




















# Distinct biological signature and modifiable risk factors underlie the comorbidity between major depressive disorder and cardiovascular disease

Received: 31 August 2023

Accepted: 8 May 2024

Published online: 17 June 2024

 Check for updates

Jacob Bergstedt <sup>1,24</sup>✉, Joëlle A. Pasman<sup>2,24</sup>, Ziyang Ma <sup>2,24</sup>, Arvid Harder<sup>2</sup>, Shuyang Yao <sup>2,3</sup>, Nadine Parker <sup>4</sup>, Jorien L. Treur<sup>5</sup>, Dirk J. A. Smit <sup>5</sup>, Oleksandr Frei <sup>4,6</sup>, Alexey A. Shadrin<sup>4,7</sup>, Joeri J. Meijisen <sup>8</sup>, Qing Shen<sup>1,9,10</sup>, Sara Hägg<sup>2</sup>, Per Tornvall<sup>11</sup>, Alfonso Buil <sup>8</sup>, Thomas Werge <sup>8,12</sup>, Jens Hjerling-Leffler <sup>3</sup>, Thomas D. Als <sup>13,14,15</sup>, Anders D. Børglum <sup>14,15</sup>, Cathryn M. Lewis <sup>16,17</sup>, Andrew M. McIntosh <sup>18,19</sup>, Unnur A. Valdimarsdóttir <sup>1,20,21</sup>, Ole A. Andreassen <sup>4,7</sup>, Patrick F. Sullivan <sup>2,22,23</sup>, Yi Lu <sup>2,24</sup> & Fang Fang <sup>1,24</sup>✉

Major depressive disorder (MDD) and cardiovascular disease (CVD) are often comorbid, resulting in excess morbidity and mortality. Here we show that CVDs share most of their genetic risk factors with MDD. Multivariate genome-wide association analysis of shared genetic liability between MDD and atherosclerotic CVD revealed seven loci and distinct patterns of tissue and brain cell-type enrichments, suggesting the involvement of the thalamus. Part of the genetic overlap was explained by shared inflammatory, metabolic and psychosocial or lifestyle risk factors. Our data indicated causal effects of genetic liability to MDD on CVD risk, but not from most CVDs to MDD, and showed that the causal effects were partly explained by metabolic and psychosocial or lifestyle factors. The distinct signature of MDD–atherosclerotic CVD comorbidity suggests an immunometabolic subtype of MDD that is more strongly associated with CVD than overall MDD. In summary, we identified biological mechanisms underlying MDD–CVD comorbidity and modifiable risk factors for prevention of CVD in individuals with MDD.

Major depressive disorder (MDD) and cardiovascular disease (CVD) are comorbid<sup>1,2</sup>. Several mechanisms might explain the observed comorbidity<sup>2</sup>. One explanation is that genetic risk factors for MDD and CVDs overlap<sup>3,4</sup>. While observed genetic correlations between MDD and CVD are modest<sup>2–4</sup>, this may reflect local genetic correlations of opposing directions attenuating correlation on the genome-wide level, leading to an underestimation of the genetic overlap<sup>5</sup>. The large polygenicity of MDD<sup>6</sup> might also mask subtypes with stronger genetic relationships to CVD.

The observed MDD–CVD comorbidity could also be due to nongenetic factors<sup>7</sup>. Cardiovascular risk factors such as high systolic blood pressure, high body mass index (BMI), high levels of low-density lipoprotein cholesterol, high levels of physical inactivity, presence of type II diabetes and smoking have all been associated with MDD<sup>8–10</sup>. Moreover, accumulating data show that psychosocial and/or lifestyle factors associated with MDD, such as low educational attainment, exposure to childhood maltreatment, loneliness and atypical sleep patterns, are also important risk factors for CVD<sup>11–14</sup>.

A full list of affiliations appears at the end of the paper. ✉ e-mail: [jacob.bergstedt@ki.se](mailto:jacob.bergstedt@ki.se); [fang.fang@ki.se](mailto:fang.fang@ki.se)

One common mechanism underlying MDD and CVD, as well as their shared risk factors, could be chronic inflammation. Atherosclerosis, the accumulation of fibrofatty lesions in the arterial wall, is the main cause of CVD<sup>15</sup>. The buildup of atherosclerotic plaque is a long-term inflammatory process mediated by immune components in crosstalk with arterial wall cells<sup>16</sup>. Many lines of evidence also support a role for inflammation in MDD<sup>17</sup>. Excessive or long-term psychosocial stress promote the maturation and release of inflammatory cytokines such as interleukin (IL)6, which activate the central nervous system to produce behaviors related to MDD<sup>17</sup>. Importantly, low-grade inflammation, measured by high C-reactive protein levels, has been observed in more than a quarter of patients with depression<sup>18</sup>, suggesting the presence of an inflammatory subtype of MDD<sup>19</sup>, which might be especially strongly associated with CVD.

The full extent of the genetic overlap between MDD and CVD has not been explored. It remains unknown whether the genetic overlap is associated with specific tissues or brain cell types, or how this overlap relates to shared risk factors such as blood pressure, psychosocial or lifestyle traits, metabolic traits, and inflammation. Moreover, causal effects linking these traits are not fully understood<sup>20–23</sup>.

In this Article, we dissect the genetic overlap between MDD and CVD (peripheral artery disease, heart failure, coronary artery disease, stroke and atrial fibrillation). First, we assessed genetic overlap between MDD and CVD on the genome-wide level, as well as on the level of local partitions of the genome and overlapping risk variants with MiXeR<sup>24</sup> and Local Analysis of (co)Variant Association (LAVA)<sup>25</sup>. Second, we identified genetic variants and genes that contribute to the shared genetic liability between MDD and atherosclerotic CVD (ASCVD; peripheral artery disease, heart failure, coronary artery disease and stroke) using genomic structural equation modeling (SEM)<sup>26</sup>. We mapped identified variants to brain cell types using annotations based on recent single-cell RNA sequencing in postmortem human brain samples<sup>27</sup>. Third, we assessed shared risk factors explaining the genetic correlation between MDD and CVD. Finally, we used Mendelian randomization (MR) to investigate putative causal pathways linking MDD and CVD.

## Results

### Study design

We acquired summary statistics from the largest and most recent genome-wide association studies (GWASs) so far of MDD<sup>6</sup>, MDD symptoms, CVDs and shared risk factors (Fig. 1 and Supplementary Table 1). We used summary statistics from a GWAS of MDD involving 294,322 cases<sup>6</sup>. For CVDs, we used summary statistics from GWASs of peripheral artery disease<sup>28</sup>, heart failure<sup>29</sup>, coronary artery disease<sup>30</sup>, stroke<sup>31</sup> and atrial fibrillation<sup>32</sup>, based on 12,086, 47,304, 181,523, 73,652 and 60,620 cases, respectively. We considered five categories of risk factors: blood pressure<sup>33</sup>, psychosocial or lifestyle<sup>34–38</sup>, childhood maltreatment<sup>39</sup>, metabolic<sup>40–42</sup> and inflammation<sup>43,44</sup>. These data were used together with a comprehensive set of methods to elucidate etiological pathways underlying the comorbidity between MDD and CVD (Fig. 1).

### Most genetic risk factors for CVD overlap with MDD

We found moderate genetic correlations between MDD and CVDs, as estimated with linkage disequilibrium score regression (LDSC<sup>45</sup>; Extended Data Fig. 1 and Supplementary Table 2). The strongest correlations were noted for peripheral artery disease ( $r_g = 0.30$ , 95% CI [0.22, 0.38],  $P = 1 \times 10^{-13}$ ), heart failure ( $r_g = 0.29$ , 95% CI [0.23, 0.34],  $P = 1 \times 10^{-24}$ ) and coronary artery disease ( $r_g = 0.25$ , 95% CI [0.21, 0.28],  $P = 9 \times 10^{-45}$ ), while smaller but statistically significant correlations were observed for stroke and atrial fibrillation ( $r_g = 0.18$  and  $r_g = 0.11$ ,  $P < 1 \times 10^{-7}$ ). ASCVDs showed strong correlations among each other while atrial fibrillation was only moderately genetically correlated to the ASCVDs (Extended Data Fig. 2a and Supplementary Table 2). For MDD symptoms, poor appetite or overeating showed consistently the

strongest genetic correlations with the CVDs (Extended Data Fig. 1a and Supplementary Table 3).

We estimated local genetic correlations between MDD and CVDs in each of 2,495 distinct genomic regions using LAVA<sup>25</sup>. We found 54 statistically significant, predominantly positive, local correlations between MDD and CVDs, 40 of which were noted for MDD and coronary artery disease (Fig. 2 and Supplementary Table 4). We also assessed correlations between MDD and the CVDs in 16 loci in the human leukocyte antigen (HLA) region (Extended Data Fig. 1b). Out of 50 assessed correlations, 8 were statistically significant, indicating that this region is a hotspot of correlation between MDD and CVDs.

Next, we investigated genetic overlap on the level of genetic variants using MiXeR<sup>24</sup> (Supplementary Tables 5 and 6). We identified more causal variants, that is, genetic variants with nonzero effects taking linkage disequilibrium (LD) into account, for MDD than for the CVDs, suggesting that MDD is more polygenic than CVD. To verify these results, we estimated polygenicity using a complementary approach implemented in SBayesS<sup>46</sup>, which showed estimates that were highly correlated with those of MiXeR (Extended Data Fig. 2b and Supplementary Table 7).

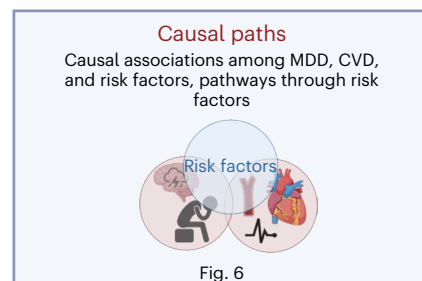
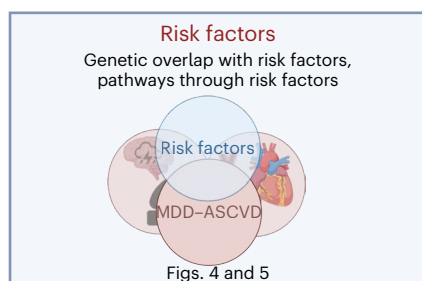
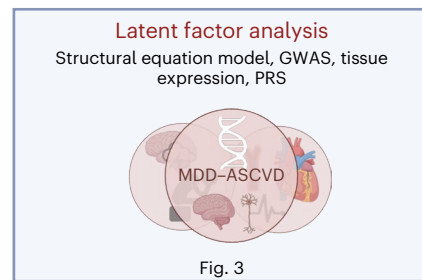
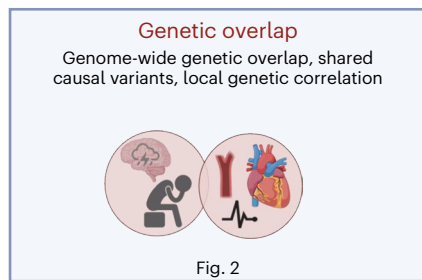
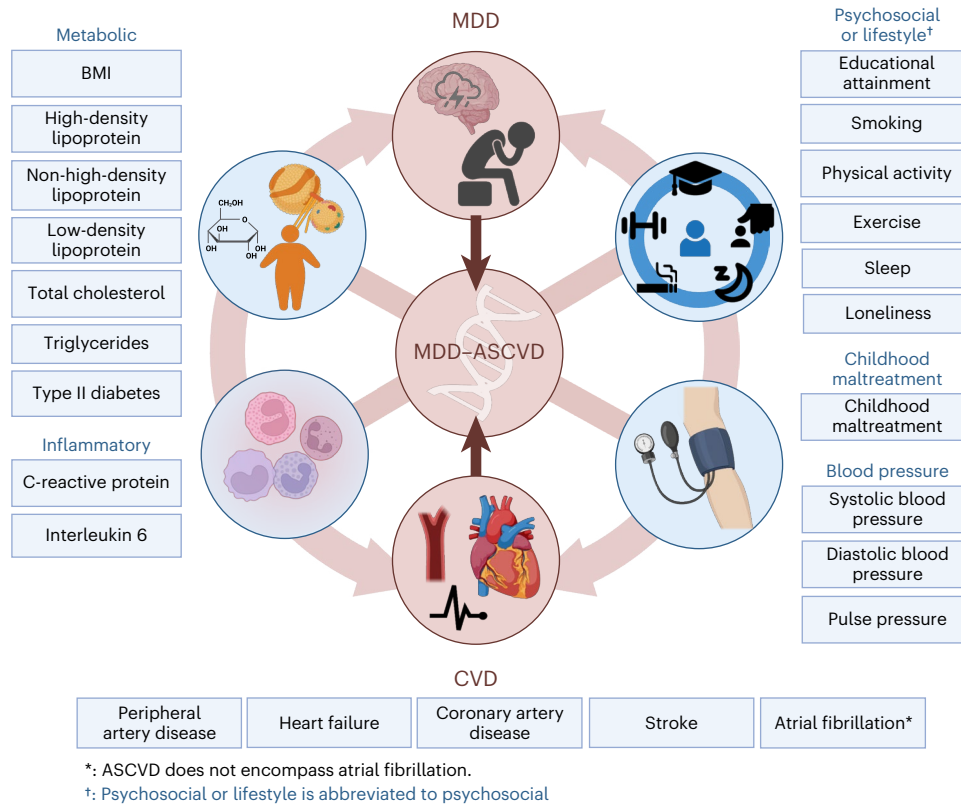
Bivariate MiXeR results showed that CVDs shared most of their causal variants with MDD (from 64% for atrial fibrillation to 92% for heart failure; Fig. 2b and Supplementary Table 6) whereas MDD shared only few of its causal variants with CVDs (<20%). Performance metrics indicated that results for peripheral artery disease should be interpreted with caution. Both shared genetic variants and local genetic correlations exhibited strong degrees of effect direction concordance (Fig. 2c), suggesting that genetic risk variants for CVDs are strongly correlated with a genetic subcomponent of MDD.

### Shared genetic liability to MDD and CVD

To further characterize the genetic overlap, we explicitly modeled the shared genetic liability between MDD and CVD as a higher-order latent factor using genomic SEM. We excluded atrial fibrillation because it deteriorated the model fit (comparative fit index (CFI) of 0.918 and standardized root mean squared residual (SRMR) of 0.072), so that the interpretation of the latent factor changed to representing ASCVD. The final model had an excellent fit (CFI of 0.999 and SRMR of 0.021). The loading of ASCVD on the MDD-ASCVD factor was  $\beta = 2.46$  (95% CI [2.00, 2.91],  $P = 1.63 \times 10^{-26}$ ). For model identification purposes, the loading of MDD was fixed to 1. Factor loadings for the ASCVDs on the ASCVD factor were  $\beta = 0.79$  (95% CI [0.68, 0.90],  $P = 6.21 \times 10^{-45}$ ) for stroke,  $\beta = 1.03$  (95% CI [0.89, 1.17],  $P = 2.58 \times 10^{-46}$ ) for peripheral artery disease, and  $\beta = 1.08$  (95% CI [0.95, 1.21],  $P = 2.10 \times 10^{-59}$ ) for heart failure, respectively (Fig. 3a; loadings are given standardized with respect to the genetic variance of the traits). For coronary artery disease, the factor loading was fixed to 1 for identification purposes. For comparison, we also fit a latent factor for the ASCVDs alone (without MDD), which showed similar fit and parameter estimates (CFI of >0.999 and SRMR of 0.013; Extended Data Fig. 3a). This shows that shared genetic liability to different ASCVDs as well as to ASCVDs and MDD (to a lesser extent) can be explained by a single underlying factor.

The GWAS on the latent MDD-ASCVD factor resulted in 205 independent genome-wide significant loci (Fig. 3b, independent at  $R^2 < 0.1$  and distance  $\geq 250$  kb; Supplementary Table 8). Almost three-quarters (74.6%) of the genome-wide statistically significant single-nucleotide polymorphisms (SNPs) showed a high  $Q$  heterogeneity, suggesting that their effects were more in line with an independent pathway than a common pathway model (Methods). Most of this heterogeneity was due to MDD, as the GWAS for latent ASCVD without MDD showed fewer genome-wide statistically significant SNPs with a high heterogeneity (30.6%; Extended Data Fig. 3b).

For the latent MDD-ASCVD factor, we filtered out variants that showed statistically significant heterogeneity and considered only variants where the latent MDD-ASCVD factor was the best model for



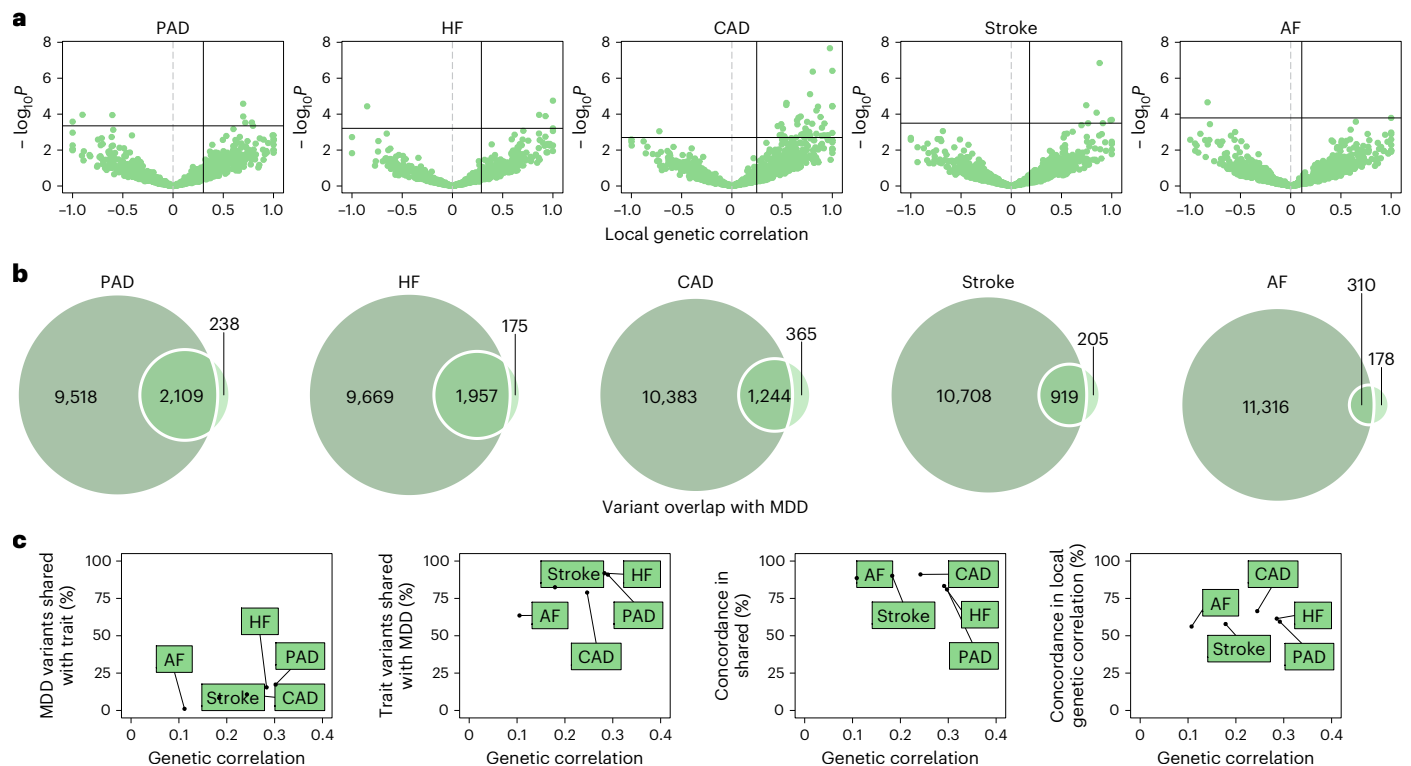
**Fig. 1 | Study design.** The comorbidity between MDD and CVD was investigated using genetic and causal inference methods, including assessing overlap with and mediation through shared risk factors (blood pressure, psychosocial or lifestyle, childhood maltreatment, metabolic and inflammation; Supplementary

Table 1). The risk factor group psychosocial or lifestyle is abbreviated to psychosocial. Created with [BioRender.com](https://www.biorender.com) (license agreements VG26BG3VTL, MK26BG46H3).

the follow-up analyses. We retained 72 independent loci underlying the shared genetic liability (Fig. 3b and Supplementary Table 9). The top SNP after filtering was rs11670056 in the *ELL* gene, which encodes a part of the transcription elongation factor complex and has previously been associated with CVDs, blood traits, BMI and educational attainment (enrichment in associations with other traits for statistically significant SNPs are shown in Extended Data Fig. 4a). There were 19 top SNPs with significant expression quantitative trait loci (eQTL) for one or multiple

genes (Fig. 3b and Supplementary Table 10). Besides *ELL*, multiple genes on chromosome 10 around *INA* and *CNNM2* were identified. *INA* and *CNNM2*, which encode a neurofilament and a protein involved in ion transportation, respectively, have previously been associated with psychiatric as well as cardiovascular traits.

