types [41]. We identified histone chromatin markers in regions associated with strong transcription ($n = 6$; Supplementary Table 6). In the 257 variants in LD with our lead variants, there were histone chromatin markers consistent with active ($n = 13$) or flanking active ($n = 21$) transcription start sites, transcription at the 3' or 5' end ($n = 7$), or in regions associated with strong transcription ($n = 50$) (Supplementary Table 8). For most of our loci, significant chromatin interactions were detected between regions containing our variants and regions overlapping gene promoters (Supplementary Table 9); for instance, between the locus on chromosome 21 (lead variant rs11702544) and regions overlapping the promoter of multiple genes, including *PDXK*, *RR1BP*, and *HSF2BP*.

Finally, to explore the potential clinical relevance of our findings, we performed an integrated druggability analysis of identified genes, as previously described [42]. We queried high and medium priority candidate gene targets (identified by FUMA and OpenTargets) using the Drug-Target Interaction database (DGIdb), which revealed 2 genes annotated as clinically actionable or members of the druggable genome (Supplementary Table 10). Several of these gene targets are implicated in ion transport (*NKAIN3*), vitamin metabolism (*PDXK*), and immune or viral response (*CD96*, *RRP1B*) pathways. We identified 1 gene, *RRP1B*, with a reported drug interaction. *RRP1B* was shown to interact with an FDA-approved drug, Atenolol, that has been evaluated in late-stage clinical trials using DrugBank and ClinicalTrials.gov databases (Supplementary Table 10). Atenolol is a well-established anti-hypertensive drug used to treat high blood pressure, heart failure, or angina in some patients. Together these results suggest a potential drug repurposing opportunity to intervene in a common pathway implicated in cardiometabolic disorders.

DISCUSSION

In this study, we investigated genome-wide variant-by-psycho-social factor interactions (G×Psy) in large, multi-ancestry meta-analyses of serum lipids. We identified nine novel lipid loci using this strategy, including four loci based on the 1df test of interaction and eight loci based on the 2df joint test of interaction and main effects (with three loci significantly associated using both strategies). Importantly, most of these associations could not have been identified in a standard GWAS that does not take interaction into account. Our inclusion of relatively large sample sizes representing diverse ancestries facilitated novel findings. Functional annotation highlights the promise of some of these identified loci for understanding the potential influence of psychosocial factors on serum lipids.

Both the 1df test of interaction and the 2df test of main effect and interaction identified statistically significant results for rs73597733 (intronic to *MACROD2*) × DEPR on HDLC, in which the main effect of the variant was near zero, with a large positive association among those with depressive symptoms. Intriguingly, an interaction between an intronic variant in *MACROD2* (not in LD with rs73597733) was previously found between thiazide diuretic use and HDLC [43]. Other intronic variants in *MACROD2* have been associated with the ceramides and sphingomyelins, suggesting a potential role in lipids pathways [44]. There is a large body of evidence for associations of intronic variants in *MACROD2* with complex psychosocial, neurological, and psychiatric traits, including: attention deficit hyperactivity disorder [45, 46], morningness (being a morning person) [47], risk-taking behavior [48], eating disorders [49], autism [50–52], and bipolar disorder [52, 53]. Infants with atypical neonatal neurobehavioral scores had differentially methylated CpG sites within the *MACROD2* gene [54]. *Macrod2*

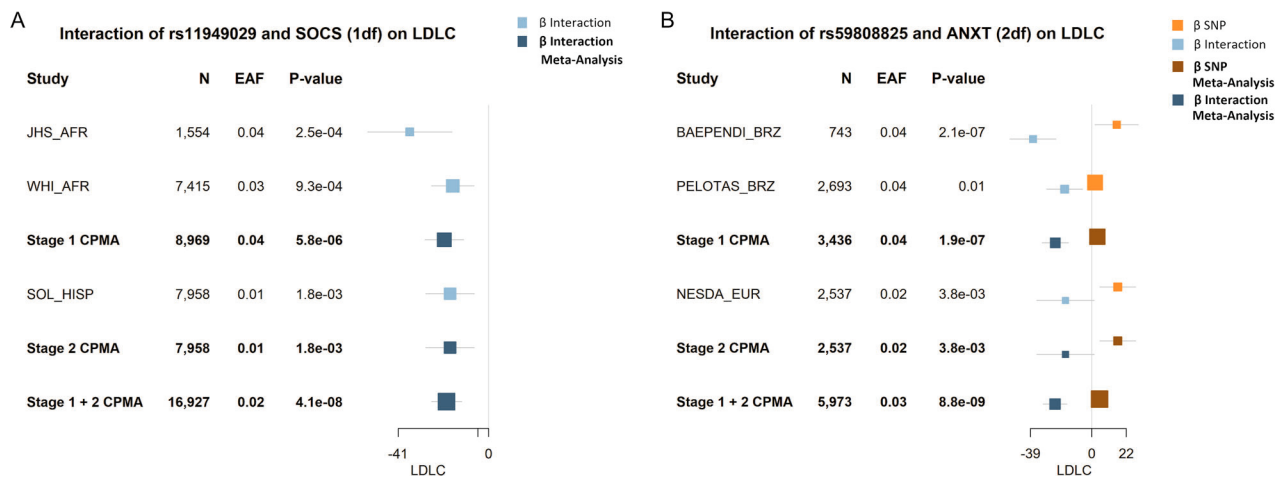


Fig. 2 Forest Plots of Key Findings. Forest plot showing all studies contributing data on an interaction of **A.** rs11949029 $\times$ social support (SOCS) on LDLC using a 1df test of the interaction term (this interaction was also statistically significant using the 2df joint test of the main effect and interaction, shown in Supplementary Figure 4D); and **B.** rs59808825 $\times$ anxiety symptoms (ANXT) on LDLC. Box size represents the precision of the estimate, with larger boxes shown for results with lower variance. Abbreviations: African ancestry (AFR), Brazilian (BRZ), Effect Allele Frequency (EAF), European Ancestry (EUR), Hispanic (HISP), cross-population meta-analysis (CPMA).

knockout mouse models displayed hyperactivity that became more pronounced with age [55]. Intronic variants in *MACROD2* have also been associated with measures of cognitive ability [56–58] and a variety of brain measurements [59–63]. Given the significant evidence for the involvement of this gene in a range of complex psychological and psychiatric phenotypes and a previous finding for an interactive effect on HDLC, our reported finding of an interaction between an intronic variant in *MACROD2* and DEPR on HDLC seems of particular interest and worthy of further investigation.

An association between rs59808825 (110 kb upstream of *GRAMD1B*) and ANXT on LDLC was $P < 5 \times 10^{-8}$ for both the 2df joint test of main effect and interaction and the 1df test of interaction, and no association was observed for this variant in an analysis without interaction modeled. *GRAMD1B* was identified as a locus for schizophrenia in multiple studies [64–69], a condition that has been linked with anxiety [70, 71]. The protein encoded by *GRAMD1B*, Gramd1b or Aster-B, has a role in cholesterol homeostasis, transporting accessible cholesterol from the plasma membrane to the endoplasmic reticulum [72, 73]. It was recently discovered that Aster proteins including Aster-B are key players in dietary lipid absorption in mice: the systemic absorption of dietary cholesterol was reduced by treatment with a small-molecule Aster inhibitor and mice without intestinal Aster proteins were protected from diet-induced hypercholesterolemia [74]. While further investigation is needed to propose a biological mechanism that might underlie the observed interaction between this variant and ANXT on LDLC, the known associations between nearby *GRAMD1B* with both complex psychiatric and psychological phenotypes and absorption of dietary lipid are intriguing.

We identified a 2df interaction of DEPR with variants on chromosome 21 (lead variant rs11702544 [*RRP1B*]) that appeared to represent a novel main effect of a common variant on HDLC. Interestingly, there was some evidence for an association of rs11702544 with HDLC using a standard GWAS model in the recent Global Lipids Genetics Consortium results ($P = 2.2 \times 10^{-6}$) [1], consistent with the contribution of a main effect of this variant contributing to the 2df joint test of main effect and interaction. FUMA annotation identified 3 independent genomic loci in this region, each of which is an eQTL for *PDXK*, *RRP1B*, and *HSF2BP*. Each of the genes has been previously associated with risk of diseases for which serum lipids concentration is a key risk factor: *PDXK* and *RRP1B* with coronary artery disease [75] and *HSF2BP* with cardiovascular disease [76]. *PDXK* encodes a protein essential

for the generation of the active form of Vitamin B₆. *PDXK* mRNA levels in adipose tissue were strongly associated with adipogenic, lipid-droplet-related, and lipogenic genes, and administration of the active form of Vitamin B₆ led to increased adipogenic markers in adipocyte precursor cells [77]. While the role for variants in this locus in HDLC concentration is not clear, they have been shown to affect *PDXK* expression, which could affect HDLC concentration through the expression of genes involved in lipogenesis. Our druggability analysis also identified *PDXK* as part of the druggable genome. *RRP1B* is a target gene that interacts with the beta-blocker drug Atenolol, which is sometimes used to treat hypertension and chronic angina.

