

# Multi-trait association analysis reveals shared genetic loci between Alzheimer's disease and cardiovascular traits

Received: 16 January 2024

Accepted: 14 October 2024

Published online: 13 November 2024



Fotios Koskeridis<sup>1,2,3</sup>✉, Nurun Fancy<sup>3,4</sup>, Pei Fang Tan<sup>5,6</sup>, Devendra Meena<sup>1</sup>, Evangelos Evangelou<sup>1,2</sup>, Paul Elliott<sup>1,3</sup>, Dennis Wang<sup>5,6,7</sup>, Paul M. Matthews<sup>3,4</sup>, Abbas Dehghan<sup>1,3</sup> & Ioanna Tzoulaki<sup>1,3,8</sup>

Several cardiovascular traits and diseases co-occur with Alzheimer's disease. We mapped their shared genetic architecture using multi-trait genome-wide association studies. Subsequent fine-mapping and colocalisation highlighted 16 genetic loci associated with both Alzheimer's and cardiovascular diseases. We prioritised rs11786896, which colocalised with Alzheimer's disease, atrial fibrillation and expression of *PLEC* in the heart left ventricle, and rs7529220, which colocalised with Alzheimer's disease, atrial fibrillation and expression of *CIQ* family genes. Single-cell RNA-sequencing data, co-expression network and protein-protein interaction analyses provided evidence for different mechanisms of *PLEC*, which is upregulated in left ventricular endothelium and cardiomyocytes with heart failure and in brain astrocytes with Alzheimer's disease. Similar common mechanisms are implicated for *CIQ* in heart macrophages with heart failure and in brain microglia with Alzheimer's disease. These findings highlight inflammatory and pleomorphic risk determinants for the co-occurrence of Alzheimer's and cardiovascular diseases and suggest *PLEC*, *CIQ* and their interacting proteins as potential therapeutic targets.

Alzheimer's disease (AD), the most common cause of dementia, is a leading health challenge for our times. More than 55 million people worldwide were estimated to be living with dementia in 2020 with 60%–70% of them being AD cases<sup>1,2</sup>. AD has been considered a brain-specific disease whose primary pathology is confined to the brain. However, accumulating evidence suggests mechanistic links between a wide range of cardiovascular (CV) abnormalities and AD<sup>3–5</sup>. Epidemiological studies and experimental data have shown consistent associations between manifestations of clinical CV diseases such as coronary artery disease (CAD), atrial fibrillation (AF) and stroke, with higher risk of AD<sup>6–8</sup>. Several hypotheses have been

proposed to explain this. Indeed, atherosclerosis, the main underlying cause of CV diseases, also has profound consequences on the cerebrovascular system. These include reduced blood flow and potential vascular damage in the brain, impaired cerebral perfusion, and associations with inflammation and oxidative stress, all of which are factors that can contribute to neurodegenerative pathology and increase the risk of AD<sup>9</sup>. Beyond atherosclerosis, other hemodynamic effects associated with hypertension, arteriosclerosis and subsequent aortic stiffening have been associated with cerebrovascular damage and cognitive function, potentially accelerating the onset and progression of AD<sup>10</sup>.

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. <sup>2</sup>Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece. <sup>3</sup>UK Dementia Research Institute, Imperial College London, London, UK. <sup>4</sup>Department of Brain Sciences, Imperial College London, London, UK. <sup>5</sup>Institute for Human Development and Potential, Agency for Science, Technology and Research (A\*STAR), Singapore, Republic of Singapore. <sup>6</sup>Bioinformatics Institute, Agency for Science, Technology and Research (A\*STAR), Singapore, Republic of Singapore. <sup>7</sup>National Heart and Lung Institute, Imperial College London, London, UK. <sup>8</sup>Systems Biology, Biomedical Research Institute of the Academy of Athens, Athens, Greece. ✉e-mail: [f.koskeridis@imperial.ac.uk](mailto:f.koskeridis@imperial.ac.uk)

AD and CV traits also share common genetic determinants<sup>11</sup>. Genome-wide association studies (GWAS) have identified genetic risk loci for both AD and pathological CV traits and identified common genetic factors that may refer to the shared underlying pathways. One example of such genes is apolipoprotein E (*APOE*), which encodes a lipid-transport protein involved in cholesterol metabolism<sup>12</sup>, that is the strongest genetic risk factor for AD<sup>13,14</sup> and a risk factor for adverse CV traits, including CAD<sup>15</sup> and myocardial infarction<sup>16</sup>. Nonetheless, the precise mechanisms and molecular processes that modulate the AD and CV link remain elusive. A deeper understanding of their shared genetic architecture will provide insights into potentially common and distinct aetiologies of these conditions. Identifying shared targets and mechanisms that confer functional effects could lead to the discovery of interventions that address both neurodegenerative and CV diseases.

Here, we further investigated the commonalities in the genetic architecture of AD and CV traits and identified potential pleiotropic loci affecting multiple traits aiming to define common targets for therapeutic modulation. We explored a wide range of CV abnormalities and two common main risk factors, atherosclerosis and blood pressure (BP), proposed to underlie both CV and AD, to investigate different molecular pathways that may link different CV manifestations to AD. We performed a large-scale multi-trait GWAS analysis on AD and CV traits, followed by genetic colocalisation analysis to highlight candidate pleiotropic genes and their tissue sites of action. To further characterise biological pathways involved in both diseases, we leveraged data from single-cell RNA-sequence for differential gene expression with disease to explore relevant gene co-expression networks and protein-protein interactions in the brain and CV tissues. A schematic overview of the study is presented in Fig. 1.

## Results

### Multi-trait genetic association analysis identifies 5 novel AD loci and 9 shared loci between AD and CV traits

We performed five pairwise multi-trait analyses of GWAS (MTAG)<sup>17</sup> on AD and CAD, AF, stroke, carotid intima-media thickness (cIMT), and systolic and diastolic blood pressure (SBP, DBP). We examined the

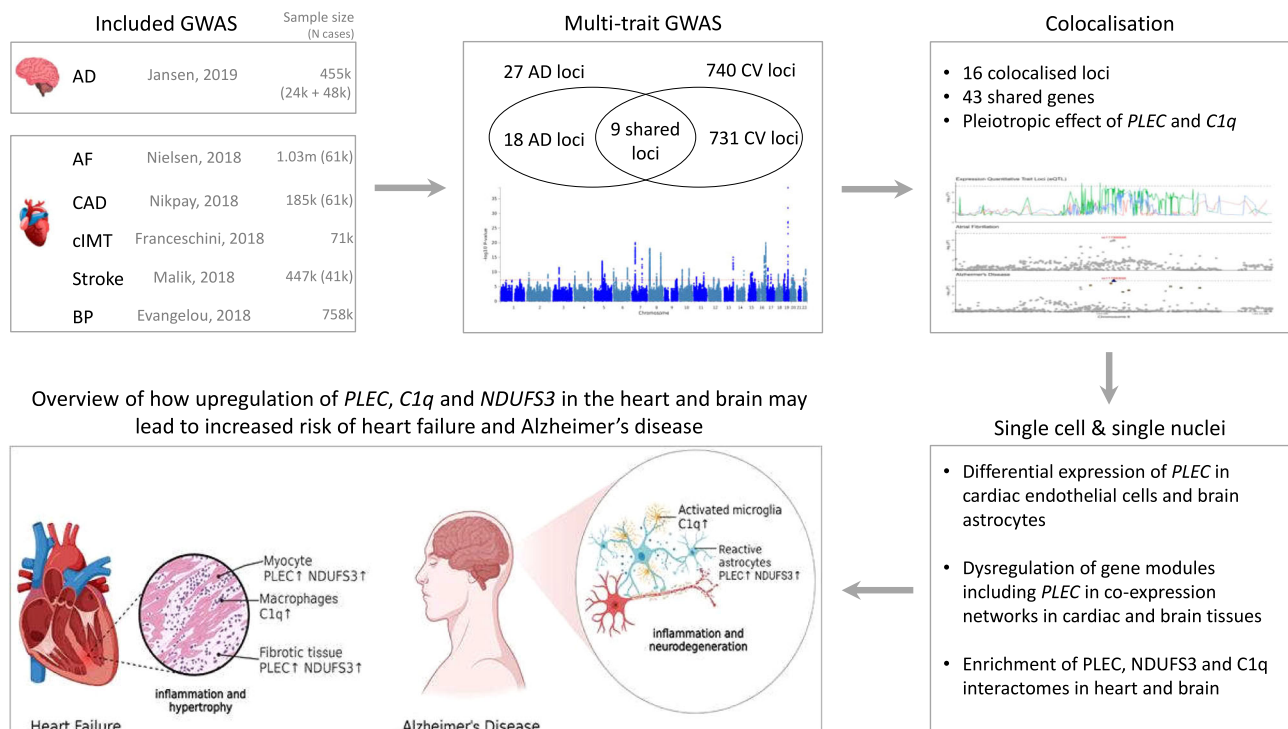
bivariate genetic correlation between AD and the examined CV traits (Supplementary Table 1) and visually illustrated the MTAG results alongside those from the original GWAS (Supplementary Fig. 1-6). Across all pairwise MTAG analyses, we identified 27 unique genetic loci associated with AD at genome-wide significance (GWS) level ( $P < 5 \times 10^{-8}$ ) corresponding to 114 unique single-nucleotide polymorphisms (SNPs) (Supplementary Data 1). Out of the 27 AD loci, 5 were novel (not within  $\pm 500$  kilobases (kb) of the previously known AD loci) and among them rs73069394 (*ULK4*) displayed the strongest association (Table 1).