We found seven novel loci in the latent MDD-ASCVD GWAS that were not among the risk loci in the MDD and ASCVD GWASs that constituted the latent factor (Extended Data Fig. 4b and Supplementary



**Fig. 2 | Genetic overlap between MDD and CVD beyond genome-wide genetic correlation.** **a**, Volcano plots based on LAVA results showing genomic loci (green dots) with the local genetic correlation between MDD and each of the CVDs (x axis) and the corresponding  $P$  value (y axis). Loci exceeding the horizontal line are significant at  $P_{FDR} < 0.05$  (Benjamini–Hochberg-adjusted  $P$  value). Multiple testing was performed separately for each trait over all considered loci. Empirical  $P$  values were obtained via a permutation procedure with partial integration, evaluating the two-sided hypothesis of no association using the estimated parameters as test statistics. **b**, Venn diagrams based on MiXeR results showing the number of causal variants (number of nonzero variants required to explain 90% of trait heritability) that are unique to MDD (left circle), unique

to CVD (nonoverlapping part of right circle) or shared between MDD and CVD (overlapping part of circles). **c**, Genetic correlation estimated by LDSC (x axis) against the percentage of MDD causal variants that are shared with the CVD trait as estimated by MiXeR (first plot), the percentage of CVD trait causal variants that are shared with MDD (second plot) and the percentage of shared variants that have concordant effect directions (third plot). The fourth plot shows the percentage of local genetic correlations from LAVA that have concordant effect directions on the y axis. In **a–c**, the sample sizes and information for underlying summary statistics GWASs are reported in Supplementary Table 1. AF, atrial fibrillation; CAD, coronary artery disease; HF, heart failure; PAD, peripheral artery disease.

Table 11). The top SNPs in these loci have not been identified in any GWAS recorded in the GWAS catalog before, but four of them have shown suggestive associations ( $P < 0.05$ ) with metabolic traits.

Using partitioned LDSC, we observed that the heritability of the latent MDD–ASCVD factor was enriched in genes with expression specific to endothelial tissues, which was also observed for latent ASCVD but not for MDD (Fig. 3c and Supplementary Table 12). To gain deeper insights into brain-specific mechanisms, we leveraged high-resolution human brain single-nucleus RNA sequencing data<sup>27</sup> and identified four human brain cell types that exhibited enriched MDD–ASCVD heritability, including deep-layer corticothalamic and 6b cells, midbrain-derived inhibitory neurons, miscellaneous neurons and vascular cells (Fig. 3d and Supplementary Table 13). Notably, except for miscellaneous neurons, these cell types displayed no enrichment for either latent ASCVD or MDD, suggesting that the genetic variance for MDD–ASCVD comorbidity has a distinct functional signature.

To externally validate the MDD–ASCVD phenotype, we computed polygenic risk scores (PRS) based on the summary statistics for the latent MDD–ASCVD factor, as well as for MDD and latent ASCVD, and found them to predict ASCVD and MDD diagnoses in UK Biobank (UKB) (all  $P < 2 \times 10^{-13}$ ; Fig. 3e and Supplementary Tables 14 and 15; note that source and target samples were overlapping and the  $R^2$  values should only be interpreted relative to one another). The PRS for latent ASCVD and the latent MDD–ASCVD factor explained similar

amounts of variance in ASCVDs. This is in line with the MiXeR findings (Fig. 2b), suggesting that most causal variants for ASCVDs are shared with MDD. In contrast, as most causal variants for MDD are not shared with ASCVDs, the PRS for the latent MDD–ASCVD factor explained less than half as much variance in MDD as the MDD PRS.

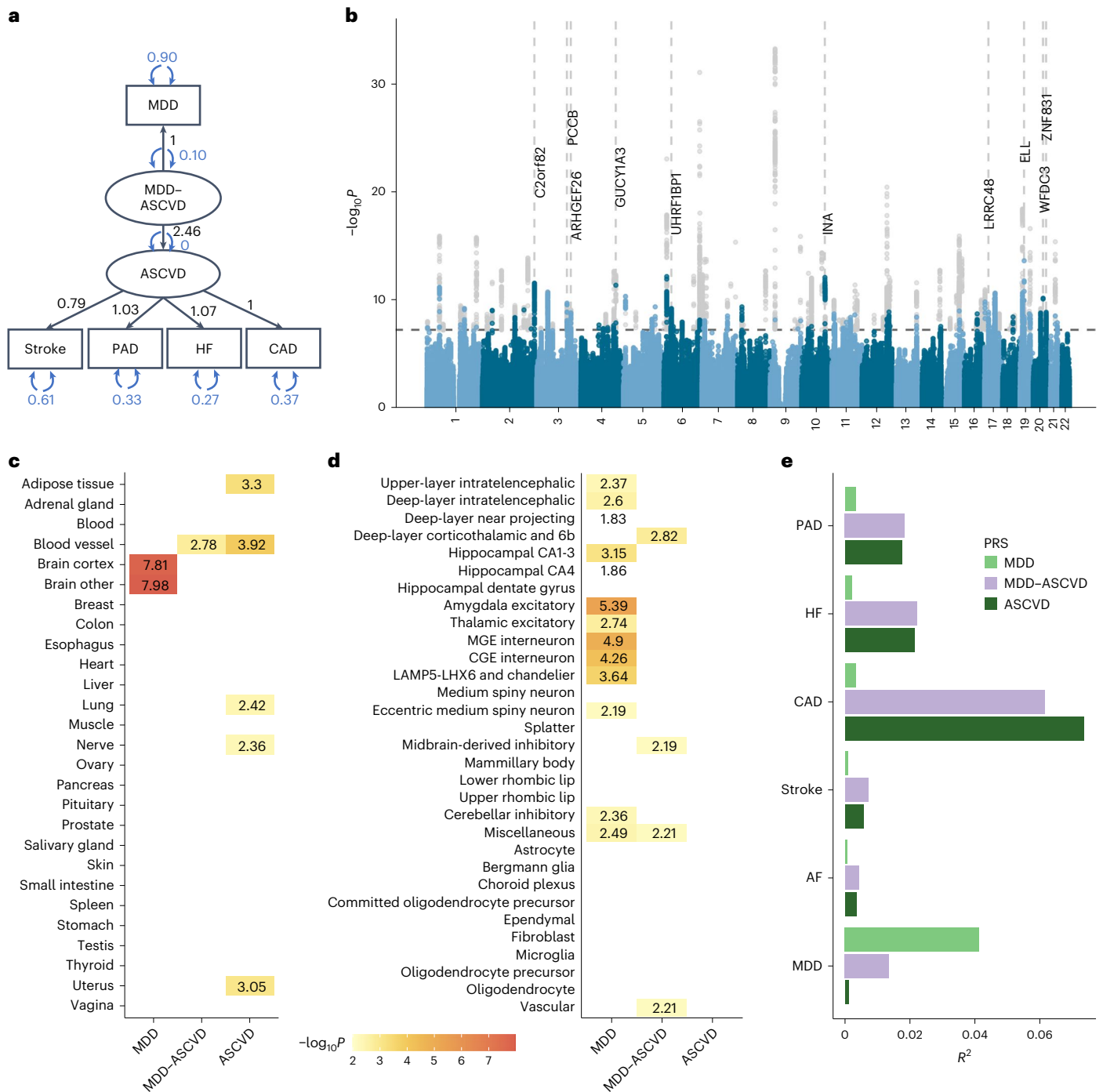
Next, we assessed genetic correlations between the latent MDD–ASCVD factor and MDD symptoms. We found that poor appetite or overeating and suicidal thoughts are the symptoms most strongly correlated with MDD–ASCVD. In contrast, poor appetite or overeating is among the least genetically correlated symptoms to MDD (Extended Data Fig. 4c).

Finally, we estimated genetic correlation of attention deficit and hyperactivity disorder (ADHD), anxiety disorders, posttraumatic stress disorder (PTSD), bipolar disorder and schizophrenia, with MDD, latent MDD–ASCVD and latent ASCVD. We found that ADHD, anxiety disorder and PTSD were genetically correlated with latent ASCVD (Extended Data Fig. 4d). In addition, PTSD and ADHD showed similar genetic correlations for MDD and latent MDD–ASCVD, suggesting that variants that are shared between MDD and ASCVD might explain most of the genetic correlation between MDD and these disorders.

### Genetic overlap between MDD and risk factors

Next, we aimed to identify risk factors linking MDD and CVD. First, we assessed genetic correlations between MDD and risk factors. We observed strong to moderate genetic correlations of MDD with





**Fig. 3 | Shared genetic liability latent factor for MDD-ASCVD. a**, The latent factor model as specified in genomic SEM with the ‘observed’ variables in rectangles and the latent variables in circles. Factor loadings (standardized with respect to the genetic variance of the traits) are given in black and variances in blue. **b**, Latent MDD-ASCVD factor GWAS results. The x axis shows the genomic position and the y axis shows statistical significance as  $-\log_{10}(P)$ . Genome-wide significant SNPs ( $P < 5 \times 10^{-8}$ ) that were filtered out because of significant heterogeneity  $Q_{SNP}$  are displayed in gray. The top ten eQTL genes are displayed with dashed vertical lines indicating their position. *P* values were computed using a two-sided Z-test. **c**, Enrichment results in GTEx tissues for the latent MDD-ASCVD factor, with latent ASCVD (without MDD) and MDD only as comparison. **d**, Enrichment results for the latent MDD-ASCVD factor, latent ASCVD and MDD

only in brain cell types. **e**, The proportion of variance explained in MDD and CVD phenotypes in the UKB (defined using ICD codes listed in Supplementary Table 14) by each of three PRSs for the latent MDD-ASCVD factor, latent ASCVD or MDD only. **c, d**, Enrichment is measured using significance testing in a one-sided Z-test displayed as  $-\log_{10}(P)$ . Only tissues with a significant association ( $P_{FDR} < 0.05$ ; Benjamini-Hochberg adjustment for multiple testing) are shown. Multiple testing was performed over tested tissues/cell-types. In **a-e**, sample sizes and information for underlying GWAS summary statistics are reported in Supplementary Table 1. MDD-ASCVD, common factor for MDD and ASCVD; ASCVD, common factor for the atherosclerotic cardiovascular diseases. CA, cornu ammonis; CGE, caudal ganglionic eminence; MGE, medial ganglionic eminence.

psychosocial or lifestyle factors, such as loneliness ( $r_g = 0.68$ , 95% CI [0.64, 0.72]), childhood maltreatment ( $r_g = 0.55$ , 95% CI [0.50, 0.60]) and exercise ( $r_g = -0.33$ , 95% CI [-0.38, -0.29]) (Fig. 4b and Extended Data Fig. 6). Among metabolic factors, MDD showed the strongest genetic correlation with type II diabetes ( $r_g = 0.19$ , 95% CI [0.16, 0.23]) as well as levels of high-density lipoprotein cholesterol ( $r_g = -0.14$ , 95% CI [-0.17, -0.11]) and triglycerides ( $r_g = 0.18$ , 95% CI [0.15, 0.21]). We observed genetic correlation of MDD with levels of the inflammatory markers IL6 ( $r_g = 0.22$ , 95% CI [0.11, 0.33]) and C-reactive protein ( $r_g = 0.15$ , 95% CI [0.10, 0.19]). We did not observe genetic correlations of MDD with blood pressure traits. As a comparison, for CVDs, the largest genetic correlations were found between heart failure and BMI ( $r_g = 0.55$ , 95% CI [0.50, 0.60]) and type II diabetes ( $r_g = 0.49$ , 95% CI [0.42, 0.55]) (Extended Data Fig. 5 and Supplementary Table 2). The results for MDD symptoms largely followed the pattern of MDD diagnosis, although poor appetite or overeating showed stronger genetic correlations with metabolic factors than did MDD diagnosis (Extended Data Fig. 6 and Supplementary Table 3).

Causal-variant and local genetic correlation analysis revealed several patterns of genetic overlap between MDD and risk factors. We found that psychosocial or lifestyle factors, childhood maltreatment and BMI showed similar levels of polygenicity to MDD (Extended Data Fig. 2b and Supplementary Tables 5 and 7). In addition, they exhibited a large degree of shared variants and many local genetic correlations with MDD (Fig. 4a, Extended Data Fig. 7 and Supplementary Tables 4 and 6). Out of these factors, loneliness and childhood maltreatment also showed high levels of effect direction concordance with MDD, both for risk variants and local genetic correlations (Fig. 4b and Supplementary Tables 4 and 6). Combined with high polygenicity, such concordance translates to large genome-wide genetic correlations. In contrast, educational attainment, smoking, exercise, physical activity, BMI and sleep duration had similar levels of polygenicity and a large degree of polygenic overlap with MDD, but low effect direction concordance, suggesting that genome-wide genetic correlations underestimate the genetic overlap with MDD for these traits.

Genetic factors for blood pressure traits showed unique patterns of genetic overlap with MDD. They were polygenic (>5,000 causal variants; Extended Data Fig. 2b) but did not overlap strongly with genetic risk factors for MDD (<0.30 of risk variants overlapping; Fig. 4a). Moreover, risk variants that did overlap showed low degree of effect direction concordance (48–57% of shared variants in the same direction; Fig. 4b). We observed 97 significant local genetic correlations with MDD for the three blood pressure traits, 60% of which were positive. These findings suggest that MDD and blood pressure share variants that exhibit both positive and negative correlations, which are canceled out in the genome-wide estimate.

Type II diabetes, lipid traits and C-reactive protein showed low polygenicity (<2,500 causal risk variants). Type II diabetes, triglyceride levels and C-reactive protein levels shared most of their risk variants with MDD, and these variants showed high degrees of concordance (>85% of shared variants in the same direction). High-density lipoprotein cholesterol shared most of its risk variants with MDD, in consistently opposite directions.

### Risk factors underlying genetic overlap between MDD and CVD

To assess risk factors explaining the genetic overlap between MDD and CVD, we estimated genetic correlations adjusted for risk factors (individually or as a group) using genomic SEM (Fig. 5a, Extended Data Fig. 8 and Supplementary Table 16). The largest reduction in point estimate was observed for the genetic correlation between MDD and peripheral artery disease after adjustment for the group of psychosocial or lifestyle factors. Similarly, the genetic correlation of MDD with coronary artery disease and the latent ASCVD factor were attenuated

after adjusting for psychosocial or lifestyle factors. The reduction was mainly driven by loneliness (Extended Data Fig. 8). Genetic correlations of MDD with peripheral artery disease and stroke were no longer statistically significant after adjusting for psychosocial or lifestyle factors. We also observed some attenuation in the genetic correlation between MDD and CVDs after adjusting for childhood maltreatment, metabolic factors or inflammatory markers, although confidence intervals (CIs) overlapped.

Next, we specified the risk factors as mediators instead of covariates in the genomic SEM model and compared the path estimates from MDD with CVD. Observing attenuation in the association between MDD and CVD when a risk factor is modeled as a mediator rather than a covariate supports the interpretation that the risk factor mediates some of the link between MDD and CVD. For all psychosocial or lifestyle factors together, the inflammation traits, and for type II diabetes, the 95% CIs did not overlap between the mediator and covariate models, suggesting that these risk factors are mediating part of the link between MDD and CVD (Supplementary Table 17).

Finally, we estimated genetic correlations between the risk factors and the latent MDD–ASCVD factor (Fig. 5b) and found that the latent MDD–ASCVD factor was substantially more genetically correlated with blood pressure traits, C-reactive protein levels and metabolic factors than MDD only, suggesting that these factors characterize the genetic liability to MDD–ASCVD rather than to MDD alone.

