Genetic risk factors underlying white matter hyperintensities and cortical atrophy

Yash Patel, Jean Shin, Eeva Sliz, Ariana Tang, Aniket Mishra, Rui Xia, Edith Hofer, Hema Sekhar Reddy Rajula, Ruiqi Wang, Frauke Beyer, et al.

nature communications

Article

https://doi.org/10.1038/s41467-024-53689-1

Genetic risk factors underlying white matter hyperintensities and cortical atrophy

Received: 1 April 2024

Accepted: 18 October 2024

Published online: 04 November 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

White matter hyperintensities index structural abnormalities in the cerebral white matter, including axonal damage. The latter may promote atrophy of the cerebral cortex, a key feature of dementia. Here, we report a study of 51,065 individuals from 10 cohorts demonstrating that higher white matter hyperintensity volume associates with lower cortical thickness. The meta-GWAS of white matter hyperintensities-associated cortical 'atrophy' identifies 20 genome-wide significant loci, and enrichment in genes specific to vascular cell types, astrocytes, and oligodendrocytes. White matter hyperintensities-associated cortical 'atrophy' showed positive genetic correlations with vascular-risk traits and plasma biomarkers of neurodegeneration, and negative genetic correlations with cognitive functioning. 15 of the 20 loci regulated the expression of 54 genes in the cerebral cortex that, together with their co-expressed genes, were enriched in biological processes of axonal cytoskeleton and intracellular transport. The white matter hyperintensities-cortical thickness associations were most pronounced in cortical regions with higher expression of genes specific to excitatory neurons with long-range axons traversing through the white matter. The meta-GWAS-based polygenic risk score predicts vascular and all-cause dementia in an independent sample of 500,348 individuals. Thus, the genetics of white matter hyperintensities-related cortical atrophy involves vascular and neuronal processes and increases dementia risk.

White matter hyperintensities (WMH) are lesions of presumed vascular origin, commonly found in the periventricular and deep white matter[1,2]. WMH are one of the markers of cerebral small vessel disease[3] and are clinically associated with a higher risk of incident stroke, dementia[4–6], and mortality[4]. Their prevalence increases with age, being present in ~20% of individuals at the age of 60 years and in >90% of individuals at the age of 80 years[7]. WMH burden also increases with modifiable vascular risk factors such as hypertension[8], type 2 diabetes, smoking, and obesity[9–11].

Radiologically, WMH are defined as areas of increased signal intensity in T2 or fluid-attenuated inversion recovery (FLAIR) imaging. The pathobiology of WMH is not fully understood, but post-mortem histology shows disrupted myelin, axons, altered water content, and mild gliosis[1,3]. Injury of axons at the site of WMH may lead to retrograde degeneration of neuronal bodies and dendritic arbour within the cerebral cortex and thus promote cortical atrophy[2]. Impaired axonal transport may contribute to this process by compromising the retrograde transport of various neurotrophic factors from the axon terminals to the cell body[12] and dendrites[13]. Previous research in smaller studies ($n < 2000$ participants) reported mostly inverse associations between WMH and cortical thickness[14–23] (and reviewed in ref. 2). Genetics and neurobiology of this relationship and its links to dementia have not been studied.

Here, we investigate the relationship between WMH and cortical thickness using a large sample of individuals with both imaging and genetic data. We observe global and regional inverse associations between WMH and cortical thickness. The genetics of WMH-associated cortical 'atrophy' is correlated positively with vascular risk factors and

e-mail: tomas.paus@umontreal.ca; zdenka.pausova@montreal.ca

enriched in vascular and glial cell types. Regions with the largest WMH-associated cortical atrophy show higher expression of genes specific to excitatory neurons projecting in the underlying white matter. Lastly, we explore the polygenic risk score of WMH-associated cortical atrophy and dementia risk.

Results

Higher WMH load associates with lower cortical thickness

Here, we studied a sample of 51,065 stroke-free and dementia-free individuals of European ancestry from 10 population-based cohorts (38–48% men, mean age of 52–77 years across cohorts; Supplementary Table 1 and Supplementary Data 1). Higher WMH volume was associated with lower cortical thickness (Fig. 1a). The association was of the same direction in all 10 cohorts (Fig. 1a). The association was adjusted for age, age^2, sex, intracranial volume, and cohort-specific covariates, and it remained significant when additionally adjusted for vascular risk factors, i.e., body mass index (BMI), hypertension, type 2 diabetes, and cigarette smoking (Fig. 1b). The WMH-cortical thickness association varied across regions of the cerebral cortex, parcellated according to the 34 regions of the Desikan-Killiany atlas; the association was of the largest effect size in the insula[24] (Fig. 1c, d). Notably, the insula serves as a key node in multimodal integration networks[25] and a point of convergence for widespread cortical and subcortical inputs[26]. Given this, it may be more vulnerable to axonal damage of multiple long-range neurons forming afferent and efferent connections.

Genetic vulnerability to WMH-associated cortical atrophy

First, to uncover genetic underpinnings of the WMH-insular cortical thickness association, we performed a meta-GWAS on the shared variance between WMH and thickness of the insula, which was calculated as the first principal component [PC1] of the two variables (Fig. 2a). PC1 was loaded positively by WMH and negatively by insular thickness, and, as such, we considered it to be an index of WMH-associated 'cortical atrophy'. The meta-GWAS of PC1 identified a total of 20 genome-wide significant loci (Fig. 2b and Supplementary Data 2). Of these 20 loci, 5 were 'unique' to WMH-associated cortical atrophy (i.e., not reported previously as being associated with either WMH or cortical thickness); additional 9 loci have been reported previously as being associated with either WMH or cortical thickness, and 6 loci have been reported previously as being associated with both WMH and cortical thickness (Supplementary Table 2).

Second, to explore the genetic overlap between PC1 (i.e., WMH-associated cortical atrophy) and various cognitive, neurodegenerative, psychiatric, and vascular-risk traits, we conducted linkage disequilibrium-score regression analysis[27]. PC1 was negatively correlated with the genetics of general cognitive ability and positively correlated with the genetics of stroke, type 2 diabetes, higher BMI, and

