Epigenetic and integrative cross-omics analyses of cerebral white matter hyperintensities on MRI

Epigenetic and integrative cross-omics analyses of cerebral white matter hyperintensities on MRI

Yunju Yang,[1,†] Maria J. Knol,[2,†] Ruiqi Wang,[3,†] Aniket Mishra,[4] Dan Liu,[5] Michelle Luciano,[6] Alexander Teumer,[7-9] Nicola Armstrong,[10] Joshua C. Bis,[11] Min A Jhun,[12] Shuo Li,[3] Hieab HH. Adams,[2,13] Nasir Ahmad. Aziz,[5,14] Mark E. Bastin,[15] Mathieu Bourgey,[16,17] Jennifer A. Brody,[11] Stefan Frenzel,[18] Rebecca F. Gottesman,[19] Norbert Hosten,[20] Lifang Hou,[21] Sharon LR. Kardia,[12] Valerie Lohner,[5] Pascale Marquis,[16,17] Susana Muñoz Maniega,[15] Claudia L. Satizabal,[22-24] Farzaneh A. Sorond,[25] Maria C. Valdés Hernández,[15] Cornelia M. van Duijn,[2,26] Meike W. Vernooij,[2,13] Katharina Wittfeld,[18,27] Qiong Yang,[3,23] Wei Zhao,[12] Eric Boerwinkle,[28,29] Daniel Levy,[23,30] Ian J. Deary,[6] Jiyang Jiang,[31] Karen A. Mather,[31,32] Thomas H. Mosley,[33] Bruce M. Psaty,[11,34] Perminder S. Sachdev,[31,35] Jennifer A. Smith,[12] Nona Sotoodehnia,[11] Charles S. DeCarli,[36] Monique MB. Breteler,[5,37] M. Arfan Ikram,[2] Hans J. Grabe,[18,27] Joanna Wardlaw,[15] WT. Longstreth Jr.,[34,38] Lenore J. Launer,[39] Sudha Seshadri,[22-24] Stephanie Debette,[4,24,40] Myriam Fornage[1,28]

[†]**These authors contributed equally to this work.**

Abstract

Cerebral white matter hyperintensities on MRI are markers of cerebral small vessel disease, a major risk factor for dementia and stroke. Despite the successful identification of multiple genetic variants associated with this highly heritable condition, its genetic architecture remains incompletely understood. More specifically, the role of DNA methylation has received little attention.

We investigated the association between white matter hyperintensity burden and DNA methylation in blood at approximately 450,000 CpG sites in 9,732 middle-aged to older adults from 14 community-based studies. Single-CpG and region-based association analyses were carried out. Functional annotation and integrative cross-omics analyses were performed to identify novel genes underlying the relationship between DNA methylation and white matter hyperintensities.

We identified 12 single-CpG and 46 region-based DNA methylation associations with white

matter hyperintensity burden. Our top discovery single CpG, cg24202936 (P=7.6x10^{-8}), was associated with *F2* expression in blood (P=6.4x10^{-5}), and colocalized with *FOLH1* expression in brain (posterior probability =0.75). Our top differentially methylated regions were in *PRMT1* and in *CCDC144NL-AS1*, which were also represented in single-CpG associations (cg17417856 and cg06809326, respectively). Through Mendelian randomization analyses cg06809326 was putatively associated with white matter hyperintensity burden (P=0.03) and expression of *CCDC144NL-AS1* possibly mediated this association. Differentially methylated region analysis, joint epigenetic association analysis, and multi-omics colocalization analysis consistently identified a role of DNA methylation near *SH3PXD2A*, a locus previously identified in genome-wide association studies of white matter hyperintensities. Gene set enrichment analyses revealed functions of the identified DNA methylation loci in the blood-brain barrier and in the immune response. Integrative cross-omics analysis identified 19 key regulatory genes in two networks related to extracellular matrix organization, and lipid and lipoprotein metabolism. A drug repositioning analysis indicated antihyperlipidemic agents, more specifically peroxisome proliferator-activated receptor alpha, as possible target drugs for white matter hyperintensities.

Our epigenome-wide association study and integrative cross-omics analyses implicate novel genes influencing white matter hyperintensity burden, which converged on pathways related to the immune response and to a compromised blood brain barrier possibly due to disrupted cell-cell and cell-extracellular matrix interactions. The results also suggest that antihyperlipidemic therapy may contribute to lowering risk for white matter hyperintensities possibly through protection against blood brain barrier disruption.

Author affiliations:

1. Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science at Houston, Houston, Texas, USA
2. Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands
3. Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, USA
4. University of Bordeaux, Inserm, Bordeaux Population Health Research Center, team VINTAGE, UMR 1219, Bordeaux, France

5. Population Health Sciences, German Centre for Neurodegenerative Diseases (DZNE), Bonn, Germany
6. Department of Psychology, University of Edinburgh, Edinburgh, UK
7. Institute for Community Medicine, University Medicine Greifswald, Greifswald 17475, Germany
8. German Centre for Cardiovascular Research (DZHK), Partner Site Greifswald, Greifswald, Germany
9. Department of Population Medicine and Lifestyle Diseases Prevention, Medical University of Bialystok, Bialystok, Poland
10. Mathematics and Statistics, Curtin University, Perth, Australia
11. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA
12. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA
13. Department of Radiology and Nuclear Medicine, Erasmus MC University Medical Center, Rotterdam, The Netherlands
14. Department of Neurology, Faculty of Medicine, University of Bonn, Bonn, Germany
15. Centre for Clinical Brain Sciences, Department of Neuroimaging Sciences, University of Edinburgh, Edinburgh, UK
16. Canadian Centre for Computational Genomics, McGill University, Montréal, Quebec, Canada
17. Department for Human Genetics, McGill University Genome Centre, McGill University, Montréal, Quebec, Canada.
18. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany
19. Stroke Branch, National Institutes of Neurological Disorders and Stroke, Bethesda, Maryland, USA
20. Department of Radiology and Neuroradiology, University Medicine Greifswald, Germany

21. Department of Preventive Medicine, Northwestern University, Chicago, Illinois, USA.

22. Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases and Department of Population Health Sciences, UT Health San Antonio, San Antonio, Texas, USA

23. The Framingham Heart Study, Framingham, Massachusetts, USA.

24. Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA.

25. Department of Neurology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA.

26. Nuffield Department of Population Health, Oxford University, Oxford, UK

27. German Center for Neurodegenerative Diseases (DZNE), Site Rostock/ Greifswald, Rostock, Germany

28. Human Genetics Center, School of Public Health, University of Texas Health Science at Houston, Houston, Texas, USA

29. Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA

30. Population Sciences Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA

31. Centre for Healthy Brain Ageing, School of Psychiatry, University of New South Wales, Sydney, New South Wales, Australia

32. Neuroscience Research Australia, Sydney, Australia

33. The Memory Impairment Neurodegenerative Dementia (MIND) Research Center, University of Mississippi Medical Center, Jackson, Mississippi, USA

34. Department of Epidemiology, University of Washington, Seattle, Washington, USA

35. Neuropsychiatric Institute, The Prince of Wales Hospital, University of New South Wales, Randwick, New South Wales, Australia

36. Department of Neurology and Center for Neuroscience, University of California at Davis, Sacramento, California, USA

37. Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of

Medicine, University of Bonn, Bonn, Germany

38. Department of Neurology, University of Washington, Seattle, Washington, USA

39. Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA

40. CHU de Bordeaux, Department of Neurology, Bordeaux, France

Correspondence to: Myriam Fornage, PhD

The Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston

1825 Pressler street, Suite 530F

Houston, Texas, USA 77030

E-mail: Myriam.Fornage@uth.tmc.edu

Running title: Epigenetics of white matter hyperintensities

Keywords

Epigenome-wide association study; white matter hyperintensities; cerebral small vessel disease; integrative cross-omics analysis; blood-brain barrier dysfunction

Abbreviations

BBB = blood-brain barrier; CADASIL = Cerebral Autosomal Dominant Arteriopathy with Sub-cortical Infarcts and Leukoencephalopathy; cSVD = cerebral small vessel disease; DNAm = DNA methylation; WMH = cerebral white matter hyperintensities

Introduction

Cerebral white matter hyperintensities (WMH) on MRI are indicative of cerebral small vessel disease (cSVD) and are part of the spectrum of brain vascular injury that impacts cognitive function, also known as vascular contributions to cognitive impairment and dementia

(VCID).[1,2] While the pathophysiology of WMH is little understood and likely heterogeneous, it likely has ischemic and neurodegenerative origins.[1] Historical pathology data suggested chronic ischemia resulting in demyelination and axonal loss as an underlying mechanism; however, neuroimaging data point to blood-brain barrier dysfunction, dysfunctional blood flow linked with impaired cerebrovascular autoregulation, vascular stiffness, periarteriolar inflammation, and more recently protein deposition (i.e. amyloid angiopathy).[2] Genetics plays a significant role in WMH with a heritability estimated from 54 to 80%[3–7]; however, the genetic variants identified in association studies to date explain only a fraction (~29%) of WMH variance.[8,9] Epigenetic changes such as DNA methylation (DNAm), which regulate gene expression, have emerged as another key component of the genetic architecture of complex traits.[10] Unlike DNA sequence variation, which remains unchanged throughout life, DNAm is plastic and highly sensitive to changes in the environment and aging.[10,11] To date, its role in cSVD has received little attention. We hypothesized that there may be patterns of DNAm associated with WMH that are common across all populations. We also hypothesized that the interplay between genotype, epigenotype, and risk factor exposure underlies cSVD etiology and used an integrated analytic framework to identify such relationships.

Materials and Methods

Overview

This study comprises five analytic parts to implicate novel genes and gene networks in WMH etiology (Figure 1). First, we performed an epigenome-wide association analysis in generally healthy, middle to late age adults to identify DNAm loci, both cytosine-phosphate-guanine (CpG) sites and differentially methylated regions (DMRs), associated with WMH burden in a multi-ancestry sample. The identified DNAm loci were then annotated for regulatory features, pathways, and association with other traits. Second, we investigated the contribution of genetic variation to variation in DNAm at the identified CpGs and used Mendelian randomization (MR) techniques to test for causal association with WMH burden and for the mediating role of expression of nearby genes. Third, we examined the role of DNAm at established WMH genome-wide association study (GWAS) loci. Fourth, we integrated gene expression and expression quantitative trait loci (eQTL) data to prioritize candidate genes associated with the identified CpGs. Lastly, we performed integrative cross-omics analyses to

derive WMH-associated genes networks and their key drivers and to reposition drug targets.

Study subjects

The sample included 9,732 middle-aged to older adults of European (EA) and African ancestry (AA) from 14 community-based studies. Our discovery sample includes 5,715 middle to late age subjects of European ancestry (EA, *n*=4,610) and of African ancestry (AA, *n*=1,105) from Atherosclerosis Risk in Communities (ARIC),[12] Biobanking and BioMolecular resources Research Infrastructure (BBMRI),[13] Cardiovascular Health Study (CHS),[14] Coronary Artery Risk Development in Young Adults (CARDIA),[15] Framingham Heart Study (FHS) offspring study,[16,17] Genetic Epidemiology Network of Arteriopathy (GENOA) study,[18] Lothian Birth Cohort (LBC) 1936,[19,20] Rotterdam Study (RS),[21,22] and Study of Health in Pomerania (SHIP) [23]. To replicate our findings, we accessed data on 3,398 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI),[24,25] FHS 3rd generation study,[26] the Older Australian Twin Study (OATS),[27,28] and the Rhineland study[29]. Additionally, we included a secondary replication sample (*n*=619) from the BRIDGET consortium.[30] Subjects with history of stroke or dementia were excluded. Details about participating studies and study-specific ethics statements are provided in Supplementary Data I. Each study obtained written informed consent from all participants and approval from the appropriate institutional review boards.

WMH burden measurements

Brain MRI was taken in the same or the closest subsequent visit to the visit in which DNAm was measured. In each study, MRI scans were performed and interpreted using standardized procedures without reference to demographic or clinical information. The field strength of the scanners used ranged mostly from 1.5 to 3.0 Tesla. T1-, T2-, and/or proton-density-weighted scans were obtained for all participants. The majority of the studies used fully automated segmentation method to quantify WMH burden. MRI procedures and WMH quantification in each study are detailed in Supplementary Data II.

DNAm profiling

DNAm levels were measured at ~450K CpGs from whole blood samples with the Illumina Infinium Human450 Methylation BeadChip in most participating cohorts. GENOA study measured methylation levels at ~27K CpGs with the Illumina Infinium HumanMethylation 27 BeadChip, entirely covered by the Human450 BeadChip. CARDIA,

SHIP-TREND, ADNI, and Rhineland Study used the Illumina MethylationEPIC BeadChip with a denser coverage of CpGs (~850K). Each study independently performed quality control (QC) for DNAm data, complying with the agreed minimum QC guidelines; CpGs with more than 95% of samples with a detection P<0.01 and samples with more than 95% of CpGs with a detection P <0.01 were selected. DNAm values were then standardized using an intra-array normalization method. The BRIDGET Consortium measured DNAm levels using Hi-seq bisulfate sequencing, and DNAm sites with sample coverage less than 95% were excluded. Details of DNAm data collection and processing in participating studies are presented in Supplementary Data III.

Cohort-level epigenome-wide association analyses

We tested association between DNAm level (untransformed beta values) and WMH burden (ln(WMH+1)) using a linear mixed regression model by ancestry group adjusted for age, sex, study site if applicable, total (intra)cranial volume (cm^3), white blood cell proportion (%),[31] and within-ancestry principal components (PCs) as fixed effects and technical covariates (i.e. plate, chip-position, row, and column) as random effects. In FHS, family structure is also adjusted as a random effect. Multi-ancestry studies with a small number of subjects in each ancestry, namely CHS and CARDIA, performed a pooled-ancestry analysis that also adjusted for ancestry group as fixed effects. Additionally, subgroup analyses by hypertension status were conducted. Hypertension was defined if either systolic or diastolic blood pressure (SBP or DBP) is greater than 140 mmHg or 90 mmHg, respectively, or if a subject was taking any anti-hypertensive medication at the time of MRI measurement. In the BRIDGET study, we tested the association of DNAm with an extreme-SVD phenotype defined as excessive WMH volume with or without brain infarcts accounting for age, sex, country, the sequencing read counts, and sample relatedness.[32] DNAm measurements and statistical models used in participating studies are described in Supplementary Data III.

Epigenome-wide meta-analysis and replication analysis

We combined EWAS results based on sample-size-weighted z-score-based fixed-effect method in METAL[33] because WMH was measured on different scales in the various cohorts and because our primary aim was to identify novel DNA methylation loci for WMH burden rather than estimate effect sizes of methylation probes.[34] Hypertensive and normotensive subgroup meta-analyses and ancestry-specific meta-analyses (excluding CHS and CARDIA)

were also performed. Study-specific results were corrected for inflation during meta-analysis if inflation was detected (genomic inflation factor (λ)>1.0). An association was considered as significant if *P* is smaller than Bonferroni threshold (approximately $1.2x10^{-7}$). A less stringent threshold was also set as $1.0x10^{-5}$ to detect suggestive associations. CpGs on sex chromosomes were not considered because our analytic plan did not account for hemi-methylation on the X chromosome due to chromosome X inactivation in women. Cross-reactive CpGs reported by Chen *et al.*[35] and those showing evidence of heterogeneity (Cochran Q P-value <0.05) were removed from the results *post hoc*. In the replication samples, associations for the identified CpGs were tested. CpGs were considered replicated if they were significant at the Bonferroni threshold (0.05/the number of the CpGs). We plotted epigenetic associations *in cis* (±50 kb) using R 'coMET' package.[36]

Annotation of regulatory features and traits

We scored genomic positions of the identified CpGs according to RegulomeDB's [37] ranking criteria ranging from one (likely to affect binding and linked to expression of a gene target) to five (minimal binding evidence) and also computed a probability score within a range from zero to one (the most likely to be a regulatory variant). CpGs at the locations with significant regulatory features (rank category one or two, and probability score ≥ 0.9) are discussed. We also identified enhancers or promoters mapped to CpGs using the database of genome-wide enhancer-to-gene or promoter-to-gene associations computed based on five elements: eQTLs, eRNA co-expression, transcription factor co-expression, capture Hi-C and gene target distance (GeneHancer DB).[38] Identified CpGs were also searched in EWAS catalog[39] and EWAS atlas[40] to identify associated traits reported in previous EWAS. Lastly, to examine possible correlations among the CpGs, Spearman correlations were calculated in 906 EA and 639 AA subjects from the ARIC study.

DMR analysis

We performed a DMR analysis to identify a group of CpGs that collectively influence WMH burden using two specific methods, Comb-p[41] and DMRcate[42], accounting for their spatial correlations. Briefly, Comb-p detects regional enrichment of low *P*s at varying distance using the Stouffer-Liptak-Kechris correction for adjacent *P*s.[41] DMRcate models Gaussian kernel smoothing within pre-defined distance (1Kbp in this study) and collapses contiguous significant CpGs (*P* <0.05) after multiple testing correction. DMR identified by both Comb-p

(Šidák P <0.05) and DMRCate (FDR <0.05) was considered significant. To replicate, individual association *Ps* were pooled at each identified DMRs using DMRCate in the replication samples.

Gene set enrichment analysis of WMH-associated epigenetic Loci

Identified CpGs and DMRs were tested for enrichment in gene sets from MSigDB c5 gene ontology database[43,44] and KEGG pathway database[45], using '*gsameth'* and '*gsaregion'* functions built in R 'missMethyl' package[46].

Shared epigenetics with BP

BP is an influential risk factor for WMH.[47–49] To investigate the shared epigenetics between WMH burden and BP, we performed a pairwise multivariate association test using summary statistics from a previous EWAS of SBP and DBP[50]. CpGs associated with both traits were tested against the null hypothesis H_0: $\beta_{WMH}=\beta_{SBP\ or\ DBP}=0$. The test uses Z-scores for each trait and estimates multivariate test statistics accounting for the trait correlation calculated based on the null associations (trait-specific $P>1x10^{-5}$). This method is implemented in the '*metaUSAT*' software.[51] To avoid false positive associations driven entirely by one trait, we included CpGs showing significance ($P<0.001$) for both traits. Bonferroni threshold was set at $8.33x10^{-3}$ (=0.05/6) based on the number of associations tested.