### Causal pathways linking MDD and CVD

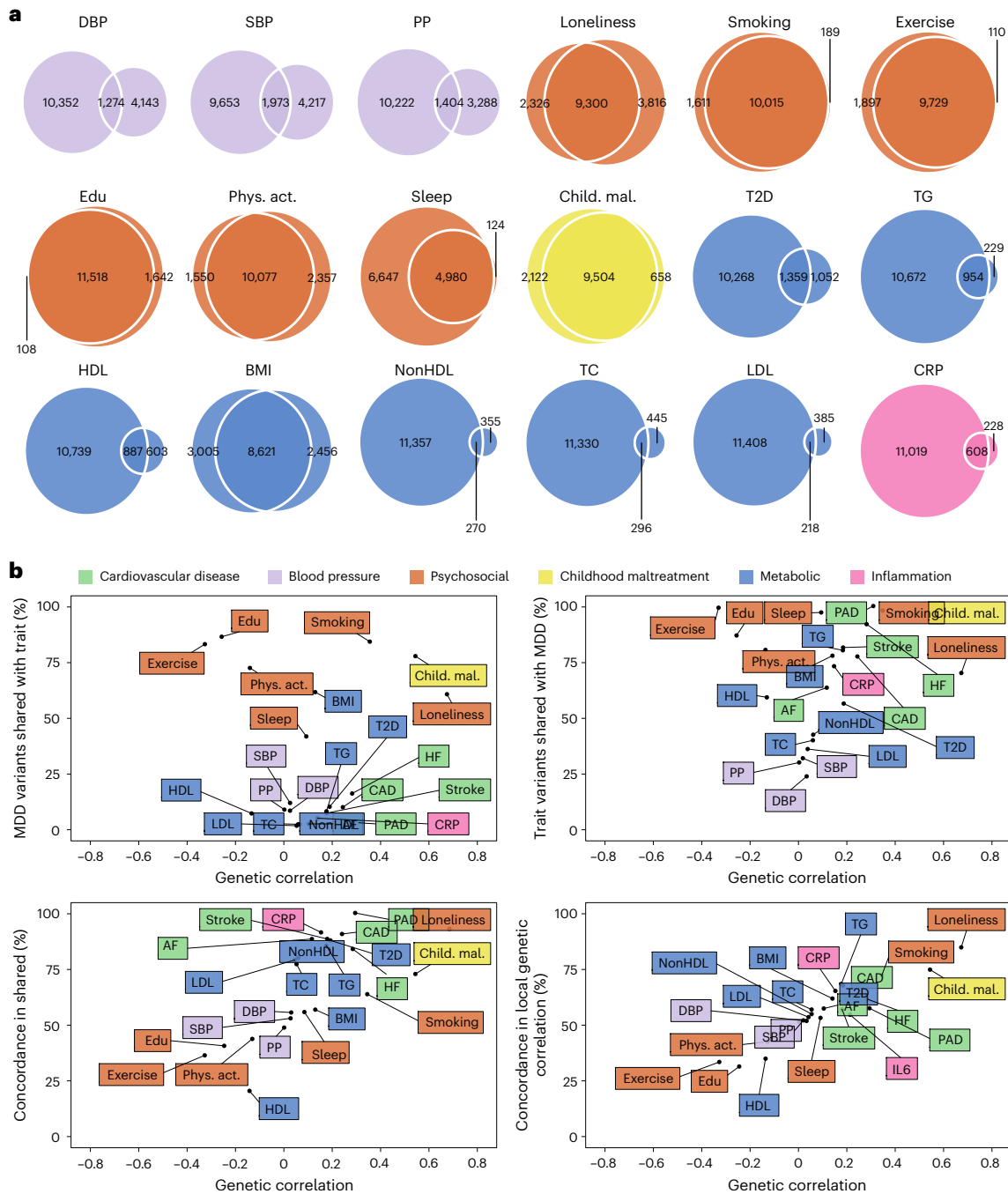
We investigated putative causal relationships between MDD and CVD using two-sample MR. Instruments were Steiger filtered, that is, SNPs explaining statistically significantly more variance in the outcome than the exposure were excluded. The results provide support for a causal effect of MDD liability on all CVDs, with the strongest effects observed for coronary and peripheral artery disease (Fig. 6a). We found no statistically significant pleiotropy and the results were consistent across weighted median, mode and Egger sensitivity analyses, providing support for the inverse variance weighted (IVW) estimates (Supplementary Table 18).

For risk factors, we observed that increased liability to MDD was associated with increased loneliness, smoking, risk of type II diabetes and levels of C-reactive protein (Fig. 6a). These results were again consistent across sensitivity analyses (Supplementary Table 18).

We also investigated potential causal effects in the other direction with genetic instruments to CVDs and risk factors as exposures and MDD as outcome (Fig. 6b). We provide evidence for a statistically significant causal effect of stroke, loneliness, smoking, exercise, educational attainment, childhood maltreatment, levels of high-density lipoprotein cholesterol and BMI on MDD risk (Fig. 6b), which was robust across sensitivity analyses (Supplementary Table 18). No robust effects were observed for other CVDs, blood pressure, other metabolic traits or inflammatory markers.

When using genetic instruments for the latent MDD–ASCVD factor to predict risk factors, statistically significant effects were observed for pulse pressure and type II diabetes (Extended Data Fig. 9d). We found statistically significant pleiotropy for systolic blood pressure.

We extended univariable results with multivariable MR to assess whether there was support for causal effects of MDD on CVD explained by the risk factors. We included only risk factors that were associated with MDD in the univariable MR analysis. The effects of MDD on CVDs were attenuated after adjusting for the psychosocial or lifestyle risk factors, driven by loneliness and smoking, and metabolic risk factors, although CIs were wide (Fig. 6c, Extended Data Fig. 10 and Supplementary Table 19). The effect of MDD liability on peripheral artery disease, heart failure, stroke and atrial fibrillation risk was no longer statistically significant after adjusting for the groups of metabolic or psychosocial or lifestyle factors (Fig. 6d).

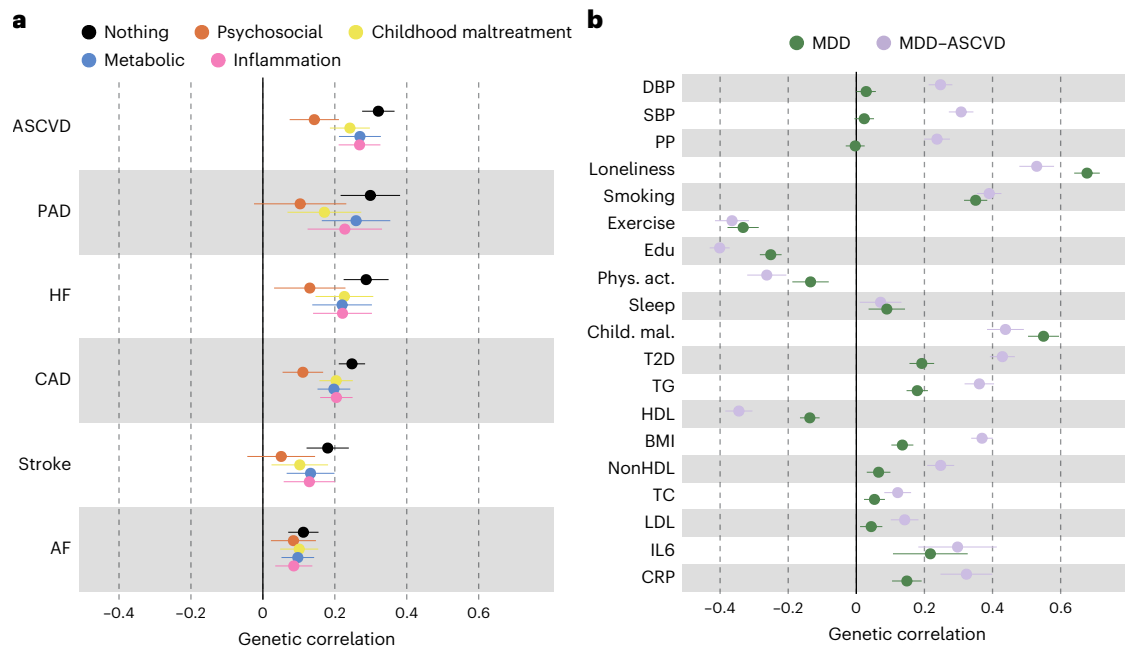


**Fig. 4 | Local and causal-variant level genetic correlations between MDD and risk factors. a**, Venn diagrams based on MiXeR results showing the number of causal variants (number of nonzero variants required to explain 90% of trait heritability) that are unique to MDD (left circle), the risk factor (nonoverlapping part of right circle) or shared between MDD and the risk factor (overlapping part of circles). **b**, Genome-wide genetic correlation estimated by LDSC ( $r_{gr}$ , x axis) against the percentage of MDD causal variants that are shared with the risk factor as estimated by MiXeR (top left), the percentage of risk factor causal variants that are shared with MDD (top right) and the percentage of shared variants that have concordant effect directions (bottom left). The bottom right plot shows the percentage of local genetic correlations from LAVA that have

concordant effect directions on the y axis. Cardiovascular traits are also shown for comparison. For **a** and **b**, standard errors for MiXeR, LAVA and LDSC results are reported in Supplementary Tables 2–6. Sample sizes for GWAS summary statistics are reported in Supplementary Table 1. Note that IL6 was excluded from MiXeR results because it failed performance checks (Methods). Child. mal., childhood maltreatment; CRP, C-reactive protein; DBP, diastolic blood pressure; edu, educational attainment; HDL, high-density lipoprotein; LDL, low-density lipoprotein; nonHDL, non-high-density lipoprotein; phys. act., physical activity; PP, pulse pressure; psychosocial, psychosocial or lifestyle; SBP, systolic blood pressure; T2D, type II diabetes; TC, total cholesterol; TG, triglycerides.

To investigate possible bias due to sample overlap, we repeated the analyses using MDD summary statistics based on GWAS excluding the UKB sample (the main source of overlap). We observed small differences in the estimates, but the interpretation remained the same for all results (Extended Data Fig. 9a,b and Supplementary Table 20).

Furthermore, we repeated the analyses using latent heritable confounder (LHC) MR, which is robust to sample overlap<sup>47</sup>. The results pattern was similar, although the point estimates were slightly attenuated for CVD risk (for example, no longer statistically significant for heart failure and atrial fibrillation) and became stronger for most



**Fig. 5 | Genetic correlation between MDD and CVD explained by risk factors.** **a**, The genetic correlation between MDD and CVD before and after adjustment for groups of risk factors (color coded). **b**, A comparison of genetic correlation between MDD (dark green) and the latent MDD-ASCVD factor (lilac) and

individual risk factors. In **a** and **b** the points and error bars represent mean genetic correlation and 95% CIs. The sample sizes for GWAS summary statistics are reported in Supplementary Table 1.

other traits (Extended Data Fig. 9c). We conclude that although sample overlap impacted the point estimates, the interpretation of results remained similar.

### Discussion

Here, we showed that genetic risk factors for CVD overlap strongly with MDD. We modeled the shared genetic liability between MDD and ASCVD as a latent factor and showed that, distinct from MDD alone, it is associated with gene expression specific to thalamic and vascular cell types in the brain and is genetically correlated with immunometabolic factors and blood pressure. Further, we showed that the association between MDD and CVD is partly explained by modifiable risk factors and provide evidence that it is probably causal in nature.

In line with previous results<sup>22,48–50</sup> we found moderate genome-wide genetic correlations between MDD and CVDs. Analysis on the level of shared risk variants showed that MDD was substantially more polygenic than the CVD traits and that most risk variants for CVD were in fact shared with MDD and had concordant effect directions. In addition, we found many positive local genetic correlations between MDD and CVD, especially for coronary artery disease, although that might reflect the fact that the coronary artery disease GWASs have larger sample size than the GWASs of the other CVDs. Interestingly, we found that the *HLA* region was a hotspot for local genetic correlation between MDD and CVDs. These findings suggest that genetic overlap between MDD and CVD is underestimated in genome-wide correlation analyses.

We modeled the shared liability between MDD and ASCVD as a latent MDD-ASCVD factor and performed a GWAS on the factor. Atrial fibrillation was excluded because it deteriorated the fit of the latent factor, which was in line with results that showed that atrial fibrillation was substantially less genetically correlated with the ASCVDs than they were with each other. Combined, these results suggest that atrial fibrillation and ASCVDs have partly distinct sources of genetic variation. We identified many loci associated with MDD-ASCVD, some of which were uniquely associated with the shared liability and not with the constituent traits. We found that heritability for the MDD-ASCVD latent factor was enriched for genes specifically expressed in vascular

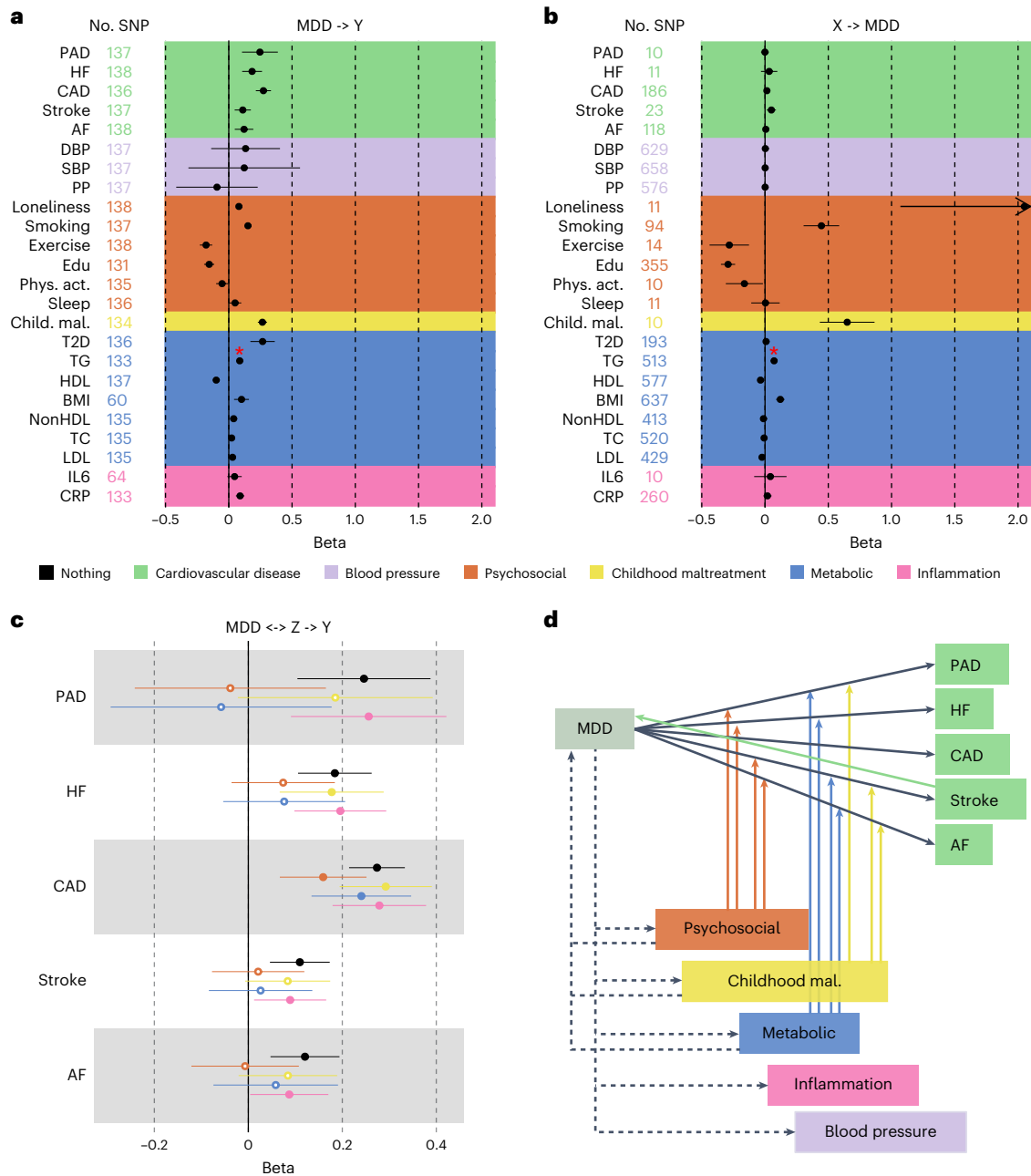
brain cells, deep-layer corticothalamic 6b (projecting to the thalamus) and midbrain-derived inhibitory neurons (predominantly located in the thalamus). This cell-type enrichment signature was not found for MDD or ASCVD alone, suggesting that a distinct mechanism involving thalamic circuits might underlie the shared liability to MDD-ASCVD. Altered thalamic function has indeed been implicated previously in CVD<sup>51–53</sup> and MDD<sup>54,55</sup>, and white matter integrity in thalamic radiations show associations with aortic area and myocardial wall thickness, suggesting that it has a role in the ‘heart–brain’ connection<sup>56</sup>.

MDD showed a high degree of genetic overlap with risk factors. We showed that MDD was substantially more polygenic than blood pressure, lipid and inflammatory traits. In contrast, psychosocial or lifestyle traits were equally polygenic to MDD and showed a large degree of overlap with MDD. Interestingly, the local and variant-level analysis indicated that blood pressure traits shared a substantial number of risk variants with MDD (in line with a previous report<sup>49</sup>), which was masked at the genome-wide level due to their opposing effect directions. Similarly, BMI and lipid traits showed discordant directions to MDD in effects of overlapping risk variants and local genetic correlations. Finally, C-reactive protein shared most of its risk variants with MDD, predominantly in the positive direction, indicating a genetic relationship between C-reactive protein and MDD that was masked in the genome-wide genetic correlation estimate. Overall, these findings refine our understanding of the polygenic overlap between MDD and risk factors shared between MDD and CVD, and indicate that it is stronger and more complex than has previously been reported.

We estimated genetic correlations between MDD and CVD adjusting for risk factors and found that psychosocial or lifestyle factors explain a substantial part of the genetic correlation between MDD and CVD and highlight loneliness as an important factor in the relationship between MDD and CVD. In addition, we found tentative support for a mediating role of C-reactive protein and IL6 levels in the association between MDD and ASCVD.

We found that, compared with MDD, the latent MDD-ASCVD factor was characterized by genetic correlations with immunometabolic factors and blood pressure, suggesting that the shared genetic





**Fig. 6 | Support for causal effects between MDD, CVD and shared risk factors.** **a**, The effect of genetic instruments for MDD (exposure) on CVD and risk factors (outcomes). **b**, The effect of genetic instruments for CVDs and risk factors on MDD. The arrow indicates that the CI for the effect of loneliness on MDD has been cut to improve readability. **c**, The effects of genetic instruments for MDD on CVD while adjusting for groups of risk factors in multivariable MR. **d**, A schematic overview of levels of evidence for causal effects, with solid lines indicating convincing evidence (consistent across sensitivity analyses) for such effects and dashed lines indicating evidence for some of the relationships tested within the

trait categories. The arrows from the risk factors to the association between MDD and the CVDs indicate that the combined risk factors attenuated the association so that it was no longer statistically significant. In **a–c**, the IVW estimate is shown. The results from the sensitivity analyses are reported in Supplementary Table 18. In **a–d**, the points and error bars represent mean effect size (regression coefficient) and 95% CIs. The sample sizes for GWAS summary statistics are reported in Supplementary Table 1. The term beta refers to the log odds ratio. The asterisk (\*) indicates that the observed statistically significant association suffered from pleiotropy; possible causal effect should not be interpreted.

liability to MDD and ASCVD is associated with an immunometabolic subtype of depression. The existence of an immunometabolic subtype of depression has been proposed previously, based on a long line of evidence of oxidative stress and neuroendocrine and inflammatory dysregulation in MDD that are preferentially associated with atypical symptoms of MDD (for example, weight gain and oversleeping)<sup>57</sup>. Indeed, we find that poor appetite or overeating is the MDD symptom with the consistently greatest genetic correlations to the CVDs, and it is among the most genetically correlated symptoms

with MDD–ASCVD, while it is among the least genetically correlated symptoms to MDD, although CIs were too wide to be conclusive. We did not find large genetic correlations between sleep duration measured using an accelerometer over 1 week and the MDD–ASCVD factor. However, statistically significant variants for the MDD–ASCVD factor were strongly enriched in statistically significant variants for self-reported short sleep duration (<6 h per night)<sup>58</sup> suggesting that short sleep duration might be more related to MDD–ASCVD comorbidity than overall sleep duration.