To further validate the associations of the novel AD loci, we applied summary data-based Mendelian randomisation (SMR)<sup>18</sup> and heterogeneity in dependent instruments (HEIDI) using gene expression data from relevant brain tissues. SMR analysis suggested a potentially causal association between *ULK4* expression in the hippocampus and AD risk ( $\beta_{SMR} = 0.04$ ,  $P_{SMR} = 3.4 \times 10^{-10}$ ,  $P_{HEIDI} = 0.22$ ). SMR was not possible for the remaining loci due to unavailability of gene expression data.

Furthermore, we found 1222 top signals associated with different CV traits at GWS level in 740 genetic loci (Supplementary Data 2). Of these, 13 novel loci were highlighted for CAD ( $N = 4$ ), cIMT ( $N = 8$ ) and stroke ( $N = 1$ ) (Table 1). Overall, 15 of the unique AD SNPs (9 loci) were additionally associated at GWS level with at least one of the examined CV traits (Supplementary Data 3).

### A colocalisation analysis defines genetic loci shared by AD and different CV traits

Using the Hypothesis Prioritisation in multi-trait Colocalization (HyPrColoc)<sup>19</sup> method on all 767 MTAG-reported loci (27 AD + 740 CV), we identified 21 loci which colocalised between AD and CV traits with a posterior probability (PP)  $> 0.5$  (Fig. 2, Supplementary Data 4). Most colocalised loci were found either between AD and AF (7 loci) or between AD and DBP (7 loci). Substantial evidence for colocalisation was observed for a locus at chr8:124,608,614  $\pm$  200 kb (*RN7SKP155*) associated with AD and cIMT (PP = 1) and a locus at chr11:47,391,948  $\pm$  200 kb (*SP11*) associated with AD and DBP



**Fig. 1 | Study design schematic overview.** AD Alzheimer's disease, AF atrial fibrillation, CAD coronary artery disease, cIMT carotid intima media thickness, BP blood pressure, CV cardiovascular, HF heart failure. Created in BioRender. Wang, D. (2023) BioRender.com/q291509.

**Table 1 | Novel genetic loci associated with the examined traits in a genome-wide significance level (two-sided  $P < 5 \times 10^{-8}$ )**

SNP	Chr	Pos	EA	OA	MAF	Beta	P	Gene
<b>Alzheimer's disease</b>								
rs7529220	1	22282619	T	C	0.139	-0.01	$1.7 \times 10^{-8}$	<i>HSPG2</i>
rs11692604	2	19947507	C	T	0.483	0.01	$4.1 \times 10^{-8}$	<i>AC019055.1</i>
rs73069394	3	41787233	A	G	0.188	0.03	$1.5 \times 10^{-29}$	<i>ULK4</i>
rs77399788	5	123003001	G	A	0.058	0.03	$1.8 \times 10^{-10}$	<i>KRT18P16</i>
rs56365761	19	39148103	G	A	0.436	-0.01	$4.6 \times 10^{-8}$	<i>ACTN4</i>
<b>Coronary artery disease</b>								
rs2552527	2	218688596	G	T	0.409	-0.02	$2.0 \times 10^{-8}$	<i>TNS1</i>
rs748431	3	14928077	G	T	0.395	0.02	$3.5 \times 10^{-8}$	<i>FGD5</i>
rs11723436	4	120901336	G	A	0.327	0.02	$2.7 \times 10^{-8}$	<i>RP11-700N1.1</i>
rs15052	19	41813375	C	T	0.158	0.02	$3.6 \times 10^{-8}$	<i>HNRNPUL1:TGFB1</i>
<b>Carotid intima-media thickness</b>								
rs10064683	5	95567760	A	G	0.352	0.02	$1.3 \times 10^{-8}$	<i>CTD-2337A12.1</i>
rs6904596	6	27491299	A	G	0.082	0.04	$1.2 \times 10^{-8}$	<i>HNRNPA1P1</i>
rs56118607	9	127898024	A	G	0.124	0.03	$3.8 \times 10^{-8}$	<i>SCAI</i>
rs1887182	10	97013497	G	T	0.467	0.02	$1.0 \times 10^{-8}$	<i>PDLIM1</i>
rs11029956	11	27355804	A	G	0.340	0.02	$7.4 \times 10^{-9}$	<i>CCDC34</i>
rs12370774	12	106510413	T	C	0.097	-0.04	$3.4 \times 10^{-8}$	<i>NUAK1</i>
rs76064118	19	2235284	T	C	0.053	0.06	$4.8 \times 10^{-8}$	<i>PLEKHJ1</i>
rs1034565	22	19984211	T	C	0.284	0.03	$9.8 \times 10^{-9}$	<i>ARVCF</i>
<b>Stroke</b>								
rs2284665	10	124226630	T	G	0.197	-0.01	$4.4 \times 10^{-8}$	<i>HTRA1</i>

SNP Single-nucleotide polymorphism; Chr Chromosome; Pos Position; EA Effect allele; OA Other allele; MAF Minor allele frequency; Beta Effect size estimate; P Two-sided P-value; Gene Mapped gene.

(PP = 0.95). Among the 21 loci with evidence for colocalisation, there were three loci for which a single candidate causal variant explained a large proportion of the association: rs11786896 (mapped in *PLEC*; colocalised with AD-AF; PP = 0.97; 86% of PP explained by SNP), rs7529220 (mapped in *HSPG2*; colocalised with AD-AF; PP = 1; 90% of PP explained by SNP) and rs429358 (mapped in *APOE*; colocalised with AD-CAD; PP = 0.57; 93% of PP explained by SNP). Although rs11786896 in *PLEC* is not classified as a novel AD locus due to its proximity to a previously reported variant (rs34173062 in *SHARPIN*), the two loci are in linkage equilibrium ( $r^2 = 0.006$ ) and the regional plots suggest it likely represents a different independent signal (Supplementary Fig. 7).

### Gene expression colocalisation analysis prioritises causal genes shared by AD and CV traits

To identify potential pleiotropic causal genes for the colocalised loci, we tested the colocalisation of AD and CV traits with the expression of nearby genes in 48 tissues using expression quantitative trait loci (eQTL) data from Genotype-Tissue Expression (GTEx) in the 21 colocalised loci. We found that AD and at least one CV trait colocalised in 16 loci with expression of one or more genes in the same tissue, for a total of 53 associations with 43 genes (Supplementary Fig. 8, Supplementary Data 5). Of these, 20 associations were found for AF (in 6 loci with 17 genes), 22 for DBP (6 loci with 16 genes), 3 for stroke (1 locus with 3 genes) and 8 for cIMT (3 loci with 6 genes).

In two loci, a single candidate causal variant explained the colocalisation of AD, CV trait and tissue-specific gene expression: rs11786896 (*PLEC*) and rs7529220 (*HSPG2*). The intronic variant rs11786896 (*PLEC*) explained the colocalisation of AD and AF with expression levels of *PLEC* in the cardiac left ventricle (PP = 0.99, %PP explained by SNP = 99%) and skeletal muscle (PP = 0.92, %PP explained by SNP = 98%) (Fig. 3). rs11786896 was associated with increased risk of AD (Odds Ratio, OR = 1.02,  $P = 5 \times 10^{-8}$ ), increased risk of AF (OR = 1.02,  $P = 1.1 \times 10^{-6}$ ) and lower expression of *PLEC* in cardiac left ventricle

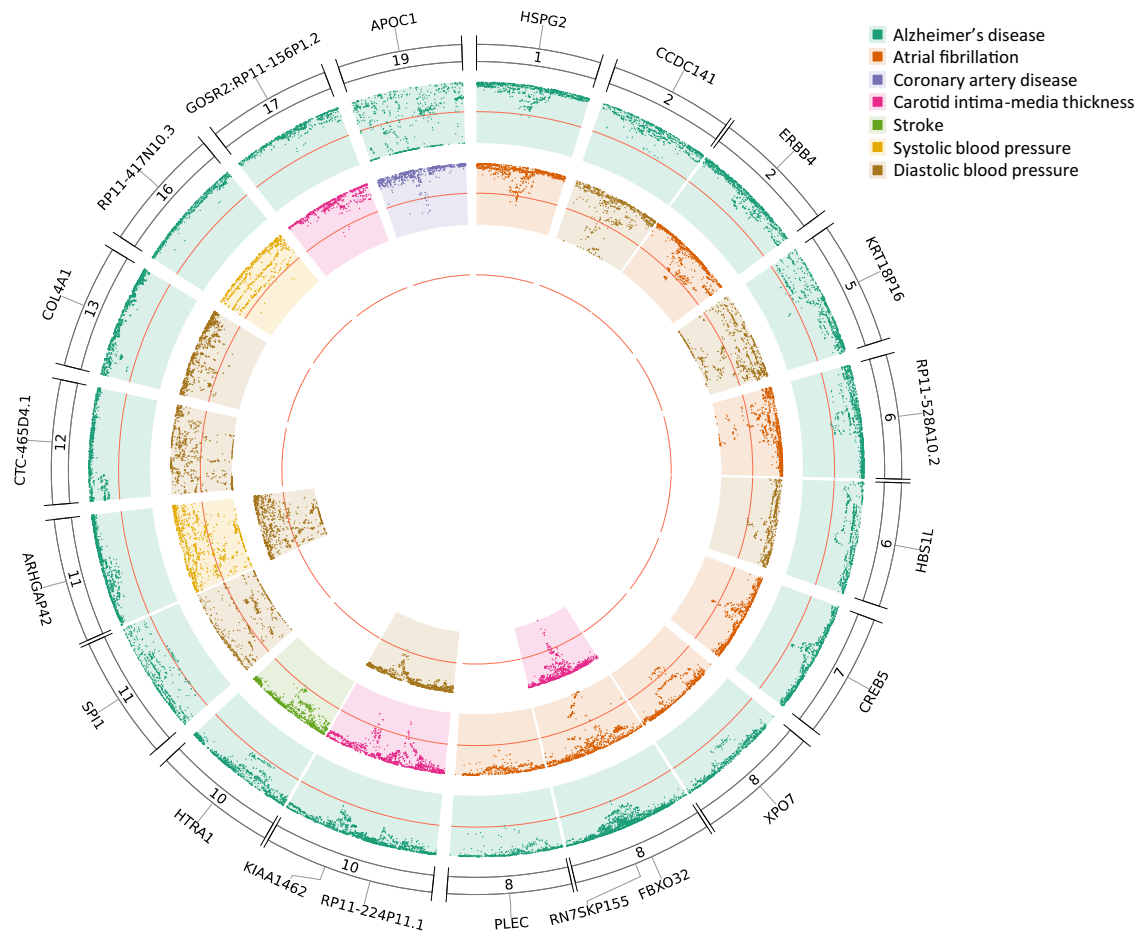
(Beta = -0.71,  $P = 5.9 \times 10^{-13}$ ), as well as in skeletal muscle (Beta = -0.3,  $P = 7.7 \times 10^{-7}$ ).