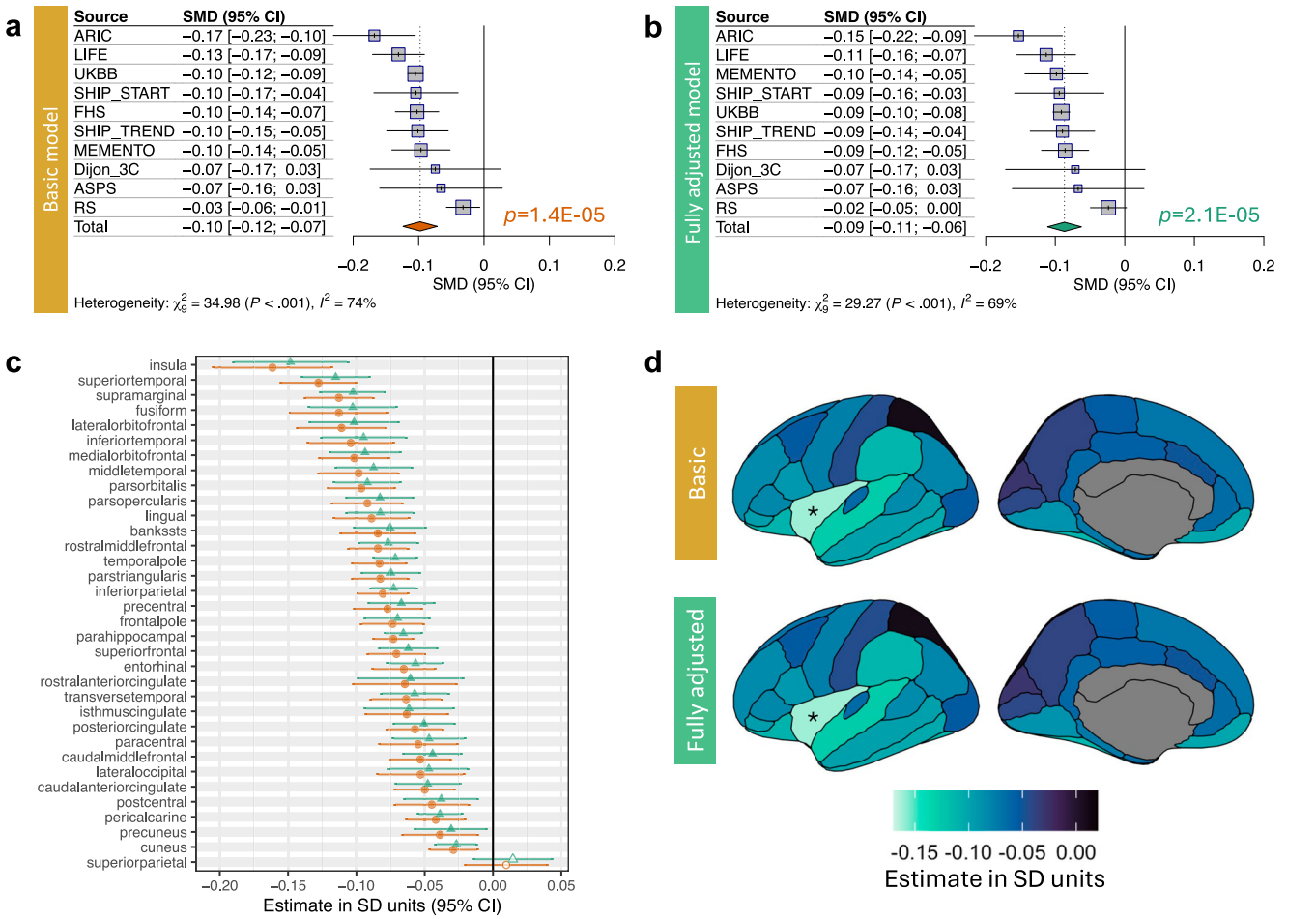


Fig. 1 | Association between white matter hyperintensities and cortical thickness. Forrest plot and meta-analytic summary statistic for the effect of WMH volume on mean cortical thickness within a baseline- (**a**) and fully-covariate adjusted model, including vascular risk factors (**b**). Meta-analytic effect sizes of WMH-cortical thickness association per cortical region are shown as a forest plot (**c**) and plotted on the surface of the cerebral cortex using ggseg[76] (**d**). For (**a**–**c**), two-sided *t* tests were used to evaluate the null hypothesis of no WMH-cortical thickness association; results with FDR-corrected *p*-values (adjusted for regions and models) less than 0.05 are marked by solid diamonds; and error bars represent 95% confidence intervals. The exact p-values are available in the Source Data.

https://doi.org/10.1038/s41467-024-53689-1

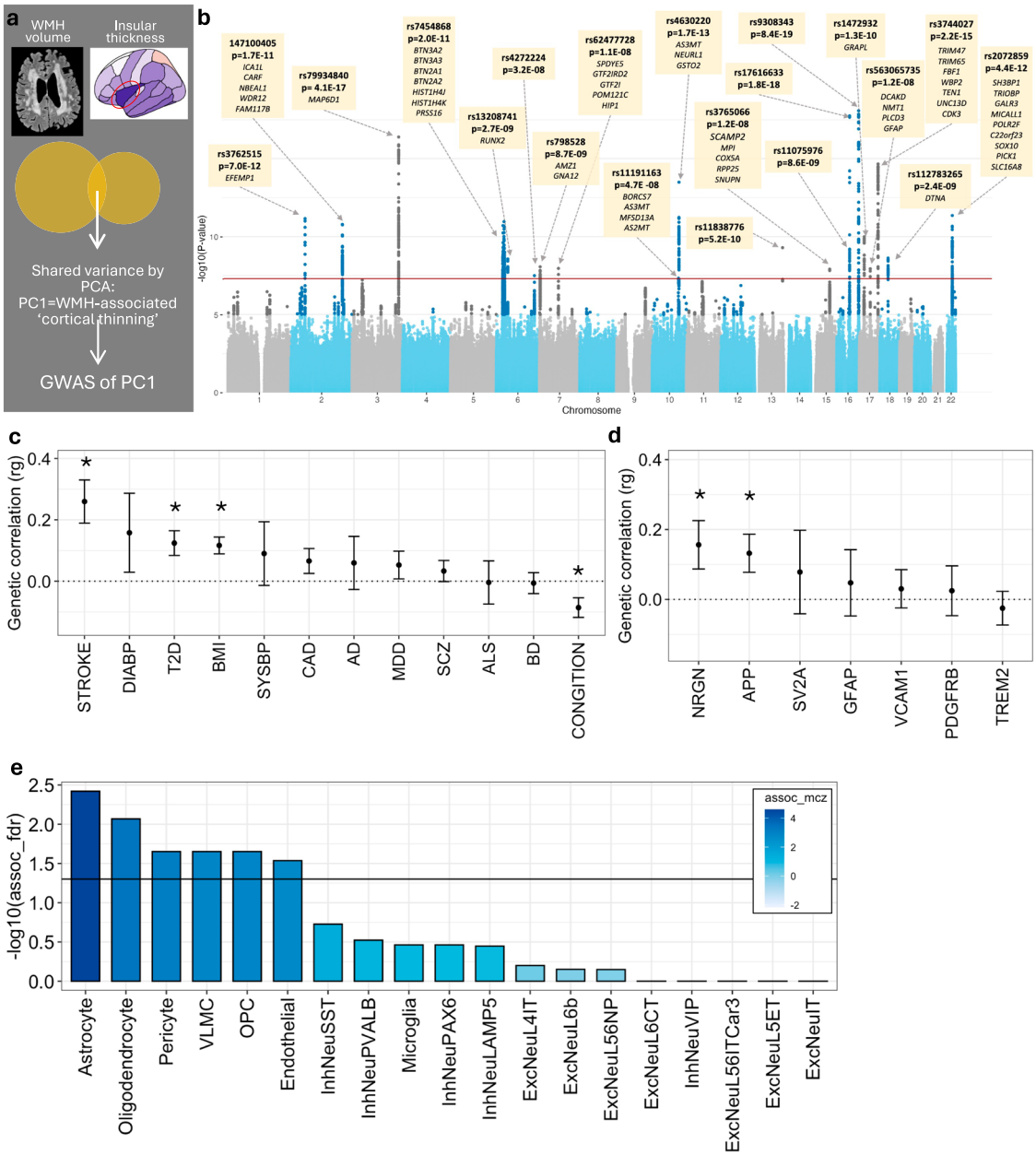


Fig. 2 | Genetic underpinnings of the WMH-cortical thickness association. **a** Schematic representation of the shared variance between WMH volume and insular cortical thickness captured via principal component 1 (PC1). **b** Summary statistics of the meta-GWAS of PC1. A two-sided Wald test was used to evaluate the null hypothesis of no SNP-PC1 association. The genome-wide significance level of 5e-08 is indicated by the red horizontal line. Each independent genome-wide significant variant is annotated, and its cis-eQTL-regulated genes are listed below variant rsID. The exact *p*-values are available in Supplementary Data 2. **c** LD-score regression estimates between the GWAS of PC1 and vascular risk factors, and neurodegenerative/psychiatric traits (error bars represent standard error, * $p < 0.05$, two-sided test). **d** LD-score regression between GWAS of PC1 and plasma protein levels of select neurodegeneration-related markers (error bars represent standard error, * nominal $p < 0.05$, two-sided test). **e** Cell-type-specific enrichment of polygenic signals from the GWAS of PC1, using single-cell disease relevance score testing[35]. Filled in colour represents Monte Carlo-based Z statistics. DIAB Diastolic Blood Pressure, T2D Type 2 Diabetes, BMI Body Mass Index, SYSBP Systolic Blood Pressure, CAD Coronary Artery Disease, AD Alzheimer's Disease, MDD Major Depressive Disorder, SCZ Schizophrenia, BD Bipolar Disorder, NRGN neurogranin, APP amyloid precursor protein, SV2A Synaptic vesicle protein 2, GFAP Glial fibrillary acidic protein, VCAM1 Vascular cell adhesion protein 1, PDGFRB platelet-derived growth factor receptor beta, TREM2 Triggering receptor expressed on myeloid cells 2. These genes were a priori selected as related to biomarkers of neurodegeneration[77]. For (**c**–**e**), the exact *p*-values are available in the Source Data.