Mental disorders that are highly comorbid with MDD, such as psychotic disorders, anxiety disorders and PTSD, have also been shown to be associated with CVD<sup>59</sup>. We find that ADHD, anxiety disorders and PTSD show genetic correlations with ASCVD and MDD–ASCVD. For ADHD and PTSD, the genetic correlation was similar between MDD and MDD–ASCVD. Future work should estimate latent factors representing shared and distinct sources of genetic covariance among MDD, ADHD, anxiety disorders and PTSD and investigate how those factors relate to ASCVD.

We found robust support for the likelihood of causal effects of MDD on CVD. Previous two-sample MR studies have observed associations between genetic liability to MDD and risk of coronary artery disease but results for heart failure and stroke have been inconsistent<sup>21,22,48,60</sup>. Using more recent GWAS data, we confirmed an effect of genetic liability to MDD on coronary artery disease and found robust associations for stroke, heart failure and peripheral artery disease.

Except for stroke, we found limited evidence for a causal effect of CVDs on MDD, which is contrary to literature suggesting that such effects exist<sup>23</sup>. The MDD sample is mainly based on large volunteer-based studies that might select against individuals with CVD. Indeed, participants in the UKB study are healthier than the general population<sup>61</sup>. In addition, interpretation of the MR estimate in this case is complicated by the fact that CVD is a time-varying exposure with late age of onset<sup>62</sup>. However, we do find that genetic instruments capture the well-established association between stroke and subsequent MDD<sup>63</sup>. Therefore, the findings in our study offer some indication that the association between CVDs and subsequent MDD might have been overestimated in previous studies, possibly due to reverse causation, surveillance bias or unmeasured confounding.

We observed effects of genetic liability to MDD on most of the risk factors. We did not observe associations between genetic liability to MDD and blood pressure traits, although the presence of correlated and anticorrelated genetic components complicates interpretation. Indeed, using genetic instruments for the latent MDD–ASCVD factor, we did observe strong associations for pulse pressure. Previous MR studies of the association of C-reactive protein and IL6 levels with MDD risk have shown inconsistent results<sup>64,65</sup>. We did not find an effect of genetic instruments for inflammatory markers on MDD. However, we did find associations between genetic liability to MDD and increased C-reactive protein levels, lipid levels and type II diabetes, offering evidence that MDD might lead to long-term dysregulated immunometabolic pathways<sup>57</sup>. Likewise, in line with previous evidence<sup>66</sup>, we find support for a causal effect of liability to MDD on smoking, which, in turn, can lead to inflammation.

On multivariable MR analysis, we observed that only the association between genetic liability to MDD and coronary artery disease remained statistically significant after adjusting for the psychosocial or lifestyle or metabolic covariates. We found that adjusting for smoking status attenuates the association between genetic liability to MDD and peripheral artery disease, for which smoking is a strong risk factor<sup>67</sup>. Interestingly, we find that loneliness is an equally important factor explaining the relationship between MDD and peripheral artery disease. This emphasizes the need for interventions and preventive policies for reducing loneliness in the population, which has further increased in prevalence during the coronavirus disease 2019 pandemic and has been described as a pandemic itself<sup>68,69</sup>.

For most CVDs, no risk factor group could fully explain the genetic association between MDD and the CVD, which suggests the existence of additional mechanisms that are not captured by the genetic data used in the study. For instance, GWAS measure lifetime genetic risk and cannot capture dynamic processes of cumulative and interactive risk. Future studies should validate our findings using longitudinal data. Follow-up studies could also investigate the clinical usefulness of PRS for the shared liability to MDD–ASCVD and evaluate their ability to identify individuals at risk for immunometabolic depression,

as we were unable to investigate this due to sample overlap in the present study. Another limitation is that the MiXeR model was not able to accurately estimate the genetic overlap between MDD and IL6. The likely reason the MiXeR model failed in this case is the markedly different genetic architectures of the two traits, with MDD being highly polygenic and IL6 being the least polygenic of the traits in the study, combined with the low sample size of the IL6 GWAS. Finally, to assess generalizability, these findings should be replicated with data from different ancestry sources. The lack of large genetic datasets from non-European populations is a crucial limitation that is widely acknowledged and yet difficult to circumvent. Observational studies have shown that MDD and CVDs could demonstrate different associations depending on ancestry<sup>70,71</sup>, and more research is needed to understand such differences<sup>22,48–50</sup>.

Our findings suggest that the shared genetic liability to MDD and ASCVD has a distinct genomic signature compared with MDD or ASCVD separately. Moreover, the shared genetic liability shows stronger genetic correlations with immunometabolic risk factors than MDD alone, in line with the idea of an inflammatory<sup>72</sup> or immunometabolic<sup>57</sup> subtype of MDD especially associated with ASCVDs, highlighting the role of inflammation in MDD–ASCVD comorbidity. Indeed, we found that the HLA region is a hotspot of local genetic correlation between MDD and the CVD traits, that genetic liability to MDD is associated with C-reactive protein levels and tentative support that inflammatory markers mediate some of the link between MDD and ASCVD. We highlight loneliness and smoking as important targets for intervention to reduce the risk of MDD and CVD, as well as CVD in individuals with MDD. Building on this work, tools can be developed to identify individuals at risk for developing immunometabolic depression (for example, using blood tests of high-density lipoprotein and C-reactive protein levels) and target them for cholesterol-lowering or anti-inflammatory medical interventions.

## Methods

### Data sources

All data sources were summary statistics from the largest and most recent GWAS so far (Supplementary Table 1). The MDD symptoms GWAS have not been published, although they have been used in other publications<sup>64</sup>. The symptom GWASs were based on Patient Health Questionnaire (PHQ-9) items measured in the UKB, that captures most, but not all, core symptoms of MDD and are available online<sup>73</sup>. The disease trait GWASs were mostly based on electronic health record diagnoses. The physical activity and sleep duration traits were measured in the UKB using an accelerometer over a 1-week period<sup>36</sup>.

The summary statistics were processed using the cleansumstats pipeline (<https://github.com/BioPsyk/cleansumstats>). SNPs were aligned and harmonized against dbSNP reference data. Analyses were conducted on the Tjenester for Sensitive Data cluster, maintained by the University of Oslo, using singularity containers. Containers are applications packaged with environmental dependencies to facilitate standardization of analyses across different sites, ensuring correct software versions and parameters<sup>74</sup>. All GWASs were ethically approved and were conducted in compliance with ethical guidelines. Ethics approval for the UKB study was given by the North West Centre for Research Ethics Committee (11/NW/0382). The work described here was approved by UKB under application number 22224.

### Genetic overlap

**Genetic correlation.** We estimated genetic overlap on the genome-wide, polygenic and local levels. For the genome-wide level, we estimated genetic correlations using LD score regression (LDSC<sup>45</sup>). We excluded the HLA region from the analysis because its complex LD structure can bias both heritability and genetic correlation results<sup>75</sup>. Note that LDSC performs well in the presence of sample overlap.

**Local genetic correlation.** We used LAVA<sup>25</sup> to assess genetic correlation in regions of the genome. We assessed local genetic correlation in 2,495 genomic regions that cover the autosomes and have been defined to minimize LD between the regions while simultaneously keeping the regions approximately equal in size. These regions are provided with the LAVA software package. We only considered local genetic correlation in loci where both traits showed marginally significant heritability ( $P < 0.05$ ). For these loci, we adjusted local genetic correlation  $P$  values for multiple testing using the Benjamini–Hochberg method. This adjustment was done separately for all pairs of traits considered. To match the results from LDSC and MiXeR (see below), we excluded the *HLA* region from the main analyses. For pairwise local genetic correlations between MDD and the CVD traits, we performed an additional analysis in the *HLA* region.

**Genetic overlap.** To investigate genetic overlap beyond genetic correlation, we used MiXeR v1.3 (ref. 24) to assess the number of shared and distinct nonzero genetic variants between MDD and another trait required to explain at least 90% of heritability in the two traits, referred to as ‘causal’ variants. As it assesses overlap regardless of the effect direction of each variant, it gives a more adequate picture of local genetic correlations of opposite directions that cancel each other out in the genome-wide genetic correlation estimate. We excluded the *HLA* region, following the software recommendations<sup>24</sup>. To assess the stability of point estimates and estimate their standard deviations, we fitted the MiXeR model 20 different times for 2 million randomly selected SNPs with minor allele frequency of at least 5% (Supplementary Table 6). The number of 20 runs follows recommendations published previously<sup>5</sup>.

The MiXeR model was evaluated for each trait by (1) comparing the Akaike information criterion (AIC) of the univariate MiXeR model with the LDSC model, which does not include a parameter representing polygenicity (Supplementary Table 5); (2) comparing the AIC of the bivariate MiXeR model with the AIC of the model with the least possible amount of polygenic overlap required to explain observed genetic correlation; (3) comparing the AIC of the bivariate MiXeR model with the AIC of the model with maximum amount of polygenic overlap (in such a model, all risk variants of the least polygenic trait are also risk variants of the other trait) and (4) evaluating the stability of the point estimates over the 20 runs (Supplementary Table 6). These metrics have been described in detail previously<sup>5</sup>. Except for the univariate test, the MiXeR model failed these checks for IL6, which was therefore excluded from the results. The MiXeR model performed poorly for the IL6 probably because it is a trait with low polygenicity and was measured in a relatively small sample (Extended Data Fig. 2 and Supplementary Table 1).

To complement the univariate polygenicity analysis in MiXeR, we estimated polygenicity using a Bayesian framework implemented in SBayesS<sup>46</sup>. We used a 15,000-sample Markov chain with a 5,000-sample burn-in. The SBayesS methods did not converge for peripheral artery disease, possibly due to low number of cases in the GWAS.

### Shared liability to MDD and ASCVD

To move beyond bivariate association to multivariate overlap, we conducted factor analysis on MDD and the CVDs using genomic SEM<sup>26</sup> to assess whether genetic latent factors could explain the genetic covariance between the traits. Genomic SEM uses LDSC to estimate the genetic covariance matrix between traits. It then uses the genetic covariance matrix in a SEM framework to identify multivariate relationships in the data. We set CFI at  $>0.90$ , SRMR at  $<0.03$  and standardized factor loadings  $\beta > 0.3$  as criteria for acceptable model fit. The best model fit was found for a factor with coronary artery disease, peripheral artery disease, heart failure and stroke as indicators, which we interpret as a latent genetic factor representing ASCVD. To model a genetic factor for shared liability to MDD and ASCVD, we defined a higher-order factor with this ASCVD factor plus MDD as indicators. The standardized

loading of the first indicator (coronary artery disease) was set to 1. The residual variance of the ASCVD factor was forced to be 0, so that all variance was forced into the MDD–ASCVD factor. For comparison, we also estimated a common factor model for ASCVD without MDD (Extended Data Fig. 3a).

We conducted a summary statistics-based GWAS on the MDD–ASCVD second-order latent factor to identify variants associated with this latent shared liability. We used the package-default diagonally weighted least squares estimator. To derive genome-wide significant independent loci we used the plink clumping procedure as implemented in FUMA<sup>76</sup>, with  $R = 0.6$  and distance of 250 kb, and reference data from 1000 Genomes. To assess the heterogeneity of the SNP effects, we fit an independent pathway model for each SNP, where each indicator was regressed on the SNP directly instead of forcing the effect through the latent factor. We compare the common pathway  $\chi^2_{\text{SNP}}^{\text{com}}$  to the independent pathway  $\chi^2_{\text{SNP}}^{\text{ind}}$  to derive the heterogeneity measure  $Q_{\text{SNP}}$ . For follow-up analyses, we filtered out all SNPs that had effects that were more consistent with an independent pathway model at  $P_{Q_{\text{SNP}}} < 0.05$ . This stringent procedure filters out SNPs with heterogeneous effects on MDD and the ASCVD factor. This way, variants are excluded that should be considered risk variants for MDD or CVDs separately rather than risk variants for MDD–ASCVD. We did not observe genomic inflation, indicating that results were not strongly affected by population stratification (LDSC intercept of 1.02).

We used FUMA to check whether genome-wide significant SNPs for the MDD–ASCVD factor were enriched in genome-wide significant SNPs for traits in the GWAS catalog. To define unique loci, we overlaid the independent genomic risk loci (now clumped at  $R^2 = 0.1$  and 3,000 kb window) for MDD–ASCVD with the risk loci for the constituent traits (MDD, peripheral artery disease, coronary artery disease, heart failure and stroke). We used the ‘intersect’ function in bedtools to assess whether risk regions were overlapping. If they were independent (according to the clump criteria) they were regarded as novel loci. We also identified genes whose regulation is significantly impacted by the top significant SNPs to interpret the biological implications of our findings. For this, we use eQTL estimates from FUMA based on PsychENCODE reference data<sup>77</sup>. We conducted a GWAS for the latent ASCVD factors using the same procedure (though without filtering out heterogeneous SNPs).

To externally validate the latent MDD–ASCVD GWAS results, we computed PRS using LDpred2 (automatic mode with HapMap3 LD reference data), as implemented in the R package bigsnpr. We compared the MDD–ASCVD PRS with a PRS based on MDD only and with a PRS based on latent CVD without MDD. As a target sample we used the UKB. Summary statistics for ASCVD traits excluding the UKB were unavailable, and we chose to leave the UKB in for all traits. Sample overlap is likely to lead to overfitting, resulting in an inflation of explained variance. However, this is less of a concern when comparing PRS among themselves, rather than assessing absolute predictive value. As target phenotypes, we extracted CVD and MDD cases according to healthcare registry International Classification of Diseases (ICD) codes from data field 41270 (Supplementary Table 14). We used logistic regression analysis to predict case status from each PRS while controlling for the first ten principal components for ancestry, sex and year of birth. Continuous variables were standardized and centered. We used Nagelkerke’s  $R^2$  to estimate explained variance in the disease traits.

### Tissue and cell-type analysis

To gain insight into the biological mechanisms underlying the shared genetic liability to MDD and ASCVD, we performed a tissue and cell-type analysis using partitioned LDSC. Cell-type identification was based on the top decile of specifically expressed genes (referred to as top decile expression proportion (TDEP) genes). The methodology has been described extensively in previous studies<sup>78–80</sup>.



We identified TDEP genes for brain cell types from single-nucleus RNA sequencing data measured in dissections of three adult human postmortem brain samples for the Adult Human Brain Atlas<sup>27</sup>, part of the Human Cell Atlas. We used the manually annotated 31 superclusters and 461 clusters provided by the atlas<sup>28</sup>. We considered a curated set of 18,090 protein-coding autosomal genes, excluding those in the *HLA* region (because the method relies on LDSC), with expression in at least one of the 461 cell clusters.

To establish TDEP genes for 16 human tissues, we utilized bulk RNA sequencing data from GTEx v8 (ref. 81). In line with previous research, we removed tissues with <100 donors and nonnatural tissues (for example, cell lines) as well as testis tissues (expression outlier)<sup>78</sup>. Before analysis in partitioned LDSC, we expanded the boundaries of TDEP genes by 100 kb to include possible enhancers or promoters.

We tested the associations between GWAS traits and tissue or cell types by estimating heritability enrichment within the TDEP genes for each tissue or cell type. To ensure that annotation enrichment could not be better explained by other overlapping annotations, we adjusted for enrichment in 53 previously defined baseline LDSC annotations (LDSC v1.0.1) of different types of genomic regions: coding, untranslated, promoter, intronic, enhancer, histone marks and other epigenetic marks<sup>80</sup>.

### Pathways linking MDD and CVDs

**Adjusted genetic correlation and mediation in genomic SEM.** We employed several different techniques to assess whether the association between MDD and CVD could be explained by shared risk factors.

First, we estimated the genetic correlations of MDD with the individual CVD traits as well as the latent ASCVD factor adjusting for the effects of risk factors using genomic SEM. To aid interpretation, we added the covariates in groups, controlling for all trait groups separately (psychosocial or lifestyle, childhood maltreatment, metabolic or inflammation traits; Supplementary Table 1). We only included traits in the covariate groups that showed statistically significant genetic correlation with MDD. We did not adjust for blood pressure traits since none of the blood pressure traits showed statistically significant genetic correlations with MDD. Attenuation of genetic correlation after adjustment was taken to indicate that shared risk factors account for some (or all) of the association between MDD and CVD.

Next, we modeled the risk factors explicitly as mediators in genomic SEM. If the direct effect is attenuated in the mediation model as compared with the covariate model, we view this as tentative support for the existence of mediation (following procedures suggested by the software developers<sup>82</sup>). Note however that this interpretation relies on untestable assumptions. Finally, we also tested the effects of individual risk factors (both as covariates and mediators) instead of grouping them.

**Univariable and multivariable MR.** To assess tentative causal associations between MDD and CVD and risk factors, we used two-sample MR. We assessed the effects of MDD on the CVDs and risk factors, the effect of the CVDs and risk factors on MDD and the effect of latent MDD–ASCVD on risk factors. MR uses genetic variants as instrumental variables to assess the presence of causal effects of an ‘exposure’ on an ‘outcome’. Core assumptions include that the instrumental variables are robustly associated with exposure and are not associated with the outcome (other than through the effect from exposure) or unmeasured confounders. We used the IVW estimate in the two-sample MR R package<sup>83</sup> as our main estimate. As instrument SNPs, we selected independent GWAS hits at  $P < 5 \times 10^{-8}$ ,  $R^2 < 0.001$  and distance  $< 5$  Mb. In the analysis of the effect of genetic instruments of peripheral artery disease, physical activity, childhood maltreatment and IL6 on MDD risk, we allowed instruments with higher  $P$  values to be able to reach a total of ten instruments ( $P < 1 \times 10^{-5}$ ). For the MDD–ASCVD exposure, we used SNPs that showed no significantly heterogeneous effects in the genomic SEM model ( $Q_{pvalue} > 0.05$ , see above), to limit

the possibility of pleiotropic effects of this, by nature, heterogeneous instrument.

We performed several sensitivity analyses to test and adjust for violation of MR assumptions. All analyses were Steiger filtered, meaning that all SNPs that explained statistically significantly more variance in the outcome than the exposure were excluded as instruments<sup>84</sup>. Weighted median and mode regression were used to correct for effect size outliers that could represent pleiotropic effects<sup>85</sup>. MR–Egger regression was used to assess pleiotropy (pleiotropy leads to a significant intercept) and correct for it (unless  $I^2$  indicated violation of the NO Measurement Error assumption, in which case we did not report MR–Egger results<sup>86,87</sup>). We detected statistically significant pleiotropy in the analysis of the association between MDD and triglycerides, rendering the IVW estimates uninterpretable. Second, to assess the strength of our instruments, we used Cochran’s  $Q$  to assess heterogeneity in the SNP effects<sup>88</sup> and the  $F$ -statistic to control for weak instrument bias<sup>89</sup>. Third, we performed sensitivity analyses to gauge the effect of sample overlap in the GWASs that were used. Although sample overlap has been suggested to not greatly impact MR results when the source GWASs have a large sample size and the overlap is limited<sup>90</sup>, we wanted to ensure sample overlap did not lead to bias in our findings. We assessed the genetic covariance intercepts for all MDD–trait pairs from the LDSC analyses and observed that most were more than 1 s.d. away from 0, indicating that sample overlap was present (Supplementary Table 2). We repeated the analyses with MDD summary statistics leaving out the UKB sample, which is the sample responsible for most of the overlap, and compared the results. Also, we repeated the analyses using LHC MR<sup>47</sup>, which aims to correct for the presence of unmeasured heritable confounders as well as sample overlap.