The intergenic variant rs7529220 (*HSPG2*) explained the colocalisation of AD and AF with expression levels of *CIQA* (PP = 0.85, %PP = 82%), *CIQB* (PP = 0.83, %PP = 97%) and *CIQC* (PP = 0.61, %PP = 99%) in breast mammary tissue. The variant was associated with higher risk of AD (OR = 1.01,  $P = 1.7 \times 10^{-8}$ ), higher risk of AF (OR = 1.01,  $P = 2.7 \times 10^{-10}$ ) and increased expression of *CIQA* (Beta = 0.19,  $P = 2.4 \times 10^{-4}$ ), *CIQB* (Beta = 0.17,  $P = 2.6 \times 10^{-4}$ ), and *CIQC* (Beta = 0.15,  $P = 8.1 \times 10^{-4}$ ) in mammary tissue.

### *PLEC* and *CIQ* are differentially expressed in the left ventricle with heart failure

Defining the cells in which target genes are expressed and directions of expression associated with disease risk and with disease is important for predicting the directions of effect for potential therapeutic modulation. Therefore, we tested the expression patterns of the 43 genes indicated by the colocalisation analysis across cell types in single-cell RNA from the heart to discover whether differences in differential expression with disease were consistent with those predicted for disease risk. *PLEC* was expressed in all cell types found in the cardiac left ventricle, while *CIQ* was expressed only in macrophages (Fig. 4A, B). *PLEC* was differentially expressed with heart failure (HF) relative to healthy controls with upregulated expression in the endothelium ( $\log_2$  fold change,  $\log_2FC = 0.40$ ,  $P = 0.015$ ) but downregulated in macrophages ( $\log_2FC = -0.59$ ,  $P = 6.7 \times 10^{-5}$ ). There was only a trend for differential expression with HF in cardiomyocytes ( $\log_2FC = 0.91$ ,  $P = 0.06$ ). All *CIQ* associated genes were downregulated in cardiac macrophages (*CIQA*,  $\log_2FC = -1.39$ ,  $P = 1 \times 10^{-6}$ ; *CIQB*,  $\log_2FC = -1.34$ ,  $P = 3.5 \times 10^{-5}$ ; *CIQC*,  $\log_2FC = -1.28$ ,  $P = 1.4 \times 10^{-4}$ ).

We explored differential expression with HF further by constructing high dimensional weighted gene co-expression networks (WGCN) for *PLEC* and *CIQ* to identify modules of highly correlated



**Fig. 2 | Circular figure visualising regional plots on the colocalised loci between Alzheimer's disease (AD) and cardiovascular traits (CV).** The figure presents the distribution of  $P$ -values ( $-\log_{10}P$ ) from MTAG with inner orientation ( $P$ -values are

derived from two-sided statistical tests). The annotations show the mapped genes of the AD/CV top lead SNPs on the colocalised loci.

genes across cell types. We generated 12 gene co-expression modules in cardiac endothelial cells, 14 in cardiomyocytes and 6 in macrophages (Supplementary Data 6). A differential module eigengene analysis indicated that the module including *PLEC* was upregulated in cardiac vascular endothelial cells and cardiomyocytes in HF cases relative to the healthy controls and downregulated in macrophages with dilated cardiomyopathy (dCM) cases relative to healthy controls (Supplementary Data 7). The *CIQ*-containing module also was downregulated in cardiomyocytes in HF cases relative to healthy controls and downregulated in macrophages in dCM cases relative to healthy controls. Gene-set enrichments for biological processes defined the pathways most highly enriched in *PLEC*-containing modules (Supplementary Data 8), which included the “vascular endothelial growth factor receptor-2 signalling” and “endothelium development” pathways in endothelial cells (Fig. 4C) and many pathways related to mitochondrial oxidative metabolism in cardiomyocytes (Supplementary Fig. 9). In macrophages, the module including *CIQ* genes was enriched for “complement activation” and “synapse pruning” pathways, among others (Supplementary Fig. 10, Supplementary Data 9).

#### **PLEC and CIQ interactomes are enriched in cardiomyocytes, cardiac vascular endothelial cells and macrophages**

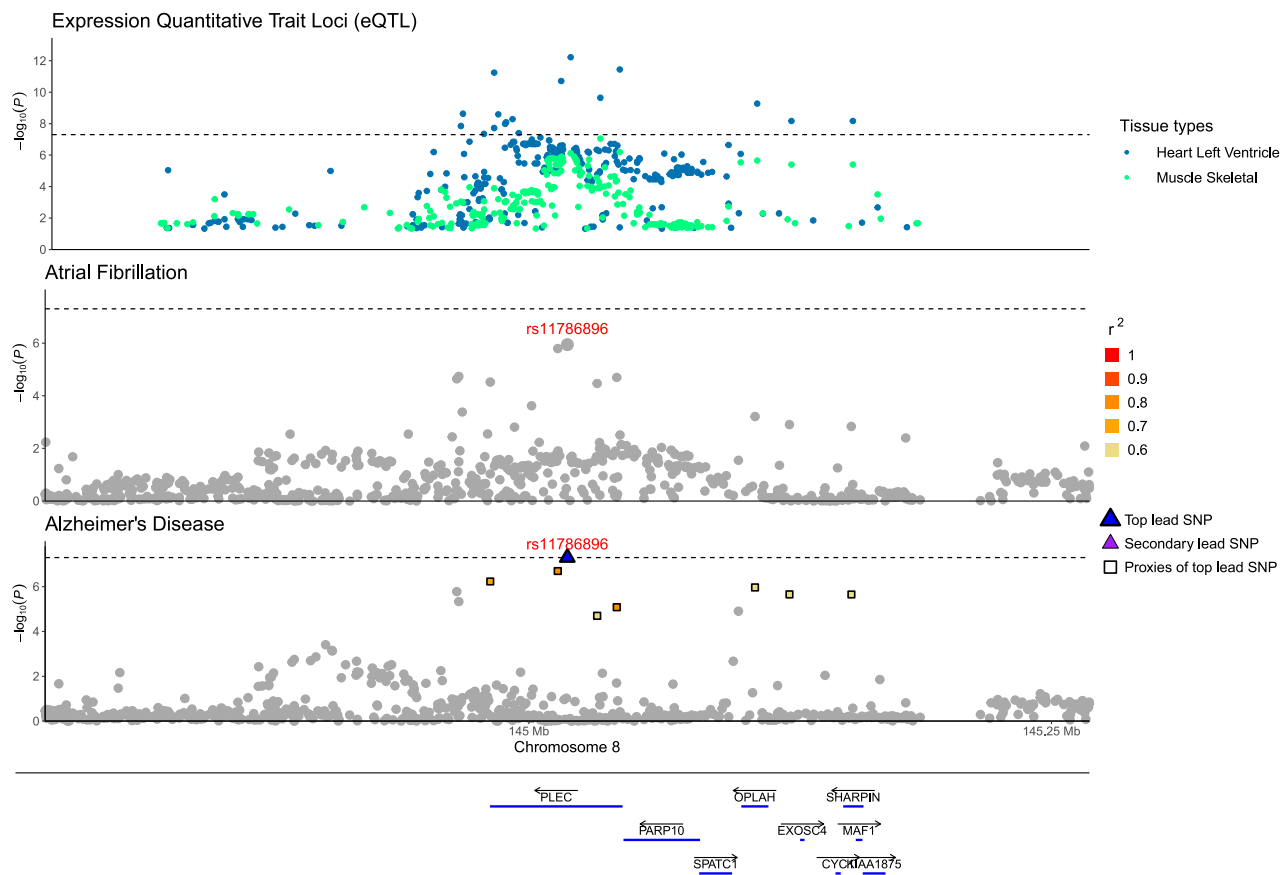
To gain insights into the potential functional roles of proteins, we performed cell-specific protein-protein interaction (PPI) analyses on the set of colocalised candidate genes by constructing their protein interactomes across cell types of human cardiovascular tissue (Supplementary Data 10 and 11). The *PLEC* interactome was upregulated in

a pathway related to “ribosomal small subunit assembly” in endothelial cells (Fig. 4D) and upregulated in a “SRP-dependent co-translational protein targeting to membrane” pathway in cardiomyocytes (Supplementary Fig. 11). Additionally, an interactome containing both *PLEC* and *NDUFS3* was enriched with HF and was upregulated in the pathways related to “aerobic electron transport chain” in endothelial cells (Fig. 4D) and “acetyl-CoA biosynthetic process from pyruvate” and “energy coupled proton transport” in cardiomyocytes (Supplementary Fig. 11). In macrophages, the *PLEC*-*NDUFS3* interactome in HF was enriched for “mitochondria electron transport of cytochrome c to oxygen” while *CIQ* interactome was enriched for “cell junction disassembly” (Supplementary Fig. 12).