two plasma markers of neurodegeneration, i.e., amyloid-precursor protein (APP) and neurogranin (Fig. 2c, d). Fragments of APP (amyloid β) are the major constituent of AD-associated amyloid plaques, and mutations or duplications of APP are implicated in familial AD[28–30]. APP and its fragments may also function as long-distance sensors of cellular activity/damage, and regulators of axonal transport (among others), which may be particularly important in large neurons[31]. Elevated levels of neurogranin in cerebrospinal fluid may indicate synaptic dysfunction[32] and the levels are also elevated in plasma following acute brain injury[33].

Third, to identify which cell types mediate the genetic vulnerability to WMH-associated cortical atrophy, we tested if genes

annotated to the meta-GWAS loci were enriched in genes specific to any of the 19 transcriptionally defined cell types derived from the Allen Institute's SMARTseqv4 single-nucleus RNA sequencing dataset[34]. Single-cell disease-relevance score testing[35] implicated astrocytes, oligodendrocytes, pericytes, vascular leptomeningeal cells (VLMC), oligodendrocyte precursor cells (OPC), and endothelial cells (Fig. 2e). These cell types include those forming small vessels (astrocytes, pericytes, and endothelial cells) and providing physical protection and tropic support to axons that constitute the deep white-matter tracks (oligodendrocytes)[1,2]. As such, these cell types may tie small vessel injury to axons traversing through the white matter and, in turn, to neuronal cell bodies and dendritic arbour in the cerebral cortex.

Cortical cell types and biological processes of WMH-associated cortical atrophy

First, to reveal which cell types in the cerebral cortex may contribute to WMH-associated cortical atrophy, we employed a cell-type enrichment method[34,36]. This approach exploits the spatial relationship – across the 34 regions of the Desikan-Killiany atlas – between the observed inter-regional variation in the effect size of the WMH-cortical thickness association (Fig. 1c, d) and inter-regional variation in the expression of genes specific to the above-specified 19 transcriptionally defined cortical cell types[34]. In this analysis, the most significant cell types were subtypes of excitatory neurons with cell bodies in cortical layers 2, 3, 5, and 6 (ExcNeuIT, EcxNeuL6b, ExcNeuL56ITCar3, and ExcNeu56NP; Fig. 3a, b). Neurons from these layers have extensive long-range axonal projections traversing through the white matter (intra- and inter-hemispheric)[37]. Genes specific to these excitatory neurons showed higher expression in cortical regions, demonstrating larger negative WMH-cortical thickness associations. In contrast, genes specific to another subtype of excitatory neurons, EcxNeuL4IT, which have cell bodies in layer 4 and have fewer axonal projections traversing through the white matter, did not show this relationship (Fig. 3a, b). Consistently, layer 4 excitatory neurons are expanded within the cortical regions receiving thalamic sensory input[37,38], such as the primary visual primary somatosensory cortices, which – in the present study – demonstrated smaller effect sizes of the WMH-cortical thickness associations (Fig. 1c, d).

Second, to uncover biological processes of WMH-associated cortical atrophy in the cerebral cortex, we analysed 54 genes whose cortical expression was regulated[39] by 15 of the 20 identified meta-GWAS loci of PC1 (Supplementary Data 2). Co-expression analysis of these 54 genes using cortical bulk RNA sequencing data from 5 independent datasets with 534 unique donors[40] showed that the most positively co-expressed genes were enriched in biological processes related to cytoskeleton and cellular transport (e.g., 'cytoskeletal organisation', 'cell polarity', 'organelle localisation', and 'intracellular transport [protein and organelle]'; Fig. 3c and Supplementary Fig. 1). Given the polarised morphology of neurons, axonal transport is essential for normal neuron functioning[41,42]. Defects in axonal transport have been implicated in multiple neurodegenerative diseases as an early pathological feature[42]. Additional enriched biological processes were 'axonogenesis', 'cellular component disassembly', 'catabolic processes', and 'cellular respiration' (Fig. 3c).

Polygenic risk score of WMH-associated cortical atrophy increases the risk for vascular and all-cause dementia

To examine whether genetic vulnerability to WMH-associated cortical atrophy was associated with a higher risk of dementia, we tested if a polygenic risk score (PRS) generated from the meta-GWAS summary-statistics of PC1 was associated with a higher risk of vascular dementia, all-cause dementia, and/or Alzheimer's disease. In an independent sample of 500,348 participants from the FinnGen study[43], we show that the PRS was associated with a higher risk of vascular dementia (3624 cases) and all-cause dementia (21,257 cases) but not late-onset Alzheimer's disease (9690 cases). Specifically, individuals in the top decile of the PRS showed a 52% higher risk of vascular dementia ($p = 8.3 \times 10^{-8}$) and an 18% higher risk of all-cause dementia ($p = 1.7 \times 10^{-6}$) compared with those in the bottom decile of the PRS. The 6% risk increase in Alzheimer's disease did not reach statistical significance ($p = 0.27$) (Fig. 4).