We also identified an association using the 2df test rs34636484 (*CD96*) $\times$ SOCS on LDLC. The main effect appeared to contribute more to the association than the interaction at this locus, as the association was also apparent in a standard GWAS model in our data and the 1df test of interaction was not significant ($P = 0.27$). Based on these results, the association at rs34636484 appears to represent a novel main effect locus; however, this result should be interpreted with caution. The association of rs34636484 and LDLC was recently evaluated in the Global Lipids Genetics Consortium with a much larger sample size ($n = 1,393,230$ at this locus) and was not statistically significant ($P = 0.029$) [1].

Some of our significant associations were fairly consistent across studies within the same population group, but with no compelling evidence of association in other population groups, despite the availability of data. For instance, rs61248562 (*UNC13C*) $\times$ DEPR on HDLC was significant only among EUR, and not AFR or HISP in whom allele frequencies were higher and sample sizes were comparable. Similarly, an interaction of rs73597733 (*MACROD2*) $\times$ DEPR on HDLC in AFR was not seen in HISP at similar sample sizes (with a slightly lower allele frequency). It is unclear why these associations may differ by population group, but this phenomenon has been reported in previous gene-lifestyle interaction publications [6, 78, 79]. Differences in gene-lifestyle interactions across populations may arise from genomic factors, such as variations in linkage disequilibrium that lead to the tagging of different variants, as well as from lifestyle factors, such as differences in the measurement of or the experience of the psychosocial factor or in the behaviors or conditions associated with that psychosocial factor.

Psychosocial factors are complex traits that are associated with a variety of other factors, including some lifestyle exposures that we have previously evaluated using the same genome-wide

Table 2. Genome-wide variants significantly associated with serum lipids using a 2df joint test of main effect and interaction in Meta-Analysis of Stages 1 & 2 and were independent of known lipids loci^a.

rsid Index Variant Nearest Gene ^b	ChrBP (GRCh37)	Tested Allele: Freq	Main Effect	SE	Interaction Effect	SE	P _{1df} ^c	P _{2df} ^c	Population, Sample Size	Lipid ^d	Psychosocial Factor
rs6730082 RNU4-73P	2:9877621	C:0.26	-0.0038	0.0040	-0.045	0.0040	3.1E-06	2.5E-08	AFR, n = 16,886	HDLc	DEPR
rs142378953 AC090043.1	3:1695235	G:0.01	0.039	0.032	0.24	0.060	6.0E-05	4.9E-08	CPMA, n = 18,911	TG	DEPR
rs34636484 CD96	3:111356092	G:0.03	-4.5	0.80	1.8	1.6	0.27	3.1E-08	CPMA, n = 42,162	LDLC	SOCs
rs11949029 CTC-207P7.1	5:163470425	A:0.02	6.6	2.2	-19.0	3.5	4.1E-08	2.4E-08	CPMA, n = 16,927	LDLC	SOCs
rs4562311 NKAIN3	8:63589187	G:0.22	0.013	0.014	-0.13	0.032	3.5E-05	8.5E-09	CPMA, n = 6,667	TG	ANXT
rs59808825 GRAMD1B	11:123152496	C:0.03	5.0	2.1	-22.9	4.2	4.3E-08	8.8E-09	CPMA, n = 5,973	LDLC	ANXT
rs73597733 MACROD2	20:15941137	A:0.04	-0.015	0.010	0.14	0.025	8.4E-09	1.0E-08	AFR, n = 11,234	HDLc	DEPR
rs11702544 RBP1B	21:45091861	C:0.39	0.0049	0.0010	0.0043	0.0027	0.1	2.3E-08	CPMA ^e /EUR, n = 100,182	HDLc	DEPR

^aListed variants were more than 500 KB from 95% Credible Sets reported in Graham et al. (<https://doi.org/10.1038/s41586-021-04064-3>).
^bIndex variant is for the most statistically significant association in the region and Nearest Gene is determined based on ANNOVAR annotations implemented in FUMA. While the "Nearest Gene" is listed for simplicity, we acknowledge that the nearest gene may not be of functional relevance for the given variant.
^cBolded values indicate $P < 5 \times 10^{-8}$.
^dAnalyses were conducted on natural log-transformed HDLC and TG values, while LDLc was untransformed (all original values in mg/dL).
^eThe most statistically significant analysis (indicated) is displayed in table. Full results for all genome-wide significant associations given in Supplementary Table S4.

interaction study approach. Overlap in the interaction results for this study and previous analyses for one of these associated lifestyle factors could be very informative for disentangling the mechanism underlying these statistical interactions. We compared our statistically significant findings with those that we have previously reported for genome-wide interactions of smoking [6], alcohol intake [8], physical activity [7], educational attainment [9], and sleep duration [10] on serum lipids; no overlap among the results was identified. If one of our loci were found to be associated with a psychosocial factor, that could provide additional context into the relationship between psychosocial factors and serum lipids. To explore this possibility, we evaluated recent GWAS for these traits [29, 33, 80–84], but did not identify any overlap with our loci of interest.

Some of the strengths of this study include the relatively large sample sizes for a study of psychosocial factors, with analyses including up to 133,157 individuals. Also notable was the particular attention to the inclusion of non-European ancestry individuals (reaching over 19,000 AFR and nearly 16,000 HISP, although the number of ASN and BRZ was smaller, <5000 per population). The sample sizes for the non-European ancestry groups, however, were relatively small in size, particularly in terms of the statistical power needed for a gene-environment interaction study. We used a two-stage design with both a 1df test of interaction and a 2df joint test of main effect and interaction, an approach that is well-established for the study of gene-lifestyle interactions [6–10, 78, 79, 85, 86]. Our study also has some limitations. First, we had a smaller sample size for Stage 2, particularly for certain populations; as a result, the power for our two-stage approach was reduced. Second, despite our best efforts to harmonize psychosocial factors, the use of different instruments to measure these outcomes may have resulted in heterogeneity among studies, which would have reduced the power to identify lipids loci. In addition, these phenotypes themselves are quite complex and heterogeneous, and that complexity is not reflected in our categorization. Moreover, although our sample size is large for a study of lipids and psychosocial factors, it is not large enough to enable correction for multiple testing with adequate statistical power, and so its results need further validation. We did not have enough statistical power to usefully evaluate differences in these interactions by sex, which may prove to be of interest, as there are differences in the pathophysiology of cardiovascular disease by sex, and women experience a greater burden of depression [87]. Additionally, the association between TG and depressive symptoms has been shown to differ by sex, with men showing a stronger association [88], and low social support had a greater adverse effect on cardiovascular disease prevention among men than women [89]. Evaluating these interactions would require much greater sample sizes than were available in the current study. Although we have organized our contributing studies into population groups, there is likely to be meaningful heterogeneity within those groups in terms of relevant environmental background. For instance, the East Asian population group included individuals living in China as well as individuals with ancestry in China living in the United States. Information regarding neuropsychiatric medication use was not collected, though it is possible that use of these medications might directly or indirectly influence serum lipid levels [90]. In silico functional annotation and druggability analyses identified loci and candidate drug-gene interactions that are of interest for further follow-up; future experimental studies are needed to validate these findings.

In summary, we identified novel lipids loci in this large, multi-ancestry meta-analyses of genome-wide interaction studies of variants and psychosocial factors. Understanding these loci may help to disentangle the complex interplay between factors such as anxiety, depression, and low social support on serum lipids, a key biomarker of cardiometabolic risk.

MATERIALS AND METHODS

Study design

We adopted a two-stage study design (Fig. 1) that was implemented according to the Gene-Lifestyle Interactions Working Group of the CHARGE consortium [5]. We included men and women aged between 18 to 80 years of age with available data on lipids and psychosocial factors, and with genotype data imputed to the 1000 Genomes reference panel.

Stage 1 included 77,413 individuals in 31 study/population groups. Each study conducted genome-wide analyses (GWAS) incorporating a variant-by-psychosocial factor multiplicative interaction term. Centralized quality control was carried out, which was followed by a meta-analysis within and across five population groups: African ancestry (AFR), Asian ancestry (ASN), Brazilian (BRZ), European ancestry (EUR), and Hispanic (HISP). Variants that showed suggestive ($P < 10^{-5}$) associations for either a 1df test of interaction or a 2df joint test of interaction and main effect were carried forward for evaluation in Stage 2. Stage 2 analyses included data on 55,744 individuals from 19 studies distributed in 4 population groups. As no BRZ samples were included in Stage 2, no population-specific Stage 1 + 2 meta-analysis was undertaken, though the BRZ samples were included in cross-population meta-analyses (CPMA). Analytical details (Supplementary Table S1) and descriptive statistics (Supplementary Table S2) of each participating study for Stages 1 and 2 are provided.