#### **PLEC is differentially expressed in brain astrocytes and is upregulated in AD**

We also tested the expression of the 43 genes from our colocalisation analysis across different cell types in post-mortem human brain samples. *PLEC* was highly expressed in astrocytes and (to a lesser degree) in neurons. *CIQ* genes were expressed primarily in microglia (Fig. 5A, B). *PLEC* was significantly upregulated in astrocytes from AD donors relative to non-diseased control donors ( $\log_2FC = 1.01$ ,  $P = 0.003$ ). *CIQ* genes were not significantly differentially expressed in microglia but consistently showed lower mean expression with AD (*CIQA*:  $\log_2FC = -0.46$ ,  $P = 0.23$ ; *CIQB*:  $\log_2FC = -0.5$ ,  $P = 0.3$ ; *CIQC*:  $\log_2FC = 0.14$ ,  $P = 0.9$ ).

We explored the cell-specific differential expression of these and co-expressed genes further with WGCNA, which identified 14 gene



**Fig. 3 | Regional plot on the colocalised locus of candidate causal variant rs11786896 (mapped in *PLEC*) for Alzheimer's disease (bottom), atrial fibrillation (middle) and the expression quantitative trait loci (top) for the associated**

**tissues.** The x-axis presents the chromosomal positions of single-nucleotide polymorphisms, while the y-axis shows the negative logarithm to base 10 of *P*-values ( $-\log_{10}(P)$ ) from two-sided statistical tests.

co-expression modules in astrocytes and 8 in microglia (Supplementary Data 12). Differential module eigengene analyses showed that the *PLEC*-containing module was upregulated in astrocytes and the *CIQ*-containing module was downregulated in microglia (Supplementary Data 13). Gene-set enrichment for biological processes identified 25 associated pathways for the *PLEC*-containing module in astrocytes including “neuron projection morphogenesis” and “extracellular structure organisation” and 25 associated pathways for the *CIQ* gene module in microglia including the “positive regulation of intrinsic apoptotic signalling” (Supplementary Fig. 13, Supplementary Data 14).

We further explored the discordant directions in *CIQ* expression between eQTLs and post-mortem brain single nuclei data in relation to the beta-amyloid pathology load in microglia. For subjects with lower beta-amyloid load in the brain, we observed an increasing expression of *CIQ* in concordance with eQTL but subjects with higher beta-amyloid load showed decreased expression (Supplementary Fig. 14).

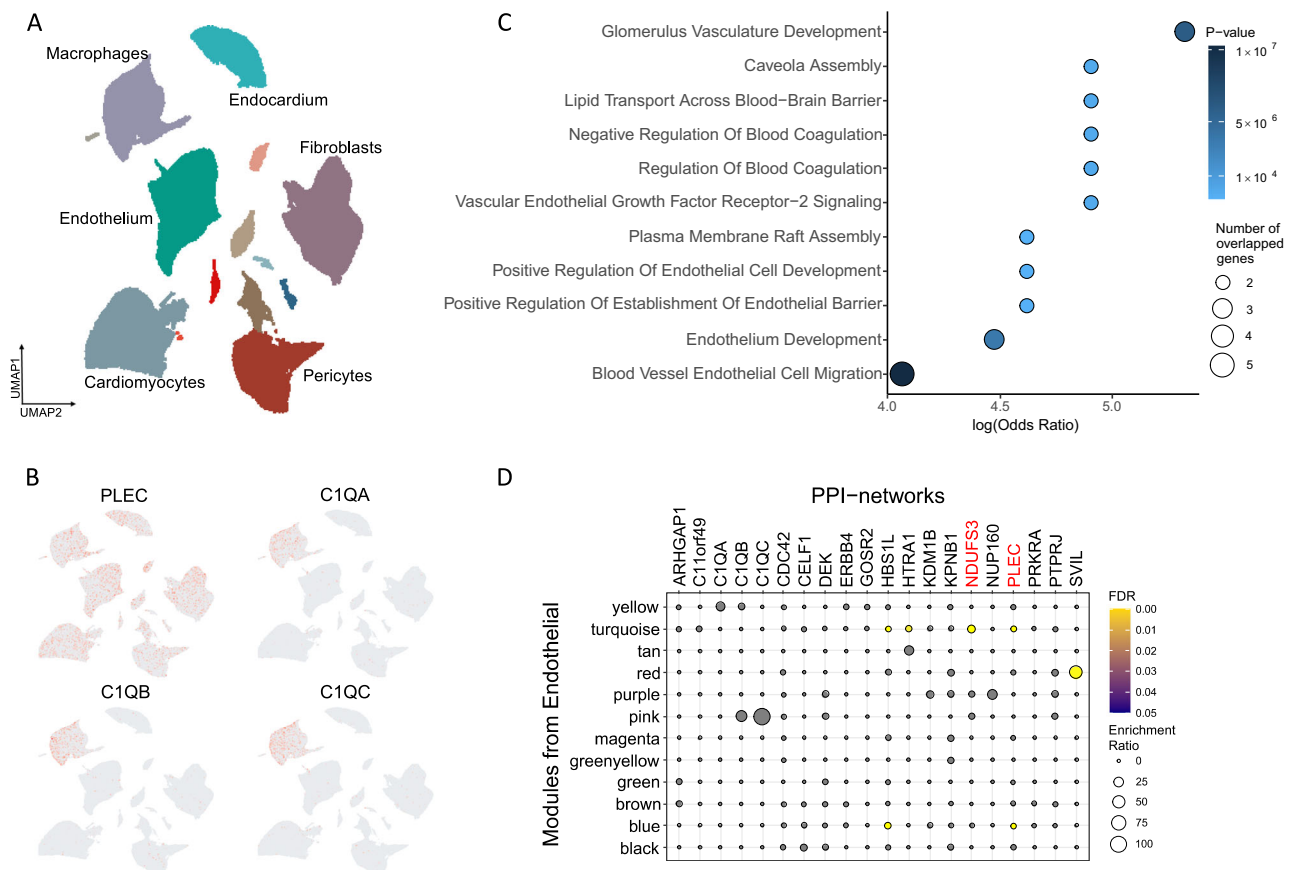
### **PLEC, NDUFS3 and CIQ interactomes are enriched in astrocytes and microglia**

A PPI analysis on candidate gene networks highlighted that the *PLEC*-*NDUFS3* interactome also was enriched in both astrocytes and microglia in AD cases compared to controls (Supplementary Data 10, 15 and 16). Associated functional pathways were enriched for the “aerobic electron transport chain” and “SRP-dependent co-translational protein targeting to membrane” pathways in the astrocytes (Fig. 5D). In microglia, the *CIQ* interactome was enriched in “SRP-dependent co-translational protein targeting to membrane” pathway (Supplementary Fig. 15).

## **Discussion**

We adopted a comprehensive approach to understand the co-occurrence between AD and various CV diseases and traits based on several multi-trait GWAS to characterise their shared genetic architecture. Convergent evidence from colocalisation between AD, AF and eQTLs prioritised two genetic regions that each included a single candidate causal variant (rs11786896 which was an eQTL for *PLEC* and rs7529220 which was an eQTL for *CIQA*, *CIQB*, and *CIQC*) shared between AD and AF. Single-cell RNA-sequence data, co-expression network and protein-protein interaction analyses together were consistent in showing that *PLEC* is upregulated in left ventricular endothelium and cardiomyocytes with HF and in brain astrocytes with AD. By contrast, while *CIQ* genes are predicted to be upregulated with greater disease risk in cardiac macrophages for HF and in brain microglia for AD, we found opposite directions of difference with disease for both. We explored differences in the direction of changes from early disease (low beta-amyloid pathology load) to late (high beta-amyloid pathology load) for microglia and found the congruence with directions predicted for disease risk in early disease that was lost with in later disease progression. Our findings provide new insights into genetic pleiotropic effects and potential shared mechanisms causally related to both AD and CV traits.

Our study highlighted 5 not previously reported AD loci including SNPs located in or near *HSPG2*, *ACO19055.1*, *ULK4*, *KRT18P16* and *ACTN4*. Of them, the intronic variant rs73069394 (*ULK4*) showed the strongest association with AD and was also GWS associated with DBP. The locus did not show evidence for colocalisation between AD and DBP suggesting this variant is not likely the causal one for both traits. Apart from DBP, GWAS studies have shown associations between this



**Fig. 4 | Single-nuclei transcriptomes of cardiac tissue for genes increasing the risk of cardiomyopathy.** **A** The transcriptomes form discrete cell-specific clusters using Uniform Manifold Approximation and Projection (UMAP). **B** Expression of *PLEC* and *C1Q* genes across cell-specific clusters from UMAP. Red indicates higher expression. **C** Enriched pathways of the *PLEC*-containing module in cardiovascular

endothelial cells between dilated cardiomyopathy cases and healthy controls (*P*-values are derived from two-sided statistical tests). **D** Enrichment of pathway genes in the protein interactomes of candidate genes in cardiovascular endothelial cells. The *PLEC*-interacting module is the brown and the *C1q*-interacting module is the pink.