Discussion

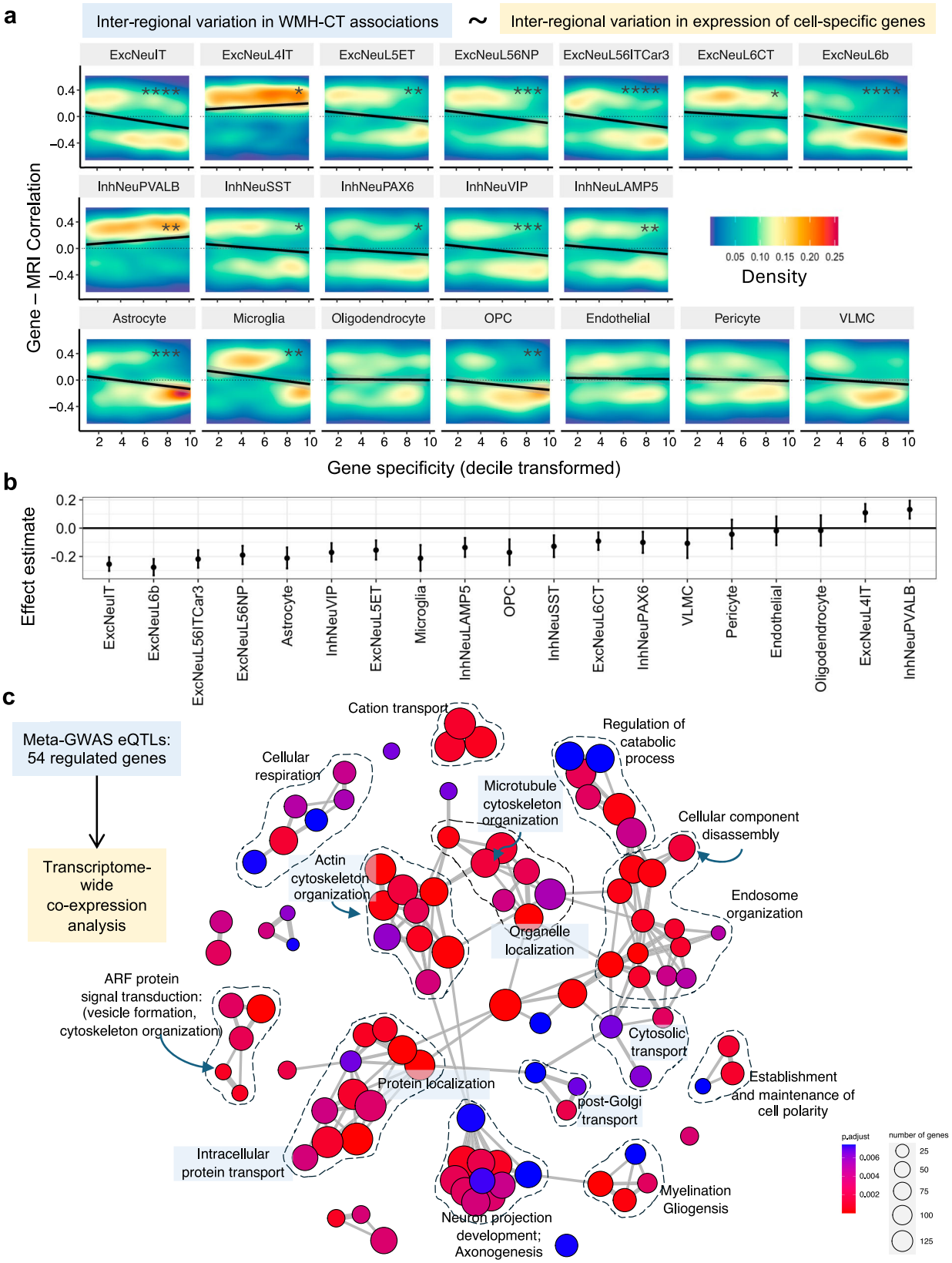
The present study of 51,065 individuals from 10 independent cohorts with brain imaging data shows that WMH burden associates inversely with thickness of the cerebral cortex. Our meta-GWAS on the shared variance between WMH and cortical thickness, together with a series of in silico analyses employing single-nuclei and bulk RNA sequencing data from the human cerebral cortex, suggests that WMH-associated cortical atrophy involves smallvesselforming cells and oligodendrocytes and impacts – in the cerebral cortex – excitatory neurons with long-range axonal projections traversing through the white matter. PRS based on this meta-GWAS predicts a higher risk of vascular dementia and all-cause dementia in an independent sample of 500,348 individuals.

Our results strongly support the relationship between the combined WMH-thickness phenotype (i.e., PC1) and processes related to the cerebral vasculature. We observed: (i) the positive genetic correlations of PC1 with vascular risk factors, (ii) the polygenic signals from the GWAS of PC1 were enriched (among others) in genes specific to cell types forming small vessels, i.e., endothelial cells, pericytes and astrocytes, and (iii) the PC1 PRS was associated most strongly with vascular dementia. Our findings can be interpreted in two (mutually non-exclusive) ways: (i) vascular changes in white matter lead to its damage and, in turn, cortical atrophy, via, for example, impaired axonal transport; and (ii) WMH are an index of impaired cerebral vasculature, including the vasculature in the cerebral cortex, which leads to cortical atrophy. Our results cannot speak to either the directionality of such possible relationships or the relative contributions of the two hypothetical pathways.

The present study was cross-sectional, and thus the observed association between WMH and cortical thickness does not imply causality. Our Mendelian Randomisation (MR) analyses were inconclusive, indicating causal effects of WMH on insular cortical thickness with the 'inverse variance-weighted' method only (Supplementary Data 3). Although the causality of the WMH-insular cortical thickness association is not proven, we did observe an inverse association between the PRS of WMH and insular cortical thickness (Supplementary Table 3). In addition, we observed some genetic overlap between the two variables: 15 of the 20 GWAS-significant loci of PC1 were associated not only with PC1 but also with WMH and cortical thickness (at $p < 0.05$, Supplementary Data 4), and 6 of these were reported previously as GWAS-significant loci of both WMH and cortical thickness (Supplementary Table 2).

All our meta-GWAS and related analyses were performed using PC1 derived from WMH and thickness of the insula, which was a region of the cerebral cortex demonstrating the largest negative effect size of the WMH-cortical association. We chose the insula based on the statistical significance and the consistency of the association across the 10 cohorts in terms of its directionality and spatial distribution across the cortex. While the exact reasons for the insula showing the strongest association with WMH remain uncertain, it may be related to its role as a key node in multimodal integration networks[25], which may make the insula's multiple long-range neurons that form its afferent and efferent fibres and traverse through the white matter more vulnerable to axonal damage. Nonetheless, we performed supplementary analyses using PC1 derived from WMH and global (mean) cortical thickness, and the results were quite similar: (i) all GWAS-significant loci of PC1 derived with insular cortical thickness were also associated with PC1 derived with the mean cortical thickness (at $p < 6 \times 10^{-3}$, Supplementary

https://doi.org/10.1038/s41467-024-53689-1



Data 5); (ii) genetic correlations with cognitive, neurodegenerative, psychiatric, and vascular-risk traits showed similar patterns (Supplementary Fig. 2); and (iii) the PRS-associated risks for dementia were similar; for vascular dementia, for example, the risk was 52% higher in individuals with the top (vs. bottom) decile of the PRS with insular cortical thickness, and it was 33% higher in individuals with the top (vs. bottom) decile of the PRS with the mean cortical thickness (Supplementary Fig. 3).

The present study was performed in individuals of European ancestry. The lack of other ethnicities is a limitation of this study, as other multi-ethnic research of complex genetic traits indicates that simple trans-ethnic transferability of the results may be limited. For

Fig. 3 | Cell types in the cerebral cortex mediating the WMH-cortical thickness association. a Results from cell-type enrichment analysis across the cerebral cortex. Association between gene specificity for a given cell type (*x*-axis) and the gene's correlation coefficient between its expression and effect sizes of the WMH-cortical thickness association across the 34 cortical regions. The black line represents linear regression fit, and the colour represents the density of data points. **b** Effect sizes and confidence intervals from this cell-type enrichment analysis using a Pearson's correlation coefficient (two-sided *p*-value). **c** Gene ontology enrichment analysis of genes co-expressed with genes regulated by genome-wide significant variants from the GWAS of PC1. Each node represents a significant biological process. Clusters of terms with high similarity are linked by edges, and manually annotated with an overarching/representative biological process ExcNeu excitatory neuron, InhNeu inhibitory neuron, OPC oligodendrocyte precursor cell, VLMC vascular leptomeningeal cells; FDR corrected *p*-values **** $< 0.5 \times 10^{-10}$, *** $< 0.5 \times 10^{-06}$, ** $< 0.5 \times 10^{-04}$, * $< 0.5 \times 10^{-2}$; and the exact *p*-values are available in the Source Data.