Heritability analysis and GWAS of WMH-associated CpGs

Inter-individual variation in DNAm may result from differences in environmental exposures, stochastic variation, or genetic influences. To examine the contribution of genetic variation to variation in DNAm at the identified CpGs, we estPimated the narrow sense heritability (h^2_{meth}) in the FHS Offspring Cohort subjects (n=2,377) adjusting for age, sex, blood cell counts, PCs and technical covariates. Body mass index (BMI) and smoking status were additionally corrected in sensitivity analyses.

To further identify genetic variants associated with DNAm levels at the WMH-associated CpGs, we performed GWAS in ARIC EA subjects (n=984). Genotypes were measured with Affymetrix 6.0 array and imputed from 1000 Genome phase one version three reference using MaCH v1.0.16. Variants were excluded if minor allele frequency (MAF)<0.01, sample call rate<95%, or imputation quality<0.3. The untransformed methylation beta value was tested for genetic association adjusting for age, sex, top 42 methylation PCs, and blood cell

components.

Bi-directional MR analysis of the identified CpGs and WMH burden

To determine if the WMH-associated DNAm level is a causal factor for WMH burden (Forward-MR) or a secondary outcome of WMH burden (Reverse-MR), we performed a bi-directional two-sample MR analysis[52] for the identified CpGs with at least three instrumental variables (IVs). We identified methylation quantitative trait loci (mQTL) associations in-cis (±1Mb) from the FHS study (n=4,170) that had been validated using ARIC data (n=963).[53] Those mQTLs were clumped at linkage disequilibrium (LD) r^2<0.05 for independence. For WMH, the UK biobank (UKBB) GWAS summary statistics (n=11,226)[54] was downloaded from the Cerebrovascular Disease Knowledge Portal (http://www.cerebrovascularportal.org/) on 01/09/2019. Reverse-MR analysis was performed using eight clumped genome-wide significant associations (LD r^2<0.05). Since the FHS mQTL study shares only significant associations *in cis*,[53] we used the mQTL association statistics from ARIC EA subjects for reverse-MR analysis.

For each CpG in both directions, causal association was tested based on the IVW method in the R package '*TwoSampleMR*'.[55] To validate the MR result, sensitivity analyses based on weighted median and MR Egger methods, and built-in tests for pleiotropy and heterogeneity were also performed. For existence of pleiotropy (MR egger intercept test P<0.05), the Egger regression estimate was assessed instead of the IVW estimate.

Mediating effect of *in cis* genes between CpGs and WMH burden

To investigate if expression of nearby genes mediates the relationship between the identified CpG and WMH burden, a two-step MR analysis was performed. We tested the directional relationships 1) from "*the exposure* (CpGs)" to "*the mediator* (gene expression)" (step 1) and 2) from "*the mediator* (gene expression)" to "the outcome (WMH burden) (step 2) using the identified mQTL IVs, the WMH GWAS associations,[54] and eQTL associations from the GTEx version eight brain eQTL data accessed via eQTL Catalogue (https://www.ebi.ac.uk/eqtl/) on 2020/11/12. Among available GTEx brain tissues, cortex (n=205), frontal cortex (n=175), cerebellum (n=209), cerebellar hemisphere (n=175), and caudate basal ganglia (n=194) were selected. MR association based on the IVW method was again tested and sensitivity analysis was also performed. Gene expression with IVW P<0.05 at

both steps was considered as a potential mediator in the association between the identified CpG and WMH burden.

Cis-acting genes associated with the identified CpGs in blood

To functionally annotate the identified CpGs, we tested associations with gene expression in blood in long-range [53]cis-regions (±5Mb) in 1,966 and 728 EA subjects from FHS and RS, respectively. Expression of the nearest gene/ mRNA was regressed on DNAm β score at the CpG adjusting for age, sex, population structure, and family structure (FHS only), blood cell counts and technical covariates. Technical covariates and family structure were modeled as random effects. In sensitivity analyses, smoking status and BMI were added to the model. Estimates from two studies were then combined for each gene using the sample-sized based meta-analysis method in METAL[33].

Genes colocalizing with the identified CpGs in Brain

To investigate cis-acting genes colocalized with the identified CpGs in the brain, we performed a multiple-trait colocalization (moloc) analysis using brain QTL data. Prior to this analysis, we examined the inter-individual correlations between DNAm levels in whole blood and in prefrontal cortex at the identified CpGs, using publicly available data.[56] For CpGs with significant correlation (P <0.05) between blood and prefrontal cortex, we tested the posterior probabilities for full-colocalization (PPFC) that multiple traits (DNAm, gene expression, and WMH burden) share causal variants at each locus, given the data. We used coloc priors of p1 (GWAS)= p2 (mQTL)= p3 (eQTL) = $1x10^{-5}$. We identified EA-specific GWAS associations[8] and brain mQTL (n=543) and eQTL (n=534) associations accessed via http://mostafavilab.stat.ubc.ca/xqtl/.[57] If PPFC is greater than 0.7, we considered the gene is significantly colocalized with CpG and WMH burden. Moloc analysis was performed using the R package "*moloc*".[58]

Epigenetic regulation of known GWAS loci

We next investigated the role of DNAm at established WMH GWAS loci, which may not have been detectable at the genome-wide significance threshold. Among 26 loci reported in the latest WMH GWAS,[8] we mapped 450K-array CpGs to 21 loci. EWAS associations at each of these 21 loci were pooled using the Brown's method (implemented in the package "poolr") adjusting for dependence among CpGs.[59] For dependency information, we calculated

correlation among CpGs located in the GWAS loci using ARIC methylation data (906 EA and 639 AA subjects). A GWAS locus with combined *P* was considered significant if *P* is smaller than Bonferroni-adjusted threshold (0.05/number of loci tested).

Alternatively, we performed a moloc analysis at the 21 GWAS loci, again using the GWAS and brain QTL data.[8,57] With the priors of $1x10^{-5}$, we considered genes with a PPFC greater than 70% as convincingly colocalized with DNAm and WMH burden.

Identification of biological pathways using multi-dimensional data integration

Integrating multi-omics associations for WMH may boost power to identify novel genes influencing WMH burden. We integrated genetic[8], transcriptomic[60], and epigenetic genome-wide association studies of WMH using the R package '*mergeomics'* (version 1.2). [61] To reduce noise in the GWAS data, the top 50% of genetic associations[8] were included and pruned at $r^2<0.5$ based on HapMap3 LD information as recommended.[61] For transcriptomic associations, we used the recent WMH transcriptome-wide association study (TWAS) results.[60] For epigenetic associations, we used our discovery EWAS. Markers were primarily mapped to the nearest genes. For CpGs, cis-acting genes reported in the MesaEpiGenomics study[62] were additionally annotated. For each GWAS, EWAS and TWAS, we tested marker-level enrichment with hierarchical permutation size of 20,000 based on biological pathways from pre-defined public databases: KEGG[45], REACTOME[63], Biocarta[64], and the gene ontology knowledgebase[65,66]. Then, we meta-analyzed the enriched gene sets from association studies and identified the WMH-associated gene sets (FDR-adjusted $P<0.05$).

To describe the regulatory network of the identified gene sets and identify its local hub genes, we performed a weighted key driver analysis (wKDA) using the web-based software *Mergeomics version 2.0.*[67] Gene regulatory network was constructed using in-house brain-specific Bayesian network (minimum hub overlap 0.33 and directed edge type) [68] and visualized via Cytoscape version 3.8.2.[69]

We also conducted an overlap-based drug repositioning analysis "*PharmOmics"* based on the identified key driver genes (FDR <0.05) to predict potential drugs or small molecules targeting WMH.[70] *PharmOmics* comprises a curated drug signature database covering 941 drugs, constructed from transcriptomic data across >20 tissues from rat, human, and mouse. For our analysis, we selected drug signatures from relevant tissues (*in vivo* human transcriptome

data in cardiovascular and nervous system, and *in vitro* transcriptome data from murine oligodendroglial precursor cells), and examined the overlap between these drug signature genes and key-driver genes from our identified WMH-associated gene sets.

Results

Identification of epigenetic changes associated with WMH burden

Study sample characteristics

In the discovery sample, the mean age ranges from 49.7 years in SHIP to 74.6 in CHS. Sex ratios are balanced in all studies except for GENOA study, which has 72.8% female. ARIC, CARDIA, and CHS have both EA and AA subjects, other studies consist of single ancestry subjects (AA or EA). In the primary replication sample, subjects from FHS 3rd generation and Rhineland Study (mean age 47.1 and 54.1 years, respectively), which compose 86.2% of the replication study, are younger than most discovery studies and show relatively smaller median WMH burden (0.34 in FHS 3rd generation study and 0.40 in Rhineland Study). All subjects in the replication studies are of EA. Demographic characteristics of participating cohorts are shown in Supplementary Table 1.

Novel DNAm loci are associated with WMH burden

In the discovery sample, we identified a novel epigenome-wide significant association between WMH burden and level of DNAm at cg24202936 (Z=5.38, P=7.58x10^{-8}), which mapped to *SEPTIN7P11*. Associations at cg24202936 in each study are presented in a forest plot (Supplementary Figure 1) and regional associations within 50 kb are presented with annotations (Supplementary Figure 2). At the suggestive significance threshold of 1x10^{-5}, we identified 11 additional loci (**Error! Reference source not found.**). The associations remained significant (P <0.05/12=4.17x10^{-3}) after adjusting for BMI, smoking status, and SBP and DBP. Quantile-quantile (QQ) and Miami plots are presented in Supplementary Figure 3 and 4. All subsequent analyses focus on these 12 CpGs, which are referred to as "target CpGs". None of the target CpGs associations were replicated in independent samples and a meta-analysis of the discovery and replication samples showed significant heterogeneity in many of the resulting associations, which was not present in the discovery cohorts (Supplementary Table 2). Target CpGs showed consistent associations with WMH in subgroup analyses by ancestry and

hypertension status (Supplementary Table 3 and 4). Cg06450373 in *CDH18* (P=6.48x10^{-8}) was identified in normotensive subjects (Supplementary Table 5); but not replicated. In a gene set enrichment analysis on discovered CpGs (P<1.0x10^{-5}), "cell-cell junction organization" was identified as the top pathway (P=1.32x10^{-3}, false discovery rate (FDR)=0.32).

Annotated regulatory functions of target CpGs

We found significant regulatory features from RegulomeDB at the genomic positions of cg24202936 (rank 2b and score 0.93), and cg06809326 (rank 2b and score 0.91) (Supplementary Table 6). Cg24202936 is located near a transcriptional starting site (0.2 Kb upstream), and identified as a transcriptional factor binding site computationally annotated with 20 genes (Supplementary Table 6). Previously reported EWAS traits associated with target CpGs are presented in Supplementary Table 7. In particular, cg24202936 was previously reported associated with HIV infection.[71] Cg06450373, cg031161214, cg01506471, and cg14547240 were correlated each other in both ancestries with weak to moderate r (0.23 to 0.55) (Supplementary Figure 5). In AA, cg23586595 showed weak but significant correlations with cg13476133 (r=0.32), cg03116124 (r=-0.42), and cg14547240 (r=-0.36). No correlated CpG (|r|>0.3) was identified for our top CpG, cg24202936, in both ancestries.

WMH-associated DMRs are enriched in immune response-related pathways

We identified 46 DMRs in associations with WMH burden (Supplementary Table 8). Notably, one DMR was located in *SH3PXD2A,* previously identified in genome-wide association studies (GWAS) of WMH.[8,72–74] Identified DMRs were enriched in several gene ontologies, including STAT (signal transducer and activator of transcription) family protein binding (FDR=4.91x10^{-3}) and defense response to virus (FDR=5.68x10^{-3}), which are related to the immune response (Supplementary Table 9). Of the 46 identified DMRs, 13 mapping to *PRMT1, ABAT, BHMT2, C11orf21, IZUMO1, C5orf66, ENPEP, SLC35F3, FBXO47, SLC45A4, KCTD16, KITLG,* and *UCN3* were replicated (Supplementary Table 8). Of note, *ENPEP, SLC35F3,* and *SLC45A4* were previously reported in BP GWAS.[75–81]

Shared epigenetic loci between WMH and BP

At the Bonferroni-corrected threshold (P <8.33x10^{-3}), we identified six CpGs associated with both WMH burden and BP (Supplementary Table 10). For WMH-DBP, cg23291754 mapping to *MOBKL1A* (P=2.38x10^{-7}) and cg24372586 mapping to *GNL1* (P=7.84x10^{-7}) were identified. For WMH-SBP, cg00711496 mapping to *CDC42BPB* (P=1.99x10^{-7}), cg04987734

mapping to *C19orf76;PRMT1* (P=3.09x10^{-7}), cg00934987 mapping to *SEPT4* (P=1.07x10^{-6}), and cg18770635 mapping to *KLHDC7B* (P=1.68x10^{-6}) were identified.

Heritability of the WMH-associated CpGs

Significant h^2_{meth} was estimated for cg17417856 (40.4%, P=1.37x10^{-8}), cg06809326 (26.5%, P=1.03x10^{-4}), cg23586595 (24.2%, P=1.47x10^{-3}), cg17577122 (14.3%, P=2.80x10^{-2}), and cg24202936 (15.5%, P=1.34x10^{-2}) (Table 2). Additional adjustment for BMI and smoking status did not significantly modify these estimates. In GWAS of the target CpGs in the ARIC EA sample, we observed significant *cis* genetic influence on cg06809326, cg13476133, and cg24202936 (Supplementary Figure 6). This result agrees with a previous publication that included the same dataset.[53]

Mendelian randomization analyses between target CpGs and WMH burden

Forward two-sample multiple IV MR analysis was performed for two target CpGs, cg06809326 and cg24202936, which have at least three independent cis-mQTL IVs in Huan T *et al.*[53] (Supplementary Table 11) We found a marginally significant causal relationship from cg06809326 to WMH burden (P=2.91x10^{-2}). Higher methylation level at the locus is associated with greater WMH burden (odds ratio (OR) [95% confidence interval (CI)]=1.39 [1.03, 1.87]). Evidence was lacking for horizontal pleiotropy (Egger intercept P=0.41) or heterogeneity (P=0.42) (Supplementary Table 12). In reverse-MR analysis, evidence that WMH causally influence methylation levels at any of the target CpGs was lacking (Supplementary Table 12 and 13).

Using the same three IVs, we also investigated whether cg6809326 is causally associated with expression of nearby genes (step 1). Two *cis* transcripts were annotated to this CpG in GTEx version eight data.[82] They both encode a long noncoding RNA designated as *CCDC144NL*, and *CCDC144NL-AS1* and we identified one IV for both transcripts. In all five brain tissues, we found evidence of causal association between cg06809326 and both *CCDC144NL* and *CCDC144NL-AS1* (Supplementary Table 14). In step two, a marginal association between *CCDC144NL* and WMH burden was observed in caudate basal ganglia and cortex (step one P=1.11x10^{-3} and step two P=3.94x10^{-2} in caudate basal ganglia; step one P=1.21x10^{-3} and step two P=4.28x10^{-2} in cortex).

DNAm at established GWAS loci and WMH burden

[8]We estimated the combined effect of DNAm at each locus from our EWAS results at the 21 established GWAS loci.[8] Consistent with our DMR results, CpGs at the GWAS locus *SH3PXD2A* were jointly associated with WMH (P=8.48x10^{-3}), but evidence of DNAm effects on WMH at other loci was lacking (Supplementary Table 15). We also conducted a multiple trait colocalization analysis (moloc)[58] of brain mQTL and expression QTL (eQTL)[57], and WMH-associated single nucleotide polymorphisms (SNPs). At 17 out of the 21 GWAS loci, we identified significant colocalization evidence (PPFC >0.7) (Supplementary Table 16 and Supplementary Figure 7). At eight loci, the SNPs with the highest PPFC were the sentinel SNPs in the GWAS.

Candidate genes implicated by gene expression associations with the target CpGs

At the Bonferroni threshold (6.93x10^{-5}=0.05/722 cis-genes in ±5Mb of the target CpGs), we identified significant associations between cg23586595 and *PLAC8* (P=2.98x10^{-7}), and between cg24202936 and *F2* (P=6.39 x10^{-5}) (Table 3). Adjusting for additional covariates (smoking status and BMI) did not change these associations.

Cg24202936, cg01506471, and cg06809326 showed significant correlation estimates (|r|>0.3) between blood and brain (r=0.33, 0.87, and 0.57, respectively) (Supplementary Table 17) and, thus, were tested for colocalization. We found that mQTLs for cg24202936 and WMH GWAS SNPs colocalize with *FOLH1* expression in dorsolateral prefrontal cortex (DLPFC) (PPFC=0.75) (Supplementary Table 18). Also, suggestive evidence existed for colocalization of cg06809326 mQTLs, *CCDC144NL-AS1* eQTLs, and WMH SNPs (PPFC=0.69).

Integrative cross-omics analysis

Integrative cross-omics analysis identifies novel gene regulatory networks

At FDR <0.05, we identified 576 WMH-associated gene sets enriched from the integrated data of GWAS, EWAS, and TWAS out of 12,303 gene sets from curated databases.[45,63–66] Top associated gene-sets includes “regulation of actin cytoskeleton” (P=1.14x10^{-45}, 211 genes), “telomeres, telomerase, cellular aging, and immortality” (P=1.10x10^{-35}, 18 genes), “integrin-mediated cell surface interactions” (P=3.17x10^{-34}, 84 genes), “thrombin signaling through proteinase activated receptors” (P=1.41x10^{-33}, 32 genes), and “Nef protein mediated CD4 down-regulation” (P=4.70x10^{-32}, nine genes). All enriched

pathways with FDR P <0.05 are listed in Supplementary Table 19.