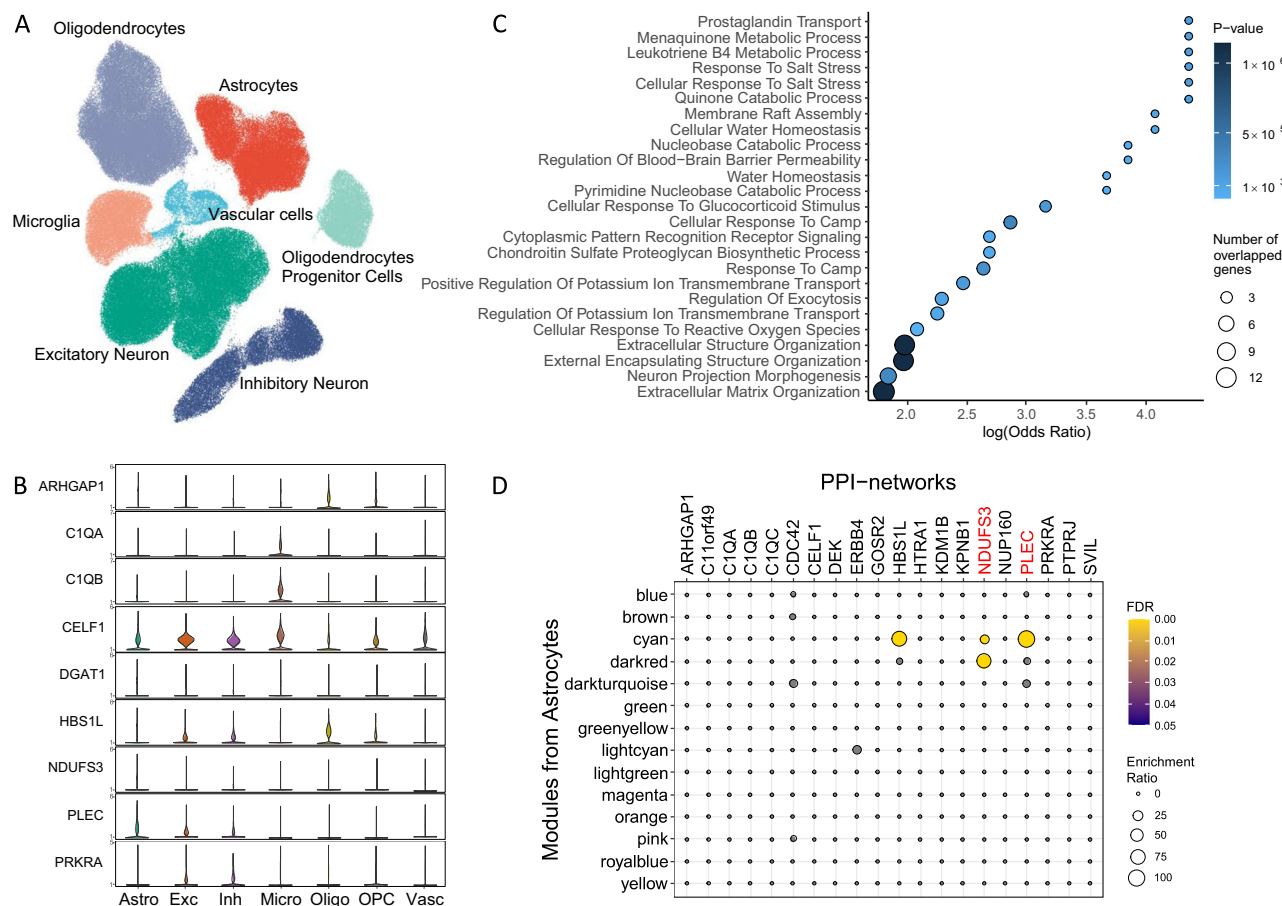
locus and schizophrenia and bipolar disorder<sup>20</sup>. The role of the *ULK4* in relation to neurodegeneration is little studied but the gene function shows biological plausibility to AD. *ULK4* protein is involved in the regulation of autophagy and plays multiple roles in brain function including neuronal growth, endocytosis and myelination which are pathways implicated to AD pathology<sup>20,21</sup>. Here, we also provided additional evidence from SMR analysis supporting a potentially causal association between expression levels of *ULK4* and AD.

Out of the several CV traits and diseases examined with AD, the largest number of pleiotropic signals with AD was observed for AF and BP highlighting the importance of pathways related to these CV traits in explaining comorbidities with AD. Numerous observational studies, provide growing evidence that BP and AF are associated with cognitive impairment, risk of AD and other dementias<sup>22,23</sup>. The suggested mechanistic links between these traits and AD involve a combination of cerebrovascular damage, neuroinflammation, amyloid-beta accumulation, oxidative stress, and endothelial dysfunction<sup>24,25</sup>. However, it is unclear whether the diseases have a shared pathophysiology or whether the relationship arises as downstream consequences of BP and AF (e.g., stroke).

Here we provide evidence suggesting a shared genetic determinant that may contribute to the pathophysiology of AF and AD. The colocalised intronic variant rs11786896 within the plectin gene (*PLEC*) was associated with lower expression of *PLEC* in the cardiac left ventricle (and skeletal muscle) and increased the risk of both AD and AF. Furthermore, *PLEC* was upregulated in left ventricular endothelium and cardiomyocytes in HF cases and in brain astrocytes in AD cases. A low-frequency missense variant in *PLEC* has been previously

associated with atrial fibrillation in a whole-genome sequencing data<sup>26</sup> whereas another missense variant has been linked to structural brain connectivity<sup>27</sup>. The intronic variant highlighted in this analysis has been previously associated with right ventricular structure and function<sup>28</sup> but has not been identified in GWAS studies as an AD or AF signal. *PLEC* is a member of a protein family, named plakins, with a crucial structural role in the cytoskeleton influencing cell architecture and tissue integrity and a partially functional role in the assembly, positioning, and regulation of signalling complexes<sup>29,30</sup>. Previous studies of human tissues or preclinical models provide independent evidence for an association of plectin with diseases including AD and AF<sup>26,31</sup>.

Plectin is highly expressed in the central nervous system, especially at the interfaces between glia and pial cells and between glia and endothelial cells and is thought to be important to blood-brain barrier and pial surface integrity<sup>32</sup>. Plectin deficiency in mice has been associated with diminished learning capabilities and reduced long-term memory compared to wild-type littermates<sup>33</sup>. Here we provide evidence that the risk of AD may be affected via functions of plectin in astrocytes<sup>34</sup>. Astrocytes play multiple roles, central to the pathology of AD, including metabolic support for neurons, modulation of brain microvascular function and, through activities associated with those of microglia, inflammatory responses<sup>34,35</sup>. We hypothesise that these functional roles are mediated in part by interactions of plectin with intermediate filaments (IFs), microtubules and actin filaments<sup>34</sup>. IFs are important structural components of the cytoskeleton with crucial roles in synaptic activity, neurogenesis and repair after brain injury<sup>36</sup>. Differences in expression of plectin modulate neuronal function and



**Fig. 5 | Single-nuclei transcriptomes of human post-mortem brains for genes increasing the risk of Alzheimer's disease.** **A** The transcriptomes form discrete cell-specific clusters using Uniform Manifold Approximation and Projection (UMAP). **B** Expression of candidate genes across cell types in human brain. **C** Enriched pathways of the *PLEC*-containing module in astrocytes between Alzheimer's disease cases and healthy controls (*P*-values are derived from two-sided

statistical tests). **D** Enrichment of pathway genes in the protein interactomes of candidate genes in astrocytes. The *PLEC*-interacting module is the cyan and the *NDUFS3*-interacting module is the dark red. Astro Astrocytes, Exc Excitatory Neuron, Inh Inhibitory Neuron, Micro Microglia, Oligo Oligodendrocytes Progenitor Cells, Vasc Vascular cells.

vesicular trafficking generally and interactions with tau suggest potential roles specific to AD<sup>33,37,38</sup>.

The role of *PLEC* in AF has been largely hypothesised to act via structural effects on the heart and cause electrophysiological abnormalities<sup>26</sup>. Here, we also show evidence for upregulation of *PLEC* in cardiomyocytes of HF patients. Therefore, in accordance with the hypothesised mechanisms linking *PLEC* to AD above, *PLEC* may play related roles in cardiomyocytes for assembling and mobilising the intermediate filaments and their networks. These effects further modulate contractile function in cardiomyocytes and inflammatory responses in macrophages which may further contribute to AF<sup>39</sup>.

Another localised variant between AD and AF, the intergenic rs7529220, which is located 19 k upstream from Heparan Sulfate Proteoglycan 2 (*HSPG2*) and 21 k downstream from Chymotrypsin Like Elastase 3B (*CELA3B*), was associated with increased risk of AD and AF and higher expression of three genes of the Complement Component 1, Q Subcomponent (*CIQ*) family (*CIQA*, *CIQB*, *CIQC*) in breast mammary tissue (and, by inference, in brain vasculature) and is a previously unreported locus for AD. The variant is located 680 kb downstream of *CIQ* genes. The complement system plays a central role in synaptic remodelling in the brain and in cellular damage response more generally in the body<sup>40,41</sup>. We hypothesise that greater expression of *CIQ* may lead to higher activity of the complement system which in turn may potentiate synapse loss in early AD<sup>42</sup>. Similarly, *CIQ* has roles in the genesis of atherosclerotic plaques<sup>43</sup> and in the regulation of early

stages of inflammatory responses to the cardiomyocyte injury associated with a range of cardiac traits<sup>44</sup>.