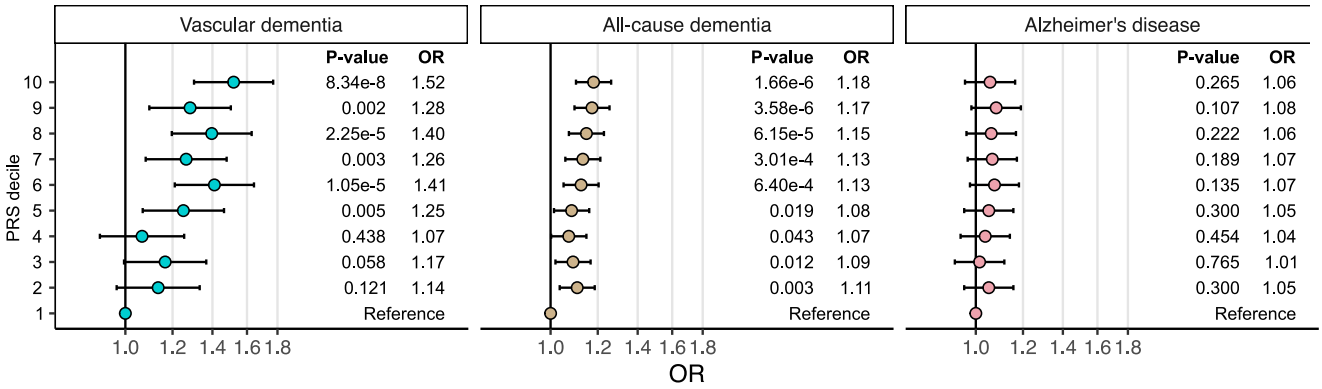


Fig. 4 | Association between polygenic risk score (PRS) of WMH and insular thickness-derived PC1 and the risk of each vascular dementia, all-cause dementia, and late-onset Alzheimer's disease. The odds ratios were calculated in FinnGen (consisting of a total of 500, 348 individuals) by comparing each of the top nine PRS deciles to the lowest decile and adjusting for age, sex, the first 10 genetic principal components and genotyping arrays. Error bars represent 95% confidence intervals. The exact *p*-values are available in the Source Data.

example, the Pearson correlation of effect sizes of BP loci between European and African ancestries was only 0.37[43]. Another potential limitation is the fact that WMH and cortical thickness were assessed with varying methods (e.g., 1.5 T or 3 T MRI scanners) across cohorts. Although the WMH-cortical thickness associations showed consistent direction of effect and spatial distribution across the cerebral cortex, we cannot entirely exclude the possibility that the varying methods impacted our findings.

In conclusion, the present study shows that the genetics of WMH-associated cortical atrophy involves vessel-forming cells, astrocytes, and oligodendrocytes. The regional variation of WMH burden on cortical thickness varied with expression specific to excitatory neurons with long-range axonal projections traversing through the white matter. Genetic vulnerability to WMH-associated cortical atrophy increases the risk for vascular and all-cause dementia.

Methods

Ethics

Ethics oversight of this study was provided by the SickKids Research Ethics Board (#1000073323). Individual cohort protocols were approved by the respective institutional review boards or equivalent organisations, and all participants provided written informed consent (see Supplement Note 1 for details).

Participants

We studied the genetics and neurobiology of the WMH-associated cortical atrophy in 51,065 participants from 10 cohort studies, including those collaborating in the Cohorts of Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium[44] and the UK Biobank[45]. All the individuals, aged between 19 and 100 years, were stroke- and dementia-free and of European ancestry (a brief description of the cohorts and basic characteristics of their participants are provided in Supplement Note 1). We constructed a polygenic score of WMH-associated cortical atrophy and studied its associations with dementia in up to 500,348 individuals from the FinnGen cohort[43].

Neuroimaging measures

Thickness of the cerebral cortex was assessed from T1-weighted magnetic resonance images of the brain using FreeSurfer[46]. Mean cortical thickness and cortical thickness at each of the 34 regions parcellated based on the Desikan-Killiany atlas[47] were analysed. The total *volume of WMH* was assessed from T2-weighted/FLAIR images (Supplementary Data 1).

Statistical analyses

Associations between WMH and cortical thickness were examined using linear regression models. Prior to model fitting, WMH and cortical thickness variables were inverse normal transformed. We fitted the regression models using a simple covariate structure (basic models) and a more complete covariate structure (fully adjusted models) to test if the associations are independent of cardiometabolic risk factors previously associated with WMHs and/or cortical thickness, namely hypertension, type 2 diabetes, BMI, and smoking. The basic covariates were linear and quadratic age, intracranial volume, and cohort-specific MRI-related covariates (e.g., MRI site). Cohorts that included related individuals used mixed-effects models that adjusted for family structure. The effect estimates from individual cohort analyses were then meta-analysed using the inverse-variance method with random effects that incorporate heterogeneity in the estimates. Sex-specific analyses were not conducted in this study.

Genome-wide association analysis of the shared variance between WMH and cortical thickness In each cohort, the shared variance between WMH and cortical thickness was derived using principal component analysis of the 2 variables. Principal component 1 (PC1), loaded positively by WMHs and negatively by cortical thickness, was considered an index of 'WMH-associated cortical thinning'. PC1 adjusted for basic covariates was used to fit linear regression models to

test the association between PC1 and allele dosage of SNPs. Additive genetic effects were assumed. The regression models were adjusted for sex, the first 10 genetic principal components (to account for genetic relatedness), and/or cohort-specific covariates (adjustments for age were done at the level of PC1 derivation). In cohorts including family members, linear mixed-effects models were used to account for family structure. The cohort-specific GWAS results were examined for quality control with Easy QC software (v9.0)[48]: genetic variants with a minor allele frequency (MAF) < 0.05, low imputation quality (R2 < 0.4), and available in < 10,000 individuals were removed from further analyses. The results were meta-analysed with METAL[49] under fixed effects models with a sample-sized method (i.e., using z-scores and their directions). The top SNPs were additionally meta-analysed under a modified random effect model[50].

Genetic correlations of PC1 GWAS We used linkage disequilibrium (LD)-score regression analysis[27] to test the genetic correlations between PC1 and vascular-risk (systolic blood pressure[51], diastolic blood pressure[51], stroke[52], BMI[53], diabetes[54], coronary artery disease[55]), psychiatric (schizophrenia[56], major depression[57], bipolar disorder[58]) and neurodegenerative (Alzheimer's disease[59]) disorders and general intelligence[60]. Plasma protein-level GWASs were obtained from the Pharma Proteomics Project using individuals from the UK Biobank[61].

Cell-type enrichment analysis of PC1 GWAS We used single-cell disease relevance score (scDRS)[35] to test the enrichment of the PC1 GWAS signals within the single-cell RNAseq dataset from the Allen Brain Institute. Briefly, this approach uses the top 1000 genes associated with PC1 (derived using MAGMA v1.09b[62]) and computes an aggregate expression of PC1-genes within each nucleus and cell type. These scores are compared with control scores (using random sets of genes and nuclei/cell types) to compute cell-level and cell-type-level *p*-values.

Functional enrichment analysis of PC1 GWAS. For this, we examined the genes whose expression in the cerebral cortex was regulated by GWAS-identified SNPs. To identify those genes, we performed (*cis*-) eQTL mapping using MetaBrain (https://www.metabrain.nl/, accessed Jan 2023), where Bonferroni FDR correction was applied to variants eQTL-eGene tested models, rather than all eQTL-eGene pairs[39]. Then, for each of the *cis*-eQTL-regulated genes, we identified the genes that were co-regulated (or co-expressed) in the human cerebral cortex, as measured in 5 independent datasets including a total of 534 donors aged 0-102 years, as we did in refs. 40,63. The co-expression strengths were measured by the effect estimates of the linear mixed models regressing the expression level of each *cis*-eQTL gene against the expression of each of the other genes, adjusting for database and doner-ID as random effects, and donor age and sex as fixed effects. Models were fitted to the gene-expression data harmonised across 5 databases, for which the expression values were log-transformed and normalised across regions and databases to adjust for differences in age ranges, sampling regions, and technical methods as described in ref. 40. We then selected the top 0.5% of the positively co-expressed (i.e., estimate > 0) genes as the final set for Gene-Ontology biological process (GO-BP) enrichment analysis, which was performed using the ClusterProfiler R package[64].