We derived two WMH burden-associated gene networks in brain. The first network is comprised of four sub-networks. Five key driver genes (*FMOD, COL3A1, SERPING1, SLC13A4,* and *ISLR*) represent a sub-network of "extracellular matrix (ECM) organization, ECM-receptor interaction, focal adhesion, and collagen formation". Additionally, three related sub-networks, "smooth muscle contraction" with key driver *TAGLN;* "G-protein-coupled receptor (GPCR) ligand binding" with key drivers *GAL, ECEL1, ESR1,* and *NTS;* and "cytokine signaling in immune system" with key drivers *IFIT1* and *RTP4*, make up the network (Figure 2 and Supplementary Table 20). We also identified an independent second network associated with "lipid and lipoprotein metabolism", with key driver gene *KNG1*. Genes included in each subnetwork are presented in Supplementary Table 21.

Overlap-based drug repositioning analysis of WMH-associated genes

Using drug signatures derived from *in vivo* cardiovascular and nervous system data, we predicted antihyperlipidemic drugs, including PPAR-α (peroxisome proliferator-activated receptor-alpha) agonist "fenofibrate", as the top therapeutic target. Using drug signatures derived from murine oligodendroglial precursor cells data, we predicted several small molecules, including a glycogen synthase kinase inhibitor and a phenylalanyl tRNA synthetase inhibitor that may have therapeutic potential for Alzheimer's disease[83] and autoimmune diseases[84], respectively (Supplementary Table 22).

Discussion

This first EWAS of WMH burden in 9,732 middle-aged to older adults from 14 community-based cohorts identified several novel epigenetic loci. Although we could not independently replicate the association of single CpGs with WMH, likely due to a limited sample size and differences between the discovery and replication sample, functional annotation and bioinformatic analyses provided strong supportive evidence. Moreover, powerful DMR analyses identified 46 DMRs of which 13 were replicated. Integrative analyses of multi-omics information also suggested novel gene networks with key drivers and potential drug targets for WMH.

We identified a novel epigenetic locus, cg24204936, mapping to a pseudogene

SEPTIN7P11. Functional genomic data integration revealed two candidate genes whose expression may be influenced by variation in DNAm at this locus: *F2* in blood and *FOLH1* in DLPFC. Prothrombin encoded by *F2* plays an essential role in blood clot formation, angiogenesis, tissue repair and vascular integrity. A prothrombotic state has been reported in patients with symptomatic cSVD.[85,86] Moreover, circulating prothrombin has been associated with WMH and stroke in older Japanese with hypertension.[87,88] However, it remains unclear whether activation of coagulation plays a major role in the etiology of WMH or is secondary to injury to the cerebral small vessels and white matter.[89] *FOLH1* encodes glutamate carboxypeptidase II that catalyzes the hydrolysis of N-acetylaspartylglutamate (NAAG). An elevated level of NAAG in the cerebrospinal fluid has been reported in two patients with almost complete absence of myelin in the central nervous system[90] and has been proposed as a diagnostic biomarker for rare diseases of the white matter.[91]

An epigenetic locus mapping to *PRMT1*, which encodes a protein arginine N-methylase, was identified in single-CpG and DMR analyses and also as a shared epigenetics locus with BP. The biological link between DNAm at *PRMT1* and WMH burden may involve pathways related to endothelial dysfunction, which have previously been implicated in WMH etiology.[92] *PRMT1,* a predominant member of the PRMT family, methylates histone and non-histone proteins to regulate various cellular functions.[93] *PRMT1* is essential for the development of neurons, astrocytes, and oligodendrocytes and is critical for myelin formation.[94] PRMTs also catalyze the formation of ADMA (asymmetric dimethylarginine), which reduces nitric oxide production, promotes endothelial dysfunction in the blood-brain barrier (BBB), and triggers the immune response in atherosclerosis.[92,95,96] Higher ADMA levels have been repeatedly associated with cSVD and its monogenic form, CADASIL (Cerebral Autosomal Dominant Arteriopathy with Sub-cortical Infarcts and Leukoencephalopathy).[97–103]

Single-CpG association combined with functional genomic analyses and DMR analyses identified a novel epigenetic locus near *CCDC144NL;CCDC144NL-AS1* (coiled-coil domain containing 144 family and its antisense RNA1). Cg06809326 is under strong cis-genetic control and brain expression of its nearest gene,*CCDC144NL;CCDC144NL-AS1*, may mediate the association between DNAm and WMH burden (Supplementary Figure 8). A TWAS of WMH using blood gene expression data[60] did not report a significant association for *CCDC144NL;CCDC144NL-AS1* expression, possibly due to its low expression in blood. *CCDC144NL-AS1* encodes a long non-coding mRNA transcript that controls expression of target genes by acting as a molecular sponge for various regulatory miRNAs.[104–109] *In vitro*

studies have uncovered several of its target genes with potentially relevant function to cSVD. These include matrix metalloproteinases MMP2 and MMP9[108]; F-actin and vimentin[110]; and transforming growth factor beta (TGF-β)-activated kinase 1 (TAK1).[106] MMP2 and MMP9can damage the BBB by triggering recruitment of immune cells[111] and have been implicated in white matter injury and cSVD.[112] F-actin plays an important role in maintaining the shape of endothelial cells and the integrity of the BBB.[113] Disturbed TGF-β signaling has been implicated in the pathogenesis of several monogenic forms of cSVD.[114–117] Deficiency of TAK1 in mouse brain endothelial cells resulted in endothelial cell death, small vessel rarefaction, and disruption of the BBB.[118]

A central role of endothelial dysfunction, possibly resulting in a compromised BBB, in WMH burden[119] is further suggested by identified DNAm associations in genes involved in cell junctions. Endothelial cells are tightly anchored to each other by layers of junction proteins.[120] Claudins are integral membrane proteins that comprise tight junctions specifically in brain microvascular endothelial cells and that regulate BBB permeability.[121] Claudin-5 mapped to cg17577122 is the most enriched tight junction protein in the BBB, and its dysfunction has been implicated in neurodegenerative and neuroinflammatory diseases, and cSVD.[122–126] A recent DMR analysis using DLPFC DNAm levels also identified *CLDN5* to be associated with cognitive decline.[127] In normotensive subjects, we identified a CpG mapping to cadherin 18 (*CDH18*) that encodes an adherens junction protein, which mediates calcium-dependent cell-cell adhesion. *CDH18* is also involved in cell junction organization process and in cell signaling pathways including G-proteins signaling together with *F2.*[128]

Our DMR analysis, aggregated epigenetic associations using Brown's method, [59] and *moloc* analysis using brain QTL data consistently identified an epigenetic association at a known WMH GWAS locus, *SH3PXD2A* (SH3 and PX-domain-containing protein 2A).[8,72–74] Several genome-wide associations with WMH-related traits have also been reported at *SH3PXD2A,* including white matter microstructure[129] and stroke[130,131]. *SH3PXD2A* encodes an adaptor protein (TKS5) involved in the formation of podosomes that act as sites of close contact to as well as degradation of ECM.[132] Gene set enrichment of identified DNAm loci and integrative cross-omics analyses collectively point to a central role of the ECM in WMH burden. One of the two WMH burden networks identified through the Mergeomics approach centered around key driver genes involved in ECM organization and function and the top associated module was "regulation of actin cytoskeleton". Notably, actin polymerization and disassembly of junctional proteins within microvascular endothelial cells were shown to play a

key role in early BBB disruption in a murine model.[133]

Another network includes genes that function in lipid and lipoprotein metabolism and our overlap-based drug repositioning analyses suggested antihyperlipidemic drugs as potential drug targets. A recent MR analysis showed that genetically increased high-density lipoprotein cholesterol (HDL-C) level was associated with lower WMH volume and lower risk of small vessel stroke.[134] Statin therapy for cSVD has also been regarded as promising since individuals with high WMH burden typically carry higher vascular risk factors. Few randomized clinical trials assessing the effect of lipid lowering on WMH progression have been conducted and they have generally provided mixed results.[135–137] While they suggest a possible role of statins, in particular rosuvastatin, in preventing WMH progression, the lack of high-quality data prevents strong evidence-based recommendation at this time.[138] It has been postulated that statins improve endothelial function and stabilize the BBB in cSVD.[139,140] Studies that investigated membrane proteins including phospholipid flippase (ATP11B) and aquaporin-4 showed that the loss of these proteins cause pathological features of cSVD including endothelial cell dysfunction with reduced tight junctions, nitric oxide, oligodendrocyte progenitor cell maturation block and microglial activation.[126,141]

Finally, this study provides further emphasis concerning the long-observed perivascular inflammation as an additional crucial player in cSVD pathology and provides a possible explanation. Interestingly, gene set enrichment analyses identified a possible role of the defense response to viral infection with several DMR-associated genes related to interferon gamma signaling and the innate immune response (*DTX3L-PARP9, BNIP3,* and *IFITM1)*. Our top associated CpG has been previously reported in an EWAS of chronic HIV infection[71] and our drug-repositioning analysis also identified a HIV antiviral as a possible drug target. Several studies have reported that people with HIV are at higher risk of an increased burden of WMH compared to uninfected controls.[142,143]

Several limitations of our study must be acknowledged. First, many of our EWAS discoveries were not independently confirmed. Since a series of functional analyses showed biological relevance, we suspect that the lack of replication may stem from the limited size of the replication sample and from differences between the discovery and replication samples as hinted by the increased heterogeneity in the DNAm association observed in the meta-analysis (Supplementary Table 2). Indeed, variation in WMH burden was smaller in the replication studies than in the discovery studies perhaps due to the younger age of the participants. The

younger cohorts, CARDIA (*n*=277) with a mean age 53.9 years and SHIP (*n*=214) with a mean age 49.7 years, make up only 8.59% of the discovery sample; whereas the Rhineland Study with a mean age 54.1 years and FHS 3rd generation cohort with a mean age of 47.1 years, make up over 86% of the replication sample. Replication of several WMH-associated loci identified through more powerful DMR analyses further underscore an underpowered replication study for single CpG associations. Additional studies are needed to confirm the findings presented here. Second, we conducted a subgroup analysis stratified by hypertension status, but statistical power in each stratum was limited. A more ideal design to study this and other modifiable risk factors of cSVD will be a longitudinal study or a stratified association study with a larger sample size. Similarly, our study was not sufficiently powered to examine ancestry-specific associations of DNAm with WMH and possible ancestry difference in epigenetic patterns could not be investigated. Third, we did not adjust for additional lifestyle factors or comorbidities as we chose to maximize our sample size by minimizing the number of covariates in the models. Our primary goal was to identify novel DNAm loci associated with WMH burden and we cannot exclude the possibility that the identified loci may reflect, in part, variation in those risk factors. Fourth, the currently publicly available brain QTL data are limited to cis-regions of omics markers and, thus, our *in-silico* bioinformatics analyses were restricted only to the CpGs with substantial cis-acting genetic influence. For example, cg17417856 in *PRMT1* had a strong heritability estimate (h^2=0.40, *P*=1.37x10^{-8}) but was not followed-up because it was under polygenic control. Lastly, the study was conducted in blood and cell type-specific associations, most notably in brain, may have been missed. To extrapolate the findings in blood to brain, we assessed the correlation with DNAm in brain, and utilized available brain QTL data. Due to the difficulties of getting both brain DNAm and MRI data from a large population-based sample, an EWAS of WMH burden using brain DNAm may not be easy to achieve. However, findings from this large blood-based study may provide a basis for an epigenetic candidate gene study in the brain.

Data availability

The data that support the findings of this study are included in this manuscript. Full EWAS summary statistics are available in dbGaP at phs000930.v9.p1.

Acknowledgements

The authors thank the staff and participants of ARIC study, BBMRI, CARDIA, CHS, FHS, GENOA, LBC1936, RS, SHIP, ADNI, Rhineland Study, OATS, and BRIDGET for their pivotal contributions. We acknowledge collaborative contributions from the neuroCHARGE working group. We also acknowledge Dr. Dan Levy's team (Drs. Roby Joehanes and Tianxiao Huan) for the mQTL and eQTL data used in this study. A full set of study-specific acknowledgments is provided in the Supplementary Data I**.**

Funding

We thank all study participants for contributing to this work. This project was largely supported by grant R01NS087541 from US National Institute of Neurological Disorders and Stroke with additional support from U01AG052409 and R01HL105756. A full set of study-specific funding sources is provided in the Supplementary Data I**.**

Competing interests

Dr. DeCarli is a consultant of Novartis pharmaceuticals on a safety study for heart failure. Dr. Wardlaw is supported by the UK Dementia Research Institute which is funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK. Dr. Psaty serves on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. Dr. Grabe has received travel grants and speaker's honoraria from Fresenius Medical Care, Neuraxpharm, Servier and Janssen Cilag as well as research funding from Fresenius Medical Care.

Supplementary Materials

Supplementary material is available at Brain online.

Supplementary Data I-IV

Supplementary Figures 1-8

Supplementary Tables 1-22

Supplementary References 1-43

References

1. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol*. 2010;9(7):689-701.

2. Wardlaw JM, Benveniste H, Williams A. Cerebral Vascular Dysfunctions Detected in Human Small Vessel Disease and Implications for Preclinical Studies. *Annu Rev Physiol*. 2022;84(1):409-434.

3. Carmelli D, DeCarli C, Swan GE, *et al*. Evidence for genetic variance in white matter hyperintensity volume in normal elderly male twins. *Stroke*. 1998;29(6):1177-1181.

4. Atwood LD, Wolf PA, Heard-Costa NL, *et al*. Genetic variation in white matter hyperintensity volume in the Framingham Study. *Stroke*. 2004;35(7):1609-1613.

5. Kochunov P, Glahn D, Winkler A, *et al.* Analysis of genetic variability and whole genome linkage of whole-brain, subcortical, and ependymal hyperintense white matter volume. *Stroke*. 2009;40(12):3685-3690.

6. Turner ST, Jack CR, Fornage M, Mosley TH, Boerwinkle E, de Andrade M. Heritability of leukoaraiosis in hypertensive sibships. *Hypertension*. 2004;43(2):483-487.

7. Duperron MG, Tzourio C, Sargurupremraj M, *et al*. Burden of Dilated Perivascular Spaces, an Emerging Marker of Cerebral Small Vessel Disease, Is Highly Heritable. *Stroke*. 2018;49(2):282-287.

8. Sargurupremraj M, Suzuki H, Jian X, *et al*. Cerebral small vessel disease genomics and its implications across the lifespan. *Nat Commun.* 2020;11(1):6285.

9. Jian X, Satizabal CL, Smith A v, *et al*. Exome Chip Analysis Identifies Low-Frequency and Rare Variants in MRPL38 for White Matter Hyperintensities on Brain Magnetic Resonance Imaging. *Stroke.* 2018;49(8):1812-1819.

10. Handel AE, Ebers GC, Ramagopalan S v. Epigenetics: molecular mechanisms and implications for disease. *Trends Mol Med.* 2010;16(1):7-16.

11. Jaenisch R, Bird A. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nat Genet*. 2003;33(3S):245-254.

12. Wright JD, Folsom AR, Coresh J, *et al.* The ARIC (Atherosclerosis Risk In Communities) Study: JACC Focus Seminar 3/8. *J Am Coll Cardiol*. 2021;77(23):2939-2959.

13. BBMRI Stakeholder's Forum. BBMRI: A Step Closer - Stakeholder's Forum Report.; 2010. Accessed February 14, 2022. https://www.bbmri-eric.eu/wp-content/uploads/2016/07/stakeholders-forum-report-a-step-closer-a4.pdf

14. Fried LP, Borhani NO, Enright P, *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol.* 1991;1(3):263-276.

15. Friedman GD, Cutter GR, Donahue RP, *et al.* CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol.* 1988;41(11):1105-1116.

16. Dawber TR, Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation.* 1966;34(4):553-555. doi:10.1161/01.cir.34.4.553

17. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med.* 1975;4(4):518-525.

18. Daniels PR, Kardia SLR, Hanis CL*, et al.* Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med*. 2004;116(10):676-681.

19. Wardlaw JM, Bastin ME, Valdés Hernández MC, *et al.* Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. *Int J Stroke*. 2011;6(6):547-559.

20. Deary IJ, Gow AJ, Taylor MD, *et al*. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr*. 2007;7(1):1-12.

21. Ikram MA, Brusselle G, Ghanbari M*, et al*. Objectives, design and main findings until 2020 from the Rotterdam Study. *Eur J Epidemiol*. 2020;35(5):483-517.

22. Ikram MA, van der Lugt A, Niessen WJ, *et al*. The Rotterdam Scan Study: design update 2016 and main findings. *Eur J Epidemiol*. 2015;30(12):1299-1315. doi:10.1007/s10654-015-0105-7

23. Volzke H, Alte D, Schmidt CO, *et al*. Cohort profile: the study of health in Pomerania. *Int J Epidemiol.* 2011;40(2):294-307.

24. Khachaturian ZS. Perspective on the Alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimers Dement.* 2010;6(3):199-201.

25. Petersen RC, Aisen PS, Beckett LA, *et al*. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology.* 2010;74(3):201-209.

26. Splansky GL, Corey D, Yang Q, *et al*. The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol.* 2007;165(11):1328-1335.

27. Sachdev PS, Lee T, Lammel A, *et al*. Cognitive functioning in older twins: The Older Australian Twins Study. *Australas J Ageing.* 2011;30(s2):17-23.

28. Sachdev PS, Lammel A, Trollor JN, *et al.* A comprehensive neuropsychiatric study of elderly twins: the Older Australian Twins Study. *Twin Res Hum Genet*. 2009;12(6):573-582.

29. Breteler MMB, Wolf H. P2-135: The Rhineland study: a novel platform for epidemiologic research into Alzheimer disease and related disorders. *Alzheimers Dement*. 2014;10:P520.

30. University of Bordeaux. BRIDGET: BRain Imaging, cognition, Dementia and next generation GEnomics. Accessed February 14, 2022. https://bridget.u-bordeaux.fr/Project

31. Houseman EA, Accomando WP, Koestler DC, *et al*. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86.

32. Mishra A, Chauhan G, Violleau MH, *et al.* Association of variants in HTRA1 and NOTCH3 with MRI-defined extremes of cerebral small vessel disease in older subjects. *Brain.* 2019;142(4):1009-1023.

33. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26(17):2190-2191.

34. Pereira T v, Patsopoulos NA, Salanti G, Ioannidis JPA. Discovery properties of genome-wide association signals from cumulatively combined data sets. *Am J Epidemiol.* 2009;170(10):1197-1206.

35. Chen Y an, Lemire M, Choufani S, *et al*. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8(2):203-209.

36. Martin TC, Yet I, Tsai PC, Bell JT. coMET: visualisation of regional epigenome-wide association scan results and DNA co-methylation patterns. *BMC Bioinformatics*. 2015;16(1):131.

37. Boyle AP, Hong EL, Hariharan M, *et al*. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 2012;22(9):1790-1797.

38. Fishilevich S, Nudel R, Rappaport N, *et al*. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database.* 2017:bax028.

39. Battram T, Yousefi P, Crawford G, *et al*. The EWAS Catalog: A Database of Epigenome-wide Association Studies. *OSF Preprints*. [Preprint] doi:10.31219/osf.io/837wn

40. Li M, Zou D, Li Z, *et al.* EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res*. 2019;47(D1):D983-D988. doi:10.1093/nar/gky1027

41. Pedersen BS, Schwartz DA, Yang I v, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics.* 2012;28(22):2986-2988.

42. Peters TJ, Buckley MJ, Statham AL, *et al.* De novo identification of differentially methylated regions in the human genome. *Epigenetics Chromatin*. 2015;8(1):6.

43. Subramanian A, Tamayo P, Mootha VK, *et al*. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci*. 2005;102(43):15545.

44. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1(6):417-425.

45. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27-30.

46. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics.* 2016;32(2):286-288.

47. Debette S, Schilling S, Duperron MG, Larsson SC, Markus HS. Clinical Significance of Magnetic Resonance Imaging Markers of Vascular Brain Injury: A Systematic Review and Meta-analysis. *JAMA Neurol*. 2019;76(1):81-94.

48. Prins ND, Scheltens P. White matter hyperintensities, cognitive impairment and dementia: an update. *Nat Rev Neurol.* 2015;11:157.

49. Prins ND, van Dijk EJ, den Heijer T, *et al*. Cerebral small-vessel disease and decline in information processing speed, executive function and memory. *Brain*. 2005;128(9):2034-2041.

50. Richard MA, Huan T, Ligthart S, *et al.* DNA Methylation Analysis Identifies Loci for Blood Pressure Regulation. *Am J Hum Genet*. 2017;101(6):888-902.

51. Ray D, Boehnke M. Methods for meta-analysis of multiple traits using GWAS summary statistics. *Genet Epidemiol.* 2018;42(2):134-145.

52. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol.* 2013;178(7):1177-1184.

53. Huan T, Joehanes R, Song C, *et al*. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nat Commun*. 2019;10(1):4267.

54. Traylor M, Tozer DJ, Croall ID, *et al*. Genetic variation in PLEKHG1 is associated with white matter hyperintensities (n = 11,226). *Neurology*. 2019;92(8):e749-e757.

55. Hemani G, Zheng J, Elsworth B, *et al*. The MR-Base platform supports systematic causal inference across the human phenome. *Elife.* 2018;7.

56. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10(11):1024-1032.

57. Ng B, White CC, Klein HU, *et al*. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. *Nat Neurosci.* 2017;20(10):1418-1426.

58. Giambartolomei C, Zhenli Liu J, Zhang W, *et al.* A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics.* 2018;34(15):2538-2545.

59. Brown MB. 400: A method for combining non-independent, one-sided tests of significance. *Biometrics.* 1975;31:987-992.

60. Lin H, Satizabal C, Xie Z, *et al.* Whole blood gene expression and white matter Hyperintensities. *Mol Neurodegener.* 2017;12(1):67.

61. Shu L, Zhao Y, Kurt Z, *et al*. Mergeomics: multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genomics*. 2016;17(1):874.

62. Liu Y, Ding J, Reynolds LM, *et al*. Methylomics of gene expression in human monocytes. *Hum Mol Genet*. 2013;22(24):5065-5074.

63. Jassal B, Matthews L, Viteri G, *et al*. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2020;48(D1):D498-D503.

64. Nishimura D. BioCarta. *Biotech Software & Internet Report: The Computer Software Journal for Scient.* 2001;2(3):117-120.

65. Carbon S, Douglass E, Good BM, *et al*. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* 2021;49(D1):D325-D334.

66. Ashburner M, Ball CA, Blake JA, *et al.* Gene ontology: tool for the unification of biology. *Nat Genet.* 2000;25(1):25-29.

67. Arneson D, Bhattacharya A, Shu L, Makinen VP, Yang X. Mergeomics: a web server for identifying pathological pathways, networks, and key regulators via multidimensional data integration. *BMC Genomics.* 2016;17(1):722.

68. Zhu J, Lum PY, Lamb J, *et al*. An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet Genome Res.* 2004;105(2-4):363-374.

69. Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498-2504.

70. Chen YW, Diamante G, Ding J, *et al*. PharmOmics: A Species- and Tissue-specific Drug Signature Database and Online Tool for Drug Repurposing. *bioRxiv.* [Preprint] doi:10.1101/837773

71. Gross AM, Jaeger PA, Kreisberg JF, *et al*. Methylome-wide Analysis of Chronic HIV Infection Reveals Five-Year Increase in Biological Age and Epigenetic Targeting of HLA. *Mol Cell.* 2016;62(2):157-168.

72. Persyn E, Hanscombe KB, Howson JMM, Lewis CM, Traylor M, Markus HS. Genome-wide association study of MRI markers of cerebral small vessel disease in 42,310 participants. *Nat Commun*. 2020;11(1):2175.

73. Verhaaren BFJ, Debette S, Bis JC, *et al*. Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI. *Circ Cardiovasc Genet*. 2015;8(2):398-409.

74. Armstrong NJ, Mather KA, Sargurupremraj M, *et al*. Common Genetic Variation Indicates Separate Causes for Periventricular and Deep White Matter Hyperintensities. *Stroke*. 2020;51(7):2111-2121.

75. Giri A, Hellwege JN, Keaton JM, *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat Genet*. 2019;51(1):51-62. doi:10.1038/s41588-018-0303-9

76. Kichaev G, Bhatia G, Loh PR, *et al*. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet*. 2019;104(1):65-75.

77. Sakaue S, Kanai M, Tanigawa Y, *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*. 2021;53(10):1415-1424.

78. Warren HR, Evangelou E, Cabrera CP, *et al.* Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet.* 2017;49(3):403-415.

79. Takeuchi F, Akiyama M, Matoba N, *et al.* Interethnic analyses of blood pressure loci in populations of East Asian and European descent. *Nat Commun*. 2018;9(1):5052.

80. Kato N, Takeuchi F, Tabara Y, *et al.* Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat Genet.* 2011;43(6):531-538.

81. Zhu Z, Wang X, Li X, *et al*. Genetic overlap of chronic obstructive pulmonary disease and cardiovascular disease-related traits: a large-scale genome-wide cross-trait analysis. *Respir Res.* 2019;20(1):64.

82. The GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369(6509):1318-1330.

83. Lauretti E, Dincer O, Praticò D. Glycogen synthase kinase-3 signaling in Alzheimer's disease. Biochim Biophys *Acta Mol Cell Res.* 2020;1867(5):118664.

84. Kwon NH, Fox PL, Kim S. Aminoacyl-tRNA synthetases as therapeutic targets. *Nat Rev Drug Discov.* 2019;18(8):629-650.

85. Tomimoto H, Akiguchi I, Ohtani R, *et al*. The Coagulation-Fibrinolysis System in Patients With Leukoaraiosis and Binswanger Disease. *Arch Neurol.* 2001;58(10):1620-1625.

86. Wiseman S, Marlborough F, Doubal F, Webb DJ, Wardlaw J. Blood Markers of Coagulation, Fibrinolysis, Endothelial Dysfunction and Inflammation in Lacunar Stroke versus Non-Lacunar Stroke and Non-Stroke: Systematic Review and Meta-Analysis. *Cerebrovasc Dis.* 2014;37(1):64-75.

87. Nagai M, Hoshide S, Kario K. Association of Prothrombotic Status With Markers of Cerebral Small Vessel Disease in Elderly Hypertensive Patients. *Am J Hypertens.* 2012;25(10):1088-1094.

88. Kario K, Yano Y, Matsuo T, Hoshide S, Eguchi K, Shimada K. Additional impact of morning haemostatic risk factors and morning blood pressure surge on stroke risk in older Japanese hypertensive patients. *Eur Heart J.* 2011;32(5):574-580.

89. Markus HS, Hunt B, Palmer K, Enzinger C, Schmidt H, Schmidt R. Markers of Endothelial and Hemostatic Activation and Progression of Cerebral White Matter Hyperintensities. *Stroke.* 2005;36(7):1410-1414.

90. Wolf NI, Willemsen MAAP, Engelke UF, *et al.* Severe hypomyelination associated with increased levels of N-acetylaspartylglutamate in CSF. *Neurology.* 2004;62(9):1503.

91. Mochel F, Boildieu N, Barritault J, *et al.* Elevated CSF N-acetylaspartylglutamate suggests specific molecular diagnostic abnormalities in patients with white matter diseases. *Biochim Biophys Acta*. 2010;1802(11):1112-1117.

92. Wardlaw JM, Smith C, Dichgans M. Small vessel disease: mechanisms and clinical implications. *Lancet Neurol.* 2019;18(7):684-696.

93. Wei H, Mundade R, Lange KC, Lu T. Protein arginine methylation of non-histone proteins and its role in diseases. *Cell Cycle.* 2014;13(1):32-41.

94. Hashimoto M, Murata K, Ishida J, Kanou A, Kasuya Y, Fukamizu A. Severe Hypomyelination and Developmental Defects Are Caused in Mice Lacking Protein Arginine Methyltransferase 1 (PRMT1) in the Central Nervous System. *J Biol Chem*. 2016;291(5):2237-2245.

95. Dowsett L, Higgins E, Alanazi S, Alshuwayer NA, Leiper FC, Leiper J. ADMA: A Key Player in the Relationship between Vascular Dysfunction and Inflammation in Atherosclerosis. *J Clin Med.* 2020;9(9):3026.

96. Watson CP, Pazarentzos E, Fidanboylu M, Padilla B, Brown R, Thomas SA. The transporter and permeability interactions of asymmetric dimethylarginine (ADMA) and L-arginine with the human blood-brain barrier in vitro. *Brain Res*. 2016;1648(Pt A):232-242.

97. Janes F, Cifù A, Pessa ME, *et al.* ADMA as a possible marker of endothelial damage. A study in young asymptomatic patients with cerebral small vessel disease. *Sci Rep.* 2019;9(1):14207.

98. Pikula A, Böger RH, Beiser AS, *et al.* Association of plasma ADMA levels with MRI markers of vascular brain injury: Framingham offspring study. *Stroke.* 2009;40(9):2959-2964.

99. Guan J, Yan C, Gao Q, *et al.* Analysis of risk factors in patients with leukoaraiosis. *Medicine.* 2017;96(8):e6153-e6153.

100. Khan U, Hassan A, Vallance P, Markus HS. Asymmetric Dimethylarginine in Cerebral Small Vessel Disease. *Stroke.* 2007;38(2):411-413.

101. Notsu Y, Nabika T, Bokura H, *et al.* Evaluation of Asymmetric Dimethylarginine and Homocysteine in Microangiopathy-Related Cerebral Damage. *Am J Hypertens*. 2009;22(3):257-262.

102. Gao Q, Fan Y, Mu LY, Ma L, Song ZQ, Zhang YN. S100B and ADMA in cerebral small vessel disease and cognitive dysfunction. *J Neurol Sci.* 2015;354(1):27-32.

103. Rufa A, Blardi P, de Lalla A, *et al.* Plasma Levels of Asymmetric Dimethylarginine in Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarct and Leukoencephalopathy. *Cerebrovasc Dis.* 2008;26(6):636-640.

104. Guo L, Peng Y, Meng Y, *et al.* Expression profiles analysis reveals an integrated miRNA-lncRNA signature to predict survival in ovarian cancer patients with wild-type BRCA1/2. *Oncotarget.* 2017;8(40):68483-68492.

105. He J, Guan J, Liao S, *et al*. Long Noncoding RNA CCDC144NL-AS1 Promotes the Oncogenicity of Osteosarcoma by Acting as a Molecular Sponge for microRNA-490-3p and Thereby Increasing HMGA2 Expression. *Onco Targets Ther.* 2021;14:1-13.

106. Fan H, Ge Y, Ma X, *et al.* Long non-coding RNA CCDC144NL-AS1 sponges miR-143-3p and regulates MAP3K7 by acting as a competing endogenous RNA in gastric cancer. *Cell Death Dis.* 2020;11(7):521.

107. Zhang L, Chi B, Chai J, *et al*. LncRNA CCDC144NL-AS1 Serves as a Prognosis Biomarker for Non-small Cell Lung Cancer and Promotes Cellular Function by Targeting miR-490-3p. *Mol Biotechnol.* 2021;63(10):933-940.

108. Zhang Y, Zhang H, Wu S. LncRNA-CCDC144NL-AS1 Promotes the Development of Hepatocellular Carcinoma by Inducing WDR5 Expression via Sponging miR-940. *J Hepatocell Carcinoma.* 2021;8:333-348.

109. Niu P, Yao B, Wei L, Zhu H, Fang C, Zhao Y. Construction of prognostic risk prediction model based on high-throughput sequencing expression profile data in childhood acute myeloid leukemia. *Blood Cells Mol Dis.* 2019;77:43-50.

110. Zhang C, Wu W, Zhu H, *et al.* Knockdown of long noncoding RNA CCDC144NL-AS1 attenuates migration and invasion phenotypes in endometrial stromal cells from endometriosis. *Biol Reprod.* 2019;100(4):939-949.

111. Anwar MM, Özkan E, Gürsoy-Özdemir Y. The role of extracellular matrix alterations in mediating astrocyte damage and pericyte dysfunction in Alzheimer's disease: A comprehensive review. *Eur J Neurosci*. 2021:1-23.

112. Montaner J, Ramiro L, Simats A, *et al.* Matrix metalloproteinases and ADAMs in stroke. *Cell Mol Life Sci.* 2019;76(16):3117-3140.

113. Prasain N, Stevens T. The actin cytoskeleton in endothelial cell phenotypes. *Microvasc Res.* 2009;77(1):53-63.

114. Hara K, Shiga A, Fukutake T, *et al.* Association of HTRA1 Mutations and Familial Ischemic Cerebral Small-Vessel Disease. *New Engl J Med.* 2009;360(17):1729-1739.

115. Kast J, Hanecker P, Beaufort N, *et al.* Sequestration of latent TGF-β binding protein 1 into CADASIL-related Notch3-ECD deposits. *Acta Neuropathol Commun*. 2014;2:96.

116. Hamel E. Cerebral Circulation: Function and Dysfunction in Alzheimer's Disease. *J Cardiovasc Pharm.* 2015;65(4).

117. Smahi A, Courtois G, Vabres P, *et al.* Genomic rearrangement in NEMO impairs NF-κB activation and is a cause of incontinentia pigmenti. *Nature.* 2000;405(6785):466-472.

118. Ridder DA, Wenzel J, Müller K, *et al*. Brain endothelial TAK1 and NEMO safeguard the neurovascular unit. *J Exp Med.* 2015;212(10):1529-1549.

119. Wardlaw JM, Makin SJ, Valdés Hernández MC, *et al.* Blood-brain barrier failure as a core mechanism in cerebral small vessel disease and dementia: evidence from a cohort study. *Alzheimers Dement.* 2017;13(6):634-643.

120. Sukriti S, Tauseef M, Yazbeck P, Mehta D. Mechanisms regulating endothelial permeability. *Pulm Circ.* 2014;4(4):535-551.

121. Lippmann ES, Azarin SM, Kay JE, *et al*. Derivation of blood-brain barrier endothelial cells from human pluripotent stem cells. *Nat Biotechnol.* 2012;30(8):783-791.

122. Greene C, Hanley N, Campbell M. Claudin-5: gatekeeper of neurological function. *Fluids Barriers CNS.* 2019;16(1):3.

123. Yang Y, Kimura-Ohba S, Thompson JF, *et al.* Vascular tight junction disruption and angiogenesis in spontaneously hypertensive rat with neuroinflammatory white matter injury. *Neurobiol Dis.* 2018;114:95-110.

124. Berndt P, Winkler L, Cording J, *et al.* Tight junction proteins at the blood-brain barrier: far more than claudin-5. *Cell Mol Life Sci.* 2019;76(10):1987-2002.

125. Bailey EL, Wardlaw JM, Graham D, Dominiczak AF, Sudlow CLM, Smith C. Cerebral small vessel endothelial structural changes predate hypertension in stroke-prone spontaneously hypertensive rats: a blinded, controlled immunohistochemical study of 5- to 21-week-old rats. *Neuropathol Appl Neurobiol*. 2011;37(7):711-726.

126. Rikesh RM, Sophie Q, Silvie RR, *et al.* Reversal of endothelial dysfunction reduces white matter vulnerability in cerebral small vessel disease in rats. *Sci Transl Med.* 2018;10(448):eaam9507.

127. Hüls A, Robins C, Conneely KN, *et al.* Brain DNA Methylation Patterns in CLDN5 Associated With Cognitive Decline. *Biol Psychiat.* 2021;91(4):389-398.

128. Belinky F, Nativ N, Stelzer G, *et al.* PathCards: multi-source consolidation of human biological pathways. *Database.* 2015:bav006.

129. Zhao B, Zhang J, Ibrahim JG, *et al.* Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n = 17,706). *Mol Psychiatr.* 2021;26(8):3943-3955.