The variant rs429358, located within *APOE*, showed evidence for colocalisation between AD and CAD. However, further investigation using eQTL could not prioritise any shared gene within the locus expressed in the examined tissues. Given *APOE*'s well-established role in AD and its potential involvement in other CV traits, including CAD<sup>45</sup>, it is plausible that the locus affects the two traits independently through different pathways (horizontal pleiotropy) or through a shared gene expressed in tissues other than those examined in our study.

Our study had several strengths. First, we secured high statistical power for our study by including GWAS with substantial sample sizes ranging from 185,000 to 1,000,000 participants and we boosted the power even higher by performing suitable multivariate methods. Second, we combined advanced methods of genetic epidemiology and basic sciences and sought to provide supporting evidence from a variety of data. Third, exploring a wide spectrum of CV diseases pairwise enabled a comprehensive investigation into the diverse shared mechanisms underlying the relationship between different manifestations of CV conditions and AD.

However, a number of limitations also must be acknowledged. We restricted our analyses to a population of European ancestry. The lack of genetic diversity may have hampered the possibility of detecting other relevant variants. Additionally, we did not investigate a

considerable portion of the genetic predisposition coming from rare variants ( $MAF < 1\%$ ) as we excluded them from our analyses. This exclusion is a restriction of the MTAG method, in order to mitigate the risk of false-positive findings and biased results. Another limitation of using MTAG is that genetic variants that are not present in at least one GWAS dataset of each AD-CV pair were excluded from the corresponding pairwise MTAG analysis, and therefore some variants tested in one pairwise MTAG may not be tested in another. GWAS compares the cases we are studying with controls, but due to comorbidity, the number of other diseases might be higher in cases than in controls. As we used summary statistics without access to individual-level data, we couldn't examine or address this heterogeneity. Despite this limitation, the identified shared genes remain valid and important for understanding their genetic connections. Moreover, we used statistical methods to detect pleiotropy, and therefore considered a genetic locus pleiotropic if it was statistically significantly associated with two or more phenotypes. However, this approach for identification of pleiotropic genes may not always highlight shared biological pathways, as the identified genes could affect the traits independently via different pathways (horizontal pleiotropy), or they could even be expressed in different tissues in response to different signals<sup>46,47</sup>. The GWAS for AD combined clinically diagnosed cases of AD with AD-proxy status (inferred based on whether the parent was diagnosed with dementia). The correlation between the GWAS with and without the 'AD-by-proxy' cases was high<sup>13</sup> adding to the validity of this approach. Nonetheless, the inclusion of individuals with parental AD diagnoses in the original GWAS may have introduced greater heterogeneity and increased probability of misclassification, suggesting that some of our observed associations might be misestimated or influenced by other types of dementia. Considering the systemic nature of AD, which encompasses multiple pathways and various tissues, our study conducted a thorough investigation across all tissues, assuming correlations among many eQTLs across different tissues. However, it is important to acknowledge that some eQTLs from tissues not directly linked to AD or CV traits could represent false positive associations. To mitigate this, we validated our findings using single-cell and single-nuclei data obtained from brain tissue samples of AD cases and cardiac tissue samples from individuals with HF. Furthermore, due to a limited number of cells for specific cell types, we had to combine single-cell data from multiple samples. We focused on tissue samples that were already enriched for cardiomyocytes, endothelium, and macrophages. Moreover, the expression for some candidate genes in our data was limited and thus additional sequencing data and reads are needed to investigate them further. Additional RNA sequencing data of different AD and CV conditions would probably be even more informative. Finally, epidemiological data frequently encounter challenges in clarifying pathogenic mechanisms. Although we provided evidence from mechanistic experiments, the exact biological processes involved cannot be inferred. Further work is needed to validate our findings and the suggested disease pathways.

In conclusion, we performed a multi-trait analysis on AD and CV traits and a subsequent colocalisation analysis detecting 16 shared genetic loci and further prioritising two shared causal variants between the aforementioned traits. Our findings define shared mechanisms for AD and different CV diseases. The complement system has been explored as a target for preventive or disease-modifying therapies in CV disease<sup>48</sup> and AD<sup>49</sup>. Our work suggests that plectin or members of its interactome could offer new and potentially promising targets for preventive and therapeutic medicines with benefits across these common comorbid disorders.

## Methods

### Study population

We restricted our study to a population of European ancestry. We used the summary statistics from seven GWAS on the following diseases:

AD<sup>13</sup>, AF<sup>50</sup>, CAD<sup>51</sup>, cIMT (see Supplementary Methods), stroke<sup>52</sup>, SBP<sup>53</sup>, and DBP<sup>53</sup>. We selected these datasets based on the sample size (with preference for larger GWAS), the quality of the phenotypic data, the date of publication (favouring more recent studies), and their relevance to our research questions. Supplementary Table 2 lists basic characteristics of all included GWAS datasets.

### Genotypic quality control

The seven initial datasets contained genotyped and imputed SNPs ranging from 7 to 34 million SNPs. We included in the analysis only SNPs that were present in both datasets (AD and the examined CV trait). Furthermore, we excluded all insertions, deletions, and rare variants (minor allele frequency;  $MAF < 0.01$ ), variants with sample sizes  $< 2/3$  of the 90<sup>th</sup> percentile and palindromic SNPs. Finally, more than 5.75 million SNPs were included in the analysis.

### Multi-trait association analysis

We performed five bivariate analyses on AD and a different each time CV trait (1. AF, 2. CAD, 3. cIMT, 4. Stroke, 5. SBP-DBP) using MTAG<sup>17</sup>. We calculated the genetic correlation between the traits and further corrected our data for sample overlap using bivariate linkage disequilibrium (LD) score regression as implemented in MTAG. Each MTAG analysis generated distinct trait-specific datasets (11 in total: 5 with AD plus 6 with CV traits) containing the trait-specific effect estimates for the included SNPs after leveraging for genetic correlation of the examined traits. As a result, the summary statistics from MTAG can be interpreted as similar to those from a univariate single-trait GWAS<sup>17</sup>.

### Functional mapping and annotation

We used Functional Mapping and Annotation of GWAS (FUMA)<sup>54</sup> to functionally analyse all the generated summary results from MTAG. All the genome-wide significant (GWS) SNPs ( $P < 5 \times 10^{-8}$ ) were initially clumped ( $r^2 < 0.6$ ) to determine the coordinates of the genomic risk loci and then clumped again ( $r^2 < 0.1$ ) to define independent signals. SNPs in pairwise LD at  $0.1 \leq r^2 < 0.6$  or SNPs located closer than 500 kb were assigned to the same LD block. SNPs that survived the second clumping were the independent signals. Independent SNPs with the smallest  $P$ -value in each LD block were defined as the top signals while the remaining were secondary signals. We further performed annotation and gene prioritisation analysis including all SNPs that survived the first clumping. We used the European sample of 1000 Genome Project Phase 3<sup>55</sup> to calculate pairwise LD between SNPs. SNPs were positionally mapped to their nearest protein-coding genes (Ensembl build v92).

To ensure the robustness of our findings, we considered only MTAG loci if the respective top variants were also associated with the examined trait ( $P < 0.01$ ) in the corresponding original GWAS with a concordant direction of effect (between MTAG and the original GWAS). To identify the unique AD top and secondary independent signals for AD, we gathered independent signals from all pairwise MTAG analyses and excluded duplicate signals or proxies (within  $\pm 500$  kb or LD  $r^2 > 0.1$ ), keeping the strongest signal with the smallest  $P$ -value.

### Novel loci definition and replication

A signal indicated by MTAG was considered novel if its top variant achieved genome-wide significance ( $P < 5 \times 10^{-8}$ ) in the MTAG results and was also significant ( $P < 0.01$ ) in the included original univariate GWAS with a concordant direction of effect (between MTAG and the original GWAS). Additionally, the top variant should not be located within  $\pm 500$  kb or in LD ( $r^2 > 0.1$ ) with previously reported loci. For AD, in addition to the included GWAS study<sup>13</sup>, novel loci were compared to two previously published key GWAS studies<sup>56,57</sup>.

We applied SMR<sup>18</sup> and HEIDI analysis to investigate whether expressions of the novel AD MTAG-identified genes in relevant brain

tissues were causally associated with AD risk. SMR was performed by integrating summary eQTL data from GTEx version 8<sup>58</sup> in hippocampus and cortex (SNPs within 1Mb of the transcription start site with  $P < 1 \times 10^{-5}$ ) and AD summary statistics from MTAG.

### Trait-trait and trait-eQTL colocalisation analysis

We used HyPrColoc R package<sup>19</sup> to perform colocalisation analysis. HyPrColoc is a Bayesian divisive clustering algorithm for identifying shared genetic associations between traits in a genomic region using GWAS summary statistics. We performed this method to identify colocalised loci between AD and CV traits and prioritise causal variants explaining the shared association.

We performed a trait-trait colocalisation analysis for each top signal indicated from MTAG in a region  $\pm 200$  kb from the top SNP. We considered variant-specific priors for our analyses, which assumes that the probability of a variant being colocalised with a set of traits decreases as the number of the set of traits increases. The variant-specific priors model requires the specification of two priors. We specified the prior probability that a variant is associated with a single trait only at  $P = 1 \times 10^{-4}$  and a conditional prior probability that a variant is associated with an additional trait given that it is already associated with another trait at  $P_c = 0.02$ .

A  $PP > 0.5$  was considered adequate evidence that the examined traits colocalise in the locus. We divided the evidence of colocalisation into two categories: (1) considerable evidence ( $0.5 < PP < 0.75$ ) and (2) strong evidence ( $PP \geq 0.75$ ). To deal with spurious pleiotropy, we restricted the analyses to regions with at least one SNP with  $P < 5 \times 10^{-4}$  in the respective univariate GWAS. Additionally, we visually inspected the colocalised loci by constructing suitable regional plots. The variants explaining at least 80% of the shared association were considered candidate causal. To limit the probability of false positive findings, we considered as causal the variants that were associated ( $P < 0.01$ ) with both AD and the respective CV trait in the respective included univariate GWAS.