Cell-type enrichment analysis of the WMH-cortical thickness association This was performed using the cortical, single-nucleus gene-expression data from Allen Human Brain Atlas[65]. The gene expression data were mapped to the 34 cortical regions of the Desikan-Killiany Atlas[66], and genes were filtered for consistency using an internal donor-to-median filter, and an external interregional-correlation filter, using gene expression data from the BrainSpan[67]. From the initial 20,737 genes, 2511 genes remained for further analyses. Next, for each of the filtered genes, cell-type specificity was calculated for 19 cell types using the human cortical single nuclei RNAse data from the Allen Brain Institute[34]. An aggerate measure of cell specificity was quantified using the CELLEX package (v 1.2.2)[68]. For each cell type, we calculated the Pearson correlation coefficients between the interregional profiles of WMH-cortical thickness association estimates and interregional profiles of gene expression for all the filtered genes. Then we assessed the relationship between a gene's specificity for a given cell type and the correlation of the gene's expression with WMH-cortical thickness association by two methods: linear association and top cell-specific gene enrichment. The linear enrichment test evaluates the linear association between increasing cell-type gene specificity and its effect on the correlation between expression and WMH-cortical thickness. The top-specific gene method evaluates the mean of the top 100 cell type-specific gene's correlation (between expression and WMH-cortical thickness) against randomly bootstrapped sets of genes. Both methods converged on the same significant cell types of interest, and only the effect estimate from the linear approach is shown in Fig. 3b. *P*-values were adjusted for multiple testing of 19 cell types using the Benjamini-Hochberg false-discovery procedure[69].

Polygenic risk score (PRS) of PC1 (i.e., WMH-associated cortical atrophy) We built a genome-wide PRS of PC1 (derived from WMH and thickness of the insular cortex) using PRS-CS (v1.1.0)[70] as implemented in the FinnGen PRS-CS pipeline (described at https://github.com/FINNGEN/CS-PRS-pipeline). PRS-CS computes SNP effect sizes using high-dimensional Bayesian regression with continuous shrinkage (CS) priors, leveraging GWAS summary statistics and a linkage disequilibrium (LD) reference panel. Here, we used GWAS summary statistics of both PC1 of WMH and insular thickness and PC1 of WMH and the mean cortical thickness along with the European LD reference panel from the 1000 Genomes Project. Both PRSs were categorised into deciles, and the associations with vascular dementia, all-cause dementia, and late-onset Alzheimer's disease were evaluated in R (v4.4.0) using logistic regression by comparing each of the upper nine PRS deciles with the lowest decile and adjusting for age, sex, the first 10 genetic principal components, and genotyping arrays. The genotyping details are available in the Supplement.

Mendelian randomisation (MR) to test the causal effects of WMH on the thickness of the insular cortex Two-sample MR analyses were performed, where genetic variants were used as instrumental variables to determine if higher WMH (exposure) causes lower cortical thickness (outcome) by using 2 sets of genetic association summary statistics. For exposures, we used summary statistics from the GWAS of WMH by Sargurupermraj et al.[6] and from the GWAS conducted in the UK Biobank participants included in this study (UKBB). For outcomes, we used summary statistics for each insular thickness and mean cortical thickness that were obtained from the meta-GWAS of all cohorts in the present study and from the GWAS of the UKBB participants included in the present study. For the MR, three sets of exposure-outcome summary statistics were considered: Sargurupremraj-UKBB, Sargurupremraj-all cohorts, and UKBB-UKBB. The causal effects were evaluated utilising 2-sample MR methods implemented in the 'TwoSampleMR' package[71,72].

PRS of WMH We constructed PRS of WMH utilising the pruning and p-value thresholding (P + T) shrinkage strategy implemented in PRSice-2 (v2.3.5)[73]. The base-data for WMH were from the GWAS by Sargurupremraj, et al.[6], and target genotype and phenotype data were from the UKBB participants in this study. Sets of independent SNP were obtained by pruning based on the LD structure of the target genotype. To avoid bias due to sample overlap (i.e., some UKBB participants were included in both datasets), the degree of sample overlap was estimated and used to adjust the base-data summary statistics with EraSOR[74]. PRS was computed based on SNPs with p-values below 11 thresholds (5e-08, 1e-07, 1e-06, 1e-05, 1e-04, 1e-03, 0.05, 0.01, 0.5, 0.1, and 1). To enhance the statistical power in detecting PRS of WMH vs. cortical thickness associations, we used the first principal component (PRS_{PCA}) of the resulting set of PRS to test these associations as proposed in ref. 75. Associations were

evaluated by regressing PRS or PRS_{PCA} on thickness values adjusted for the basic-model covariates, MRI site, and the first 10 genetic principal components in R (v4.4.0).

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 meta-GWAS generated in this study have been deposited in the following database: https://figshare.com/articles/dataset/White_matter_hyperintensities_and_cortical_atrophy_genetic_risk_factors_and_underlying_neurobiology/27038320. Allen Human Brain Atlas gene expression, as parcellated in the Desikan-Killiany atlas is shared here: https://figshare.com/articles/dataset/Cell-specific_gene-expression_profiles_and_cortical_thickness_in_the_human_brain/4752955. Allen Human multiple cortical snRNAseq data was downloaded from here: https://portal.brain-map.org/atlases-and-data/rnaseq/human-multiple-cortical-areas-smart-seq. GWAS for plasma protein levels were downloaded from: https://www.synapse.org/Synapse:syn51364943/wiki/. Source Data are provided as a source data file. Source data are provided in this paper.

References

1. Wardlaw, J. M., Smith, C. & Dichgans, M. Small vessel disease: mechanisms and clinical implications. *Lancet Neurol.* **18**, 684–696 (2019).
2. Ter Telgte, A. et al. Cerebral small vessel disease: from a focal to a global perspective. *Nat. Rev. Neurol.* **14**, 387–398 (2018).
3. Wardlaw, J. M., Valdés Hernández, M. C. & Muñoz-Maniega, S. What are White Matter Hyperintensities Made of? *J. Am. Heart Assoc.* **4**, e001140 (2015).
4. Debette, S., Schilling, S., Duperron, M.-G., Larsson, S. C. & Markus, H. S. Clinical significance of magnetic resonance imaging markers of vascular brain injury: a systematic review and meta-analysis. *JAMA Neurol.* **76**, 81–94 (2019).
5. Sargurupremraj, M. et al. Genetic complexities of cerebral small vessel disease, blood pressure, and dementia. *JAMA Netw. Open* **7**, e2412824 (2024).
6. Sargurupremraj, M. et al. Cerebral small vessel disease genomics and its implications across the lifespan. *Nat. Commun.* **11**, 6285 (2020).
7. Debette, S. & Markus, H. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *Bmj* **341**, https://doi.org/10.1136/bmj.c3666 (2010).
8. The SPRINT MIND Investigators for the SPRINT Research Group. Association ofintensive vs standard blood pressure control with cerebral white matter lesions. *JAMA* **322**, 524–534 (2019).
9. Murray, A. D. et al. Brain white matter hyperintensities: Relative importance of vascular risk factors in nondemented elderly people. *Radiology* **237**, 251–257 (2005).
10. Debette, S. et al. Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. *Neurology* **77**, 461–468 (2011).
11. Staals, J., Makin, S. D. J., Doubal, F. N., Dennis, M. S. & Wardlaw, J. M. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. *Neurology* **83**, 1228–1234 (2014).
12. Ito, K. & Enomoto, H. Retrograde transport of neurotrophic factor signaling: implications in neuronal development and pathogenesis. *J. Biochem.* **160**, 77–85 (2016).
13. Yamashita, N., Yamane, M., Suto, F. & Goshima, Y. TrkA mediates retrograde semaphorin 3A signaling through plexin A4 to regulate dendritic branching. *J. Cell Sci.* **129**, 1802–1814 (2016).
14. Lambert, C. et al. Characterising the grey matter correlates of leukoaraiosis in cerebral small vessel disease. *NeuroImage Clin.* **9**, 194–205 (2015).
15. Lambert, C. et al. Longitudinal patterns of leukoaraiosis and brain atrophy in symptomatic small vessel disease. *Brain* **139**, 1136–1151 (2016).
16. Tuladhar, A. M. et al. Relationship between white matter hyperintensities, cortical thickness, and cognition. *Stroke* **46**, 425–432 (2015).
17. Dickie, D. A. et al. Progression of white matter disease and cortical thinning are not related in older community-dwelling subjects. *Stroke* **47**, 410–416 (2016).
18. Godin, O. et al. Association of white-matter lesions with brain atrophy markers: the three-city Dijon MRI study. *Cerebrovasc. Dis.* **28**, 177–184 (2009).
19. Rossi, R. et al. Topographic correspondence between white matter hyperintensities and brain atrophy. *J. Neurol.* **253**, 919–927 (2006).
20. Raji, C. A. et al. White matter lesions and brain gray matter volume in cognitively normal elders. *Neurobiol. Aging* **33**, 834.e7–16 (2012).
21. Knopman, D. S. et al. Vascular imaging abnormalities and cognition. *Stroke* **46**, 433–440 (2015).
22. Rizvi, B. et al. The effect of white matter hyperintensities on cognition is mediated by cortical atrophy. *Neurobiol. Aging* **64**, 25–32 (2018).
23. Dickie, D. A. et al. Cortical thickness, white matter hyperintensities, and cognition after stroke. *Int J. Stroke* **15**, 46–54 (2020).
24. Benarroch, E. E. Insular cortex: functional complexity and clinical correlations. *Neurology* **93**, 932–938 (2019).
25. Sepulcre, J., Sabuncu, M. R., Yeo, T. B., Liu, H. & Johnson, K. A. Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. *J. Neurosci.* **32**, 10649–10661 (2012).
26. Gogolla, N. The insular cortex. *Curr. Biol.* **27**, R580–R586 (2017).
27. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
28. Rovelet-Lecrux, A. et al. APP locus duplication in a Finnish family with dementia and intracerebral haemorrhage. *J. Neurol. Neurosurg. Psychiatry* **78**, 1158 (2007).
29. Kasuga, K. et al. Identification of independent APP locus duplication in Japanese patients with early-onset Alzheimer disease. *J. Neurol. Neurosurg. Psychiatry* **80**, 1050–1052 (2009).
30. TCW, J. & Goate, A. M. Genetics of β-Amyloid Precursor Protein in Alzheimer's Disease. *Cold Spring Harb. Perspect. Med.* **7**, a024539 (2017).
31. van der Kant, R. & Goldstein, L. S. B. Cellular functions of the amyloid precursor protein from development to dementia. *Dev. Cell* **32**, 502–515 (2015).
32. Casaletto, K. B. et al. Neurogranin, a synaptic protein, is associated with memory independent of Alzheimer biomarkers. *Neurology* **89**, 1782–1788 (2017).
33. Yang, J., Korley, F. K., Dai, M. & Everett, A. D. Serum neurogranin measurement as a biomarker of acute traumatic brain injury. *Clin. Biochem.* **48**, 843–848 (2015).
34. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
35. Zhang, M. J. et al. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat. Genet.* **54**, 1572–1580 (2022).
36. Shin, J., Patel, Y., Parker, N., Paus, T. & Pausova, Z. Prediabetic HbA1c and Cortical Atrophy: Underlying Neurobiology. *Diabetes Care* **46**, 2267–2272 (2023).
37. Peng, H. et al. Morphological diversity of single neurons in molecularly defined cell types. *Nature* **598**, 174–181 (2021).