We also investigated associations of genetic liability to MDD and CVD adjusted for the effect of risk factors using multivariable MR (MVMR) analyses. The difference with the mediation test in genomic SEM is that the MVMR analysis use instrument variables for MDD instead of all genetic variants, which yields results more in line with a causal interpretation. As MVMR relies on regression analysis, it cannot formally test mediation; instead, the causal estimate is adjusted for the risk factor. To support a directional interpretation, we included only risk factors as mediators that were statistically significantly affected by MDD according to the results from the univariable analysis. For the MVMR analysis, we only derive an IVW estimate. Steiger filtering was performed on the exposure–outcome association. These analyses were not replicated in LHC, which does not accommodate multivariable analyses. For ease of interpretation, we again grouped the mediators and adjusted for all mediators in a group concurrently. We only included a mediator in a group if it was significant in the univariable analysis. Additionally, we performed analyses adjusting for single mediators. We selected instruments for each variable in a model by a clumping step with the same parameters as in the univariate MR case ( $P < 5 \times 10^{-8}$ ,  $R^2 < 0.001$ ). We then combined instruments for all variables in a model into a single set of instruments and performed another clumping step with the same parameters. These instruments were then aligned to the same effect allele. We estimated the effect of genetic liability to MDD on CVD traits adjusting for covariates using multivariate MR with the MVMR R package<sup>91</sup>.

### Statistics and reproducibility

To ensure robust results, we considered only summary statistics of GWAS involving more than 10,000 cases. We maximized sample sizes available for individual analyses instead of performing replication analyses. In addition, we ensure robust results by triangulating evidence from multiple methods and statistical frameworks. These results rely on GWASs of population samples, which are not randomized. Processing and analysis was performed using the R language for statistical computing (version 4.1.0) and the Python programming language (version 2.7.13).



## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Links to download publicly available published GWAS summary statistics data used as inputs in this study are listed in Supplementary Table 1. Single-nucleus RNA sequencing data in the adult human brain can be found at <https://github.com/linnarsson-lab/adult-human-brain>. Researchers can request access to the UKB data resources at <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>; data for PRS analysis described in this study were accessed under accession number 22224. Gene expression data from human tissues can be found at <https://www.gtexportal.org/home/datasets>. Summary statistics for GWAS of the MDD-ASCVD and ASCVD latent factors are available via figshare at <https://doi.org/10.6084/m9.figshare.25737537> (ref. 92).

## Code availability

Code used for processing and analyzing data in this work can be found at [https://github.com/jacobbergstedt/MDDCVD\\_genetics](https://github.com/jacobbergstedt/MDDCVD_genetics). Code for singularity containers can be found at <https://github.com/comorment>.

## References

- Han, X. et al. Disease trajectories and mortality among individuals diagnosed with depression: a community-based cohort study in UK Biobank. *Mol. Psychiatry* **26**, 6736–6746 (2021).
- Nemeroff, C. B. & Goldschmidt-Clermont, P. J. Heartache and heartbreak—the link between depression and cardiovascular disease. *Nat. Rev. Cardiol.* **9**, 526–539 (2012).
- Kendler, K. S., Gardner, C. O., Fiske, A. & Gatz, M. Major depression and coronary artery disease in the Swedish twin registry. *Arch. Gen. Psychiatry* **66**, 857–863 (2009).
- Xian, H. et al. Genetic vulnerability and phenotypic expression of depression and risk for ischemic heart disease in the Vietnam era twin study of aging. *Psychosom. Med.* **72**, 370–375 (2010).
- Hindley, G. et al. Charting the landscape of genetic overlap between mental disorders and related traits beyond genetic correlation. *Am. J. Psychiatry* **179**, 833–843 (2022).
- Als, T. D. et al. Depression pathophysiology, risk prediction of recurrence and comorbid psychiatric disorders using genome-wide analyses. *Nat. Med.* **29**, 1832–1844 (2023).
- Raynor, D. A., Pogue-Geile, M. F., Kamarck, T. W., McCaffery, J. M. & Manuck, S. B. Covariation of psychosocial characteristics associated with cardiovascular disease: genetic and environmental influences. *Psychosom. Med.* **64**, 191–203 (2002).
- Vaccarino, V. et al. Depression and coronary heart disease: 2018 position paper of the ESC working group on coronary pathophysiology and microcirculation. *Eur. Heart J.* **41**, 1687–1696 (2020).
- Penninx, B. W. J. H. Depression and cardiovascular disease: epidemiological evidence on their linking mechanisms. *Neurosci. Biobehav. Rev.* **74**, 277–286 (2017).
- Roth, G. A. et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *J. Am. Coll. Cardiol.* **76**, 2982–3021 (2020).
- Fan, M. et al. Sleep patterns, genetic susceptibility, and incident cardiovascular disease: a prospective study of 385,292 UK biobank participants. *Eur. Heart J.* **41**, 1182–1189 (2020).
- Valtorta, N. K., Kanaan, M., Gilbody, S., Ronzi, S. & Hanratty, B. Loneliness and social isolation as risk factors for coronary heart disease and stroke: systematic review and meta-analysis of longitudinal observational studies. *Heart* **102**, 1009–1016 (2016).
- Suglia, S. F. et al. Childhood and adolescent adversity and cardiometabolic outcomes: a scientific statement from the American Heart Association. *Circulation* **137**, e15–e28 (2018).
- Schultz, W. M. et al. Socioeconomic status and cardiovascular outcomes: challenges and interventions. *Circulation* **137**, 2166–2178 (2018).
- Libby, P. et al. Atherosclerosis. *Nat. Rev. Dis. Primers* **5**, 56 (2019).
- Libby, P. et al. Inflammation, immunity, and infection in atherothrombosis: JACC review topic of the week. *J. Am. Coll. Cardiol.* **72**, 2071–2081 (2018).
- Miller, A. H. & Raison, C. L. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat. Rev. Immunol.* **16**, 22–34 (2016).
- Osimo, E. F., Baxter, L. J., Lewis, G., Jones, P. B. & Khandaker, G. M. Prevalence of low-grade inflammation in depression: a systematic review and meta-analysis of CRP levels. *Psychol. Med.* **49**, 1958–1970 (2019).
- Suneson, K., Lindahl, J., Chamli Hårsmar, S., Söderberg, G. & Lindqvist, D. Inflammatory depression-mechanisms and non-pharmacological interventions. *Int. J. Mol. Sci.* **22**, 1640 (2021).
- Carney, R. M. & Freedland, K. E. Depression and coronary heart disease. *Nat. Rev. Cardiol.* **14**, 145–155 (2017).
- Lu, Y., Wang, Z., Georgakis, M. K., Lin, H. & Zheng, L. Genetic liability to depression and risk of coronary artery disease, myocardial infarction, and other cardiovascular outcomes. *J. Am. Heart Assoc.* **10**, e017986 (2021).
- Li, G. H.-Y. et al. Evaluation of bi-directional causal association between depression and cardiovascular diseases: a Mendelian randomization study. *Psychol. Med.* **52**, 1765–1776 (2022).
- Hare, D. L., Toukhsati, S. R., Johansson, P. & Jaarsma, T. Depression and cardiovascular disease: a clinical review. *Eur. Heart J.* **35**, 1365–1372 (2014).
- Frei, O. et al. Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nat. Commun.* **10**, 2417 (2019).
- Werme, J., van der Sluis, S., Posthuma, D. & de Leeuw, C. A. An integrated framework for local genetic correlation analysis. *Nat. Genet.* **54**, 274–282 (2022).
- Grotzinger, A. D. et al. Genomic SEM provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* **3**, 513–525 (2018).
- Siletti, K. et al. Transcriptomic diversity of cell types across the adult human brain. *Science* **382**, eadd7046 (2023).
- van Zuydam, N. R. et al. Genome-wide association study of peripheral artery disease. *Circ. Genom. Precis. Med.* **14**, e002862 (2021).
- Shah, S. et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat. Commun.* **11**, 163 (2020).
- Aragam, K. G. et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat. Genet.* **54**, 1803–1815 (2022).
- Mishra, A. et al. Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature* **611**, 115–123 (2022).
- Roselli, C. et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat. Genet.* **50**, 1225–1233 (2018).
- Evangelou, E. et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet.* **50**, 1412–1425 (2018).
- Lee, J. J. et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
- Liu, M. et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019).
- Doherty, A. et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat. Commun.* **9**, 5257 (2018).

37. Wang, Z. et al. Genome-wide association analyses of physical activity and sedentary behavior provide insights into underlying mechanisms and roles in disease prevention. *Nat. Genet.* **54**, 1332–1344 (2022).
38. Day, F. R., Ong, K. K. & Perry, J. R. B. Elucidating the genetic basis of social interaction and isolation. *Nat. Commun.* **9**, 2457 (2018).
39. Warriar, V. et al. Gene–environment correlations and causal effects of childhood maltreatment on physical and mental health: a genetically informed approach. *Lancet Psychiatry* **8**, 373–386 (2021).
40. Mahajan, A. et al. Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. *Nat. Genet.* **54**, 560–572 (2022).
41. Graham, S. E. et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675–679 (2021).
42. Yengo, L. et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700,000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
43. Ahluwalia, T. S. et al. Genome-wide association study of circulating interleukin 6 levels identifies novel loci. *Hum. Mol. Genet.* **30**, 393–409 (2021).
44. Said, S. et al. Genetic analysis of over half a million people characterises C-reactive protein loci. *Nat. Commun.* **13**, 2198 (2022).
45. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
46. Zeng, J. et al. Widespread signatures of natural selection across human complex traits and functional genomic categories. *Nat. Commun.* **12**, 1164 (2021).
47. Darrous, L., Mounier, N. & Kutalik, Z. Simultaneous estimation of bi-directional causal effects and heritable confounding from GWAS summary statistics. *Nat. Commun.* **12**, 7274 (2021).
48. Zhang, F., Cao, H. & Baranova, A. Shared genetic liability and causal associations between major depressive disorder and cardiovascular diseases. *Front. Cardiovasc. Med.* **8**, 735136 (2021).
49. Torgersen, K. et al. Shared genetic loci between depression and cardiometabolic traits. *PLoS Genet.* **18**, e1010161 (2022).
50. Hagenaaers, S. P. et al. Genetic comorbidity between major depression and cardio-metabolic traits, stratified by age at onset of major depression. *Am. J. Med. Genet. B* **183**, 309–330 (2020).
51. Wei, H.-L. et al. Disrupted resting-state functional connectivity of the thalamus in patients with coronary heart disease. *Heliyon* **9**, e13423 (2023).
52. Rosen, S. D. Silent ischemia as a central problem: regional brain activation compared in silent and painful myocardial ischemia. *Ann. Intern. Med.* **124**, 939–949 (1996).
53. Rosen, S. D. et al. Central nervous pathways mediating angina pectoris. *Lancet* **344**, 147–150 (1994).
54. Greicius, M. D. et al. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. *Biol. Psychiatry* **62**, 429–437 (2007).
55. Paul, E. R., Östman, L., Heilig, M., Mayberg, H. S. & Hamilton, J. P. Towards a multilevel model of major depression: genes, immuno-metabolic function, and cortico-striatal signaling. *Transl. Psychiatry* **13**, 171 (2023).
56. Zhao, B. et al. Heart–brain connections: phenotypic and genetic insights from magnetic resonance images. *Science* **380**, abn6598 (2023).
57. Milaneschi, Y., Lamers, F., Berk, M. & Penninx, B. W. J. H. Depression heterogeneity and its biological underpinnings: toward immunometabolic depression. *Biol. Psychiatry* **88**, 369–380 (2020).
58. Dashti, H. S. et al. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nat. Commun.* **10**, 1100 (2019).
59. Shen, Q. et al. Psychiatric disorders and subsequent risk of cardiovascular disease: a longitudinal matched cohort study across three countries. *EClinicalMedicine* **61**, 102063 (2023).
60. Tang, B., Yuan, S., Xiong, Y., He, Q. & Larsson, S. C. Major depressive disorder and cardiometabolic diseases: a bidirectional Mendelian randomisation study. *Diabetologia* **63**, 1305–1311 (2020).
61. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
62. Morris, T. T. et al. Interpretation of Mendelian randomization using a single measure of an exposure that varies over time. *Int. J. Epidemiol.* **51**, 1899–1909 (2022).
63. Taylor, W. D., Aizenstein, H. J. & Alexopoulos, G. S. The vascular depression hypothesis: mechanisms linking vascular disease with depression. *Mol. Psychiatry* **18**, 963–974 (2013).
64. Kappelmann, N. et al. Dissecting the association between inflammation, metabolic dysregulation, and specific depressive symptoms: a genetic correlation and 2-sample mendelian randomization study. *JAMA Psychiatry* **78**, 161–170 (2021).
65. Khandaker, G. M. et al. Shared mechanisms between coronary heart disease and depression: findings from a large UK general population-based cohort. *Mol. Psychiatry* **25**, 1477–1486 (2020).
66. Wootton, R. E. et al. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychol. Med.* **50**, 2435–2443 (2020).
67. Behrooz, L., Abumoawad, A., Rizvi, S. H. M. & Hamburg, N. M. A modern day perspective on smoking in peripheral artery disease. *Front. Cardiovasc. Med.* **10**, 1154708 (2023).
68. Office of the Surgeon General (OSG). *Our Epidemic of Loneliness and Isolation: The U.S. Surgeon General's Advisory on the Healing Effects of Social Connection and Community* (US Department of Health and Human Services, 2023).
69. Ernst, M. et al. Loneliness before and during the COVID-19 pandemic: a systematic review with meta-analysis. *Am. Psychol.* **77**, 660–677 (2022).
70. González, H. M. & Tarraf, W. Comorbid cardiovascular disease and major depression among ethnic and racial groups in the United States. *Int. Psychogeriatr.* **25**, 833–841 (2013).
71. Watkins, D. C., Assari, S. & Johnson-Lawrence, V. Race and ethnic group differences in comorbid major depressive disorder, generalized anxiety disorder, and chronic medical conditions. *J. Racial Ethn. Health Disparities* **2**, 385–394 (2015).
72. Felger, J. C. & Miller, A. H. Identifying immunophenotypes of inflammation in depression: dismantling the monolith. *Biol. Psychiatry* **88**, 136–138 (2020).
73. Howrigan, D. P. et al. Nealelab/UK\_Biobank\_GWAS: v2. *Zenodo* <https://doi.org/10.5281/ZENODO.8011558> (2023).
74. Akdeniz, B. C. et al. COSGAP: COntainerized statistical genetics analysis pipelines. *Bioinformatics Adv.* **4**, vbae067 (2024).
75. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
76. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
77. Wang, D. et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).

78. Bryois, J. et al. Genetic identification of cell types underlying brain complex traits yields insights into the etiology of Parkinson's disease. *Nat. Genet.* **52**, 482–493 (2020).
79. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
80. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
81. Aguet, F. et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
82. Nivard, M. Mediation model in GenomicSEM. *GitHub* <https://gist.github.com/MichelNivard/04bf4ddcf3c32f905175de3058ca967a> (2018).
83. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* **27**, R195–R208 (2018).
84. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081 (2017).
85. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* **46**, 1985–1998 (2017).
86. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
87. Bowden, J. et al. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR–Egger regression: the role of the  $I^2$  statistic. *Int. J. Epidemiol.* **45**, 1961–1974 (2016).
88. Bowden, J., Hemani, G. & Davey Smith, G. Invited commentary: detecting individual and global horizontal pleiotropy in Mendelian randomization—a job for the humble heterogeneity statistic? *Am. J. Epidemiol.* **187**, 2681–2685 (2018).
89. Burgess, S. & Thompson, S. G. Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* **40**, 755–764 (2011).
90. Mounier, N. & Kutalik, Z. Bias correction for inverse variance weighting Mendelian randomization. *Genet. Epidemiol.* **47**, 314–331 (2023).
91. Sanderson, E., Spiller, W. & Bowden, J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat. Med.* **40**, 5434–5452 (2021).
92. Bergstedt, J. MDD-ASCVD genome-wide association study summary statistics. *figshare* <https://doi.org/10.6084/m9.figshare.25737537> (2024).

## Acknowledgements

This work was supported by European Union's Horizon 2020 Research and Innovation Programme (CoMorMent project; grant no. 847776). J.A.P. was supported by the US National Institutes of Mental Health (R01MH123724) and the Amsterdam University Medical Center Postdoc Career Bridging grant (27527). N.P. was supported by the Marie Skłodowska-Curie Actions Grant 801133 (Scientia fellowship) and RCN 300309. J.L.T. is supported by a European Research Council Starting grant (UNRAVEL-CAUSALITY, grant number 101076686) and a Senior Scientist Dekker Grant from the Dutch Heart Foundation (project number 03-004-2022-0055). Q.S. was supported by the Outstanding Clinical Discipline Project of Shanghai Pudong (no. PWYgy2021-02), the Talent training project of the three-year action plan (2023–2025) for strengthening the construction of the public health system in Shanghai (no. GWVI-11.2-YQ14), the Interdisciplinary Joint Research Project of Tongji University (2023-3-YB-05) and the Fundamental Research Funds for the Central Universities. C.M.L. is part-funded by the NIHR Maudsley Biomedical Research Centre at

South London and Maudsley NHS Foundation Trust and King's College London. A.M.M. is supported by the Wellcome Trust (220857/Z/20/Z) and UKRI (MR/W014386/1). A.D.B. was supported by grants from the Lundbeck Foundation (R102-A9118, R155-2014-1724 and R248-2017-2003) and NIH/NIMH (1R01MH124851-01). O.A.A. was supported by the European Union's Horizon 2020 research and innovation program (grant agreement nos. 847776 and 964874), Research Council of Norway (223273, 296030, 326813 and 324499). P.F.S. gratefully acknowledges support from the Swedish Research Council (Vetenskapsrådet, award D0886501), the US National Institutes of Mental Health (R01s MH124871, MH121545 and MH123724) and the Horizon 2020 Programme of the European Union (grant agreement no. 847776). Y.L. was supported by the European Research Council (grant agreement ID 101042183). We acknowledge HERMES and CARDIoGRAMplusC4D consortia for contributing data and guidance. The PRS analyses have been conducted using the UKB resource under application number 22224. They were enabled by resources in project sens2017519 provided by the National Academic Infrastructure for Supercomputing in Sweden at UPPMAX, funded by the Swedish Research Council through grant agreement no. 2022-06725. This work was performed on Services for Sensitive Data, University of Oslo, with resources provided by UNINETT Sigma2, the national infrastructure for high performance computing and data storage in Norway.