Phenotypes and lifestyle variables studied

Analyses were conducted separately for three lipid parameters: HDLC, LDLC, and TG. HDLC and TG were directly assayed and natural log-transformed prior to analysis. LDLC was either directly assayed or derived using the Friedewald equation: $LDLC = TG - HDLC - (TG / 5)$, if $TG \leq 400$ mg/dL [91]. If a sample was drawn from an individual who had not been fasting for at least 8 h, then neither TG nor derived LDLC values were used. LDLC values were adjusted for lipid-lowering medication use (defined as the use of a statin or of any unspecified lipid-lowering medication after 1994, when statin usage became common). If LDLC was directly assayed, adjustment for lipid-lowering drugs was performed by dividing the LDLC value by 0.7. If LDLC was derived using the Friedewald equation, total cholesterol was first adjusted for lipid-lowering drug use (total cholesterol/0.8) before calculation of LDLC. No adjustments were made for any other lipid medication, nor were adjustments made to HDLC or triglycerides for medication use. For longitudinal studies where multiple lipid measurements were available, analysts selected the measurement with the largest sample size for analysis.

The three psychosocial variables (elevated depressive symptoms [DEPR], low social support [SOCS], and elevated anxiety symptoms [ANXT]) were measured within each cohort using validated screening questionnaires and coded as binary (yes/no) variables. A standard cut point was used for DEPR and ANXT, and SOCS was defined based on the lowest quartile of perceived social support. Further details regarding the instruments used within each study are given (Supplementary Table 3). Where multiple measurements of psychosocial factors were available, we used the questionnaire administered concomitantly with the measurement of serum lipids.

Genotyping and imputation

To harmonize data across studies, all studies imputed to 1000 Genomes data. Details on genotyping and imputation for each of the included studies are given in Supplementary Table 1. Most studies used Affymetrix (Santa Clara, CA, USA) or Illumina (San Diego, CA, USA) arrays and imputed to the cosmopolitan reference panel of the 1000 Genomes Project Phase I Integrated Release Version 3 Haplotypes. Prior to analysis, studies excluded all variants with minor allele frequency < 0.01 or those that mapped to the X and Y chromosomes or the mitochondria.

Study-level genome-wide analysis

Each cohort participating in Stage 1 analysis regressed serum lipids (Y) on the variant (G), psychosocial factor (E), and their interaction (GxE), with adjustment for covariates (C) including age, sex, principal components, and study-specific variables (listed for each study in Supplementary Table 1):

$$Y = \beta_0 + \beta_G G + \beta_E E + \beta_{G \times E} G \times E + \beta_C C$$

The 1df test was based on the null hypothesis $H_0: \beta_{G \times E} = 0$, while the 2df test was based on $H_0: \beta_G = \beta_{G \times E} = 0$. [92] To ensure robust estimates of covariance matrices and robust standard errors, studies of unrelated

subjects used either the sandwich R package or ProbABEL genetic software [93]. Family studies used Mixed Model Analysis for Pedigrees and populations (MMA), a comprehensive mixed model program that provides an optimized and flexible platform incorporating a wide range of covariance structures. Stage 2 studies carried out the same regressions, but only on the variants that reached suggestive significance ($P < 10^{-5}$) in Stage 1 for any trait in population-specific or cross-population meta-analysis. For comparison, stage 1 studies also ran a main effect model (serum lipids as a function of the variant with adjustment for covariates and study-specific variables) and a main effect model additionally adjusted for the psychosocial factor.

Population groups

Appropriate selection of population descriptors is a matter of considerable discussion in the field and consensus regarding optimal terms has not yet emerged. In this work, contributing studies were subdivided into population groups based on where individuals included in those studies were expected to cluster genetically to reduce the potential for spurious findings due to population structure and to maximize the potential for discovery within a population. Inclusion of samples within a particular cluster of genetic similarity was based on consultation with study teams given their expertise and understanding of the study population. Our approach includes African ancestry (AFR), Asian ancestry (ASN), Brazilian (BRZ), European ancestry (EUR), and admixed Hispanic/Latino and Native American participants (HISP). The AFR population group includes sub-Saharan Africans as well as participants with predominantly African ancestry living in the United States. The ASN population group includes participants of predominantly Asian ancestry living in East Asia, Singapore, or the United States. The EUR population group includes participants with predominantly European ancestry living in Europe or the United States. The HISP population group includes admixed Hispanic/Latino and Native American participants living in the United States. Brazilian individuals (BRZ) were analyzed separately after consultation with local researchers regarding the genetic clustering of these participants.

Quality control and cross-population meta-analysis

We performed extensive study- and population-level quality control (QC) using the R package EasyQC for all GWAS results [94]. In study-level QC, allele frequencies for each study were compared visually to an ancestry-matched 1000 genomes reference panel to identify systematic errors in data preparation (no variants were excluded), and marker names were harmonized to ensure consistency across studies. Any resulting concerns were resolved in consultation with the contributing study. Variants were excluded if the imputation quality score was less than 0.5 or if $2 \times \text{MAF} \times N_{\text{exposed}} \times \text{imputation quality score}$ was less than 20. Population-level QC was also conducted prior to meta-analysis to check for any outliers among included studies, which might suggest improper trait transformation or model specification, among other things.

We then conducted population-specific and cross-population meta-analysis in Stage 1 using the approach developed by Manning et al. [95] and implemented in METAL [95, 96]. This method performs a joint meta-analysis of the variant and the GxPsy exposure regression coefficients and then uses a 2df test to identify genetic variants driven jointly by main and interaction effects. Additionally, we used the inverse-variance weighted meta-analysis implemented in METAL to meta-analyze GxPsy interaction coefficients alone using a 1df test. Variants in the Stage 1 meta-analysis had to be present in at least 2 cohorts or at least 3000 individuals for AFR and EUR, with a lower threshold ($n = 2000$) set for ASN, BRZ, and HISP because of the smaller number of individuals available in these ancestries. In Stage 2, we used the same approach as in Stage 1 to perform population-specific and cross-population meta-analyses. After combining results from Stages 1 and 2, variants with $P < 5.0 \times 10^{-8}$ for either the 2df joint test of the main effect and the interaction or the 1df test of the interaction were considered significant. Results with a heterozygosity p -value < 0.05 were evaluated further and excluded if results were driven by a single cohort.

The novelty of associated loci was determined by comparison to the recent Global Lipids Genetic Consortium results for GWAS meta-analyses including approximately 1.65 million individuals with notable inclusion of those of diverse ancestral backgrounds [1]. The 95% credible sets from the meta-analyses of all lipids traits (available at http://csg.sph.umich.edu/willer/public/glgc-lipids2021/results/credible_sets/) were compiled. Variants were considered novel if they were 500 kb from all variants listed in this list, as well as those reported in other major publications [34–39].

Identification of independent genomic loci and functional annotation

Identification of genetic loci related to each of the three serum lipids and functional annotation was accomplished using Functional Mapping and Annotation of GWAS (FUMA) v1.5.6 (<http://fuma.ctglab.nl/>) [97]. Variants were grouped into genomic loci using an $R^2 < 0.6$ (1000 Genomes, Phase 3 ALL as the reference population) and a merge distance of 250 kb. Functional annotation was conducted using output from the set of tools incorporated within FUMA, including RegulomeDB score, Combined Annotation Dependent Deletion (CADD) score [98], 15-core chromatin state (ChromHMM) [41, 99, 100], and expression Quantitative Trait Loci (eQTL) on the variants from lead associations as well as those in LD with those variants ($R^2 > 0.1$), using all tissues and all included databases (including GTex, Blood eQTL, BIOS, and BRAINEAC).