38. Harris, K. D. & Shepherd, G. M. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18**, 170–181 (2015).
39. de Klein, N. et al. Brain expression quantitative trait locus and network analyses reveal downstream effects and putative drivers for brain-related diseases. *Nat. Genet.* **55**, 377–388 (2023).
40. Patel, Y. et al. Virtual histology of cortical thickness and shared neurobiology in 6 psychiatric disorders. *JAMA Psychiatry* **78**, 47–63 (2021).
41. Guedes-Dias, P. & Holzbaur, E. L. Axonal transport: Driving synaptic function. *Science* **366**, eaaw9997 (2019).
42. Coleman, M. Axon degeneration mechanisms: commonality amid diversity. *Nat. Rev. Neurosci.* **6**, 889–898 (2005).
43. Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**, 508–518 (2023).
44. Psaty, B. M. et al. Cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ. Cardiovasc. Genet.* **2**, 73–80 (2009).
45. Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
46. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).
47. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).
48. Winkler, T. W. et al. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–1212 (2014).
49. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
50. Han, B. & Eskin, E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am. J. Hum. Genet.* **88**, 586–598 (2011).
51. Surendran, P. et al. Discovery of rare variants associated with blood pressure regulation through meta-analysis of 1.3 million individuals. *Nat. Genet.* **52**, 1314–1332 (2020).
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. Pulit, S. L. et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–174 (2019).
54. Loh, P.-R., Kichaev, G., Gazal, S., Schoech, A. P. & Price, A. L. Mixed-model association for biobank-scale datasets. *Nat. Genet.* **50**, 906–908 (2018).
55. 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).
56. Trubetskoy, V. et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
57. Wray, N. R. et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
58. Mullins, N. et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).
59. 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).
60. Savage, J. E. et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
61. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
62. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
63. Parker, N. et al. Assessment of neurobiological mechanisms of cortical thinning during childhood and adolescence and their implications for psychiatric disorders. *JAMA Psychiatry* **77**, 1127–1136 (2020).
64. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *J. Integr. Biol.* **16**, 284–287 (2012).
65. Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
66. French, L. & Paus, T. A FreeSurfer view of the cortical transcriptome generated from the Allen Human Brain Atlas. *Front. Neurosci.* **9**, 323 (2015).
67. Shin, J. et al. Cell-specific gene-expression profiles and cortical thickness in the human brain. *Cereb. Cortex* **28**, 3267–3277 (2018).
68. Timshel, P. N., Thompson, J. J. & Pers, T. H. Genetic mapping of etiologic brain cell types for obesity. *Elife* **9**, e55851 (2020).
69. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).
70. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).
71. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, e34408 (2018).
72. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081 (2017).
73. Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience* **8**, giz082 (2019).
74. Choi, S. W., Mak, T. S. H., Hoggart, C. J. & O'Reilly, P. F. EraSOR: a software tool to eliminate inflation caused by sample overlap in polygenic score analyses. *Gigascience* **12**, giad043 (2022).
75. Coombes, B. J., Ploner, A., Bergen, S. E. & Biernacka, J. M. A principal component approach to improve association testing with polygenic risk scores. *Genet. Epidemiol.* **44**, 676–686 (2020).
76. Mowinckel, A. M. & Vidal-Piñeiro, D. Visualization of brain statistics with R packages ggseg and ggseg3d. *Adv. Methods Pract. Psychol. Sci.* **3**, 466–483 (2020).
77. Hansson, O. Biomarkers for neurodegenerative diseases. *Nat. Med.* **27**, 954–963 (2021).

Acknowledgements

We want to acknowledge the participants and investigators of all cohorts included in the present study. This study was funded by the National Institutes of Health (R01AG056726, P30 AG066546, NS017950, and AG059421). The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (75N92022D00001, 75N92022D00002, 75N92022D00003, 5N92022D00004, 75N92022D00005). The ARIC Neurocognitive Study is supported by U01HL096812, U01HL096814, U01HL096899, U01HL096902, and U01HL096917 from the National Institutes of Health (NHLBI, NINDS, NIA and NIDCD). The authors thank the staff and participants of the ARIC study for their important contributions. The MEMENTO cohort is sponsored by Bordeaux University Hospital (coordination: CIC1401-EC, Bordeaux) and was funded through research grants from the Foundation Plan Alzheimer (Alzheimer Plan 2008–2012), the French Ministry of Research and Higher Education (Plan Maladies Neurodégénératives (2016–2019)). Computations were performed on the Bordeaux Bioinformatics Centre computer resources, University of Bordeaux. Funding support for additional computer resources at the CREDIM, University of Bordeaux, has been provided to Dr Debette by the Fondation Claude Pompidou. The Three-City (3 C)