130. Malik R, Chauhan G, Traylor M, *et al.* Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* 2018;50(4):524-537.

131. Ishigaki K, Akiyama M, Kanai M, *et al*. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet.* 2020;52(7):669-679.

132. Murphy DA, Courtneidge SA. The "ins" and "outs" of podosomes and invadopodia: characteristics, formation and function. *Nat Rev Mol Cell Bio*. 2011;12(7):413-426.

133. Shi Y, Zhang L, Pu H, *et al.* Rapid endothelial cytoskeletal reorganization enables early blood-brain barrier disruption and long-term ischaemic reperfusion brain injury. *Nat Commun.* 2016;7:10523.

134. Georgakis MK, Malik R, Anderson CD, Parhofer KG, Hopewell JC, Dichgans M. Genetic determinants of blood lipids and cerebral small vessel disease: role of high-density lipoprotein cholesterol. *Brain.* 2020;143(2):597-610.

135. Zhang H, Cui Y, Zhao Y, *et al.* Effects of sartans and low-dose statins on cerebral white matter hyperintensities and cognitive function in older patients with hypertension: a randomized, double-blind and placebo-controlled clinical trial. *Hypertens res.* 2019;42(5):717-729.

136. Guo Y, Li Y, Liu X, *et al*. Assessing the effectiveness of statin therapy for alleviating cerebral small vessel disease progression in people ≥75 years of age. *BMC Geriatr.* 2020;20(1):292.

137. ten Dam VH, van den Heuvel DMJ, van Buchem MA, *et al.* Effect of pravastatin on cerebral infarcts and white matter lesions. *Neurology.* 2005;64(10):1807-1809.

138. Wardlaw JM, Debette S, Jokinen H, *et al.* ESO Guideline on covert cerebral small vessel disease. *Eur Stroke J.* 2021;6(2):CXI-CLXII.

139. Zhang CE, Wong SM, van de Haar HJ, *et al.* Blood–brain barrier leakage is more widespread in patients with cerebral small vessel disease. *Neurology.* 2017;88(5):426.

140. Giannopoulos S, Katsanos AH, Tsivgoulis G, Marshall RS. Statins and cerebral hemodynamics. *J Cereb Blood Flow Metab.* 2012;32(11):1973-1976.

141. Holland PR, Searcy JL, Salvadores N, *et al.* Gliovascular disruption and cognitive deficits in a mouse model with features of small vessel disease. *J Cereb Blood Flow Metab.* 2015;35(6):1005-1014.

142. Moulignier A, Savatovsky J, Assoumou L, *et al.* Silent Cerebral Small-Vessel Disease Is Twice as Prevalent in Middle-Aged Individuals With Well-Controlled, Combination Antiretroviral

Therapy–Treated Human Immunodeficiency Virus (HIV) Than in HIV-Uninfected Individuals. *Clin Infect Dis.* 2018;66(11):1762-1769.

143. Mina Y, Wu T, Hsieh HC, *et al*. Association of White Matter Hyperintensities With HIV Status and Vascular Risk Factors. *Neurology.* 2021;96(14):e1823-e1834.

Figure 1 Overview of the study analytic scheme

Figure 2 WMH-associated Gene Networks. WMH-associated genes based on multi-molecular evidence are organized around the 19 key driver genes. a. WMH-associated network consisting of four sub-networks - extracellular matrix (ECM) organization (*FMOD, COL3A1, SEPING1, SLC13A4,* and *ISLR*); smooth muscle contraction (*TAGLN*); G-protein-coupled receptor (GPCR) ligand binding (*GAL, ECEL1, ESR1,* and *NTS*) and cytokine signaling in immune system (*IFIT1* and *RTP4*) b. WMH-associated network of lipid and lipoprotein metabolism (*KNG1*). Key drivers and associated gene networks identified in the Mergeomics analysis are colored in orange (RGB: 255,166,127). Neighboring genes are grouped into networks and labelled in random colors.

Identification of Epigenetic Changes Associated with WMH Burden

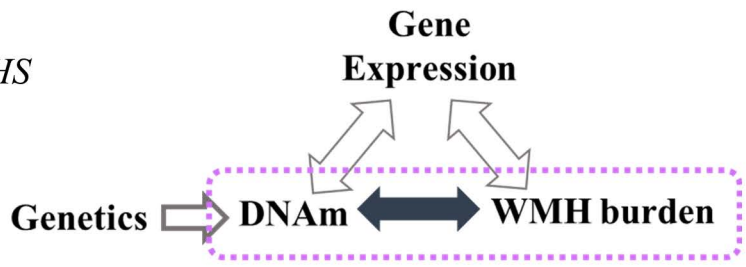
1) meta-EWAS of WMH burden (n=5,715): DNAm (β) ~ WMH burden + covariates in *ARIC, BBMRI, CARDIA, CHS, FHS offspring, GENOA, LBC1936, RS, and SHIP-TREND*

+ replication study in the younger population (n=4,017) *in ADNI, FHS 3rd generation, OATS, Rhineland study, and BRIDGET study.*

2) Differentially methylated regions associated with WMH burden regional associations accounting for correlations among neighboring CpGs

3) Functional annotations Gene set enrichment, RegulomeDB 2.0, GeneHancerDB, EWAS Atlas, and EWAS catalog

4) Mutivariate metaUSAT analysis to identify pleiotropic DNAm also associated with blood pressures

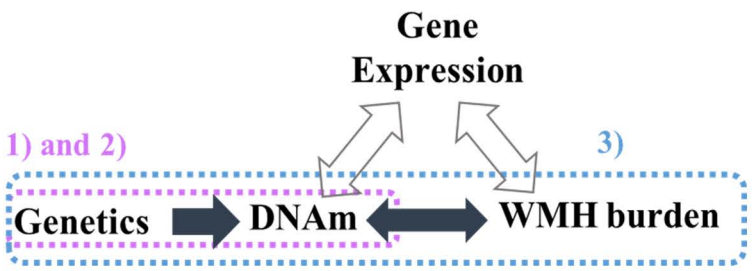


Genetic Contribution on the WMH-associated CpGs and Mendelian Randomization

1) Heritability of target CpGs in *FHS* (n=2,377)

2) Identification of mQTLs for target CpGs in *ARIC* (n=984) and *FHS* (n=4,170)

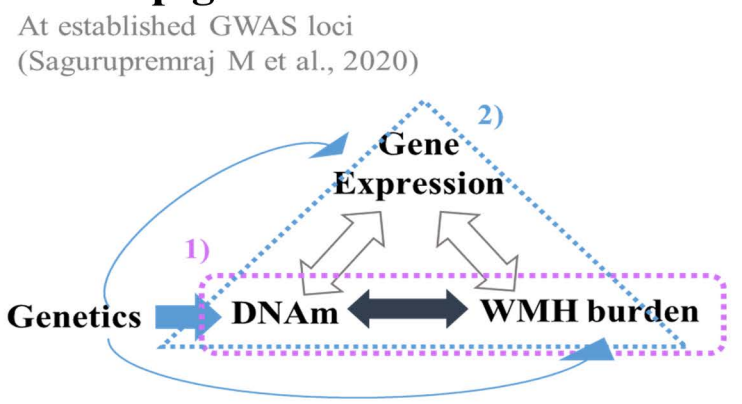
3) Mendelian randomization analysis to assess the directionality of CpG-WMH burden association, and to identify mediating effects of expression in brain of nearby genes



Evidence that Established GWAS Loci are under Epigenetic Control

1) Combined EWAS associations at GWAS loci to estimate the joint association of CpGs in established genetic loci using Brown's method

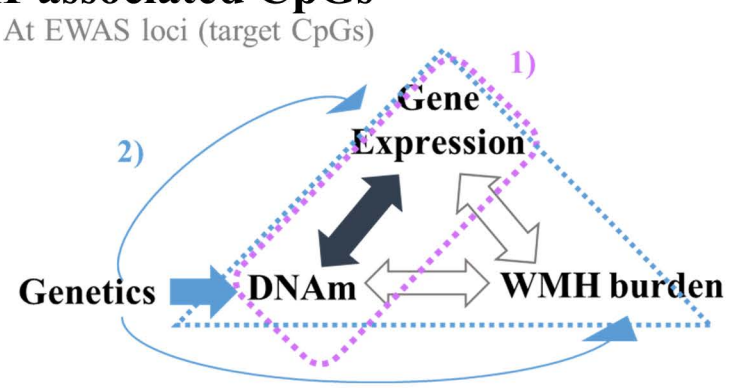
2) Moloc analysis to identify genetic variants with effects on DNAm, gene expression, and WMH burden at known GWAS loci using *ROS/MAP* DLPFC DNAm and gene expression QTL data and WMH GWAS data



Genes Functionally Linked to WMH-associated CpGs

1) Association between DNAm and gene expression in blood FHS (n=1,966) and RS (n=728)

2) Moloc analysis to identify gene expression in brain associated with target CpGs using ROS/MAP dorsolateral prefrontal cortex DNAm and gene expression QTL data and WMH GWAS data



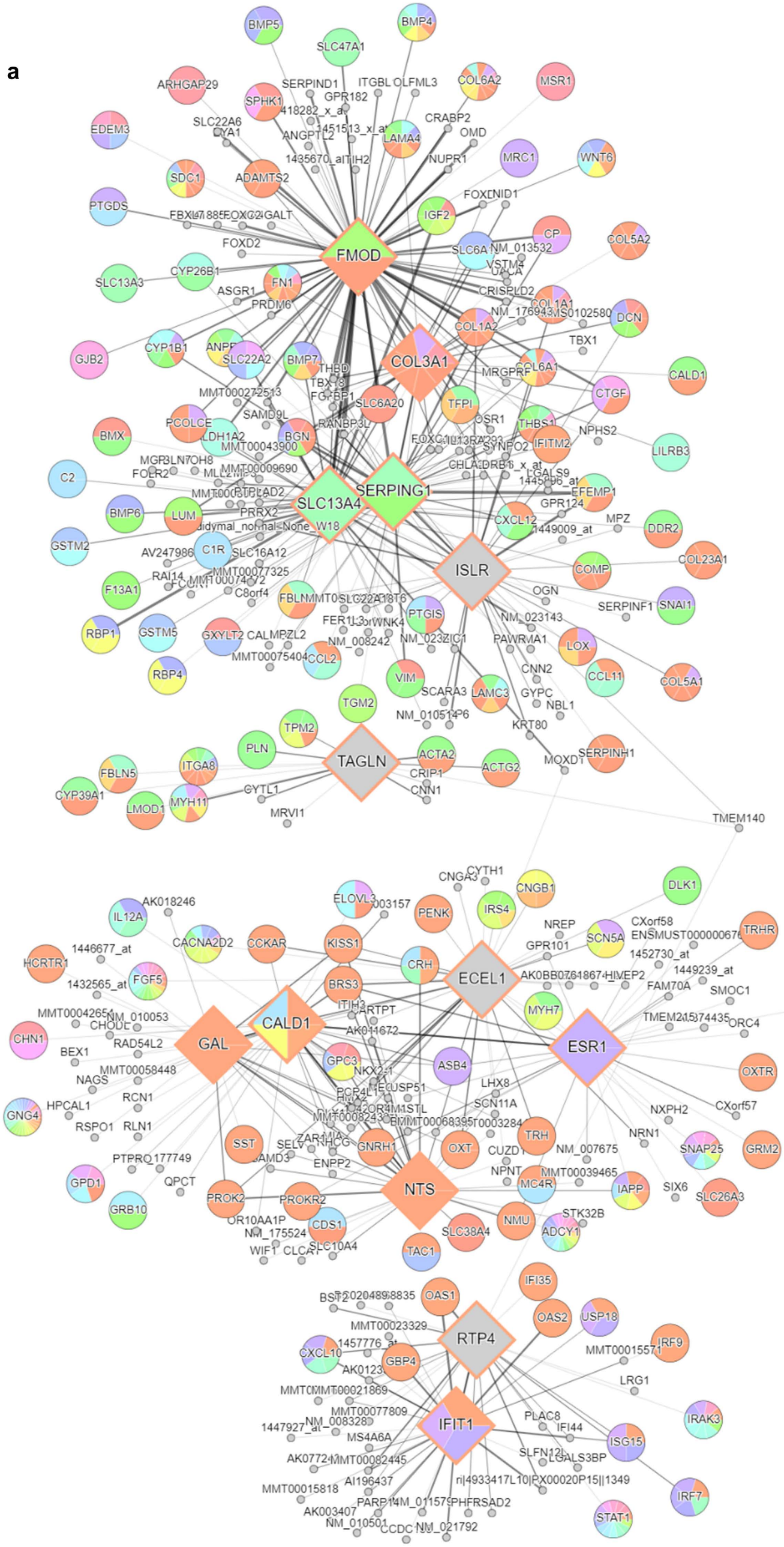
Integrative Cross-Omics Analysis

1) Mergeomics to identify gene-sets empowered by GWAS, EWAS and TWAS associations

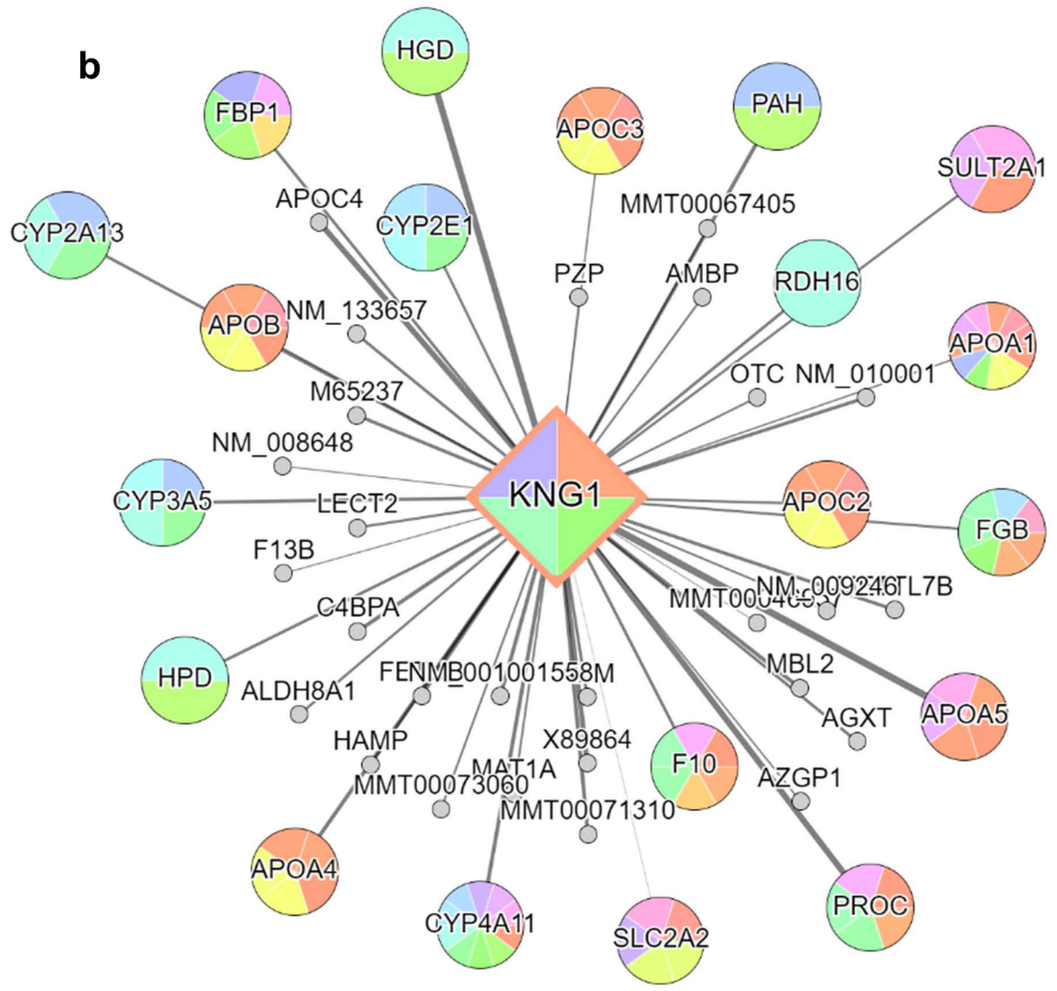
2) Weighted key driver analysis to identify WMH-associated networks and their key driver genes

3) Pharmomics to reposition drugs based on the key driver genes

a



b



Legends

Shape

Key Driver Node

Node

Color

Non module member

Module member

Multiple module member

KD border color

Top key driver module

Table 1 Single CpG associations with white matter hyperintensities burden ($P<1x10^{-5}$)

			Reduced model					Full model		
CpG	Chr:Position (hg19)	Nearest Gene	N	Z	P	Q	FDR	N	Z	P
cg24202936	11:50257256	*SEPTIN7P11*	5,359	5.38	$7.58x10^{-8}$	0.03	0.04	4,930	5.28	$1.30x10^{-7}$
cg17417856	19:50191637	*PRMT1;ADM5*	4,917	-4.95	$7.42x10^{-7}$	0.15	0.28	4,526	-4.40	$1.11x10^{-5}$
cg01506471	7:3990479	*SDK1*	5,359	-4.81	$1.52x10^{-6}$	0.21	0.3	4,930	-4.00	$6.41x10^{-5}$
cg14547240	4:15428750	*C1QTNF7*	5,359	-4.71	$2.48x10^{-6}$	0.25	0.3	4,930	-4.17	$3.10x10^{-5}$
cg21547371	3:52869521	*MUSTN1*	5,359	-4.65	$3.30x10^{-6}$	0.25	0.3	4,930	-4.06	$4.95x10^{-5}$
cg03116124	1:231293208	*TRIM67*	5,129	-4.64	$3.54x10^{-6}$	0.25	0.31	4,700	-4.58	$4.63x10^{-6}$
cg06809326	17:20799526	*CCDC144NL-AS1*	5,359	4.57	$4.80x10^{-6}$	0.28	0.34	4,930	3.44	$5.88x10^{-4}$
cg13476133	7:44185646	*GCK*	5,359	4.55	$5.46x10^{-6}$	0.28	0.36	4,930	4.03	$5.65x10^{-5}$
cg14133539	9:104568	*FOXD4*	4,917	-4.53	$5.98x10^{-6}$	0.28	0.38	4,526	-4.45	$8.41x10^{-6}$
cg17577122	22:19511967	*CLDN5*	5,359	4.50	$6.88x10^{-6}$	0.29	0.4	4,930	4.79	$1.68x10^{-6}$
cg23586595	4:84034390	*PLAC8*	5,359	4.45	$8.45x10^{-6}$	0.32	0.43	4,930	3.93	$8.36x10^{-5}$
cg23054394	3:140784675	*SPSB4*	5,359	-4.42	$9.88x10^{-6}$	0.34	0.45	4,930	-4.01	$6.07x10^{-5}$

Chr: chromosome, EA: European ancestry, FDR: local false discovery rate value, N: number of subjects tested for the CpG, P: *P* value, Q: *Q* value, SE: standard error, Z: Z-score.
The reduce model is adjusted for age, sex, study site (if applicable), total (intra)cranial volume (cm^3), white blood cell proportion (%), technical covariates and genetic principal components. The full model is additionally adjusted for body mass index, smoking status, and systolic and diastolic blood pressure measures. Associations in replication studies are based on the reduced model.