Druggability analysis

We first used the Drug-Gene Interaction database (DGIdb; v4.2.0) to query psychosocial factors-lipid interacting genes to determine the potential druggability of the candidate gene targets. We annotated genes for implicated pathways and functions using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. We annotated the druggability target categories and queried all interacting drugs reported in 41 databases (BaderLabGenes, CarisMolecularIntelligence, dGene, FoundationOneGenes, GO, HingoraniCasas, HopkinsGroom, HumanProteinAtlas, IDG, MskImpact, Oncomine, Pharos, RussLampel, Tempus, CGI, CIVIC, COSMIC, CancerCommons, ChEMBL, ChEMBLDrugs, ChEMBLInteractions, ClarityFoundationBiomarkers, ClarityFoundationClinicalTrial, DTC, DoCM, DrugBank, Ensembl, Entrez, FDA, GuideToPharmacology, JAX-CKB, MyCancerGenome, MyCancerGenomeClinicalTrial, NCI, OncoKB, PharmGKB, TALC, TEND, TTD, TdgClinicalTrial, Wikidata). We queried protein targets for available active ligands in ChEMBL. We queried gene targets in the druggable genome using the most recent druggable genome list established by the NIH Illuminating the Druggable Genome Project (<https://github.com/druggablegenome/IDGTargets>) available through the Pharos web platform. We also queried FDA-approved drugs, late-stage clinical trials, and disease indications in the DrugBank, ChEMBL, ClinicalTrials.gov databases and provided results for the top MESH and DrugBank indications and clinical trials.

DATA AVAILABILITY

All summary results are available in the GWAS Catalog with the following Accession IDs: AFR.HDLC.ANXT.2df (GCST90570645); AFR.HDLC.ANXT.1df (GCST90570646); AFR.HDLC.DEPR.2df (GCST90570647); AFR.HDLC.DEPR.1df (GCST90570648); AFR.HDLC.SOCS.2df (GCST90570649); AFR.HDLC.SOCS.1df (GCST90570650); AFR.LDLC.ANXT.2df (GCST90570651); AFR.LDLC.ANXT.1df (GCST90570652); AFR.LDLC.DEPR.2df (GCST90570653); AFR.LDLC.DEPR.1df (GCST90570654); AFR.TG.ANXT.2df (GCST90570655); AFR.TG.ANXT.1df (GCST90570656); AFR.TG.DEPR.2df (GCST90570657); AFR.TG.DEPR.1df (GCST90570658); ASN.HDLC.DEPR.2df (GCST90570659); ASN.HDLC.DEPR.1df (GCST90570660); ASN.LDLC.DEPR.2df (GCST90570661); ASN.LDLC.DEPR.1df (GCST90570662); ASN.TG.DEPR.2df (GCST90570663); ASN.TG.DEPR.1df (GCST90570664); EUR.HDLC.ANXT.2df (GCST90570665); EUR.HDLC.ANXT.1df (GCST90570666); EUR.HDLC.DEPR.2df (GCST90570667); EUR.HDLC.DEPR.1df (GCST90570668); EUR.HDLC.SOCS.2df (GCST90570669); EUR.HDLC.SOCS.1df (GCST90570670); EUR.LDLC.ANXT.2df (GCST90570671); EUR.LDLC.ANXT.1df (GCST90570672); EUR.LDLC.DEPR.2df (GCST90570673); EUR.LDLC.DEPR.1df (GCST90570674); EUR.LDLC.SOCS.2df (GCST90570675); EUR.LDLC.SOCS.1df (GCST90570676); EUR.TG.ANXT.2df (GCST90570677); EUR.TG.ANXT.1df (GCST90570678); EUR.TG.DEPR.2df (GCST90570679); EUR.TG.DEPR.1df (GCST90570680); EUR.TG.SOCS.2df (GCST90570681); EUR.TG.SOCS.1df (GCST90570682); HISP.HDLC.DEPR.2df (GCST90570683); HISP.HDLC.DEPR.1df (GCST90570684); HISP.HDLC.SOCS.2df (GCST90570685); HISP.HDLC.SOCS.1df (GCST90570686); HISP.LDLC.DEPR.2df (GCST90570687); HISP.LDLC.DEPR.1df (GCST90570688); HISP.LDLC.SOCS.2df (GCST90570689); HISP.LDLC.SOCS.1df (GCST90570690); HISP.TG.DEPR.2df (GCST90570691); HISP.TG.DEPR.1df (GCST90570692); HISP.TG.SOCS.2df (GCST90570693); HISP.TG.SOCS.1df (GCST90570694); CPMA.HDLC.ANXT.2df (GCST90570695); CPMA.HDLC.ANXT.1df (GCST90570696); CPMA.HDLC.DEPR.2df (GCST90570697); CPMA.HDLC.DEPR.1df (GCST90570698); CPMA.HDLC.SOCS.2df (GCST90570699); CPMA.HDLC.SOCS.1df (GCST90570700); CPMA.LDLC.ANXT.2df (GCST90570701); CPMA.LDLC.ANXT.1df (GCST90570702); CPMA.LDLC.DEPR.2df (GCST90570703); CPMA.LDLC.DEPR.1df (GCST90570704); CPMA.LDLC.SOCS.2df (GCST90570705); CPMA.LDLC.SOCS.1df (GCST90570706); CPMA.TG.ANXT.2df (GCST90570707); CPMA.TG.ANXT.1df (GCST90570708); CPMA.TG.DEPR.2df (GCST90570709); CPMA.TG.DEPR.1df (GCST90570710); CPMA.TG.SOCS.2df (GCST90570711); CPMA.TG.SOCS.1df (GCST90570712)

REFERENCES

- Graham SE, Clarke SL, Wu K-HH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature*. 2021;600:675–9. <https://doi.org/10.1038/s41586-021-04064-3>.
- Hoffmann TJ, Theusch E, Halder T, Ranatunga DK, Jorgenson E, Medina MW, et al. A large electronic-health-record-based genome-wide study of serum lipids. *Nat Genet*. 2018;50:401–13. <https://doi.org/10.1038/s41588-018-0064-5>.
- Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet*. 2018;50:1514–23. <https://doi.org/10.1038/s41588-018-0222-9>.
- Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2:73–80. <https://doi.org/10.1161/CIRCGENETICS.108.829747>.
- Rao DC, Sung YJ, Winkler TW, Schwander K, Borecki I, Cupples LA, et al. Multi-ancestry study of gene-lifestyle interactions for cardiovascular traits in 610 475 individuals from 124 cohorts: design and rationale. *Circ Cardiovasc Genet*. 2017;10:e001649 <https://doi.org/10.1161/CIRCGENETICS.116.001649>.
- Bentley AR, Sung YJ, Brown MR, Winkler TW, Kraja AT, Ntalla I, et al. Multi-ancestry genome-wide gene-smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. *Nat Genet*. 2019;51:636–48. <https://doi.org/10.1038/s41588-019-0378-y>.
- Kilpeläinen TO, Bentley AR, Noordam R, Sung YJ, Schwander K, Winkler TW, et al. Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nat Commun*. 2019;10:376 <https://doi.org/10.1038/s41467-018-08008-w>.
- de Vries PS, Brown MR, Bentley AR, Sung YJ, Winkler TW, Ntalla I, et al. Multi-ancestry genome-wide association study of lipid levels incorporating gene-alcohol interactions. *Am J Epidemiol*. 2019;188:1033–54. <https://doi.org/10.1093/aje/kwz005>.
- de Las Fuentes L, Schwander KL, Brown MR, Bentley AR, Winkler TW, Sung YJ, et al. Gene-educational attainment interactions in a multi-population genome-wide meta-analysis identify novel lipid loci. *Front Genet*. 2023;14:1235337 <https://doi.org/10.3389/fgene.2023.1235337>.
- Noordam R, Bos MM, Wang H, Winkler TW, Bentley AR, Kilpeläinen TO, et al. Multi-ancestry sleep-by-SNP interaction analysis in 126,926 individuals reveals lipid loci stratified by sleep duration. *Nat Commun*. 2019;10:5121 <https://doi.org/10.1038/s41467-019-12958-0>.
- Carney RM, Freedland KE. Depression and coronary heart disease. *Nat Rev Cardiol*. 2017;14:145–55. <https://doi.org/10.1038/nrcardio.2016.181>.
- Juonala M, Pulkki-Råback L, Elovainio M, Hakulinen C, Magnussen CG, Sabin MA, et al. Childhood psychosocial factors and coronary artery calcification in adulthood: the cardiovascular risk in young finns study. *JAMA Pediatrics*. 2016;170:466–72. <https://doi.org/10.1001/jamapediatrics.2015.4121>.
- Harshfield EL, Pennells L, Schwartz JE, Willett P, Kaptoge S, Bell S, et al. Association between depressive symptoms and incident cardiovascular diseases. *JAMA*. 2020;324:2396–405. <https://doi.org/10.1001/jama.2020.23068>.
- Li GH, Cheung CL, Chung AK, Cheung BM, Wong IC, Fok MLY, et al. Evaluation of bi-directional causal association between depression and cardiovascular diseases: a Mendelian randomization study. *Psychol Med*. 2022;52:1765–76. <https://doi.org/10.1017/s0033291720003566>.
- Wei YG, Cai DB, Liu J, Liu RX, Wang SB, Tang YQ, et al. Cholesterol and triglyceride levels in first-episode patients with major depressive disorder: a meta-analysis of case-control studies. *J Affect Disord*. 2020;266:465–72. <https://doi.org/10.1016/j.jad.2020.01.114>.
- Tkachev A, Stekolshchikova E, Vanyushkina A, Zhang H, Morozova A, Zozulya S, et al. Lipid alteration signature in the blood plasma of individuals with schizophrenia, depression, and bipolar disorder. *JAMA Psychiatry*. 2023;80:250–9. <https://doi.org/10.1001/jamapsychiatry.2022.4350>.
- Bot M, Milaneschi Y, Al-Shehri T, Amin N, Garmaeva S, Onderwater GLJ, et al. Metabolomics profile in depression: a pooled analysis of 230 metabolic markers in 5283 cases with depression and 10,145 controls. *Biol Psychiatry*. 2020;87:409–18. <https://doi.org/10.1016/j.biopsych.2019.08.016>.
- Brydges CR, Bhattacharyya S, Dehkordi SM, Milaneschi Y, Penninx B, Jansen R, et al. Metabolomic and inflammatory signatures of symptom dimensions in major depression. *Brain, Behavior, Immun*. 2022;102:42–52. <https://doi.org/10.1016/j.bbi.2022.02.003>.
- Xu C, Padilla V, Lozano S, Gamez D, Su BB, Wang X, et al. APOE gene associated with dementia-related traits, depression, and anxiety in the hispanic population. *Genes*. 2023;14:1405.
- Wang WW, Liu XL, Ruan Y, Wang L, Bao TH. Depression was associated with apolipoprotein E $\epsilon 4$ allele polymorphism: A meta-analysis. *Iran J Basic Med Sci*. 2019;22:112–7. <https://doi.org/10.22038/ijbms.2018.30825.7436>.