Study is conducted under a partnership agreement among the Institut National de la Santé et de la Recherche Médicale (INSERM), the University of Bordeaux, and Sanofi-Aventis. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3 C Study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, Mutuelle Générale de l'Education Nationale (MGEN), Institut de la Longévité, Conseils Régionaux of Aquitaine and Bourgogne, Fondation de France, and Ministry of Research–INSERM Programme "Cohortes et collections de données biologiques." This work is supported by France 2030 with a grant overseen by the French National Research Agency (ANR) ANR-18-RHUS-0002 and the IHU Precision and Global Vascular Brain Health Institute (ANR-23-IAHU-0001), and the BRIDGET JPND project, with funding from the EU Horizon 2020 research and innovation programme under grant agreements 643417, 640643, 667375, and 754517. It also received support from a grant (EADB) from the EU Joint Programme - Neurodegenerative Disease Research, the National Foundation for Alzheimer's Disease and Related Disorders, the Institut Pasteur de Lille, the Labex DISTALZ, and the Centre National de Recherche en Génomique Humaine. LIFE-Adult is funded by the Leipzig Research Centre for Civilisation Diseases (LIFE). LIFE is an organisational unit affiliated to the Medical Faculty of the University of Leipzig. LIFE is funded by means of the European Union, by the European Regional Development Fund (ERDF) and by funds of the Free State of Saxony within the framework of the excellence initiative. The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie Inc., AstraZeneca UK Ltd, Biogen MA Inc., Bristol Myers Squibb (and Celgene Corporation & Celgene International II Sàrl), Genentech Inc., Merck Sharp & Dohme LCC, Pfizer Inc., GlaxoSmithKline Intellectual Property Development Ltd., Sanofi US Services Inc., Maze Therapeutics Inc., Janssen Biotech Inc, Novartis AG, and Boehringer Ingelheim International GmbH. Following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.ppshp.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki.fi/en), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaalle/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta), Terveystalo Biobank (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/) and Arctic Biobank (https://www.oulu.fi/en/university/faculties-and-units/faculty-medicine/northern-finland-birth-cohorts-and-arctic-biobank). All Finnish Biobanks are members of BBMRI.fi infrastructure (https://www.bbmri-eric.eu/national-nodes/finland/). Finnish Biobank Cooperative -FINBB (https://finbb.fi/) is the coordinator of BBMRI-ERIC operations in Finland. The Finnish biobank data can be accessed through the Fingenious® services (https://site.fingenious.fi/en/) managed by FINBB. M.F. is supported in part by grants U01AG058589 and U01AG070112. AM is supported by a generic grant from the Fondation Vaincre Alzheimer, and ANR-23-CE12-0029-01 and OPE-2023-0112 grants overseen by the French National Research Agency (ANR). E.S. is supported by the Research Council of Finland (338229). K.H. is supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (SYMPATH, grant # 01ZX1906B).

Author contributions

Design and conceptualisation: Z.P., T.P., S.D., S.S., Y.P. and J.S. Formal data analysis: Y.P., J.S. and E.S. (PGS analysis). Cohort level data curation and analysis: A.T., A.M., R.X, E.H., H.S.R.R., R.W., F.B., K.H., M.R., J.Y., H.V., R.B., U.V., S.F., K.W., S.V.A., T.H.M., V.B., J.C.L., G.C., C.D., C.T., J.F.M., R.F.G., M.F., R.S., Q.Y., V.W., M.S., M.K., G.V.R., M.A.I., H.J.G., S.S., S.D., T.P. and Z.P. Writing: Y.P., J.S., Z.P. and T.P. Review and editing: A.T., A.M., R.X, E.H., H.S.R.R., R.W., F.B., K.H., M.R., J.Y., H.V., R.B., U.V., S.F., K.W., S.V.A., T.H.M., V.B., J.C.L., G.C., C.D., C.T., J.F.M., R.F.G., M.F., R.S., Q.Y., V.W., M.S., M.K., G.V.R., M.A.I., H.J.G., S.S., S.D., T.P. and Z.P.

Competing interests

H.J.G. has received travel grants and speaker honoraria from Neuraxpharm, Servier, Indorsia and Janssen Cilag. All the other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-53689-1.

Correspondence and requests for materials should be addressed to Tomas Paus or Zdenka Pausova.

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

Yash Patel[1,2,32], Jean Shin[1,2,32], Eeva Sliz[3], Ariana Tang[1,2], Aniket Mishra[4], Rui Xia[5], Edith Hofer[6,7], Hema Sekhar Reddy Rajula[4], Ruiqi Wang[8], Frauke Beyer[4,9], Katrin Horn[10], Max Riedl[10], Jing Yu[11,12], Henry Völzke[13], Robin Bülow[14], Uwe Völker[15], Stefan Frenzel[16], Katharina Wittfeld[16], Sandra Van der Auwera[16,17], Thomas H. Mosley[18], Vincent Bouteloup[4,19], Jean-Charles Lambert[20], Geneviève Chêne[4,21], Carole Dufouil[4],

Christophe Tzourio[4,21], Jean-François Mangin[22], Rebecca F. Gottesman[23], Myriam Fornage[5], Reinhold Schmidt[7], Qiong Yang[8], Veronica Witte[9], Markus Scholz[10], Markus Loeffler[10,24], Gennady V. Roshchupkin[11,12], M. Arfan Ikram[11], Hans J. Grabe[15,16], Sudha Seshadri[25], Stephanie Debette[4,26], Tomas Paus[27,28,29,30,33] ✉ & Zdenka Pausova[1,2,27,30,31,33] ✉

[1]The Hospital for Sick Children, Toronto, Ontario, Canada. [2]Departments of Physiology and Nutritional Sciences, University of Toronto, Toronto, Ontario, Canada. [3]Research Unit of Population Health, Faculty of Medicine, University of Oulu, Oulu, Finland. [4]University of Bordeaux, INSERM, Bordeaux Population Health research center, UMR1219, Bordeaux, France. [5]The Brown Foundation Institute of Molecular Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, USA. [6]Institut für Medizinische Informatik, Statistik und Dokumentation, Graz, Austria. [7]Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria. [8]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [9]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany. [10]Institute for Medical Informatics, Statistics and Epidemiology; Leipzig University, Leipzig, Germany. [11]Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands. [12]Department of Radiology and Nuclear Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands. [13]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. [14]Institute of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany. [15]Interfaculty Institute of Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany. [16]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany. [17]German Centre for Neurodegenerative Diseases (DZNE), Site Rostock/Greifswald, Greifswald, Germany. [18]The MIND Center, The University of Mississippi Medical Center, Jackson, MS, USA. [19]CHU Bordeaux, CIC 1401 EC, Pôle Santé Publique, Bordeaux, France. [20]U1167-RID-AGE facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, INSERM, CHU Lille, Institut Pasteur de Lille, University of Lille, Lille, France. [21]Department of Public Health, CHU de Bordeaux, Bordeaux, France. [22]Université Paris-Saclay, CEA, CNRS, Neurospin, Baobab, Saclay, France. [23]National Institute of Neurological Disorders and Stroke Intramural Research Program, Bethesda, Maryland, USA. [24]Leipzig Research Centre for Civilization Diseases; Leipzig University, Leipzig, Germany. [25]University of Texas, San Antonio, USA. [26]Bordeaux University Hospital, Department of Neurology, Institute for Neurodegenerative Diseases, Bordeaux, France. [27]Centre hospitalier universitaire Sainte-Justine, University of Montreal, Montreal, Canada. [28]Departments of Psychiatry and Neuroscience, Faculty of Medicine, University of Montreal, Montreal, Canada. [29]Department of Psychiatry, McGill University, Montreal, Canada. [30]ECOGENE-21, Chicoutimi, Canada. [31]Department of Pediatrics, Faculty of Medicine, University of Montreal, Montreal, Canada. [32]These authors contributed equally: Yash Patel, Jean Shin. [33]These authors jointly supervised this work: Zdenka Pausova, Tomas Paus. ✉e-mail: tomas.paus@umontreal.ca; zdenka.pausova@montreal.ca