## Author contributions

This study was conceived by J.B., P.F.S., Y.L. and F.F. It was supervised by Y.L. and F.F. J.B., J.A.P. and Z.M. were responsible for the main analyses, visualization and the drafting of the paper. A.H. performed the tissue expression analyses with help and supervision from S.Y., J.H.-L. and P.F.S., who also helped with the writing and interpretation of these results. N.P. was involved in providing and cleaning GWAS summary statistic data. D.J.A.S. supervised the genomic SEM analysis and J.L.T. supervised the MR analyses. O.F. and A.S. were involved in providing the containerized software used in this project and supervising the MiXeR analyses. J.J.M., Q.S., S.H., P.T., A.B. and T.W. contributed to interpreting findings. C.M.L., A.M.M., T.D.A. and A.D.B. were responsible for the MDD GWAS summary statistics and were involved in the interpretation of findings. S.H. helped with UKB data access and curation for the PRS analyses. S.H., U.A.V., O.A.A., P.F.S., Y.L. and F.F. provided suggestions and expert supervision and contributed to the writing. All authors have read and provided feedback on the paper.

## Funding

Open access funding provided by Karolinska Institute

## Competing interests

C.M.L. sits on the SAB of Myriad Neuroscience and has received speaker/consultancy fees from SYNLAB and UCB. A.D.B. has received speaker's honoraria from Lundbeck. O.A.A. is a consultant to Cortechs.ai and Precision Health, and has received speaker's honoraria from Lundbeck, Janssen, Otsuka and Sunovion. P.F.S. has received consulting fees from and is a shareholder of Neumora Therapeutics. The other authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s44161-024-00488-y>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44161-024-00488-y>.

**Correspondence and requests for materials** should be addressed to Jacob Bergstedt or Fang Fang.

**Peer review information** *Nature Cardiovascular Research* thanks Gabriëlla Blokland, Albert Henry and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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

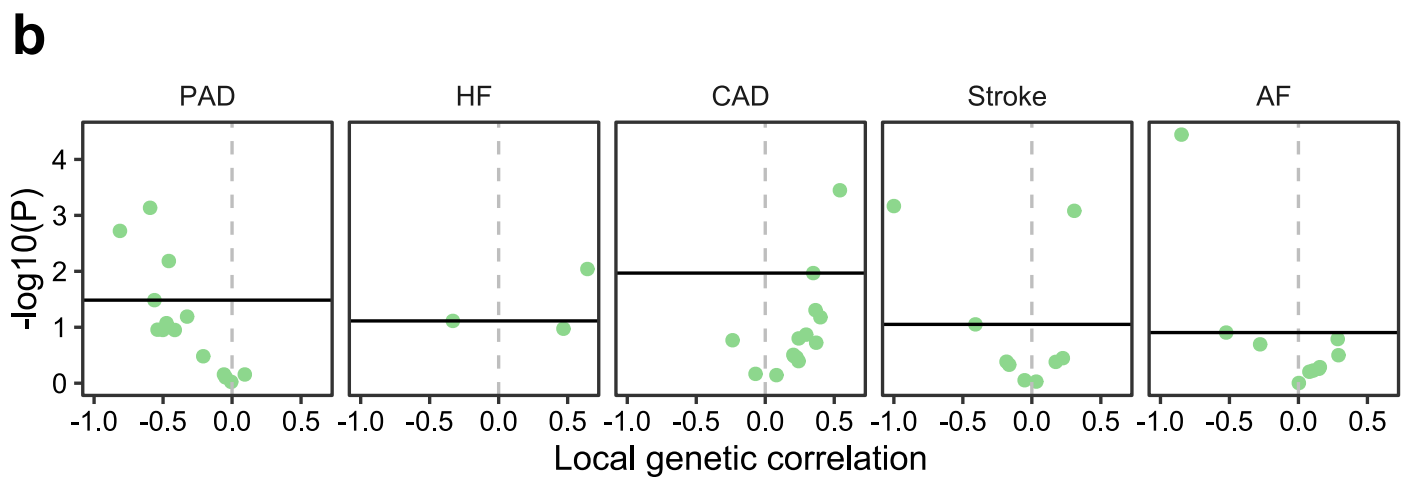
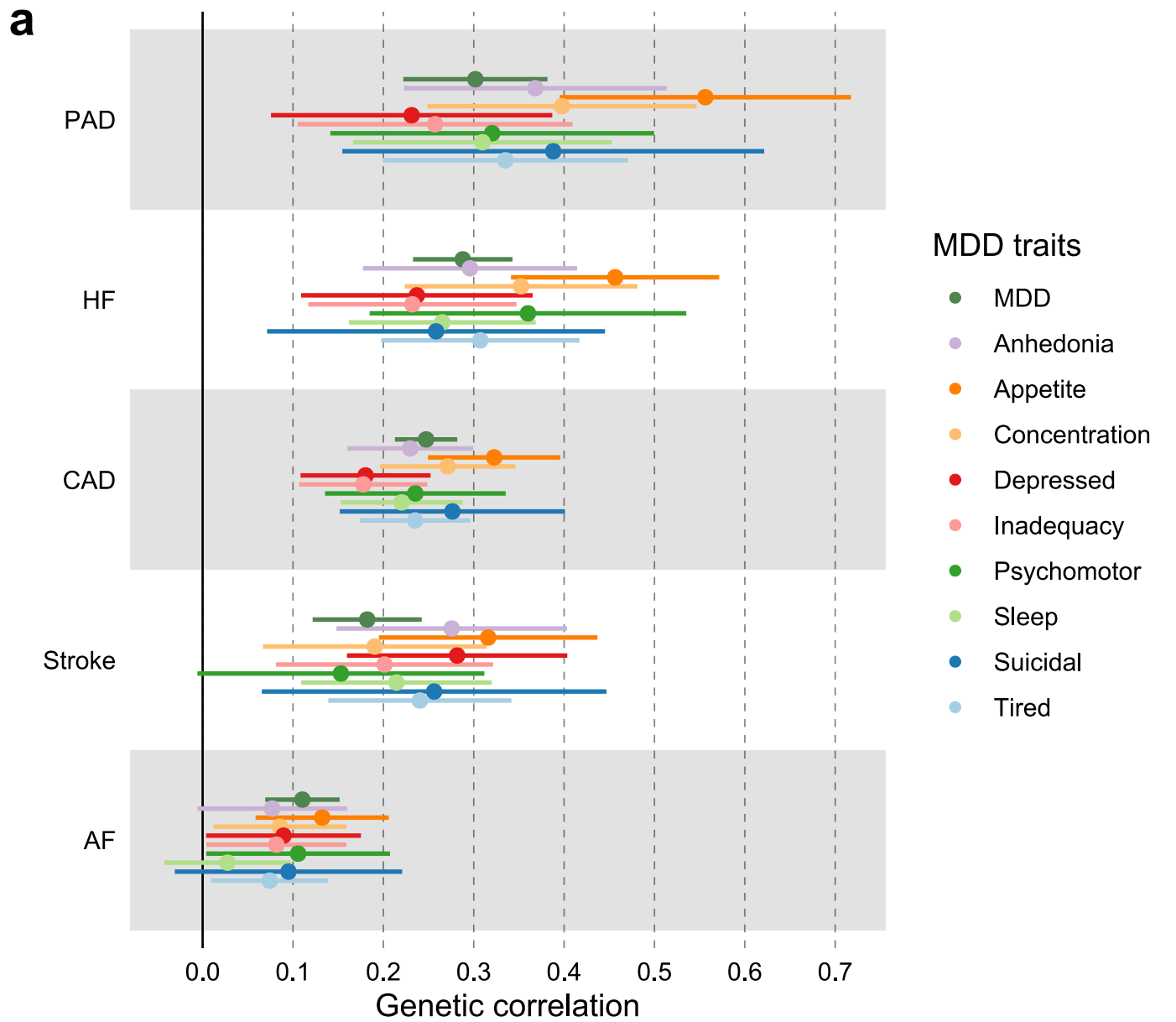
**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing,

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/>.

© The Author(s) 2024

<sup>1</sup>Unit of Integrative Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. <sup>2</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. <sup>3</sup>Department Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. <sup>4</sup>Centre for Precision Psychiatry, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway. <sup>5</sup>Genetic Epidemiology, Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands. <sup>6</sup>Centre for Bioinformatics, Department of Informatics, University of Oslo, Oslo, Norway. <sup>7</sup>K.G. Jebsen Centre for Neurodevelopmental Disorders, University of Oslo and Oslo University Hospital, Oslo, Norway. <sup>8</sup>Institute of Biological Psychiatry, Mental Health Center Sct. Hans, Mental Health Services Copenhagen, Roskilde, Denmark. <sup>9</sup>Clinical Research Center for Mental Disorders, Shanghai Pudong New Area Mental Health Center, Tongji University School of Medicine, Shanghai, China. <sup>10</sup>Institute for Advanced Study, Tongji University, Shanghai, China. <sup>11</sup>Department of Clinical Science and Education Södersjukhuset, Karolinska Institutet, Stockholm, Sweden. <sup>12</sup>Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. <sup>13</sup>Department of Molecular Medicine (MOMA), Molecular Diagnostic Laboratory, Aarhus University Hospital, Aarhus, Denmark. <sup>14</sup>Department of Biomedicine, Aarhus University, Aarhus, Denmark. <sup>15</sup>Center for Genomics and Personalized Medicine, Aarhus, Denmark. <sup>16</sup>Social, Genetic and Developmental Psychiatry Centre, King's College London, London, UK. <sup>17</sup>Department of Medical and Molecular Genetics, King's College London, London, UK. <sup>18</sup>Centre for Clinical Brain Sciences, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK. <sup>19</sup>Centre for Genomics and Experimental Medicine, University of Edinburgh, Edinburgh, UK. <sup>20</sup>Centre of Public Health Sciences, Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland. <sup>21</sup>Department of Epidemiology, Harvard TH Chan School of Public Health, Harvard University, Boston, MA, USA. <sup>22</sup>Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>23</sup>Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>24</sup>These authors contributed equally: Jacob Bergstedt, Joëlle A. Pasman, Ziyang Ma, Yi Lu, Fang Fang. ✉ e-mail: [jacob.bergstedt@ki.se](mailto:jacob.bergstedt@ki.se); [fang.fang@ki.se](mailto:fang.fang@ki.se)

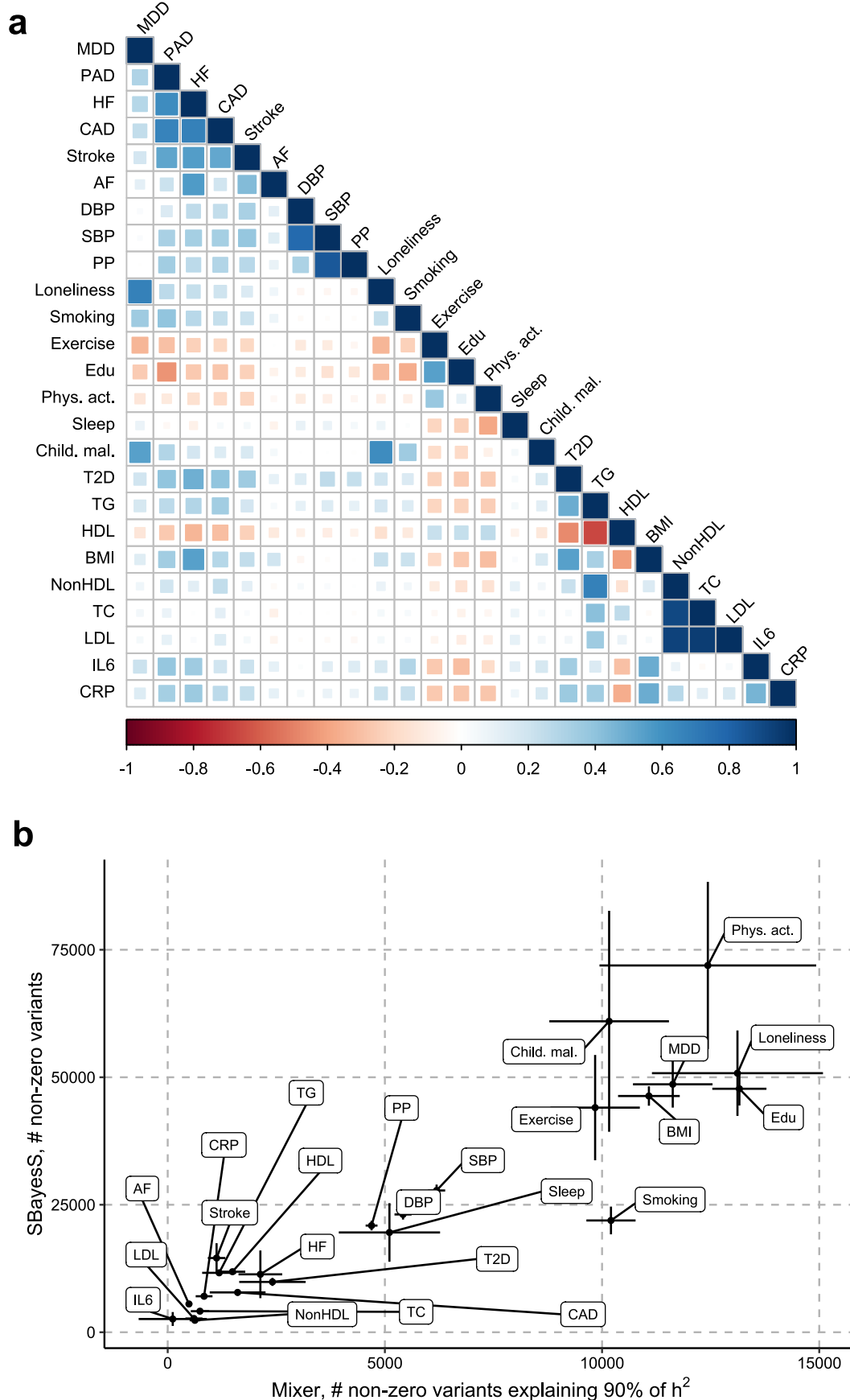




Extended Data Fig. 1 | See next page for caption.

**Extended Data Fig. 1 | Genetic correlation of MDD and MDD symptoms with CVDs. a.** Results are based on LD Score Regression analysis. Points and error bars represent mean genetic correlation and 95% CIs. **b.** Local genetic correlation between MDD and CVDs in 16 loci in the HLA region. Only loci with marginally statistically significant local heritability for both traits are shown. Points above the vertical line are significant based on multiple testing adjustment for considered loci performed for each CVD trait separately. Multiple testing was

adjusted for using the Benjamini-Hochberg procedure. Empirical *P*-values were obtained via a permutation procedure with partial integration, evaluating the two-sided hypothesis of no association using the estimated parameters as test statistics **a-b** Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; HF=Heart Failure; MDD=Major depressive disorder; PAD=Peripheral Artery Disease.

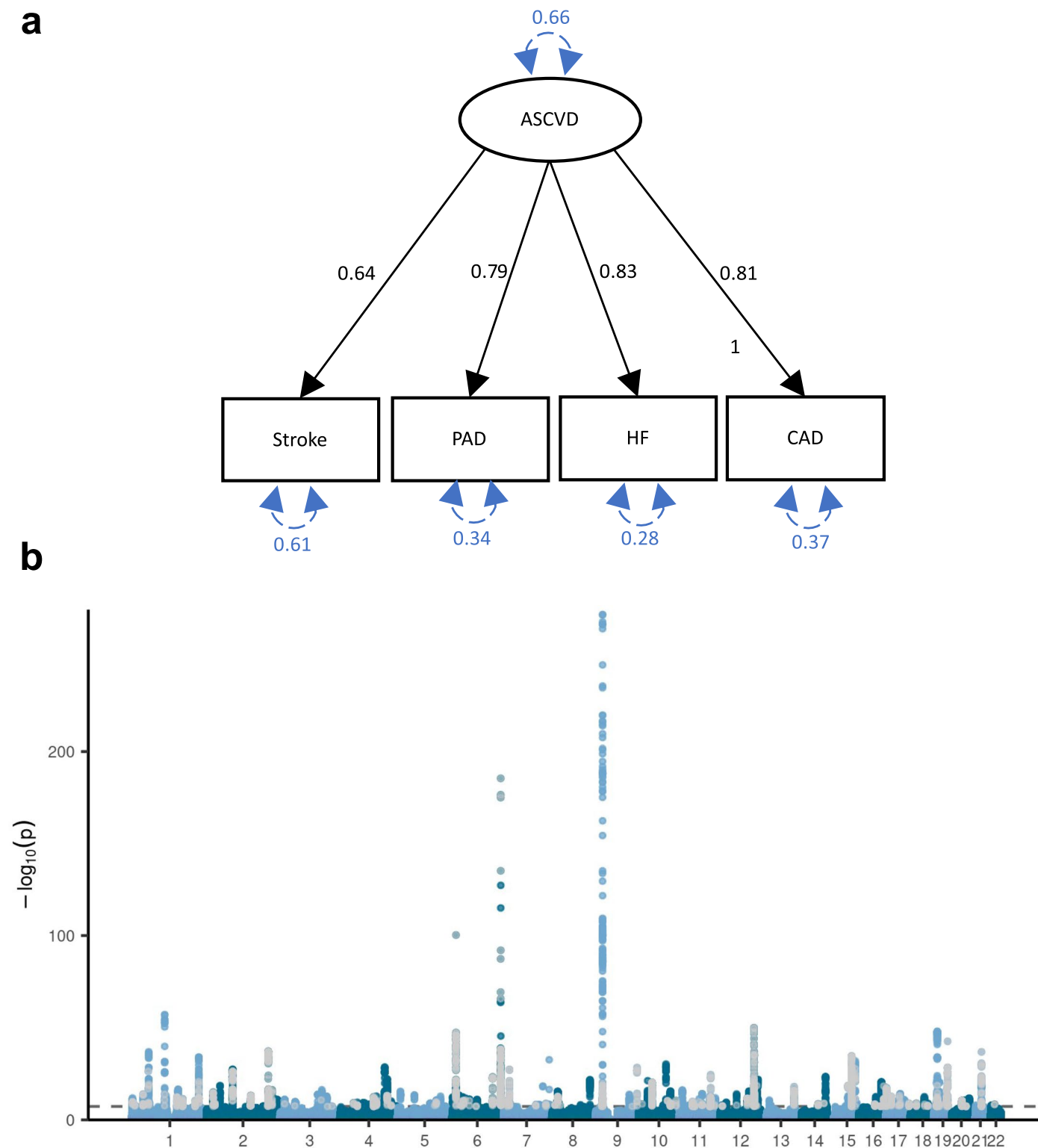


Extended Data Fig. 2 | See next page for caption.