21. Zhang M, Chen J, Yin Z, Wang L, Peng L. The association between depression and metabolic syndrome and its components: a bidirectional two-sample Mendelian randomization study. *Transl Psychiatry*. 2021;11:633 <https://doi.org/10.1038/s41398-021-01759-z>.
22. Dregan A, Rayner L, Davis KAS, Bakolis I, Arias de la Torre J, Das-Munshi J, et al. Associations between depression, arterial stiffness, and metabolic syndrome among adults in the UK Biobank Population Study: a mediation analysis. *JAMA Psychiatry*. 2020;77:598–606. <https://doi.org/10.1001/jamapsychiatry.2019.4712>.
23. Datta BK, Coughlin SS, Gummadi A, Mehrabian D, Ansa BE. Perceived social support and cardiovascular risk among nonelderly adults in the United States. *Am J Cardiol*. 2023;209:146–53. <https://doi.org/10.1016/j.amjcard.2023.09.106>.
24. García-Ulloa AC, Miranda-Gil V, Díaz-Pineda M, Garnica-Carrillo MF, Serrano-Pérez NH, Tron-Gomez MS, et al. Comparison of metabolic, lifestyle and mental health parameters in people with diabetes and relatives with and without family support. *Diabetes Metab Syndr Obes: Targets Ther*. 2024;17:239–45. <https://doi.org/10.2147/dms0.5445890>.
25. Skogberg N, Castaneda AE, Agyemang C, Koponen P, Lilja E, Laatikainen T. The association of depressive and anxiety symptoms with the metabolic syndrome and its components among Russian, Somali, and Kurdish origin adults in Finland: a population-based study. *J Psychosom Res*. 2022;159:110944 <https://doi.org/10.1016/j.jpsychores.2022.110944>.
26. Milaneschi Y, Simmons WK, van Rossum EFC, Penninx BWJH. Depression and obesity: evidence of shared biological mechanisms. *Mol Psychiatry*. 2019;24:18–33. <https://doi.org/10.1038/s41380-018-0017-5>.
27. Zou L, Tian Y, Wang Y, Chen D, Lu X, Zeng Z, et al. High-cholesterol diet promotes depression- and anxiety-like behaviors in mice by impact gut microbe and neuroinflammation. *J Affect Disord*. 2023;327:425–38. <https://doi.org/10.1016/j.jad.2023.01.122>.
28. Yu H, Qin X, Yu Z, Chen Y, Tang L, Shan W. Effects of high-fat diet on the formation of depressive-like behavior in mice. *Food Funct*. 2021;12:6416–31. <https://doi.org/10.1039/D1FO00044F>.
29. Cheng B, Cheng S, Jia Y, Liu L, Meng P, Qi X, et al. Genome-wide association studies in non-anxiety individuals identified novel risk loci for depression. *Eur Psychiatry*. 2022;65:e38 <https://doi.org/10.1192/j.eurpsy.2022.32>.
30. Meng P, Ye J, Chu X, Cheng B, Cheng S, Liu L, et al. Associations between genetic loci, environment factors and mental disorders: a genome-wide survival analysis using the UK Biobank data. *Transl Psychiatry*. 2022;12:17 <https://doi.org/10.1038/s41398-022-01782-8>.
31. Pinakhina D, Yermakovich D, Vergasova E, Kasyanov E, Rukavishnikov G, Rezapova V, et al. GWAS of depression in 4520 individuals from the Russian population highlights the role of MAGI2 (S-SCAM) in the gut-brain axis. *Front Genet*. 2023;13:972196 <https://doi.org/10.3389/fgene.2022.972196>.
32. Schoeler T, Speed D, Porcu E, Pirastu N, Pingault J-B, Kutalik Z. Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nat Hum Behav*. 2023;7:1216–27. <https://doi.org/10.1038/s41562-023-01579-9>.
33. Silveira PP, Pokhvisneva I, Howard DM, Meaney MJ. A sex-specific genome-wide association study of depression phenotypes in UK Biobank. *Mol Psychiatry*. 2023;28:2469–79. <https://doi.org/10.1038/s41380-023-01960-0>.
34. Do R, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM, et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet*. 2013;45:1345–52. <https://doi.org/10.1038/ng.2795>.
35. Lanktree MB, Elbers CC, Li Y, Zhang G, Duan Q, Karczewski KJ, et al. Genetic meta-analysis of 15,901 African Americans identifies variation in EXOC3L1 is associated with HDL concentration[S]. *J Lipid Res*. 2015;56:1781–6. <https://doi.org/10.1194/jlr.P059477>.
36. Peloso GM, Auer PL, Bis JC, Voorman A, Morrison AC, Stitzel NO, et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet*. 2014;94:223–32. <https://doi.org/10.1016/j.ajhg.2014.01.009>.
37. Spracklen CN, Chen P, Kim YJ, Wang X, Cai H, Li S, et al. Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. *Hum Mol Genet*. 2017;26:1770–84. <https://doi.org/10.1093/hmg/ddx062>.
38. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature*. 2010;466:707–13. <https://doi.org/10.1038/nature09270>.
39. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet*. 2013;45:1274–83. <https://doi.org/10.1038/ng.2797>.
40. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348:648–60. <https://doi.org/10.1126/science.1262110>.
41. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348:648–60. <https://doi.org/10.1126/science.1262110>.
42. Kavousi M, Bos MM, Barnes HJ, Lino Cardenas CL, Wong D, Lu H, et al. Multi-ancestry genome-wide study identifies effector genes and druggable pathways for coronary artery calcification. *Nat Genet*. 2023;55:1651–64. <https://doi.org/10.1038/s41588-023-01518-4>.
43. de Las Fuentes L, Sung YJ, Sitlani CM, Avery CL, Bartz TM, Keyser C, et al. Genome-wide meta-analysis of variant-by-diuretic interactions as modulators of lipid traits in persons of European and African ancestry. *Pharmacogenomics J*. 2020;20:482–93. <https://doi.org/10.1038/s41397-019-0132-y>.
44. Montasser ME, Aslibekyan S, Srinivasasainagendra V, Tiwari HK, Patki A, Bagheri M, et al. An Amish founder population reveals rare-population genetic determinants of the human lipidome. *Commun Biol*. 2022;5:334 <https://doi.org/10.1038/s42003-022-03291-2>.
45. Rao S, Baranova A, Yao Y, Wang J, Zhang F. Genetic relationships between attention-deficit/hyperactivity disorder, autism spectrum disorder, and intelligence. *Neuropsychobiology*. 2022;81:484–96. <https://doi.org/10.1159/000525411>.
46. Karlsson Linnér R, Mallard TT, Barr PB, Sanchez-Roige S, Madole JW, Driver MN, et al. Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nat Neurosci*. 2021;24:1367–76. <https://doi.org/10.1038/s41593-021-00908-3>.
47. Jansen PR, Watanabe K, Stringer S, Skene N, Bryois J, Hammerschlag AR, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat Genet*. 2019;51:394–403. <https://doi.org/10.1038/s41588-018-0333-3>.
48. Karlsson Linnér R, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat Genet*. 2019;51:245–57. <https://doi.org/10.1038/s41588-018-0309-3>.
49. Kushima I, Imaeda M, Tanaka S, Kato H, Oya-Ito T, Nakatochi M, et al. Contribution of copy number variations to the risk of severe eating disorders. *Psychiatry Clin Neurosci*. 2022;76:423–8. <https://doi.org/10.1111/pcn.13430>.
50. Anney RJL, Ripke S, Anttila V, Grove J, Holmans P, Huang H, et al. Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol Autism*. 2017;8:21 <https://doi.org/10.1186/s13229-017-0137-9>.
51. Bacchelli E, Cameli C, Viggiano M, Iglizios R, Mancini A, Tancredi R, et al. An integrated analysis of rare CNV and exome variation in Autism Spectrum Disorder using the Infinium PsychArray. *Sci Rep*. 2020;10:3198 <https://doi.org/10.1038/s41598-020-59922-3>.
52. Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, et al. Multi-trait analysis for genome-wide association study of five psychiatric disorders. *Transl Psychiatry*. 2020;10:209 <https://doi.org/10.1038/s41398-020-00902-6>.
53. Bigdeli TB, Fanous AH, Li Y, Rajeevan N, Sayward F, Genovese G, et al. Genome-wide association studies of schizophrenia and bipolar disorder in a diverse cohort of US veterans. *Schizophrenia Bull*. 2021;47:517–29. <https://doi.org/10.1093/schbul/sbaa133>.
54. Everson TM, Marsit CJ, Michael O'Shea T, Burt A, Hermetz K, Carter BS, et al. Epigenome-wide analysis identifies genes and pathways linked to neurobehavioral variation in preterm infants. *Sci Rep*. 2019;9:6322 <https://doi.org/10.1038/s41598-019-42654-4>.
55. Crawford K, Oliver PL, Agnew T, Hunn BHM, Ahel I. Behavioural characterisation of MacroD1 and MacroD2 knockout mice. *Cells*. 2021;10:368.
56. Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM, et al. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Mol Psychiatry*. 2019;24:169–81. <https://doi.org/10.1038/s41380-017-0001-5>.
57. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat Commun*. 2018;9:2098 <https://doi.org/10.1038/s41467-018-04362-x>.
58. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018;50:1112–21. <https://doi.org/10.1038/s41588-018-0147-3>.
59. Fan CC, Loughnan R, Makowski C, Pecheva D, Chen C-H, Hagler DJ, et al. Multivariate genome-wide association study on tissue-sensitive diffusion metrics highlights pathways that shape the human brain. *Nat Commun*. 2022;13:2423 <https://doi.org/10.1038/s41467-022-30110-3>.
60. van der Meer D, Kaufmann T, Shadrin AA, Makowski C, Frei O, Roelfs D, et al. The genetic architecture of human cortical folding. *Sci Adv*. 2021;7:eabj9446 <https://doi.org/10.1126/sciadv.abj9446>.
61. Sha Z, Schijven D, Carrion-Castillo A, Joliet M, Mazoyer B, Fisher SE, et al. The genetic architecture of structural left-right asymmetry of the human brain. *Nat Hum Behav*. 2021;5:1226–39. <https://doi.org/10.1038/s41562-021-01069-w>.