Table 2 Heritability estimates of WMH-associated CpGs

		Reduced model			Full Model		
CpG	**Nearest Gene**	h^2_{meth}	**S.E.**	***P***	h^2_{meth}	**S.E.**	***P***
cg0150647	*SDK1*	0.02	0.07	0.38	0.01	0.07	0.42
cg03116124	*TRIM67*	0.01	0.07	0.45	0.01	0.07	0.47
cg06809326	*CCDC144NL*	0.26	0.07	1.03×10^{-4}*	0.27	0.07	9.51×10^{-5}*
cg13476133	*GCK*	0.09	0.07	0.11	0.09	0.07	0.12
cg14133539	*FOXD4*	0.08	0.07	0.14	0.07	0.07	0.17
cg14547240	*C1QTNF7*	0.06	0.07	0.20	0.06	0.07	0.18
cg17417856	*PRMT1;ADM5*	0.40	0.08	1.37×10^{-8}*	0.40	0.08	3.06×10^{-8}*
cg17577122	*CLDN5*	0.14	0.08	2.80×10^{-2}	0.15	0.08	2.27×10^{-2}
cg21547371	*MUSTN1*	0.00		0.50	0.00		0.50
cg23054394	*SPSB4*	0.00		0.50	0.00		0.50
cg23586595	*PLAC8*	0.24	0.08	1.47×10^{-3}*	0.23	0.08	2.51×10^{-3}*
cg24202936	*LOC441601*	0.15	0.07	1.34×10^{-2}	0.16	0.07	1.17×10^{-2}

*indicates significant after adjustment for multiple testing burden (12 single-CpGs, $0.05/12=4.17\times10^{-3}$)

h^2_{meth}: the narrow-sense heritability in an additive genetic model. S.E.: standard error

Reduced model is adjusted for age, sex, blood cell counts, principal components of the ancestry and technical covariates. Full model is additionally adjusted for BMI and smoking.

Table 3 Cis-Genes (±5Mb) whose expression is associated with identified CpGs

CpG	Gene	Region (hg19)	n	$Z_{redueced}$	$P_{reduced}$	Z_{full}	P_{full}
cg23586595	*PLAC8*	4:84011211-84138405	2,687	-5.13	2.98×10^{-7}	-5.11	3.27×10^{-7}
cg24202936	*F2*	11:46740749-46761054	1,963	-4.00	6.39×10^{-5}	-4.01	6.04×10^{-5}

Z score and P value from the reduced and full model are presented. Reduced model is adjusted for age, sex, blood cell counts, principal components of the ancestry and technical covariates. Full model is additionally adjusted for BMI and smoking. *PLAC8*: placenta-associated 8 and *F2*: coagulation factor II.

TABLE OF CONTENTS

Supplementary Data

Supplementary Figures

Supplementary Tables

Supplementary References

SUPPLEMENTARY DATA
SUPPLEMENTARY DATA I. DESCRIPTIONS OF PARTICIPATING STUDIES

Alzheimer's Disease Neuroimaging Initiative (ADNI)

ADNI is a longitudinal, multi-site observational study including Alzheimer's disease (AD), mild cognitive impairment (MCI), and elderly individuals with normal cognition assessing clinical and cognitive measures, MRI and PET scans (FDG and 11C PIB) and blood and CNS biomarkers.[1,2] In 2004, the initial phase, ADNI-1, recruited 400 subjects with MCI, 200 subjects with early AD and 200 cognitively normal elderly from multiple site in North America. The study was extended with ADNI-GO in 2009 including 500 ADNI-1 subjects (normal controls and MCI cases) and 200 new early MCI cases. In 2011, one more extension, ADNI-2, started to study approximately 450-500 ADNI-1 subjects, 200 ADNI-GO subjects and 650 new cases (150 controls, 150 early MCI cases, 150 late MCI cases and 200 mild AD cases). In this study, 387 normal or early MCI subjects who provided brain MRI and methylation data were included. Brain MRI for WMH burden assessment was taken at ADNI-1 in 73 subjects and at ADNI-2 in 314 subjects. Samples for DNA methylation were drawn in ADNI-GO and ADNI-2 studies. Days between blood sampling and brain MRI were within 365 days in 85% of subjects (n=329), between a year and two years in 8% (n=31), between two and three years in 16 subjects and more than 3 years in 11 subjects.

Ethic Statement

The study protocols for the ADNI studies were approved by the Institutional Review Boards of all the participating institutions.

Acknowledgements

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Atherosclerosis Risk in Communities (ARIC) Study

The ARIC Study is a prospective longitudinal investigation of the development of atherosclerosis and its clinical sequelae in which 15,792 individuals aged 45 to 64 years were enrolled at baseline.[3] At the inception of the study in 1987-1989, participants were selected by probability sampling from four communities in the United States: Forsyth County, North Carolina; Jackson, Mississippi (African-Americans only); suburbs of Minneapolis, Minnesota; and Washington County, Maryland. First four clinical examinations were carried out approximately at three-year intervals (exam 1, 1987-1989; exam 2, 1990-1992; exam 3, 1993-1995; exam 4, 1996-1998) and three additional examinations were completed on survivors in 2011-2013, 2016-2017, and 2018-2019. During the first 2 years of the exam 3, participants aged 55 and older from the Forsyth County and Jackson sites were invited to undergo a cranial MRI. DNA methylation was measured with the Infinium HumanMethylation450 BeadChip (HM450) array (Illumina Inc., San Diego, CA) on genomic DNA extracted from blood samples collected at visit 2 or 3. Besides, subjects were contacted semi-annually to update their medical histories between examinations since 2012. After QC, 906 European-American and 639 African-American subjects with methylation data and brain MRI were included in this study.

Ethic Statement

Written informed consent was provided by all study participants, and the study design and methods were approved by institutional review boards at the collaborating institutions: University of Mississippi Medical Center Institutional Review Board (Jackson Field Center); Wake Forest University Health Sciences Institutional Review Board (Forsyth County Field Center); University of Minnesota Institutional Review Board (Minnesota Field Center); the Johns Hopkins School of Public Health Institutional Review Board (Washington County Field Center); and the University of North Carolina Non-Biomedical Institutional Review Board.

Acknowledgments

The ARIC Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I, HHSN268201700005I). Neurocognitive data is collected by U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, 2U01HL096917 from the NIH (NHLBI, NINDS, NIA and NIDCD), and with previous brain MRI examinations funded by R01-HL70825 from the NHLBI. Funding for methylation (HM450) data was supported by 5RC2HL102419 and R01NS087541. The authors thank the staff and participants of the ARIC Study for their important contributions.

The Brain Imaging, cognition, Dementia and next generation Genomics: a Transdisciplinary approach to search for risk and protective factors of neurodegenerative disease (BRIDGET) study

The BRIDGET consortium comprises samples from Three City Dijon (3C-Dijon), Sydney Memory and Ageing Study (MAS), Older Australian Twins Study (OATS) and The Study of Health in Pomerania - TREND (SHIP-TREND) cohorts. Brief description and ethics statement of each participating cohort is described below:

A) Three City Dijon study (3C-Dijon): The Three City Dijon (3C-Dijon) study is a population-based cohort of 4931 French non-institutionalized individuals aged 65 years and older.[4] A total of 2,763 individuals aged ≤80 years were invited to undergo a brain MRI between June 1999 and September 2000. Participation rate was high (83%, *n*= 2,285) but because of financial restrictions, only 1,924 MRI scans were performed. The Ethical Committee of the University Hospital of Kremlin-Bicêtre approved the study protocol. All participants signed an informed consent to participate in the study.

B) Sydney Memory and Ageing Study (MAS): The Sydney Memory and Ageing Study began in 2005 and is a longitudinal community-based study investigating mild cognitive impairment and the rate of cognitive change over time. Participants aged 70-90 years were randomly recruited from the compulsory electoral roll in Sydney, Australia. Exclusion criteria included limited English or a medical/psychological condition that would prevent them from completing assessments, dementia diagnosis and an age, non-English speaking background and education-adjusted MMSE score of <24. More information can be found in Sachdev et al., 2010.[5] Written informed consent was provided by all participants and the study was approved by the Human Research Ethics Committees of UNSW Sydney and the South Eastern Sydney and Illawarra Area Health Service.

C) Older Australian Twins Study (OATS): see OATS section below.

D) Study of Health in Pomerania (SHIP-TREND): see SHIP-TREND section below.

Definition of Extreme-SVD

In individual participating cohorts, the composite extreme phenotype of SVD (extreme-SVD) was defined based on the distribution of WMH volume, presence or absence of lacunes, presence or absence of brain infarct.[6] Prior to defining extreme-SVD we also removed participants with brain tumors, stroke, or dementia at baseline. The objective was to define a group with extensive SVD severity (individuals in the upper quartile of WMH distribution and having one or more brain infarcts) and a group with minimal SVD severity (lower quartile of WMH distribution and without any brain infarcts). As smoking affects methylation in blood, we also removed current smokers from the subset of extreme-SVD participants. Following table shows the number of extreme-SVD participants defined in individual cohorts:

Cohort	3C-Dijon	MAS	OATS	SHIP-TREND	Total
N with MRI	1,924	530	412	2,189	5,055
N extensive SVD "cases"	90	75	56	90	311
N minimal SVD "controls"	90	76	53	90	309
N extreme-SVD	180	151	109	180	620

Methyl C-Sequencing and data processing

The DNA samples of defined extreme-SVD participants in individual cohorts was sent to the McGill Genome Center, Montreal, Canada for high depth Methyl C-Seq sequencing. In 3C-Dijon and SHIP cohort the DNA was extracted from buffy coat using Qiagen protocol, whereas in MAS and OATS cohorts for majority of samples the DNA was extracted from whole blood using the Qiagen Autopure method. The samples were profiled through targeted methylation sequencing as previously described.[7] Briefly, in MCC-Seq a whole-genome sequencing library is prepared and bisulfite converted, amplified and a capture enriching for targeted bisulfite-converted DNA

fragments is carried out following which the captured DNA was amplified and sequenced on the Illumina HiSeq2000/2500 system using 100-bp paired-end sequencing.
Sequenced reads were aligned to reference human genome build GRCh38 using BWA v.0.6 software employing the BISMARK pipeline.[8] We removed (i) clonal reads, (ii) reads with low mapping quality score (<20), (iii) reads with more than 2% mismatch to converted reference over the alignment length, (iv) reads mapping on the forward and reverse strand of the bisulfite converted genome, (v) read pairs not mapped at the expected distance based on library insert size, and (vi) read pairs that mapped in the wrong direction. At each CpG sites we computed Methylation score using following formula:
Methylation score = Total non-converted C reads / Total reads
We applied following benchmark filtering to avoid potential biases in downstream analyses: ≥5 total reads, no overlap with SNPs (dbSNP 137), ≤20% methylation difference between strands, no off-target reads and no overlap with DAC Blacklisted Regions (DBRs) or Duke Excluded Regions (DERs) generated by the ENCODE project: (http://hgwdev.cse.ucsc.edu/cgi-bin/hgFileUi?db=hg19&g=wgEncodeMapability).

Acknowledgments
This work is supported by: grants overseen by the French National Research Agency (ANR) as part of the "Investments d'Avenir" Programme ANR-18-RHUS-002; the EU Joint Programme - Neurodegenerative Disease Research (JPND) project, through the following funding organisations under the aegis of JPND -www.jpnd.eu. Australia, National Health and Medical Research Council, Austria, Federal Ministry of Science, Research and Economy; Canada, Canadian Institutes of Health Research; France, French National Research Agency; Germany, Federal Ministry of Education and Research; Netherlands, The Netherlands Organisation for Health Research and Development; United Kingdom, Medical Research Council; funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 643417 and No 640643. Part of the computations for this work were performed at the Bordeaux Bioinformatics Center (CBiB) and the CREDIM (Centre de recherche et de développement en informatique) of the University of Bordeaux, for which funding supported was provided to Dr. Debette by the Fondation Claude Pompidou. Part of the data analyses were enabled by compute and storage resources provided by Compute Canada and Calcul Québec.

Three City Dijon study (3C-Dijon): The Three City (3C) 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 3C 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." We thank Dr. Anne Boland (CNG) for her technical help in preparing the DNA samples for analyses. This work was supported by the National Foundation for Alzheimer's disease and related disorders, the Institut Pasteur de Lille, the labex DISTALZ and the Centre National de Génotypage.

Sydney Memory and Ageing Study (MAS): We thank the participants and their informants for their time and generosity in contributing to this research. We also acknowledge the MAS research team:

https://cheba.unsw.edu.au/research-projects/sydney-memory-and-ageing-study. The Sydney Memory and Ageing Study has been funded by three National Health & Medical Research Council (NHMRC) Program Grants (ID No. ID350833, ID568969, and APP1093083). Collection of the methylation data was supported by the NHMRC National Institute for Dementia Research Grant APP1115462.

Older Australian Twins Study (OATS): see OATS – acknowledgements below.

Study of Health in Pomerania (SHIP-TREND): see SHIP-TREND – acknowledgements below.

Coronary Artery Risk Development in Young Adults (CARDIA)

Coronary Artery Disease Risk Development in Young Adults (CARDIA) is a longitudinal study designed to trace the development of risk factors for coronary heart disease in 5,115 relatively young individuals aged 18-30 (52% African American and 55% women).[9] The study collected information on anthropometric, lifestyle factors as well as biomarkers including brain MRI, genetic and DNA methylation data. Baseline measurements were repeated, and additional measurements performed, at Years 2, 5, 7, 10, 15, 20, 25, and 30. DNA methylation level was measured in 1,089 subjects at the examination year 15 and 1,092 subjects at the year 20. In this study, 167 subjects of European and 110 subjects of African-ancestry with the brain MRI image were contributed from CARDIA.

Ethic Statement

The CARDIA study is multicenter with recruitment in Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA. The study protocol was approved by the institutional review boards of the coordinating center and the 4 participating field centers, and written informed consent was obtained from participants at all examinations.

Acknowledgments

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (HHSN268201300025C & HHSN268201300026C), Northwestern University (HHSN268201300027C), University of Minnesota (HHSN268201300028C), Kaiser Foundation Research Institute (HHSN268201300029C), and Johns Hopkins University School of Medicine (HHSN268200900041C). CARDIA is also partially supported by the Intramural Research Program of the National Institute on Aging (NIA) and an intra-agency agreement between NIA and NHLBI (AG0005).

Cardiovascular Health Study (CHS)

The CHS is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers in the United States: Sacramento County, California; Washington County, Maryland; Forsyth County, North Carolina; and Pittsburgh, Allegheny County, Pennsylvania. The original cohort of 5,201 persons, predominantly

of European ancestry, was recruited in 1989-1990 from random samples of the Medicare eligibility lists. Later, 687 African-Americans were added to the cohort, making the total cohort of 5,888.[10] DNA methylation was measured on a randomly selected subjects without presence of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack at the exam year 5. Post QC, 378 European ancestry and African-American participants with methylation data and brain MRI were included in this study.

Ethic Statement

Written informed consent to genetic research has been obtained on all individuals included in this study. CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease. The DNA methylation dataset from CHS can be requested at https://chs-nhlbi.org/node/6222.

Acknowledgements

Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. The CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, N01HC15103; and NHLBI grants U01HL080295, R01HL087652, R01HL092111, R01HL105756, R01HL103612, R01HL111089, R01HL116747 and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 and R01AG033193 from the National Institute on Aging (NIA) as well as Laughlin Family, Alpha Phi Foundation, and Locke Charitable Foundation. A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Framingham Heart Study (FHS) offspring cohort

The FHS is a single-site, community-based, ongoing cohort study that was initiated in 1948 to investigate prospectively the risk factors for CVD. The Original cohort (n=5,209) were randomly recruited in the town of Framingham, MA, USA, and examined every 2 year since 1948; their children and spouses of the children (n=5,124), the Offspring cohort, were enrolled in 1971 and examined approximately every 4 year since 1971.[11,12] DNA methylation was measured on 2,846 Offspring participants around 8th examination cycle (2005-2008). The 4,095 adult children of the Offspring cohort subjects were recruited from 2002 to 2005 as the FHS-3rd Gene cohort.[13] DNA methylation was measured in 1,549 3rd Gen cohort subjects during their second visit (2005-2008). Brain MRI in Gen 3 was collected from 2009 to 2012 in 1963 subjects.

The Offspring cohort with both DNA methylation and MRT measures (n=1,726) was included in the discovery set. We performed the association analyses separately in 403 subjects with the double-spin echo MRI scan and 1,323 subjects with the flair MRI scan. The 865 3rd Gen cohort subjects with DNA methylation and MRI scan were included in a replication study.