For the loci found to colocalise in the trait-trait colocalisation analysis, we further performed trait-expression quantitative trait loci (eQTL) colocalisation using AD, CV trait and eQTL from 48 tissues, retrieved from Genotype-Tissue Expression version 7 (GTEx v7), implementing the same parameters and approach as described in trait-trait colocalisation. The trait-eQTL colocalisation analysis was conducted to detect shared genes between the traits and investigate the tissues they are expressed.

### Single-cell and single-nuclei data acquisition

We used Gene Expression Omnibus<sup>59</sup> (GEO) to retrieve data for single cells of the left ventricle from 6 heart failure (HF) cases and 7 healthy controls<sup>60</sup>, left ventricular single nuclei from 13 dilated cardiomyopathy (dCM) cases and 25 healthy controls<sup>61</sup>, and post-mortem brain single nuclei from 9 AD cases and 8 healthy controls<sup>62</sup>. To deal with the small sample size in cardiac single-cell data, individuals with either coronary heart failure or dCM were considered HF cases.

### Quality control of single-cell and single-nuclei data

To quality control (QC) the data we implemented the scFlow pipeline<sup>63</sup>. Samples with  $< 100$  cells in cardiac data and  $< 200$  cells in brain data were removed. For brain data, Ambient RNA profiles were performed using EmptyDrops<sup>64</sup>. We restricted the minimum number of expressive features to 300 for cardiac data and 100 for brain data. For cardiomyocyte-enriched samples, we set the minimum library size per cell to 1000 while keeping the default value for the rest cell types. Only genes with a minimum of 2 counts in at least 3 cells were included. Doublet cells and non-annotated genes were removed. After QC, there was adequate sample to analyse cardiomyocyte-enriched (2,989 cells from 2 HF cases and 5 controls) and endothelium-enriched single cells (2,269 cells from 4 HF cases and 4 controls), macrophage single nuclei (207,345 nuclei from 13 dHF and 25 controls), astrocyte (34060 nuclei

from 25 AD cases and 24 controls), and microglia single nuclei (15292 nuclei from 25 AD cases and 24 controls).

### Single-cell and single-nuclei data integration, clustering and cell-type annotation

Cells that successfully passed the QC were integrated across samples using the linked inference of genomic experimental relationships (LIGER) method<sup>65</sup> defining a parameter  $\lambda = 5$  and selecting a sample-specific optimum value for parameter K: cardiomyocyte-enriched single cells ( $K = 25$ ), endothelium-enriched single cells ( $K = 30$ ), macrophage single nuclei ( $K = 40$ ), and brain single nuclei ( $K = 20$ ). A dimensionality reduction was performed by implementing the uniform manifold approximation and projection (UMAP)<sup>66</sup> algorithm to generate two-dimensional embeddings of the LIGER integrated factors using the first 10 principal components (PCs) on heart single cells, the first 60 PCs on heart single nuclei and the first 30 PCs on brain single nuclei. We subsequently detected cell clusters of the UMAP embeddings implementing the Leiden community detection algorithm<sup>67</sup> using a sample-specific parameter k: cardiomyocyte-enriched ( $k = 9$ ) and endothelium-enriched single cells ( $k = 10$ ), heart single nuclei ( $k = 45$ ), and brain single nuclei ( $k = 50$ ). Following clustering, we used the Expression Weighted Celltype Enrichment (EWCE)<sup>68</sup> algorithm to perform a cell-type prediction on cell clusters using previously reported reference datasets for cardiac single cells<sup>69</sup>, cardiac single nuclei<sup>61</sup> and brain single nuclei<sup>70</sup>.

### Differential gene expression analysis

A differential gene expression (DGE) analysis was performed separately for each cell-specific sample on all candidate genes detected from trait-eQTL colocalisation analysis. We investigated the expression of the genes in the RNAseq data and included only genes expressed in at least one cell type of the examined tissues. We followed a generalised linear mixed model approach as implemented in MAST<sup>71</sup> after excluding genes expressed in  $< 10\%$  of cells. Units for differential expression were defined as  $\log_2$  fold change ( $\log_2FC$ ) per unit change of the respective contrast. We considered as meaningfully differentially expressed genes those with a  $\log_2FC \geq 0.25$  and a nominal  $P$ -value  $< 0.05$ .

### Weighted gene co-expression network

We further constructed cell type-specific co-expression networks on selected candidate genes using high dimensional weighted gene co-expression network analysis (hdWGCNA)<sup>72</sup> R package. We applied the K-Nearest Neighbours algorithm to identify groups of similar cells by means of transcriptomics (metacells) and constructed a metacell gene expression matrix. We constructed the co-expression network using the lowest soft power threshold that has a Scale Free Topology Model Fit  $\geq 0.8$ . Genes that were not grouped into any co-expression module were excluded ("grey" module). We also excluded modules with  $< 20$  genes. We obtained the module eigengene values, which describe the expression patterns of entire co-expression modules, and performed a differential module eigengenes analysis applying a Mann-Whitney U test. To reduce false-positive findings due to multiple testing inflation, we implemented the Benjamini-Hochberg false discovery rate (FDR) method<sup>73</sup>. We also conducted a pathway enrichment analysis using Enrichr v3.0 R package<sup>74</sup> and analysed only gene-sets with at least 20 genes.

### Protein-protein interaction analysis

We used the STRINGdb<sup>75</sup> R package to analyse the full protein-protein interaction network data from STRING v11 database<sup>76</sup>. We expanded the candidate set of genes from the trait-eQTL colocalisation analysis by incorporating genes with protein-protein interactions (experimental evidence  $\geq 700$ ) with the candidates. Using the previously constructed cell type-specific modules from hdWGCNA, we performed

an enrichment analysis per module on candidate genes between the module and reference set using Fisher's exact tests. The Benjamini-Hochberg FDR approach was used to correct for type I error inflation due to the multiple testing error.

### Reporting summary

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

### Data availability

The summary statistics from the GWAS included in this study are publicly available and can be retrieved from GWAS Catalog under the accession codes [GCST007320](#) (AD), [GCST006414](#) (AF), [GCST003116](#) (CAD), [GCST006906](#) (stroke), [GCST006624](#) (SBP) and [GCST006630](#) (DBP). Heart single-cell data from the left ventricle tissue were downloaded from Gene Expression Omnibus (GEO) under accession codes [GSE109816](#) (cardiomyocytes-enriched samples) and [GSE121893](#) (normal digested samples). Single-nuclei data for the left ventricular tissue were retrieved from GEO under accession code [GSE109816](#). The single-nuclei RNA sequencing data for human post-mortem brain samples from AD and Control samples were retrieved from GEO under accession code [GSE160936](#). The MTAG summary statistics generated in this study have been deposited in NHGRI-EBI GWAS Catalog under accession codes [GCST90449053](#) (AD from AD-AF MTAG), [GCST90449054](#) (AF), [GCST90449055](#) (AD from AD-BP MTAG), [GCST90449056](#) (SBP), [GCST90449057](#) (DBP), [GCST90449058](#) (AD from AD-CAD MTAG), [GCST90449059](#) (CAD), [GCST90449060](#) (AD from AD-cIMT MTAG), [GCST90449061](#) (cIMT), [GCST90449062](#) (AD from AD-stroke MTAG), and [GCST90449063](#) (stroke). All other data generated in this study are provided with this published article (and its supplementary information files).

### Code availability

No previously unreported custom computer code or mathematical algorithm was used to generate results central to the conclusions.