62. Smith SM, Douaud G, Chen W, Hanayik T, Alfaro-Almagro F, Sharp K, et al. An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nat Neurosci*. 2021;24:737–45. <https://doi.org/10.1038/s41593-021-00826-4>.
63. Shadrin AA, Kaufmann T, van der Meer D, Palmer CE, Makowski C, Loughnan R, et al. Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology. *Neuroimage*. 2021;244:118603. <https://doi.org/10.1016/j.neuroimage.2021.118603>.
64. Lam M, Chen C-Y, Li Z, Martin AR, Bryois J, Ma X, et al. Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat Genet*. 2019;51:1670–8. <https://doi.org/10.1038/s41588-019-0512-x>.
65. Periyasamy S, John S, Padmavati R, Rajendren P, Thirunavukkarasu P, Gratten J, et al. Association of schizophrenia risk with disordered niacin metabolism in an Indian genome-wide association study. *JAMA Psychiatry*. 2019;76:1026–34. <https://doi.org/10.1001/jamapsychiatry.2019.1335>.
66. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet*. 2018;50:381–9. <https://doi.org/10.1038/s41588-018-0059-2>.
67. Ikeda M, Takahashi A, Kamatani Y, Momozawa Y, Saito T, Kondo K, et al. Genome-wide association study detected novel susceptibility genes for schizophrenia and shared trans-populations/diseases genetic effect. *Schizophrenia Bull*. 2019;45:824–34. <https://doi.org/10.1093/schbul/sby140>.
68. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*. 2022;604:502–8. <https://doi.org/10.1038/s41586-022-04434-5>.
69. Li Z, Chen J, Yu H, He L, Xu Y, Zhang D, et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat Genet*. 2017;49:1576–83. <https://doi.org/10.1038/ng.3973>.
70. Hall J. Schizophrenia — an anxiety disorder? *Br J Psychiatry*. 2017;211:262–3. <https://doi.org/10.1192/bjp.bp.116.195370>.
71. Bulbena A, Bulbena-Cabre A. Schizophrenia and anxiety: yes, they are relatives not just neighbours. *Br J Psychiatry*. 2018;213:498–498. <https://doi.org/10.1192/bjp.2018.126>.
72. Naito T, Saheki Y. GRAMD1-mediated accessible cholesterol sensing and transport. *Biochimica et Biophysica Acta - Mol Cell Biol Lipids*. 2021;1866:158957. <https://doi.org/10.1016/j.bbalip.2021.158957>.
73. Ercan B, Naito T, Koh DHZ, Dharmawan D, Saheki Y. Molecular basis of accessible plasma membrane cholesterol recognition by the GRAM domain of GRAMD1b. *EMBO J*. 2021;40:e106524. <https://doi.org/10.15252/emboj.2020106524>.
74. Ferrari A, Whang E, Xiao X, Kennelly JP, Romartinez-Alonso B, Mack JJ, et al. Aster-dependent nonvesicular transport facilitates dietary cholesterol uptake. *Science*. 2023;382:eadf0966. <https://doi.org/10.1126/science.adf0966>.
75. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat Genet*. 2022;54:1803–15. <https://doi.org/10.1038/s41588-022-01233-6>.
76. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am J Hum Genet*. 2019;104:65–75. <https://doi.org/10.1016/j.ajhg.2018.11.008>.
77. Moreno-Navarrete JM, Jove M, Ortega F, Xifra G, Ricart W, Obis E, et al. Metabolomics uncovers the role of adipose tissue PDXK in adipogenesis and systemic insulin sensitivity. *Diabetologia*. 2016;59:822–32. <https://doi.org/10.1007/s00125-016-3863-1>.
78. Sung YJ, de las Fuentes L, Winkler TW, Chasman DI, Bentley AR, Kraja AT, et al. A multi-ancestry genome-wide study incorporating gene–smoking interactions identifies multiple new loci for pulse pressure and mean arterial pressure. *Hum Mol Genet*. 2019;28:2615–33. <https://doi.org/10.1093/hmg/ddz070>.
79. Sung YJ, Winkler TW, de las Fuentes L, Bentley AR, Brown MR, Kraja AT, et al. A large-scale multi-ancestry genome-wide study accounting for smoking behavior identifies multiple significant loci for blood pressure. *Am J Hum Genet*. 2018;102:375–400. <https://doi.org/10.1016/j.ajhg.2018.01.015>.
80. Amare AT, Vaez A, Hsu YH, Direk N, Kamali Z, Howard DM, et al. Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. *Mol Psychiatry*. 2020;25:1420–9. <https://doi.org/10.1038/s41380-018-0336-6>.
81. Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD, et al. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature*. 2021;599:628–34. <https://doi.org/10.1038/s41586-021-04103-z>.
82. Li QS, Tian C, Hinds D. Genome-wide association studies of antidepressant class response and treatment-resistant depression. *Transl Psychiatry*. 2020;10:360. <https://doi.org/10.1038/s41398-020-01035-6>.
83. Mitchell BL, Campos AI, Whiteman DC, Olsen CM, Gordon SD, Walker AJ, et al. The Australian genetics of depression study: new risk loci and dissecting heterogeneity between subtypes. *Biol Psychiatry*. 2022;92:227–35. <https://doi.org/10.1016/j.biopsych.2021.10.021>.
84. Yakovchik AY, Tolynyova DV, Kashtanova DA, Sutulova ER, Ivanov MV, Mamchur AA, et al. Genetics of psycho-emotional well-being: genome-wide association study and polygenic risk score analysis. *Front Psychiatry*. 2023;14:1188427. <https://doi.org/10.3389/fpsy.2023.1188427>.
85. Wang H, Noordam R, Cade BE, Schwander K, Winkler TW, Lee J, et al. Multi-ancestry genome-wide gene–sleep interactions identify novel loci for blood pressure. *Mol Psychiatry*. 2021;26:6293–304. <https://doi.org/10.1038/s41380-021-01087-0>.
86. Feitosa MF, Kraja AT, Chasman DI, Sung YJ, Winkler TW, Ntalla I, et al. Novel genetic associations for blood pressure identified via gene–alcohol interaction in up to 570K individuals across multiple ancestries. *PLoS ONE*. 2018;13:e0198166. <https://doi.org/10.1371/journal.pone.0198166>.
87. Bucciarelli V, Caterino AL, Bianco F, Caputi CG, Salerni S, Sciomer S, et al. Depression and cardiovascular disease: The deep blue sea of women's heart. *Trends Cardiovasc Med*. 2020;30:170–6. <https://doi.org/10.1016/j.tcm.2019.05.001>.
88. Xu L, Wang K, Wang S, Liu L, Lv X, Song Y. Sex differences in the association between serum lipids and depressive symptoms: a longitudinal population-based study. *J Affect Disord*. 2021;291:154–62. <https://doi.org/10.1016/j.jad.2021.05.011>.
89. Blakoe M, Petrova D, Garcia-Retamero R, Gonçalves K, Catena A, Ramírez Hernández JA, et al. Sex moderates the relationship between social support and cardiovascular prevention behaviors in middle-aged and older adults. *Ann Behav Med*. 2023;57:877–87. <https://doi.org/10.1093/abm/kaad030>.
90. Olsson M, Marcus SC, Corey-Lisle P, Tuomari AV, Hines P, L'Italiani GJ. Hyperlipidemia following treatment with antipsychotic medications. *Am J Psychiatry*. 2006;163:1821–5. <https://doi.org/10.1176/ajp.2006.163.10.1821>.
91. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem*. 1972;18:499–502.
92. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene–environment interaction to detect genetic associations. *Hum Heredity*. 2007;63:111–9. <https://doi.org/10.1159/000099183>.
93. Aulchenko YS, Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinforma*. 2010;11:134. <https://doi.org/10.1186/1471-2105-11-134>.
94. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, et al. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc*. 2014;9:1192–212. <https://doi.org/10.1038/nprot.2014.071>.
95. Manning AK, LaValley M, Liu CT, Rice K, An P, Liu Y, et al. Meta-analysis of gene–environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet Epidemiol*. 2011;35:11–18. <https://doi.org/10.1002/gepi.20546>.
96. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinformatics*. 2010;26:2190–1. <https://doi.org/10.1093/bioinformatics/btq340>.
97. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8:1826. <https://doi.org/10.1038/s41467-017-01261-5>.
98. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46:310–5. <https://doi.org/10.1038/ng.2892>.
99. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods*. 2012;9:215–6. <https://doi.org/10.1038/nmeth.1906>.
100. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48:481–7. <https://doi.org/10.1038/ng.3538>.