**Extended Data Fig. 2 | Polygenicity and genetic correlation for MDD, CVD, and risk factors. a,** Heatmap of the genetic correlations between MDD, the CVDs, and the risk factors, with the color indicating the effect direction (negative: red, positive: blue) and the size and shade of the square illustrating the size of the correlation. Results are based on LD Score Regression analysis. **b,** Polygenicity estimates from SBayesS (y-axis) and MiXeR (x-axis) for MDD, the CVDs, and the risk factors. Note that for PAD polygenicity estimates did not converge for SBayesS, possibly because of the few number of cases. Points and error bars represent the mean number of estimated non-zero variants

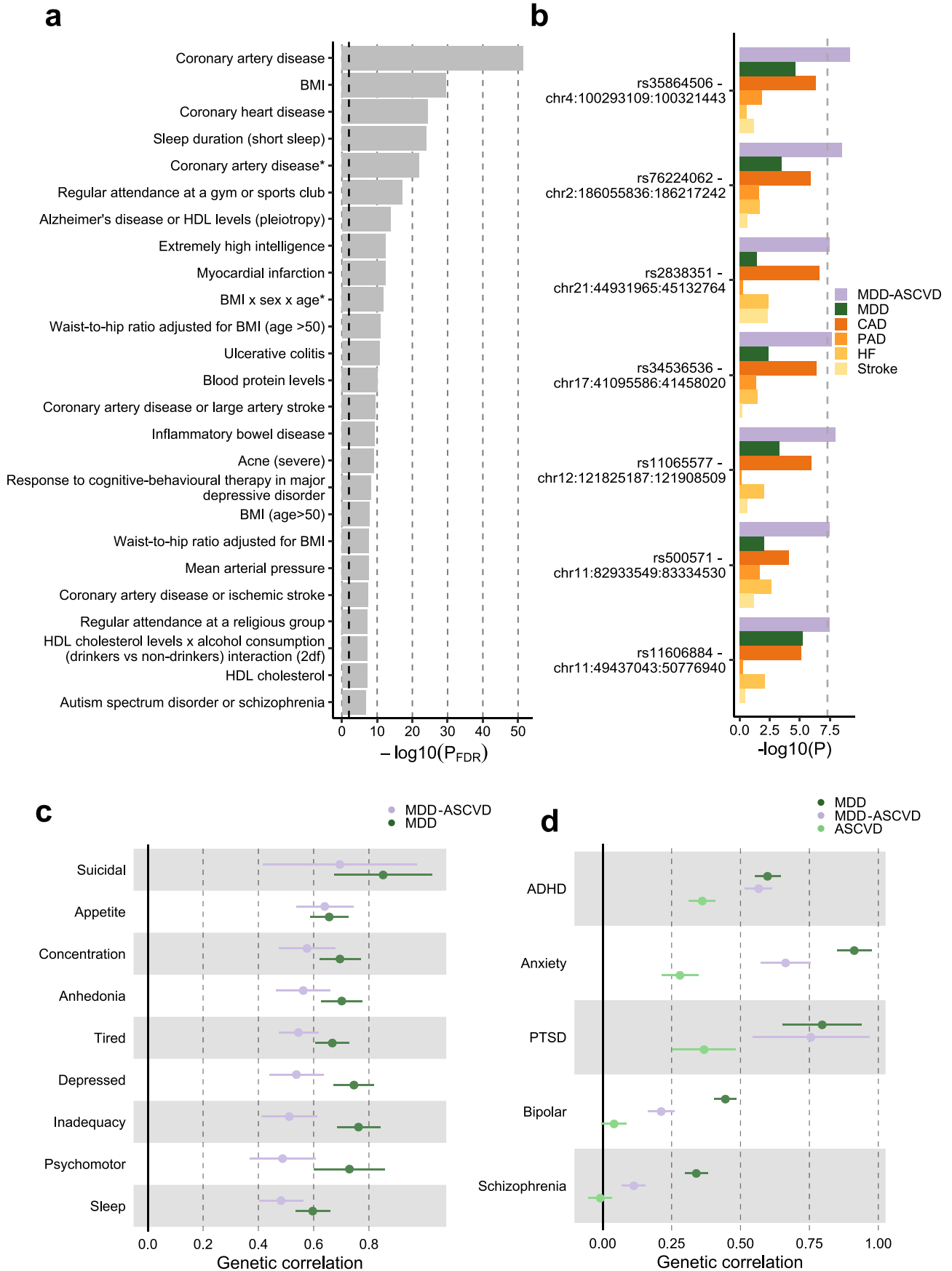
and 95% CIs. **a, b** Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; Child. Mal.=Childhood Maltreatment; CRP=C-Reactive Protein; DBP=Diastolic Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; HF=Heart Failure; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; PAD=Peripheral Artery Disease; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.





**Extended Data Fig. 3 | Genetic liability to ASCVD latent factor. a**, Latent atherosclerotic CVD (ASCVD) model, defined by stroke, peripheral artery disease (PAD), heart failure (HF), and coronary artery disease (CAD). All ‘observed’ traits are based on GWAS summary statistics. Results are from confirmatory factor analysis in Genomic SEM, and standardized factor loadings are given for each path. Circular dashed arrows give the trait variance. **b**, Manhattan plot of the GWAS on ASCVD, with each dot representing a SNP with its position on the x-axis and its  $P$ -value on the y-axis. Genome-wide significant SNPs with a significant

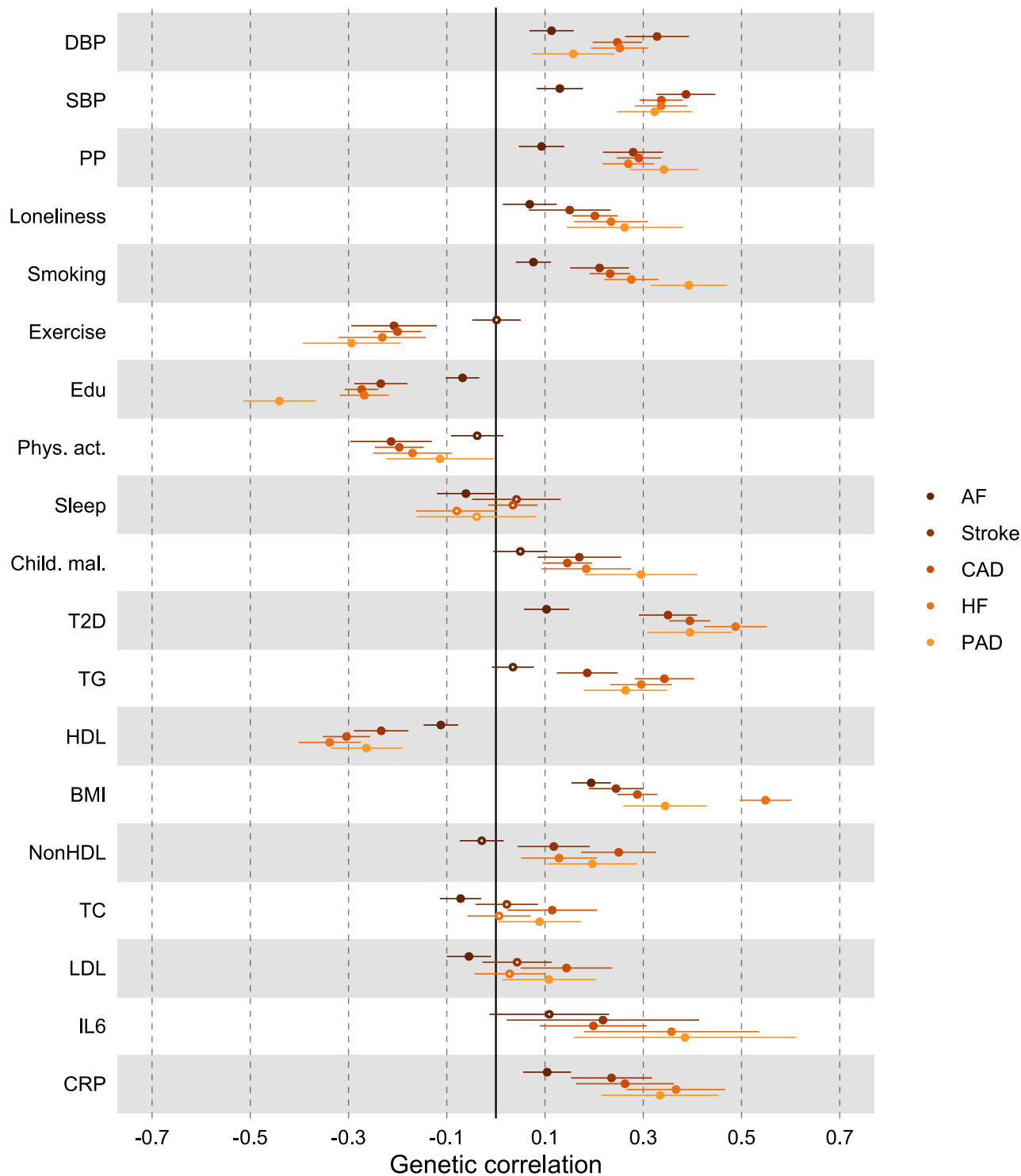
heterogeneity  $Q$  (with a strong effect on one or some of the indicators that was not well explained through the common latent factor) are displayed in grey. The dashed line indicates the genome-wide significance threshold ( $P < 5e-8$ ).  $P$ -values were computed using a two-sided Z-test. **a-b** Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; HF=Heart Failure; MDD=Major depressive disorder; PAD=Peripheral Artery Disease.



Extended Data Fig. 4 | See next page for caption.

**Extended Data Fig. 4 | The genetic signature of MDD-ASCVD. a**, Enrichment for the MDD-ASCVD GWAS SNPs in genome-wide significant SNPs for traits in the GWAS catalog computed using FUMA. The traits are as reported in the original study. Note that sleep duration here is a dichotomization of self-reported sleep. The dashed grey line indicates the significance threshold after FDR-adjustment. Adjustment for multiple testing was conducted over all traits in the GWAS catalog using the Benjamini-Hochberg procedure. *P*-values were computed using a one-sided hypergeometric test. **b**, The five SNPs that were significantly associated with MDD-ASCVD, but not with any of the constituent traits, with their *P*-value in the GWAS of the constituent traits. The dashed line indicates the genome-wide

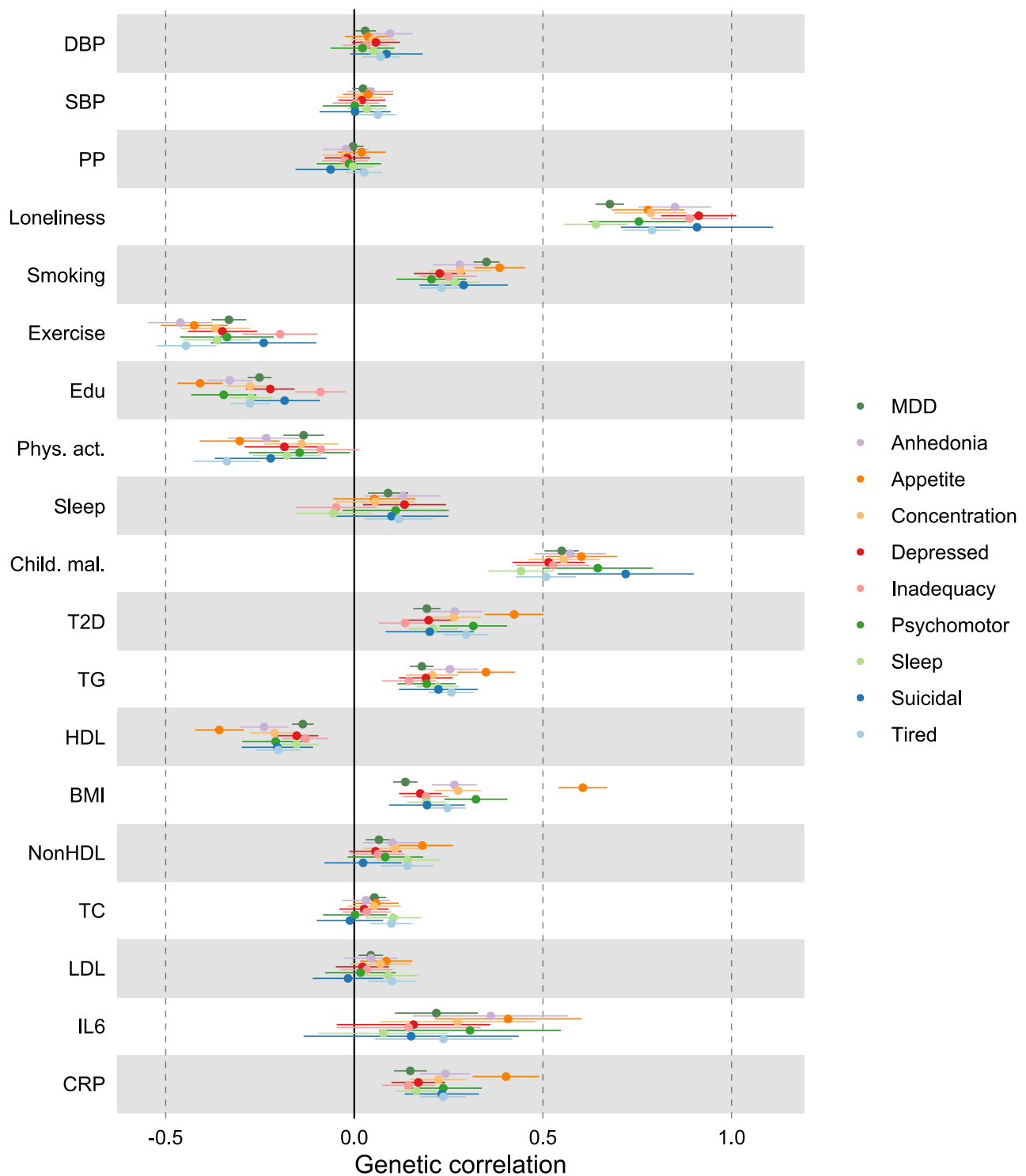
significance threshold ( $P < 5e-8$ ). *P*-values from constituent traits are taken from the original GWAS, see Supplementary Table 1. *P*-values for MDD-ASCVD were computed using a two-sided Z-test **c**, Genetic correlation of MDD symptoms with MDD, and MDD-ASCVD. **d** Genetic correlation of MDD, MDD-ASCVD, ASCVD, with five mental disorders. **c-d** Points and error bars represent mean genetic correlation and 95% CIs. Results are computed using LD Score Regression analysis. **a-d** Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. ADHD = Attention Deficit and Hyperactivity Disorder; PTSD = Posttraumatic Stress Disorder.



**Extended Data Fig. 5 | Genetic correlation between CVDs and risk factors.**

Results are based on LD Score Regression analysis. Points and error bars represent mean genetic correlation and 95% CIs. Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. Open dots indicate a non-significant genetic correlation. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; Child. Mal.=Childhood Maltreatment;

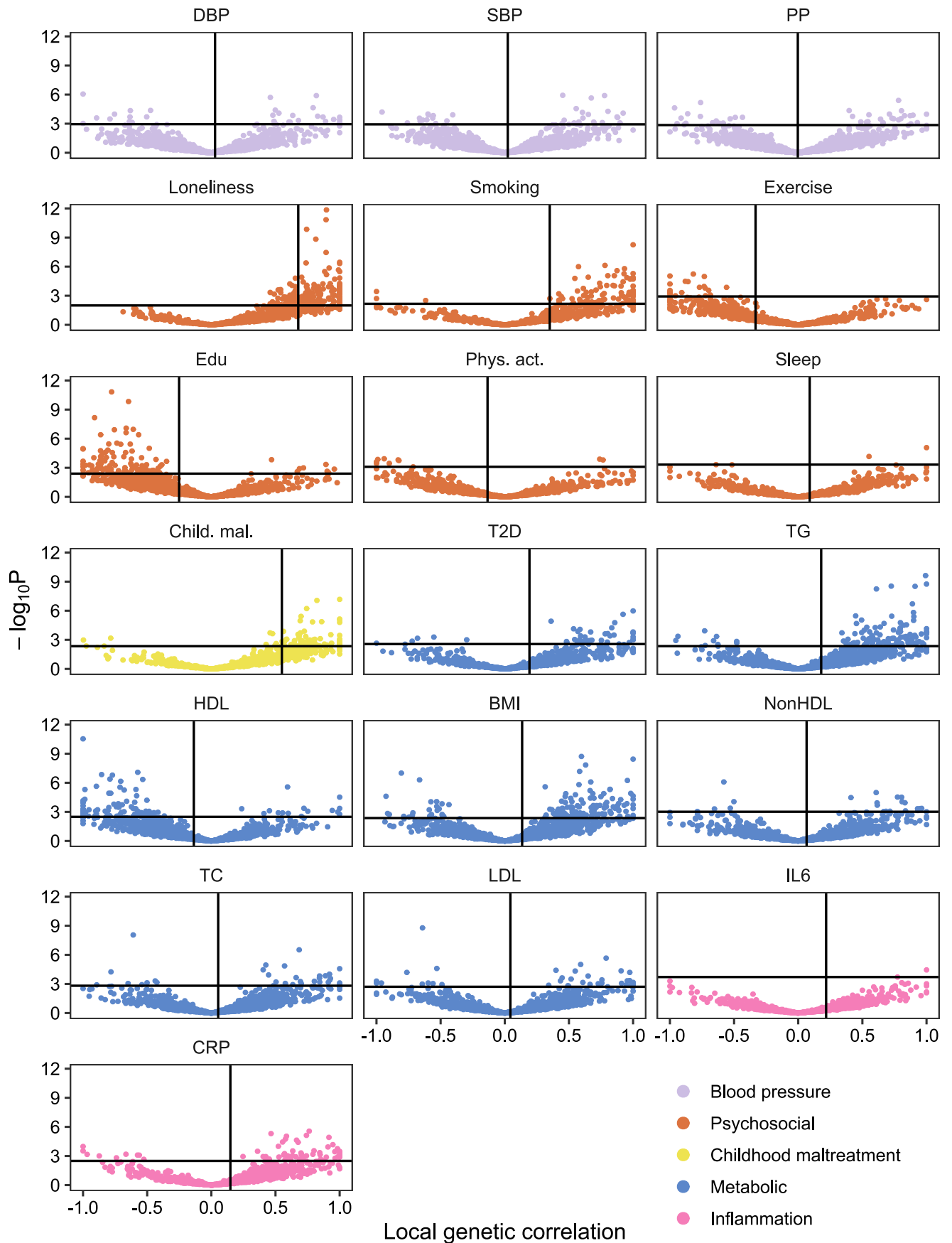
CRP = C-Reactive Protein; DBP=Diastolic Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; HF=Heart Failure; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; PAD=Peripheral Artery Disease; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.



**Extended Data Fig. 6 | Genetic correlation between MDD traits and risk factors.** Results are based on LD Score Regression analysis. Points and error bars represent mean genetic correlation and 95% CIs. Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. Open dots indicate a non-significant genetic correlation. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; Child. Mal.=Childhood Maltreatment;

CRP = C-Reactive Protein; DBP=Diastolic Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; HF=Heart Failure; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; PAD=Peripheral Artery Disease; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.