### References

- Rizzi, L., Rosset, I. & Roriz-Cruz, M. Global epidemiology of dementia: Alzheimer's and vascular types. *Biomed. Res Int.* **2014**, 908915 (2014).
- Gauthier S., W. C., Servaes S., Morais J. A. & Rosa-Neto P. *World Alzheimer Report 2022: Life After Diagnosis: Navigating Treatment, Care and Support*. <https://www.alzint.org/u/World-Alzheimer-Report-2022.pdf> (2022).
- Newman, A. B. et al. Dementia and Alzheimer's disease incidence in relationship to cardiovascular disease in the cardiovascular health study cohort. *J. Am. Geriatr. Soc.* **53**, 1101–1107 (2005).
- Breteler, M. M., Claus, J. J., Grobbee, D. E. & Hofman, A. Cardiovascular disease and distribution of cognitive function in elderly people: the rotterdam study. *BMJ* **308**, 1604–1608 (1994).
- Bleckwenn, M. et al. Impact of coronary heart disease on cognitive decline in Alzheimer's disease: a prospective longitudinal cohort study in primary care. *Br. J. Gen. Pr.* **67**, e111–e117 (2017).
- Santiago, J. A. & Potashkin, J. A. The impact of disease comorbidities in Alzheimer's disease. *Front Aging Neurosci.* **13**, 631770 (2021).
- Tini, G. et al. Alzheimer's disease and cardiovascular disease: a particular association. *Cardiol. Res. Pr.* **2020**, 2617970 (2020).
- Andra, K. et al. Plectin-isoform-specific rescue of hemidesmosomal defects in plectin (-/-) keratinocytes. *J. Invest. Dermatol.* **120**, 189–197 (2003).
- Tublin, J. M., Adelstein, J. M., Del Monte, F., Combs, C. K. & Wold, L. E. Getting to the heart of Alzheimer disease. *Circ. Res.* **124**, 142–149 (2019).
- Iadecola, C. & Gottesman, R. F. Cerebrovascular alterations in Alzheimer disease. *Circ. Res.* **123**, 406–408 (2018).
- Broce, I. J. et al. Dissecting the genetic relationship between cardiovascular risk factors and Alzheimer's disease. *Acta Neuropathol.* **137**, 209–226 (2019).
- Mahley, R. W. Central nervous system lipoproteins: ApoE and regulation of cholesterol metabolism. *Arterioscler Thromb. Vasc. Biol.* **36**, 1305–1315 (2016).
- Jansen, I. E. et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
- Schwartzentruber, J. et al. Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nat. Genet.* **53**, 392–402 (2021).
- van der Harst, P. & Verweij, N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ. Res.* **122**, 433–443 (2018).
- Hartiala, J. A. et al. Genome-wide analysis identifies novel susceptibility loci for myocardial infarction. *Eur. Heart J.* **42**, 919–933 (2021).
- Turley, P. et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
- Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).
- Foley, C. N. et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat. Commun.* **12**, 764 (2021).
- Luo, S., Zheng, N. & Lang, B. ULK4 in neurodevelopmental and neuropsychiatric disorders. *Front Cell Dev. Biol.* **10**, 873706 (2022).
- Liu, M. et al. Ulk4 deficiency leads to hypomyelination in mice. *Glia* **66**, 175–190 (2018).
- Rivard, L. et al. Atrial fibrillation and dementia: a report from the AF-SCREEN international collaboration. *Circulation* **145**, 392–409 (2022).
- Livingston, G. et al. Dementia prevention, intervention, and care: 2020 report of the lancet commission. *Lancet* **396**, 413–446 (2020).
- Iturria-Medina, Y. et al. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. *Nat. Commun.* **7**, 11934 (2016).
- Sweeney, M. D. et al. Vascular dysfunction-the disregarded partner of Alzheimer's disease. *Alzheimers Dement* **15**, 158–167 (2019).
- Thoroldsdottir, R. B. et al. A missense variant in PLEC increases risk of atrial fibrillation. *J. Am. Coll. Cardiol.* **70**, 2157–2168 (2017).
- Wainberg, M. et al. Genetic architecture of the structural connectome. *Nat. Commun.* **15**, 1962 (2024).
- Aung, N. et al. Genome-wide association analysis reveals insights into the genetic architecture of right ventricular structure and function. *Nat. Genet.* **54**, 783–791 (2022).
- Sonnenberg, A. & Liem, R. K. Plakins in development and disease. *Exp. Cell Res.* **313**, 2189–2203 (2007).
- Leung, C. L., Green, K. J. & Liem, R. K. Plakins: a family of versatile cytolinker proteins. *Trends Cell Biol.* **12**, 37–45 (2002).
- Lagisetty, Y. et al. Identification of risk genes for Alzheimer's disease by gene embedding. *Cell Genom.* <https://doi.org/10.1016/j.xgen.2022.100162> (2022).
- Lie, A. A. et al. Plectin in the human central nervous system: predominant expression at pia/glia and endothelia/glia interfaces. *Acta Neuropathol.* **96**, 215–221 (1998).
- Valencia, R. G. et al. Plectin dysfunction in neurons leads to tau accumulation on microtubules affecting neuritogenesis, organelle trafficking, pain sensitivity and memory. *Neuropathol. Appl Neurobiol.* **47**, 73–95 (2021).
- Potokar, M. & Jorgacevski, J. Plectin in the central nervous system and a putative role in brain astrocytes. *Cells* <https://doi.org/10.3390/cells10092353> (2021).

35. Nedergaard, M., Ransom, B. & Goldman, S. A. New roles for astrocytes: redefining the functional architecture of the brain. *Trends Neurosci.* **26**, 523–530 (2003).
36. Potokar, M., Morita, M., Wiche, G. & Jorgacevski, J. The diversity of intermediate filaments in astrocytes. *Cells* <https://doi.org/10.3390/cells9071604> (2020).
37. Deak, F. Neuronal vesicular trafficking and release in age-related cognitive impairment. *J. Gerontol. A Biol. Sci. Med. Sci.* **69**, 1325–1330 (2014).
38. Fuchs, P. et al. Targeted inactivation of a developmentally regulated neural plectin isoform (plectin 1c) in mice leads to reduced motor nerve conduction velocity. *J. Biol. Chem.* **284**, 26502–26509 (2009).
39. Kamal, A. H. M., Chakrabarty, J. K., Udden, S. M. N., Zaki, M. H. & Chowdhury, S. M. Inflammatory proteomic network analysis of statin-treated and lipopolysaccharide-activated macrophages. *Sci. Rep.* **8**, 164 (2018).
40. Gomez-Arboledas, A., Acharya, M. M. & Tenner, A. J. The role of complement in synaptic pruning and Neurodegeneration. *Immunotargets Ther.* **10**, 373–386 (2021).
41. Carpanini, S. M. et al. Terminal complement pathway activation drives synaptic loss in Alzheimer's disease models. *Acta Neuropathol. Commun.* **10**, 99 (2022).
42. Dejanovic, B. et al. Complement C1q-dependent excitatory and inhibitory synapse elimination by astrocytes and microglia in Alzheimer's disease mouse models. *Nat. Aging* **2**, 837–850 (2022).
43. Haskard, D. O., Boyle, J. J. & Mason, J. C. The role of complement in atherosclerosis. *Curr. Opin. Lipidol.* **19**, 478–482 (2008).
44. Mihlan, M. et al. Monomeric C-reactive protein modulates classic complement activation on necrotic cells. *FASEB J.* **25**, 4198–4210 (2011).
45. Liu, C. C., Liu, C. C., Kanekiyo, T., Xu, H. & Bu, G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat. Rev. Neurol.* **9**, 106–118 (2013).
46. Gratten, J. & Visscher, P. M. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Med* **8**, 78 (2016).
47. Pickrell, J. K. et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet* **48**, 709–717 (2016).
48. Miyamoto, T. et al. C1 esterase inhibitor in pediatric cardiac surgery with cardiopulmonary bypass plays a vital role in activation of the complement system. *Heart Vessels* **35**, 46–51 (2020).
49. Spurrier, J. et al. Reversal of synapse loss in Alzheimer mouse models by targeting mGluR5 to prevent synaptic tagging by C1Q. *Sci. Transl. Med.* **14**, eabi8593 (2022).
50. Nielsen, J. B. et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat. Genet* **50**, 1234–1239 (2018).
51. Nikpay, M. et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet* **47**, 1121–1130 (2015).
52. Malik, R. et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet* **50**, 524–537 (2018).
53. 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).
54. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
55. Genomes Project, C. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
56. Kunkle, B. W. et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet* **51**, 414–430 (2019).
57. Bellenguez, C. et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat. Genet* **54**, 412–436 (2022).
58. Consortium, G. T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
59. Barrett, T. et al. NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
60. Litvinukova, M. et al. Cells of the adult human heart. *Nature* **588**, 466–472 (2020).
61. Koenig, A. L. et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat. Cardiovasc. Res.* **1**, 263–280 (2022).
62. Smith, A. M. et al. Diverse human astrocyte and microglial transcriptional responses to Alzheimer's pathology. *Acta Neuropathol.* **143**, 75–91 (2022).
63. Khozoe, C. et al. scFlow: A scalable and reproducible analysis pipeline for single-cell RNA sequencing data. *bioRxiv* <https://doi.org/10.1101/2021.08.16.456499> (2021).
64. Lun, A. T. L. et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
65. Welch, J. D. et al. Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887 e1817 (2019).
66. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv* <https://doi.org/10.48550/arXiv.1802.03426> (2018).
67. Traag, V. A., Waltman, L. & van Eck, N. J. From louvain to leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
68. Skene, N. G. & Grant, S. G. Identification of vulnerable cell types in major brain disorders using single cell transcriptomes and expression weighted cell type enrichment. *Front. Neurosci.* **10**, 16 (2016).
69. Wang, L. et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat. Cell Biol.* **22**, 108–119 (2020).
70. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
71. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
72. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E. & Swarup, V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep. Methods* **3**, 100498 (2023).
73. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Met.* **57**, 289–300 (1995).
74. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
75. Szklarczyk, D. et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).
76. Szklarczyk, D. et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

## Acknowledgements

The Genotype-Tissue Expression version 7 (GTEx v7) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from dbGaP accession number phs000424.v7.p2 on 30/03/2022. DW is supported by the Academy of Medical Sciences

Professorship (APR7\_1002). PMM and PE acknowledge personal support from the UK Dementia Research Institute, which is funding by the UKRI Medical Research Council, Alzheimer's Society and Alzheimer's Research UK, Edmond J. Safra Foundation and Lily Safra, an NIHR Senior Investigator Award, and the Imperial College Healthcare Trust (ICHT) NIHR Biomedical Research Centre. The authors acknowledge support from the Trustees of the Sir Michael Uren Foundation for this work. The authors also acknowledge support by the British Heart Foundation Research Excellence Award (4) (RE/24/130023 and the NIHR Imperial Biomedical Research Centre (BRC).

## Author contributions

A.D., D.W., P.M., F.K. and I.T. designed the research. F.K., N.F., P.F.T., and D.M. conducted the analyses and visualised the results. F.K., N.F., P.F.T., E.E., A.D., D.W., P.M., and I.T. interpreted the results. F.K. wrote the manuscript. N.F., P.F.T., D.M., P.E., A.D., D.W., P.M., and I.T. critically revised the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at

<https://doi.org/10.1038/s41467-024-53452-6>.

**Correspondence** and requests for materials should be addressed to Fotios Koskeridis.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

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

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

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

© The Author(s) 2024