ACKNOWLEDGEMENTS

This project was largely supported by two grants from the U.S. National Heart, Lung, and Blood Institute (NHLBI), the National Institutes of Health, R01HL118305 and R01HL156991. This research was supported in part by the Intramural Research Program of the National Human Genome Research Institute (NHGRI) in the Center for Research on Genomics and Global Health (CRGGH—Z01HG200362). CRGGH is also supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), the Center for Information Technology, and the Office of the Director at the National Institutes of Health. Tuomas O. Kilpeläinen was supported by grants from the Novo Nordisk Foundation (NNF18CC0034900, NNF21SA0072102,

NNF22OC0074128, NNF23SA0084103). W. Jim Gauderman was supported by the National Cancer Institute (NCI; P01 CA196569). The views expressed in this manuscript are those of the authors and do not necessarily represent the views of NHLBI, NHGRI, the NCI; the National Institutes of Health; or the U.S. Department of Health and Human Services. Study-specific acknowledgments are included in the Supplementary Note. The authors gratefully acknowledge the invaluable contributions of Adolfo Correa, Lisa W. Martin, Annette Peters, and Astrid Petersmann. The authors would also like to recognize L. Adrienne Cupples, Diane M. Becker, and John M. Starr, who participated in this effort, but passed away prior to the completion of the manuscript.

AUTHOR CONTRIBUTIONS

ARB and SKM drafted the manuscript with substantial contributions from MRB, KLS, TWW, CLM, DCR, CNR, MF, and ERF. HA, ARB, MRB, LFB, MF, MFF, ERF, LdIF, JG, ATK, TOK, C-TL, CLM, AKM, SKM, RN, JRO'C, MProvince, BMP, PAP, DCR, KR, SSR, CNR, YJS, KLS, PSdV, TWW, XZhu contributed to the conception and design of the work. HA, NA, DEA, DKA, ARB, MRB, LB, TMB, CB, EB, Y-DIC, J-FC, KNC, MLD, KWvD, JACD, JD, CMvD, LE, MKE, JDF, AMF, NF, MF, ERF, MFF, VG, LCG, MG, XG, CG, HJG, CAH, DvH, FPH, GH, ARVRH, SEH, SH, MAI, CRI, JBJ, LK-J, PK, SLRK, CK, JEK, C-TL, JL, KL, TAL, TL, LL, L-PL, EL, YL, D-OMK, CLM, ACM, AKM, AWM, SKM, PjvdM, PM-V, SEM, YM, IMN, RN, KEN, AJO, BWJHP, ACP, PAP, MPreisig, GP, CNR, MAR, JIR, RR, DCR, XS, PJS, LS, JMS, X-oS, JAS, HS, NS, KLS, MS, AVS, KDT, LT, MYT, AGU, DV, EW, MW, GW, SW, Y-XW, RBW, H-LW, DRW, W-BW, RW, WW, YW, JY, LRY, XZhang, XZhu, WZheng, WZhao, ABZ contributed in the acquisition, analysis, or interpretation of the data. MRB conducted centralized data cleaning and meta-analyses. TWW designed the software pipeline used for harmonization, visualization, and quality control of study data. DEA, HA, TMB, LFB, EB, KNC, JACD, LE, MKE, MFF, NF, AMF, LdIF, MG, JG, LG, HJG, ARVRH, DvH, SEH, CRI, MAI, ATK, C-TL, AIL, L-PL, PM-V, YM, SEM, RN, JRO'C, AJO, PAP, GP, MPreisig, MProvince, BP, KR, SSR, JIR, JAS, JMS, PJS, X-oS, EST, LT, PSdV, RBW, SW, H-LW, KLY, LRY, WZheng, WZhao, XZhu made substantial revisions to the manuscript. All authors approved the manuscript.

FUNDING

Open access funding provided by the National Institutes of Health.

COMPETING INTERESTS

The authors declare the following competing interests: BMP serves on the Data and Safety Monitoring Board of a clinical trial funded by the manufacturer (Zoll LifeCor) and on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson; CMvD was sponsored by GlaxoSmithKline; HJG received travel grants and speakers honoraria from Fresenius Medical Care,

Neuraxpharm and Janssen Cilag as well as research funding from Fresenius Medical Care; JBJ serves in the Advisory Board Novartis; Patent holder with Biocompatibles UK Ltd. (Fransham, Surrey, UK) (Title: Treatment of eye diseases using encapsulated cells encoding and secreting neuroprotective factor and / or anti-angiogenic factor; Patent number: 20120263794), and Europäische Patentanmeldung 16 720 043.5 and Patent application US 2019 0085065 A1 (Agents for use in the therapeutic or prophylactic treatment of myopia or hyperopia).

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

For all participating cohorts, the appropriate institutional review boards approved research activities and informed consent was obtained from each participant. Further details on all participating cohorts is given in the Supplementary Note.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-025-03418-z>.

Correspondence and requests for materials should be addressed to Amy R. Bentley.