Ethic Statement

Written informed consent to genetic research has been obtained on all individuals included in this study. Ethics permission for FHS and genetic research in FHS was obtained from the Institutional Review Board of Boston University Medical Campus (IRB number H-32132, H-26671)

Acknowledgements

This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study Contract No. N01-HC-25195 and No. HHSN268201500001I, and by grants from the National Institute of Aging (R01s AG033193, AG008122, AG054076, AG033040, AG049607, AG05U01-AG049505), and the National Heart, Lung and Blood Institute (R01 HL093029, HL096917). The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, and by NIH contract N01-HC-25195. The computational work reported in this paper was performed on the Shared Computing Cluster which is administered by Boston University's Research Computing Services. We thank all the FHS study participants and staff to make this research possible.

Genetic Epidemiology Network of Arteriopathy (GENOA)

The Genetic Epidemiology Network of Arteriopathy (GENOA) study was designed to investigate the genetics of hypertension and target organ damage in hypertensive sibships.[14] The GENOA study is a community-based study including African Americans from Jackson, Mississippi and non-Hispanic whites from Rochester, Minnesota.9 The initial GENOA study, from 1996 to 2001, recruited any members in sibships having more than two individuals with clinically diagnosed hypertension before age 60. Exclusion criteria of the GENOA study were secondary hypertension, alcoholism or drug abuse, pregnancy, insulin-dependent diabetes mellitus, or active malignancy. Eighty percent of African Americans (1,482 subjects) and 75% of non-Hispanic whites (1,213 subjects) from the initial study population returned for the second examination (Phase II: 2001-2005). Study visits were made in the morning after an overnight fast of at least eight hours. Demographic information, medical history, clinical characteristics, lifestyle factors, and blood samples were collected in each phase. DNA methylation levels were measured only in African Americans participants. After quality control, DNA methylation and brain MRI data were available for 356 African Americans.

Ethic Statement

Written informed consent was obtained from all subjects, and approval was granted by participating institutional review boards at the University of Mississippi Medical Center, the Mayo Clinic, and the University of Michigan.

Acknowledgements

Support for the Genetic Epidemiology Network of Arteriopathy (GENOA) was provided by the National Heart, Lung and Blood Institute (U01HL054457, RC1HL100185, R01HL119443, and R01HL133221) and the National Institute for Neurological Disorders and Stroke (R01NS041558).

Lothian Birth Cohort 1936 (LBC1936)

The Lothian Birth Cohorts of 1936 (LBC1936) is a longitudinal study of ageing. Participants were born in 1936 and mostly live in the Lothian region of Scotland.[15,16] The study was derived of the Scottish Mental Surveys of 1947 that conducted a general cognitive ability test on nearly all 11-year-old children in Scotland. Residents in the Lothian area of Scotland (n=1,091) were recruited in late-life at mean age 70 and followed-up at their ages 70, 73, and 76. In this study, 230 subjects who provided epigenetic information and brain imaging, blood biomarkers and anthropomorphic and lifestyle measures were included.

Ethic Statement

Following written informed consent, venesected whole blood was collected for DNA extraction. Ethics permission for the LBC1936 was obtained from the Multi-Centre Research Ethics Committee for Scotland (Wave 1: MREC/01/0/56), the Lothian Research Ethics Committee (Wave 1: LREC/2003/2/29), and the Scotland A Research Ethics Committee (Waves 2 and 3: 07/MRE00/58).

Acknowledgements

We thank the cohort participants and team members who contributed to these studies. Phenotype collection in the Lothian Birth Cohort 1936 was supported by Age UK (The Disconnected Mind project). Methylation typing was supported by the Centre for Cognitive Ageing and Cognitive Epidemiology (Pilot Fund award), Age UK, The Wellcome Trust Institutional Strategic Support Fund, The University of Edinburgh, and The University of Queensland. Analysis of these data were carried out within the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE). CCACE is supported by funding from the BBSRC, the Medical Research Council (MRC), and the University of Edinburgh as part of the cross-council Lifelong Health and Wellbeing initiative (MR/K026992/1). Funding from the Australian NHMRC (613608) supported quality control analysis of the methylation data.

Older Australian Twin Studies (OATS)

Participants were recruited from the Australian Twin Registry and also through a recruitment drive. At baseline, participants were aged 65 years and over. Inclusion criteria included an ability to consent, a co-twin who also consented to participate, completion of some education in English and residence in one of the three eastern states (Victoria, New South Wales, Queensland). Exclusion criteria included inadequate English to complete the assessment, current diagnosis of malignancy or other life-threatening medical illness and/or a current acute psychosis diagnosis. At baseline, there were 623 participants with a mean age of 70.77 years and 65.2% of the sample were women. For further details see Sachdev et al. (2009) and Sachdev et al. (2011).[17,18] For the current study, one twin from each pair with both methylation and brain MRI data were included (n=81).

Ethic Statement

Written informed consent was obtained from all participants and approval for the study was obtained from the ethics committees of the Australian Twin Registry, University of New South Wales, University of Melbourne, Queensland Institute of Medical Research and the South Eastern Sydney and Illawarra Area Health Service.

Acknowledgements

We thank the OATS participants for their time and generosity in contributing to this research. We acknowledge the contribution of the OATS research team (https://cheba.unsw.edu.au/project/older-australian-twins-study) to this study. The OATS study has been funded by a National Health & Medical Research Council (NHMRC) and Australian Research Council (ARC) Strategic Award Grant of the Ageing Well, Ageing Productively Program (ID No. 401162); NHMRC Project Grants (ID No. 1045325); and NHMRC Program Grants (ID No. 568969 and 1093083). Collection of the methylation data was supported by the NHMRC National Institute for Dementia Research Grant APP1115462. This research was facilitated through access to Twins Research Australia, a national resource supported by a Centre of Research Excellence Grant (ID No. 1079102) from the National Health and Medical Research Council.

This research was facilitated through access to Twins Research Australia, a national resource supported by a Centre of Research Excellence Grant (ID No. 1079102) from the National Health and Medical Research Council.

Rotterdam Study III (RS3)

The Rotterdam Study is a prospective, population-based cohort study that includes inhabitants of the well-defined district Ommoord, located in the city of Rotterdam in the Netherlands.[19–21] The cohort was initiated in 1990 among approximately 7,900 persons aged 55 years and older, who underwent a home interview and extensive physical examination at the baseline and during the follow-up rounds every 3-4 years (RS-I). The cohort was extended in 2000-2001 (RS-II, 3,011 individuals, aged 55 years and older) and 2006-2008 (RS-III, 3,932 individuals aged 45 and older). Collected data include genetic information, epigenetic information, imaging (of heart, blood vessels, eyes, skeleton and brain) and numerous blood biomarkers, anthropomorphic and lifestyle measures.[19,20] Post QC, DNA methylation and brain MRI used for the analyses were available in 547 subjects from the first visit of RS-III and 442 subjects from the fifth visit of RS-I, the third visit of RS-II and the second visit of RS-III. The latter dataset is part of the Biobanking and Biomolecular Resources Research Infrastructure for The Netherlands (BBMRI-NL), BIOS (Biobank-based Integrative Omics Studies) project [http://www.bbmri.nl/?p=259]. Whole blood DNA methylation was quantified in a random subset of 750 individuals with genotyping and RNA expression data available in the first cohort of RS-III.

Ethic Statement

The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC (registration number MEC 02.1015) and by the Dutch Ministry of Health, Welfare and Sport (Population Screening Act WBO, license number 1071272-159521-PG). The Rotterdam Study has been entered into the Netherlands National Trial Register (NTR; www.trialregister.nl) and into the WHO International Clinical Trials Registry Platform (ICTRP; www.who.int/ictrp/network/primary/en/) under shared catalogue number NTR6831. All participants provided written informed consent to participate in the study and to have their information obtained from treating physicians.

Acknowledgements

The generation and management of the Illumina 450K methylation array data (EWAS data) for the Rotterdam Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. The EWAS data was funded

by the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by the Netherlands Organization for Scientific Research (NWO; project number 184021007) and made available as a Rainbow Project (RP3; BIOS) of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL). We thank Mr. Michael Verbiest, Ms. Mila Jhamai, Ms. Sarah Higgins, Mr. Marijn Verkerk, and Lisette Stolk PhD for their help in creating the methylation database.

The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. HHHA is supported by ZonMW grant number 916.19.151.

Rhineland Study

The Rhineland Study is an ongoing single-center, community-based, prospective cohort study that emphasizes deep phenotyping.[22,23] Participants are 30 years and above at baseline and recruited from two geographically defined areas in the city of Bonn, Germany. Core data collection includes extensive structural and functional brain imaging, cognitive and behavioral assessments, cardiovascular investigations, sensory systems assessments, questions on medical history, medication intake, life style, nutrition and mental health, and extensive collection of biomaterials. In this study, 2,111 participants with both DNA methylation data and brain MRI were included.

Ethic Statement

Approval to undertake the study was obtained from the ethics committee of the University of Bonn, Medical Faculty. We obtained written informed consent from all participants in accordance with the Declaration of Helsinki.

Acknowledgements

The Rhineland Study is funded by
German Center for Neurodegenerative Diseases (DZNE). MMBB is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC2151 – 390873048, and by the Diet-Body-Brain Competence Cluster in Nutrition Research funded by the Federal Ministry of Education and Research (grant numbers 01EA1410C and
01EA1809C). DL is supported by a scholarship from China Scholarship Council.

The Study of Health in Pomerania - TREND (SHIP-TREND)

The Study of Health in Pomerania (SHIP-TREND) is a longitudinal population-based cohort study in West Pomerania, a region in the northeast of Germany, assessing the prevalence and incidence of common population-relevant diseases and their risk factors. Baseline examinations for SHIP-TREND were carried out between 2008 and 2012, comprising 4,420 participants aged 20 to 81 years. Study design and sampling methods were previously described.[24] The medical ethics

committee of the University of Greifswald approved the study protocol, and oral and written informed consents were obtained from each of the study participants. DNA was extracted from blood samples of n=256 SHIP-TREND participants to assess DNA methylation using the Illumina HumanMethylationEPIC BeadChip array. Samples were randomly selected based on availability of multiple OMICS data, excluding type II diabetes, and enriched for prevalent MI. The samples were taken between 07:00 AM and 04:00 PM, and serum aliquots were prepared for immediate analysis and for storage at -80 °C in the Integrated Research Biobank (Liconic, Liechtenstein). Processing of the DNA samples was performed at the Helmholtz Zentrum München. Whole body MRI was offered to all study participants. After exclusion of subjects who refused participation or who fulfilled exclusion criteria for MRI (e.g. cardiac pacemaker), 2189 subjects from SHIP-TREND-0 underwent the MRI scanning. After exclusion of scans with technical artifacts, major structural abnormalities and stroke, full data sets with EWAS data and MRI scans were available in 214 subjects in SHIP-TREND. The extreme-SVD cases and controls for BRIDGET were taken from a non-overlapping subsample of the TREND cohort without diabetes mellitus.

Ethic Statement

SHIP and SHIP TREND were approved by the local ethics committee. After complete description of the study to the subjects, written informed consent was obtained. The medical ethics committee of the University of Greifswald approved the study protocol, and oral and written informed consents were obtained from each of the study participants.

Acknowledgements

SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’ funded by the Federal Ministry of Education and Research (grant 03IS2061A). DNA methylation data have been supported by the DZHK (grant 81X3400104). The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH. The SHIP authors are grateful to Paul S. DeVries for his support with the EWAS pipeline.

SUPPLEMENTARY DATA II. WMH MEASUREMENTS IN PARTICIPATING STUDIES

Study	MRI technology	Scanning Protocol	Image Processing	WMH quantification
ADNI[25–27]	1.5T (ADNI-1) and 3T (ADNI-2) MRI scanners from General Electric Medical Systems, Philips Medical Systems, or Siemens Medical Solutions	A series of sagittal T1-weighted scans and axial proton-density (PD)/ T2-weighted fast spin echo scans. T1-, T2- and PD-weighted MRI scans were co-registered, and skull stripped, then corrected for bias field.	ADNI-1: Brain images were reviewed by a neuroradiologist to exclude infarcts and other abnormalities and graded by a trained individual. ADNI-2: the FreeSurfer software package version 4.3 and 5.1 (http://surfer.nmr.mgh.harvard.edu/) by the Schuff and Tosun laboratory at the University of California-San Francisco	ADNI-1: WMH was quantified based on T1-, T2- and PD- intensities, the prior probability of WMH, and the probability of WMH conditioned on the presence of WMH at neighboring voxels. ADNI-2: WMH was measured with a Bayesian approach using segmentation of high-resolution 3D T1 and FLAIR sequences.
ARIC[28]	1.5-Tesla scanners (General Electric Medical Systems or Picker Medical Systems)	A series of sagittal T1-weighted scans and axial PD, T2- weighted and T1-weighted scans with 5 mm thickness and no interslice gaps.	Images were interpreted directly from a PDS-4 digital workstation consisting of four 1024 X 1024-pixel monitors capable of displaying all 96 images simultaneously.	WMHs were estimated as the relative total volume of periventricular and subcortical WM signal abnormality on proton density– weighted axial images by visual comparison with eight templates that successively increased from barely detectable WM changes (Grade 1) to extensive, confluent changes (Grade 8). Individuals with no WM changes received Grade 0, and those with changes worse than Grade 8 received Grade 9.
BRIDGET-3C-Dijon[29,30]	a 1.5 T Magnetom scanner (Siemens)	T1-weighted, T2-weighted, and proton density-weighted sequences	T1- and T2-weighted images of each subject were first aligned to each other using the AIR package. These images were then further analyzed with the optimized Voxel-Based Morphometry (VBM) protocol, using Statistical Parametric Mapping 99 (SPM99) that we modified in order to consider the structural characteristics of the aged brain, as described in detail elsewhere.	Fully automated image processing software was developed to detect and localize WMH and to measure WMH volume.[29] Infarcts were rated on T_1-, T_2- and proton density-weighted images by the same examiner (Y.-C.Z.), using a standardized assessment grid, to visually review all brain scans. Lacunes were defined as infarcts 3–15 mm in diameter having the same signal characteristics as cerebrospinal fluid on all sequences, located in basal ganglia, brainstem or cerebral white matter. Characteristics of lesions were visualized simultaneously in axial, coronal, and sagittal planes to discriminate them from dilated perivascular spaces. Lesions with a typical vascular shape and following the orientation of perforating vessels were regarded as dilated perivascular spaces.[30] The nomenclature of MRI markers of SVD in our study is consistent with the recently proposed neuroimaging standards for research into SVD (STRIVE).[31,32]
BRIDGET-MAS[33]	A Philips 3T Intera Quasar scanner was used in the first half of the cohort, and a Philips 3T Achieva Quasar Dual scanner	The two scanners were set to the same parameters: T1-weighted MRI – TR = 6.39 ms, TE = 2.9 ms, flip angle = 8°, matrix size = 256 × 256, FOV (field of	Both image preprocessing and the quantification of WMH were done automatically by using UBO Detection (Ref). Briefly, the preprocessing included registering T2-FLAIR to corresponding T1 image, tissue	Briefly, for WMH extraction, a k-nearest neighbour (kNN)-based classifier was applied to FAST-segmented clusters to classify the clusters into WMH or non-WMH through considering intensity, anatomical location, and cluster size features. The segmented WMH were visually checked for segmentation accuracy.

Study	MRI technology	Scanning Protocol	Image Processing	WMH quantification
	was applied to the second half of the sample	view) = 256 × 256 × 190, and slice thickness = 1 mm with no gap in between, yielding 1 × 1 × 1 mm3 isotropic voxels. T2-weighted FLAIR – TR = 10000 ms, TE = 110 ms, TI = 2800 ms, matrix size = 512 × 512, slice thickness = 3.5 mm without gap, and in plane resolution = 0.488 × 0.488 mm.	segmentation using SPM12, warping to s sample-specific DARTEL space, non-brain tissue removal, bias field correction, and segmentation using FSL FAST.	
CARDIA[34]	3T scanners (Siemens Medical Solutions and Philips Medical Systems)	A sequence of sagittal 3DT2 with slice thickness 1 mm; sagittal 3D FLAIR with slice thickness 1 mm; and sagittal 3D MPRAGE with slice thickness 1mm.	Structural MR images were processed using previously described methods that were based on an automated multispectral computer algorithm.	We classified all supratentorial brain tissue into GM, WM, and CSF. GM and WM were characterized as normal and abnormal (ischemic).
CHS[28]	General Electric or Picker 1.5-Tesla scanners and Toshiba 0.35-Tesla scanner	A series of sagittal T1-weighted scans and axial PD, T2- weighted and T1-weighted scans with 5 mm thickness and no interslice gaps.	Images were interpreted directly from a PDS-4 digital workstation consisting of four 1024 X 1024-pixel monitors capable of displaying all 96 images simultaneously.	WMHs were estimated as the relative total volume of periventricular and subcortical WM signal abnormality on proton density– weighted axial images by visual comparison with eight templates that successively increased from barely detectable WM changes (Grade 1) to extensive, confluent changes (Grade 8). Individuals with no WM changes received Grade 0, and those with changes worse than Grade 8 received Grade 9.
FHS	1 or 1.5 T Siemens Magnetom scanner	Offspring cohort: 3D T1 and double echo proton density and T2 double spin echo (DSE) coronal images were acquired in 4-mm contiguous slices (n=403) and MRI scans with the FLAIR sequences were taken in 3-mm slices (n=1,323). 3rd Generation cohort: MRI scans with the FLAIR sequences were taken in 3-mm slices (n=865).	Analyzed with QUANTA 6.2 at the University of California–Davis Medical Center. Images were analyzed and interpreted blind to subject data and in random order. Semi-automated analysis of pixel distributions, based on mathematical modeling of MRI pixel intensity histograms for CSF, WM, and GM, were used to determine the optimal threshold of pixel intensity to best distinguish CSF from brain matter based on previously published methods. The intracranial vault above the tentorium was outlined manually to determine the TIV.	For segmentation of WMH from other brain tissues the first and second echo images from T2 sequences were summed and a lognormal distribution was fitted to the summed data (after removal of CSF and correction of image intensity non-uniformities). A segmentation threshold for WMH was determined as 3.5 standard deviations in pixel intensity above the mean of the fitted distribution of brain parenchyma.
GENOA[35]	1.5 T MRI scanner from General	FLAIR images consisting of 48 contiguous 3-mm interleaved slices	A fully automated algorithm was used to segment each slice of the edited	Interactive imaging processing steps were performed by a research associate who had no knowledge of the subjects' personal or medical