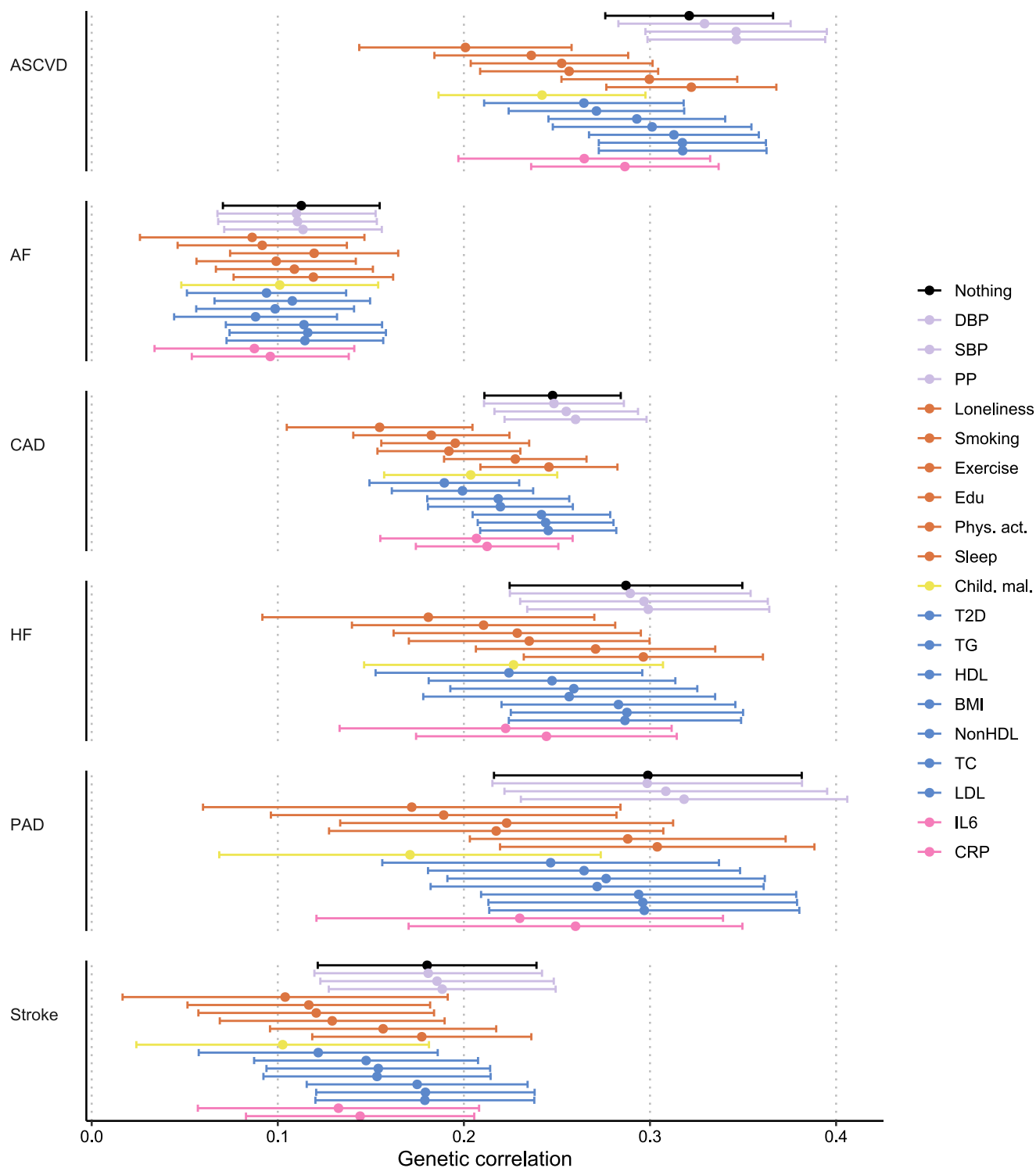




Extended Data Fig. 7 | See next page for caption.

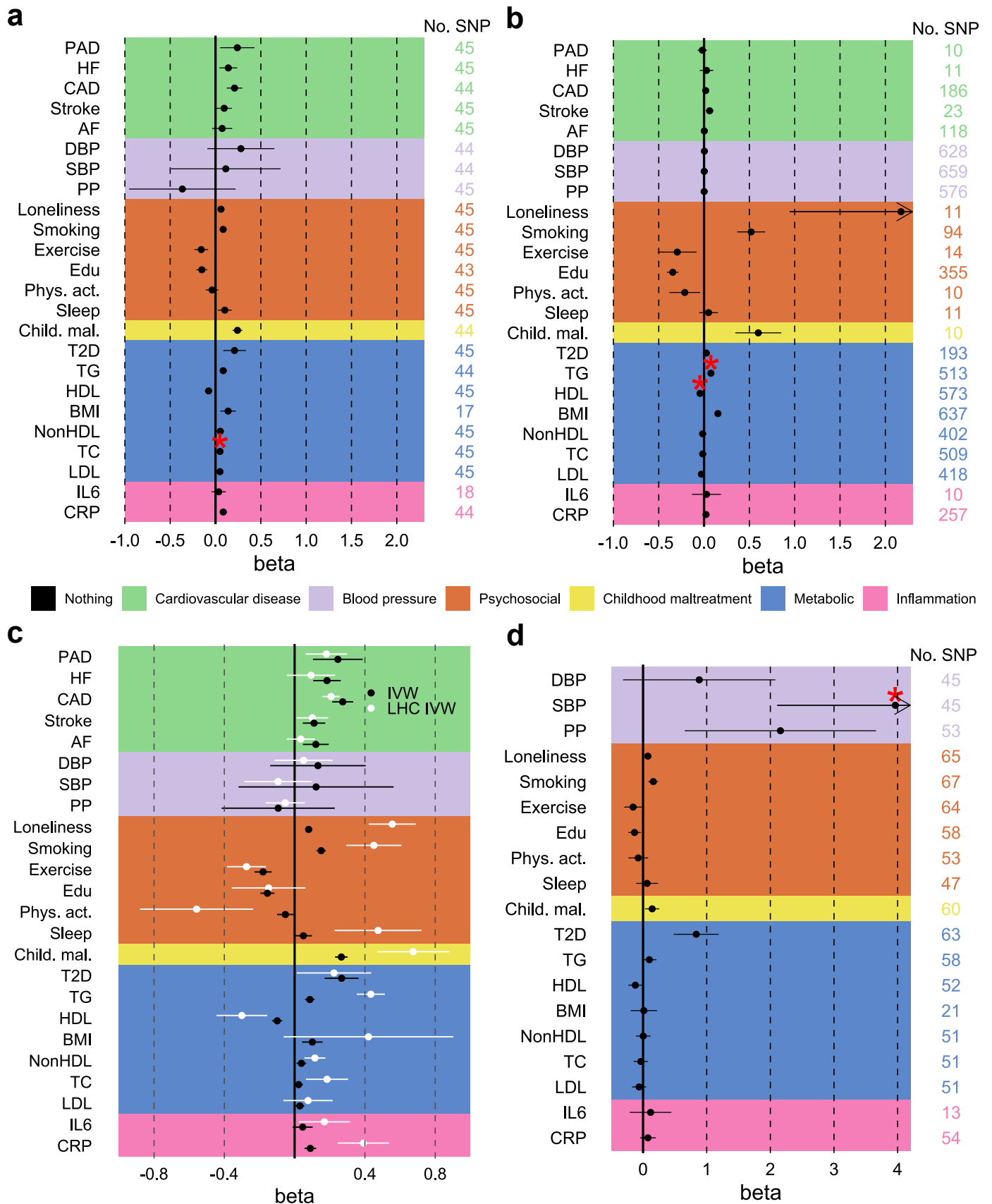
**Extended Data Fig. 7 | Local genetic correlations between MDD and risk factors.** Volcano plots based on LAVA results. Local genetic correlation between MDD and each of the risk factors (x-axis) and the corresponding  $-\log_{10}$  transformed  $P$ -value (y-axis). Empirical  $P$ -values were obtained via a permutation procedure with partial integration, evaluating the two-sided hypothesis of no association using the estimated parameters as test statistics. Correlations were estimated in the loci that showed marginally significant local heritability in 2,495 considered genomic regions. Loci exceeding the horizontal line are significant at  $P_{\text{FDR}} < .05$ . Multiple testing was adjusted individually for each trait over

considered loci using the Benjamini-Hochberg procedure. Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. Child. Mal.=Childhood Maltreatment; CRP = C-Reactive Protein; DBP=Diastolic Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.



**Extended Data Fig. 8 | Genetic correlation between MDD and CVD adjusting for individual risk factors.** Results from Genomic SEM. Points and error bars represent mean genetic correlation and 95% confidence intervals. Reference estimates of the association without any adjustment are printed in black. Sample sizes for underlying GWAS summary statistics are reported in Supplementary Table 1. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; Child. Mal.=Childhood Maltreatment; CRP=C-Reactive Protein; DBP=Diastolic

Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; HF=Heart Failure; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; PAD=Peripheral Artery Disease; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.

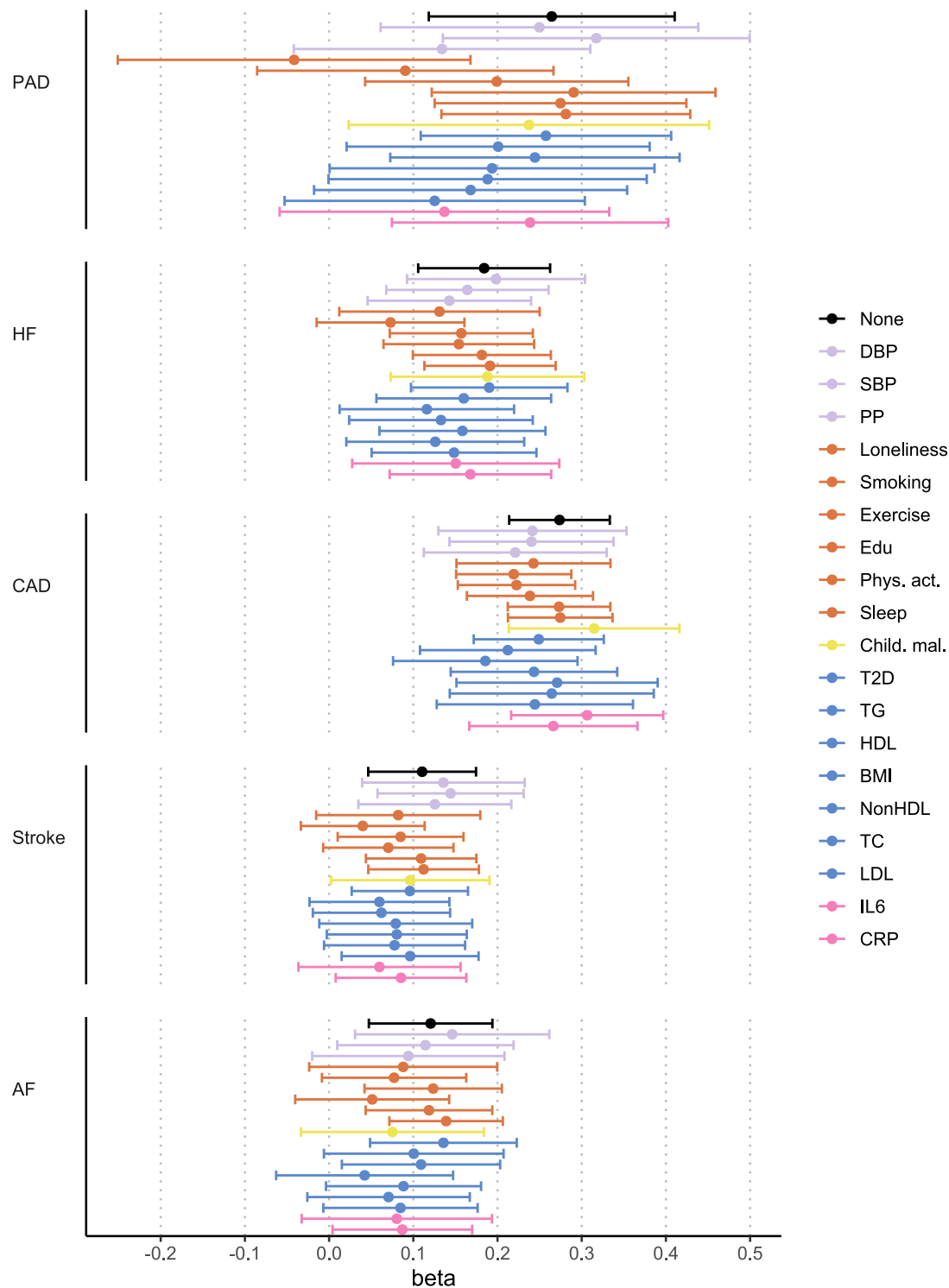


Extended Data Fig. 9 | See next page for caption.



**Extended Data Fig. 9 | Sensitivity analyses providing additional support for causal effects between MDD, CVD, and shared risk factors.** **a**, Effect of genetic instruments for MDD (exposure) on CVD and risk factors (outcomes) after excluding the UKB sample from the MDD GWAS (which was responsible for most of the sample overlap in the exposure and outcome GWAS summary statistics). **b**, Effect of genetic instruments for CVD and risk factors on MDD after excluding the UKB sample from MDD GWAS. **c**, Results from LHC MR, that corrects for heritable confounders and sample overlap, with the IVW MR estimate given as reference, for the effect of MDD on CVD and risk factors. **d** Effect of genetic instruments for MDD-ASCVD on outcomes and risk factors. **a-d** Significantly pleiotropic estimates are indicated with a red asterisk. IVW estimate shown.

Results from sensitivity analyses are reported in Supplementary Table 20. Points and error bars represent mean effect size (regression coefficient) and 95% CIs. Sample sizes for GWAS summary statistics are reported in Supplementary Table 1. The term beta refers to the log odds ratio. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; Child. Mal.=Childhood Maltreatment; CRP = C-Reactive Protein; DBP=Diastolic Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; HF=Heart Failure; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; PAD=Peripheral Artery Disease; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.



**Extended Data Fig. 10 | Support for causal effect of MDD on CVDs when adjusting for individual risk factors.** Results from multivariable MR. Points and error bars represent mean effect size (regression coefficient) and 95% CIs. IVW estimate shown. Reference estimates of the association without any adjustment are printed in black. Sample sizes for GWAS summary statistics are reported in Supplementary Table 1. The term beta refers to the log odds ratio. AF=Atrial Fibrillation; CAD=Coronary Artery Disease; Child.

Mal.=Childhood Maltreatment; CRP=C-Reactive Protein; DBP=Diastolic Blood Pressure; Edu=Educational attainment; HDL=High-Density Lipoprotein; HF=Heart Failure; IL6=Interleukin-6; LDL=Low-Density Lipoprotein; MDD=Major Depressive Disorder; NonHDL=Non-High-Density Lipoprotein; PAD=Peripheral Artery Disease; Phys. Act.=Physical activity; PP=Pulse Pressure; Psychosocial=Psychosocial/lifestyle; SBP=Systolic Blood Pressure; T2D=Type II Diabetes; TC=Total Cholesterol; TG=Triglycerides.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

R version 4.1.0  
Bash software: cleansumstats 1.6.0  
Full data collection scripts can be found at [github.com/jacobbergstedt/github.com/jacobbergstedt/MDDCVD\\_genetics](https://github.com/jacobbergstedt/github.com/jacobbergstedt/MDDCVD_genetics)  
Container data collection software can be found at <https://github.com/comorment>  
Data was processed using TSD and slurm

Data analysis

R version 4.1.0  
R packages: GenomicSEM 0.0.5; TwoSampleMR 0.5.6; MVMR 0.4; LAVA 0.1.0; bigsnpr 1.12.4  
Python version 2.7.13  
Python software: ldsc 1.0.1; MiXeR 1.3  
Bash software: plink 1.9, plink2.0, gctb 2.05  
Full data analysis scripts can be found at [github.com/jacobbergstedt/MDDCVD\\_genetics](https://github.com/jacobbergstedt/MDDCVD_genetics)  
Container data analysis software can be found at <https://github.com/comorment>  
Data was analyzed using the TSD cluster

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Links to download publicly available published GWAS summary statistics data used as inputs in this study are listed in Supplementary Table 1. Single nucleus RNA sequencing data in the adult human brain can be found at <https://github.com/linnarsson-lab/adult-human-brain>. Researchers can request access to the UK-Biobank data resources at <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>; data for PRS analysis described in this study were accessed under accession number 22224. Gene expression data from human tissues can be found at <https://www.gtexportal.org/home/datasets>. Summary statistics for GWAS of the MDD-ASCVD and ASCVD latent factors are available at <https://doi.org/10.6084/m9.figshare.25737537>.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	We adjusted on sex in the PRS analysis. Further consideration of sex is out of scope for our research aims in this project.
Reporting on race, ethnicity, or other socially relevant groupings	We did not consider race, ethnicity, or other socially relevant groupings in this study.
Population characteristics	The UK Biobank is a prospective cohort study described in Bycroft et al. 2018 Nature. Average age for participants at assessment was 58 years; 54.2% are female; and 94.7% are of European ancestry. The populations for the summary statistics GWAS were all of European ancestry. For more detailed information please see publications listed in Supplementary Table 1.
Recruitment	See Bycroft et al. 2018 Nature and publications listed in Supplementary Table 1
Ethics oversight	Ethics approval for the UK Biobank study was obtained from the North West Centre for Research Ethics Committee (11/NW/0382). The work described here was approved by UK Biobank under application number 22224

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We did not consider summary statistics from GWASs with less than 10,000 cases. We selected broad CVD diseases with large-scale GWASs available at the time of the study. For risk factors, we picked risk factors based on our hypothesis that immunometabolic factors are linked to MDD-CVD comorbidity, combined with availability of GWASs of sufficient sample size and quality.
Data exclusions	Only GWAS based on populations of European ancestry were considered to ensure comparable LD patterns underlying GWAS summary statistics.
Replication	LAVA and MiXeR analyses are complementary and show similar patterns. MDDCVD summary statistics were validated in UK Biobank. For the single nucleus brain transcriptomics, the dataset from Siletti et al. (2022) Science is to the best of the author's knowledge the first and only single-cell transcriptomic census of the entire human brain. For Mendelian randomization we conducted three analyses to ensure robust results the main version, the version without UKB in the MDD summary statistics, and a version using a complementary method latent heritable confounder Mendelian randomization. All conclusions were confirmed across sensitivity analyses and triangulation.
Randomization	We did not employ any experimental batches or groups.
Blinding	We did not do any group allocation for this study.



# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Involvement in the study                               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                        |

## Methods

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involvement in the study                        |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Plants

Seed stocks

n/a

Novel plant genotypes

n/a

Authentication

n/a