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

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



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

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2025

¹Center for Research on Genomics and Global Health, National Human Genome Research Institute, Bethesda, MD 20892, USA. ²Department of Epidemiology, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA. ³MerexAgRight Farm, Ltd., Kitale, Kenya. ⁴Division of Statistical Genomics, Washington University, St. Louis, MO, USA. ⁵Department of Epidemiology, University of Regensburg, Regensburg 93051, Germany. ⁶Department of Social Medicine, Population and Public Health, University of California, Riverside School of Medicine, Riverside, CA 92521, USA. ⁷Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen DK-2200, Denmark. ⁸The Novo Nordisk Foundation Center for Genomic Mechanisms of Disease, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. ⁹Institut Pasteur, Université Paris Cité, Department of Computational Biology, F-75015 Paris, France. ¹⁰Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA. ¹¹Department of Biostatistics, University of Washington, Seattle, WA 98101, USA. ¹²Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA. ¹³Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, Singapore 117549, Singapore. ¹⁴Department of Engineering Technology, University of Houston, Sugar Land, TX 77479, USA. ¹⁵Department of Epidemiology, University of North Carolina, Chapel Hill, NC 27516, USA. ¹⁶Department of Pediatrics, The Institute for Translational Genomics and Population Sciences, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA. ¹⁷Department of Epidemiology, Federal University of Pelotas, Pelotas, RS 96020-220, Brazil. ¹⁸Heart Institute (InCor), University of São Paulo Medical School, São Paulo, SP 05403-900, Brazil. ¹⁹Department of Biostatistics, Boston University, Boston, MA 02118, USA. ²⁰Department of Medicine, Division of Cardiology, Duke University School of Medicine, Durham, NC 27701-2047, USA. ²¹Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston, MA 02114, USA. ²²Department of Medicine, Harvard Medical School, Boston, MA 02115, USA. ²³Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 9713 GZ, The Netherlands. ²⁴Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden 2333ZA, the Netherlands. ²⁵McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX 77030, USA. ²⁶Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA. ²⁷Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA. ²⁸Division of Biostatistics, Washington University, St. Louis, MO 63110, USA. ²⁹Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, the Netherlands. ³⁰Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden 2300 RC, The Netherlands. ³¹IQVIA, 1101 CT Amsterdam, the Netherlands. ³²Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. ³³Lothian Birth Cohorts, Department of Psychology, The University of Edinburgh, Edinburgh EH8 9JZ, UK. ³⁴Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland. ³⁵Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland. ³⁶Department of Psychiatry, Lausanne University Hospital and University of Lausanne, Prilly, Vaud 1008, Switzerland. ³⁷Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio 70100, Finland. ³⁸Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA. ³⁹Interfaculty Institute of Genetics and Functional Genomics, University Medicine Greifswald, Greifswald 17489, Germany. ⁴⁰German Center for Cardiovascular Research (DZHK), partner site Greifswald, Greifswald 17475, Germany. ⁴¹Division of Epidemiology, Vanderbilt University Medical Center, Nashville, TN 37203, USA. ⁴²Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. ⁴³McKusick-Nathans Institute, Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. ⁴⁴Department of Epidemiology

and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC 29169, USA. ⁴⁵Section of Cardiovascular Research, Baylor College of Medicine, Houston, TX 77030, USA. ⁴⁶Houston Methodist DeBakey Heart and Vascular Center, Houston, TX 77030, USA. ⁴⁷Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA. ⁴⁸Institute for Minority Health Research, University of Illinois Chicago, Chicago, IL 60612, USA. ⁴⁹Center for Biostatistics and Data Science, Institute for Informatics, Data Science, and Biostatistics (I2BD), Washington University School of Medicine, St. Louis, MO 63110, USA. ⁵⁰Division of Cardiology, Washington University, St. Louis, MO 63110, USA. ⁵¹Department of Epidemiology, University of Washington, Seattle, WA, USA. ⁵²Department of Medicine, University of Washington, Seattle, WA, USA. ⁵³Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA. ⁵⁴College of Pharmacy, University of Manitoba, Winnipeg, MB R3E 0T5, Canada. ⁵⁵Department of Medicine, University of Mississippi Medical Center, Jackson, MS 39213, USA. ⁵⁶Department of Psychology, San Diego State University, San Diego, CA, USA. ⁵⁷Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio Campus, Kuopio 70100, Finland. ⁵⁸Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA. ⁵⁹Institute of Molecular and Clinical Ophthalmology Basel IOB, Basel, Switzerland. ⁶⁰Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Ophthalmology and Visual Science Key Lab, Beijing Tongren Hospital, Capital Medical University, Beijing, China. ⁶¹Department of Psychology and Logopedics, University of Helsinki, Helsinki 00014, Finland. ⁶²University of Mississippi Medical Center, Jackson, MS 39213, USA. ⁶³Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA. ⁶⁴Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore 138672, Singapore. ⁶⁵Department of Medicine, Duke Molecular Physiology Institute Duke University School of Medicine, Durham, NC 27701-2047, USA. ⁶⁶Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA. ⁶⁷Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA. ⁶⁸Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA, USA. ⁶⁹Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 46, CH- 1011 Lausanne, Switzerland. ⁷⁰Department of Psychiatry, Amsterdam Neuroscience and Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam 1081 HJ, the Netherlands. ⁷¹Veritas Management Group, Inc., Alpharetta, GA 30072, USA. ⁷²Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, MD, USA. ⁷³Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, USA. ⁷⁴Department of Health Services Research, School of Medicine and Health Services, Carl von Ossietzky Universität Oldenburg, Oldenburg 26111, Germany. ⁷⁵Institute for Health Economics and Health Care Management, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany. ⁷⁶Division of Preventive Medicine, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, AL 35205, USA. ⁷⁷Cardiovascular Health Research Unit, Division of Cardiology, University of Washington, Seattle, Washington 98195, USA. ⁷⁸Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore 119228, Singapore. ⁷⁹Department of Pediatrics, LA Biomed at Harbor-UCLA Medical Center, Torrance, CA 90502, USA. ⁸⁰Public Health Sciences, Fred Hutchinson Cancer Center, Seattle, WA 98109, USA. ⁸¹Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN 55455, USA. ⁸²Department of Internal Medicine, Erasmus MC, University Medical Center, Rotterdam, The Netherlands. ⁸³Nuffield Department of Population Health, University of Oxford, Oxford, UK. ⁸⁴Research Unit Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany. ⁸⁵Department of Epidemiology, University of Iowa College of Public Health, Iowa City, IA, USA. ⁸⁶Beijing Tongren Eye Center, Beijing Key Laboratory of Intraocular Tumor Diagnosis and Treatment, Beijing Ophthalmology and Visual Sciences Key Laboratory, Beijing Tongren Hospital, Capital Medical University, Beijing, China. ⁸⁷Department of Human Genetics, Leiden University Medical Center, Leiden 2333ZA, The Netherlands. ⁸⁸Department of Internal Medicine, Division of Endocrinology, Leiden University Medical Center, Leiden 2333ZA, The Netherlands. ⁸⁹School of Public Health, Jackson State University, Jackson, MS 39213, USA. ⁹⁰Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH 44106, USA. ⁹¹Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany. ⁹²German Center for Diabetes Research (DZD), Partner München-Neuherberg, Neuherberg, Germany. ⁹³German Centre for Neurodegenerative Diseases (DZNE), Site Rostock/Greifswald, Greifswald 17475, Germany. ⁹⁴Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald 17489, Germany. ⁹⁵Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland. ⁹⁶Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen 9713 GZ, the Netherlands. ⁹⁷Division of Biostatistics, Department of Population and Public Health Sciences, University of Southern California, Los Angeles, CA 90032, USA. ⁹⁸Icelandic Heart Association, Icelandic Heart Association, Kópavogur 201, Iceland. ⁹⁹Faculty of Medicine, University of Iceland, Reykjavik, Iceland. ¹⁰⁰Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, RS 96020-220, Brazil. ¹⁰¹Department of Clinical Epidemiology, Leiden University Medical Center, Leiden 2333ZA, the Netherlands. ¹⁰²Department of Public Health and Primary Care, Leiden University Medical Center, Leiden 2333ZA, the Netherlands. ¹⁰³Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology, and Health Systems and Population Health, University of Washington, Seattle, USA. ¹⁰⁴Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA. ¹⁰⁹These authors contributed equally: Amy R. Bentley, Michael R. Brown, Charles N. Rotimi, Myriam Fornage, Ervin R. Fox. *A list of authors and their affiliations appears at the end of the paper. ✉email: amy.bentley@nih.gov

LIFELINES COHORT STUDY

Raul Aguirre-Gamboa¹⁰⁵, Patrick Deelen¹⁰⁵, Lude Franke¹⁰⁵, Jan A. Kuivenhoven¹⁰⁶, Esteban A. Lopera Maya¹⁰⁵, Ilja M. Nolte¹⁰⁷, Serena Sanna¹⁰⁵, Harold Snieder¹⁰⁷, Morris A. Swertz¹⁰⁵, Peter M. Visscher^{107,108}, Judith M. Vonk¹⁰⁷, Cisca Wijmenga¹⁰⁵ and Naomi Wray¹⁰⁸

¹⁰⁵Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. ¹⁰⁶Department of Pediatrics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. ¹⁰⁷Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. ¹⁰⁸Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, Australia.