Study	MRI technology	Scanning Protocol	Image Processing	WMH quantification
	Electric Medical Systems	with no interslice gap. Total intracranial volume (head size) was measured from T1-weighted spin echo sagittal images, each set consisting of 32 contiguous 5 mm thick slices with no interslice gap.	multi-slice FLAIR sequence into voxels assigned to one of three categories: brain, CSF, or leukoaraiosis.	histories or biological relationships. Total leukoaraiosis volumes were determined from axial FLAIR images.
LBC1936[36]	1.5 T HDx MRI scanner from General Electric Medical Systems	T1-w coronal and T2-W, FLAIR, and T2*-weighted axial whole brain images	A semi-automatic computational program written specifically for the project, MCMxxxVI, a multispectral color fusion method that combines different pairs of sequences in red-green color space and performs minimum variance quantization to highlight different tissues.	WMH were measured in the cerebral hemispheres, cerebellum and brainstem. Intracranial volume, brain and WMH volume were extracted and manually corrected as necessary to remove false positive lesions in the insular cortex, cingulate gyrus, anterior temporal cortex, and around the floor of the third ventricle, and correct false negatives (http://www.bric.ed.ac.uk/research/imageanalysis.html). All focal stroke lesions were manually removed.
OATS and BRIDGET-OATS	New South Wales study site: Philips 1.5T Gyroscan scanner was initially used, and later replaced with a Philips 3T Achieva Quasar Dual scanner. Victoria study site: Siemens 1.5T Magnetom Avanto scanner. Queensland study site: Siemens 1.5T Sonata scanner	T1-weighted MRI – T1-weighted images acquired from 1.5T scanners in all three centres: in-plane resolution 1 × 1 mm with slice thickness of 1.5 mm, contiguous slices, TR (repetition time) = 1530 ms, TE (echo time) = 3.24 ms, TI (inversion time) = 780 ms, and flip angle = 8°. Scanning parameters for T1-weighted MRI acquired on the 3T scanner in New South Wales: TR = 6.39 ms, TE = 2.9 ms, spatial resolution = 1 × 1 × 1 mm3. T2-weighted FLAIR – FLAIR images acquired from 1.5T scanners in all three centres: TR = 10000 ms, TE = 120 ms, TI = 2800 ms, slice thickness = 3.5 mm, and in-plane resolution = 0.898 × 0.898 mm2. The scanning parameters of the 3T scanner at New South Wales: TR = 10000 ms, TE = 110 ms, TI = 2800 ms, slice thickness = 3.5 mm and in-plane resolution = 0.898 × 0.898 mm2.	Both image preprocessing and the quantification of WMH were done automatically by using UBO Detection.[33] Briefly, the preprocessing included registering T2-FLAIR to corresponding T1 image, tissue segmentation using SPM12, warping to s sample-specific DARTEL space, non-brain tissue removal, bias field correction, and segmentation using FSL FAST.[37]	Briefly, for WMH extraction, a k-nearest neighbour (kNN)-based classifier was applied to FAST-segmented clusters to classify the clusters into WMH or non-WMH through considering intensity, anatomical location, and cluster size features. The segmented WMH were visually checked for segmentation accuracy.White matter hyperintesnsities (WMH) volumes were calculated using UBO Detector
RSIII and RSI-III (BBMRI)[19]	1.5 T scanner (GE Signa Excite) using	Structural imaging is performed with T1-weighted (T1w), proton density-	The combination of different MR contrasts provided by acquired	WMH volume was quantified using two fully automated methods, which was described previously in more detail. Either the HASTE, PD

Study	MRI technology	Scanning Protocol	Image Processing	WMH quantification
	an 8-channel head coil	weighted (PDw) and fluid-attenuated inversion recovery (FLAIR) sequences.	sequences can be used for automated brain tissue and white matter lesion segmentation. For this purpose, the T1w scan is acquired in 3D at high in-plane resolution and with thin slices (voxel size <1 mm3).	and T2 sequences were used, or the FLAIR, T1 and PD sequences. Briefly, cerebrospinal fluid (CSF), gray matter (GM) and white matter (WM) are segmented by an atlas-based k-nearest neighbor classifier on multi-modal magnetic resonance imaging data. This classifier is trained by registering brain atlases to the subject. The resulting GM segmentation is used to automatically find a WMH threshold in a fluid-attenuated inversion recovery scan.
Rhineland Study	Siemens 3 Tesla Prisma MRI scanners, equipped with an 80 mT/m gradient system and a 64-channel phased-array head-neck coil	The protocol included a 3D T1-weighted Multi-Echo Magnetisation Prepared RApid Gradient-Echo (MEMPRAGE) sequence (0.8 mm isotropic); a 3D T2 Turbo-Spin-Echo (TSE) sequence (0.8 mm isotropic); a 3D FLAIR sequence (1.0 mm isotropic)	WMH were outlined using an in-house developed pipeline based on DeepMedic [38], which utilizes image information from T1-weighted, T2-weighted, and FLAIR sequences.	WMH were defined as hyperintense signal abnormalities in the white matter tracts on FLAIR and/or T2-weighted images.
SHIP-TREND and BRIDGET-SHIP-TREND [39]	1.5-T MR imager from Siemens Medical systems	A series of T1, MP-RAGE/ axial plane with the resolution of 1.0 x 1.0 x 1.0 mm^3 and T2 FLAIR / axial plane with the resolution of 0.9 x 0.9 x 3.0mm3.	The FreeSurfer 5.1.0 software	WMH quantification was based on the following sequences: T1, MP-RAGE/ axial plane, TR=1900 ms, TE=3.4 ms, Flip angle=15°, resolution of 1.0 x 1.0 x 1.0mm^3.

Abbreviations: Double spin echo (DSE). Fluid-attenuated inversion recovery (FLAIR). Proton-density (PD). Cerebrospinal fluid (CSF). White matter (WM). Grey matter (GM). Total (intra)cranial volume (TIV/TCV)

SUPPLEMENTARY DATA III. DNA METHYLATION DATA COLLECTION AND ASSOCIATION STUDY IN PARTICIPATING STUDIES

Study	Technology	Tissue source	Normalization	QC methods	Experimental/ Biological covariates	Covariate Assessment	Technical Covariates	Methods of Adjustment for Covariates
ADNI	Illumina 850K	Whole blood	BMIQ	removed probes with - a low detection rate (P>0.01 for >5% of samples) removed samples with - low call rate (P<0.01 for <95% of probes)	Model 1: ICV, age, sex, WBC, study site, visit, time between MRI and DNAm visits, early MCI status, population structure (surrogates); Model 2: + Smoking, BMI SBP, DBP	- Smoking status was coded as "Yes" or "No". - Systolic and diastolic blood pressure was measured at the MRI visit.	chip ID, chip Position	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates. Technical covariates as random effects.
ARIC	Illumina 450K	Blood leukocytes	BMIQ	removed probes with - a low detection rate (P>0.01 for >5% of samples) removed samples with - low call rate (P<0.01 for <99% of probes)	In each ancestry, Model 1: age, sex, WBC, technical covariates, population structure (surrogates); Model 2: + Smoking, BMI, SBP, DBP	- Smoking status was obtained using an interviewer-administered questionnaire and was classified as current, former, or never. - Blood pressure was assessed at exam 2 and 3, the average value of two visit was used in this study.	Chip ID, Chip Row, plate	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates. Technical covariates as random effects.
BRIDGET	MCC-Seq	Buffy coat and whole blood	Refer text above for MCC-Seq data processing.		Age, sex, cohort site, and sample non-independence (relatedness, population structure or hidden confounding factors)	NA	NA	Generalised linear mixed model of Methylation score~extreme-SVD adjusted for age, sex, cohort site and sample non-independence. We used PQLSeq software for association analysis.[40]

Study	Technology	Tissue source	Normalization	QC methods	Experimental/ Biological covariates	Covariate Assessment	Technical Covariates	Methods of Adjustment for Covariates
CARDIA	Illumina 850K	Whole blood	RCP	- excluded 6,209 CpGs with a detection rate <95% - excluded 87 samples with the percentage of low-quality methylation measurements >5% - extremely low intensity of bisulfite conversion probes (less than mean intensity – 3 × standard deviation of the intensity across samples).	Model 1: age, sex, race, WBC, technical covariates, population structure (surrogates); Model 2: + Smoking, BMI, SBP, DBP	- Smoking status was coded as "never", "ex-smoker", and "current smoker" based on a self-administered questionnaire. - Systolic and diastolic blood pressure in mmHg measured at the same examination as the DNA sampling.	chip ID, chip position	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates. Technical covariates as random effects.
CHS	Illumina 450K	Blood leukocytes	SWAN	excluded samples with - a proportion of probes falling detection of greater than 0.5% - low median intensities of below 10.5 (log2) across the methylated and unmethylated channels - QC probes falling greater than 3 standard deviation from the mean - with sex-check mismatches - failed concordance with prior genotyping or > 0.5% of probes with a detection p-value > 0.01	Model 1: age, sex, 4 genetic PCs, batch, WBC estimates, race, technical covariates Model 2:Model 1 + Smoking, BMI , SBP, DBP	- Smoking was assessed with a standard questionnaire during in-person interview at the time of blood assessment for methylation. Smoking status was categorized into never, former or current smoker. - Systolic and diastolic blood pressure in mmHg measured at the same examination as the DNA sampling.	chip ID, chip position	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates. Technical covariates as random effects.
FHS	Illumina 450K	Whole blood	DASEN (wateRmelon)	Samples were excluded if - outliers in multidimensional scaling (MDS) analysis, high missing rate (>1%), poor matching to SNP genotype Probes were excluded if - high missing rate (>20%), mapped to multiple locations, had SNP (MAF>5% in EUR 1000G) at CpG site or ≤10 bp of Single Base Extension.	Model 1: age, sex, WBC, technical covariates, population structure (surrogates); Model 2: + Smoking, BMI, SBP, DBP	- Current smoking is defined as regularly smoking in the past year which was self-reported at each examination. Using examination 8 and previous examination data, participants were classified as current smoker, ex-smoker, never smoker.	chip ID, chip position	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates. Family structure is also adjusted as a random intercept effect, its variance-covariance matrix proportional to the kinship matrix.

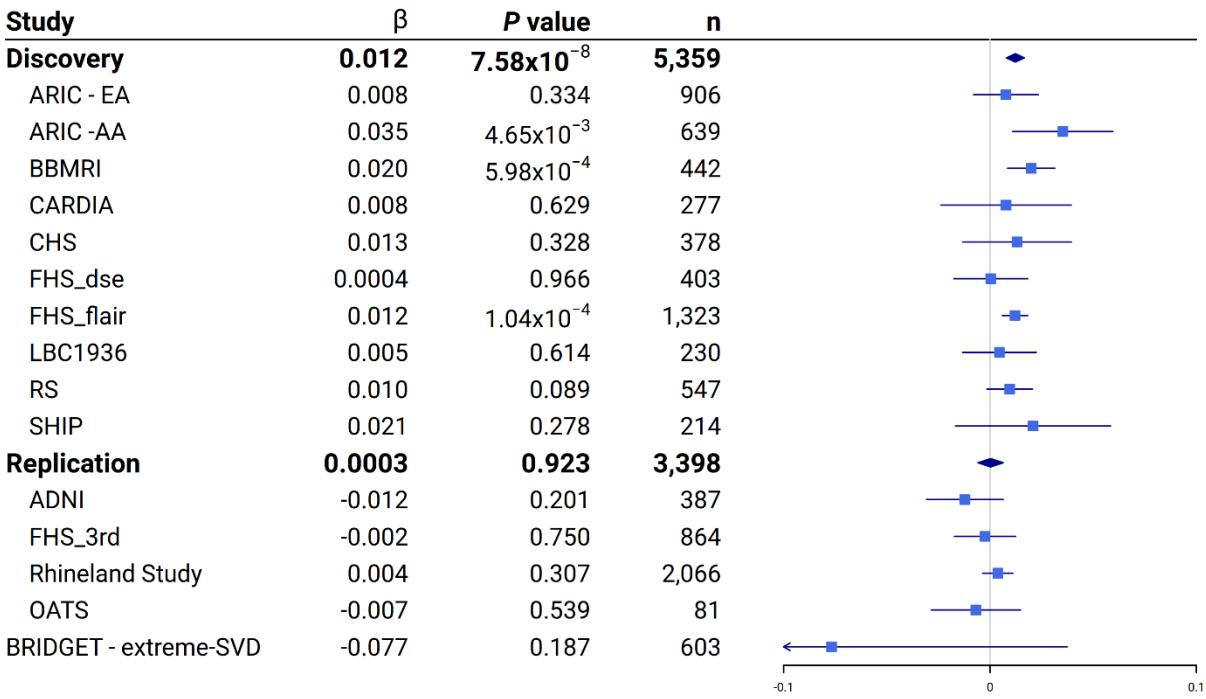
Study	Technology	Tissue source	Normalization	QC methods	Experimental/ Biological covariates	Covariate Assessment	Technical Covariates	Methods of Adjustment for Covariates
GENOA	Illumina 27K	Blood leukocytes	SWAN	- Remove probe if >5% of samples have a detection p-value of >0.01; - Remove sample if >5% of probes have a detection p-value of >0.01;	Model 1: age, sex, total intracranial volume, 4 genetic PCs, WBC, technical covariates, familial relationship; Model 2: + smoking, BMI, SBP, DBP	- Smoking status: current smoker (smokes currently or quit smoking within the past year), ex-smoker (formerly smoked, but quit more than 1 year ago), and never-smoker (smoked less than 100 cigarettes in his/her entire life). - Blood Pressure: sitting systolic and diastolic blood pressure in mmHg were measured three times with a random zero sphygmomanometer. The average of the last two measurement was used.	plate, row, column	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates. Technical covariates and familial relationship were treated as random effects.
LBC1936	Illumina 450K	Whole blood	minfi	Remove probes with - a low detection rate (P>0.01 for >5% of samples) - low quality (manual inspection) Remove samples with - low call rate (P<0.01 for <95% of probes) - a poor match between genotypes and SNP control probes - incorrect predicted sex.	Model 1: age + Sex + ICV+ 4 genetic PCs + WBC + technical covariates, Model 2: BMI + current smoking + dbp3sit + sbp3sit + age + Sex + ICV+ 4 genetic PCs + WBC + technical covariates	- Smoking status was self-reported in three categories: current smoker, ex-smoker, never smoker. - Blood Pressure: dbp3sit + sbp3sit (3rd reading)	Chip, position, sample plate, hybridization date	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates.
OATS	Illumina 450K	Blood leukocytes	SWAN	a) Exclude samples with p-value <0.01 for <95% of probes b) Exclude probes with p-value >0.01 in >5% of samples	Model 1: age, sex, site, population structure (first 2 PCs), WBC; Model 2: + smoking, BMI, SBP, DBP	- Information on cigarette smoking was obtained in wave 1 using an interviewer-administered questionnaire and was classified as current, former, or never. - BMI was calculated as weight in kilograms/(height in meters)2 from height and weight measurements obtained at wave 1. - Seated blood pressure was assessed in wave 1, the average value of three individual measurements was used in this study.	Site	Generalised estimating equations of DNAm~WMH burden adjusted for biological covariates. Family being a random effect.
RSIII and RSI-III (BBMRI)	Illumina 450K	Whole blood	DASEN (wateRmelon)	- Excluded samples with incomplete bisulfite treatment, a low detection rate (<99%), or gender mismatches.	Model 1: ICV, age, sex, array number, array position, WBC	- Smoking status: Smoking status was coded as self-reported “never smoker”, “ever smoker”, and “current	Array number, array position.	Linear mixed models of DNAm~WMH burden adjusted for

Study	Technology	Tissue source	Normalization	QC methods	Experimental/ Biological covariates	Covariate Assessment	Technical Covariates	Methods of Adjustment for Covariates
				- Excluded probes with a detection p-value>0.01 in >1% samples	counts. Model 2: model 1 + smoking, BMI, SBP, DBP	smoker" based on a home interview. - Blood Pressure: Blood pressure was measured twice at the center visit, and the average of the two measurements was used.		biological and technical covariates. Technical covariates as random effects.
Rhineland Study	Illumina 850K	buffy coat	minfi	Remove probes with - a low detection rate (detection p-value >0.01 for >1% of samples) Remove samples with - low median intensities of below 10.5 (log2) across the methylated and unmethylated channels - a low call rate (detection p-value >0.01 for <99% of probes) - sex mismatch	Model 1: age, sex,ICV, WBC subtypes, technical covariates; Model 2: + smoking, BMI, SBP, DBP	Smoking status was coded as "Yes" or "No". Body mass index was computed as weight in kilogram divided by squared height in meter from the anthropometric exams at the visit. Systolic and diastolic blood pressure was measured at the MRI visit.	chip ID, chip position, first 5 PCs of the control probe	Linear mixed models of DNAm~WMH burden adjusted for biological and technical covariates.
SHIP-TREND	Illumina 850K	Whole blood	CPACOR	- probes with detection p-value≥1E-16 were set to missing - arrays with call rates ≤0.95 were excluded - arrays with observed technical problems during steps like bisulfite conversion, hybridization or extension, as well as arrays with mismatch between sex of the proband and sex determined by the chr X and Y probe intensities were excluded.	Model 1: ICV, age, sex, WBC subtypes technical covariates; Model 2: + smoking, BMI, SBP, DBP	- Smoking status: current smoker determined by questionnaire - Blood Pressure: measured (mean of 2nd and 3rd measurement)	First 6 control probes-based PCs	linear regression adjusted for all covariates listed

SUPPLEMENTARY DATA IV. GENE EXPRESSION DATA DESCRIPTION

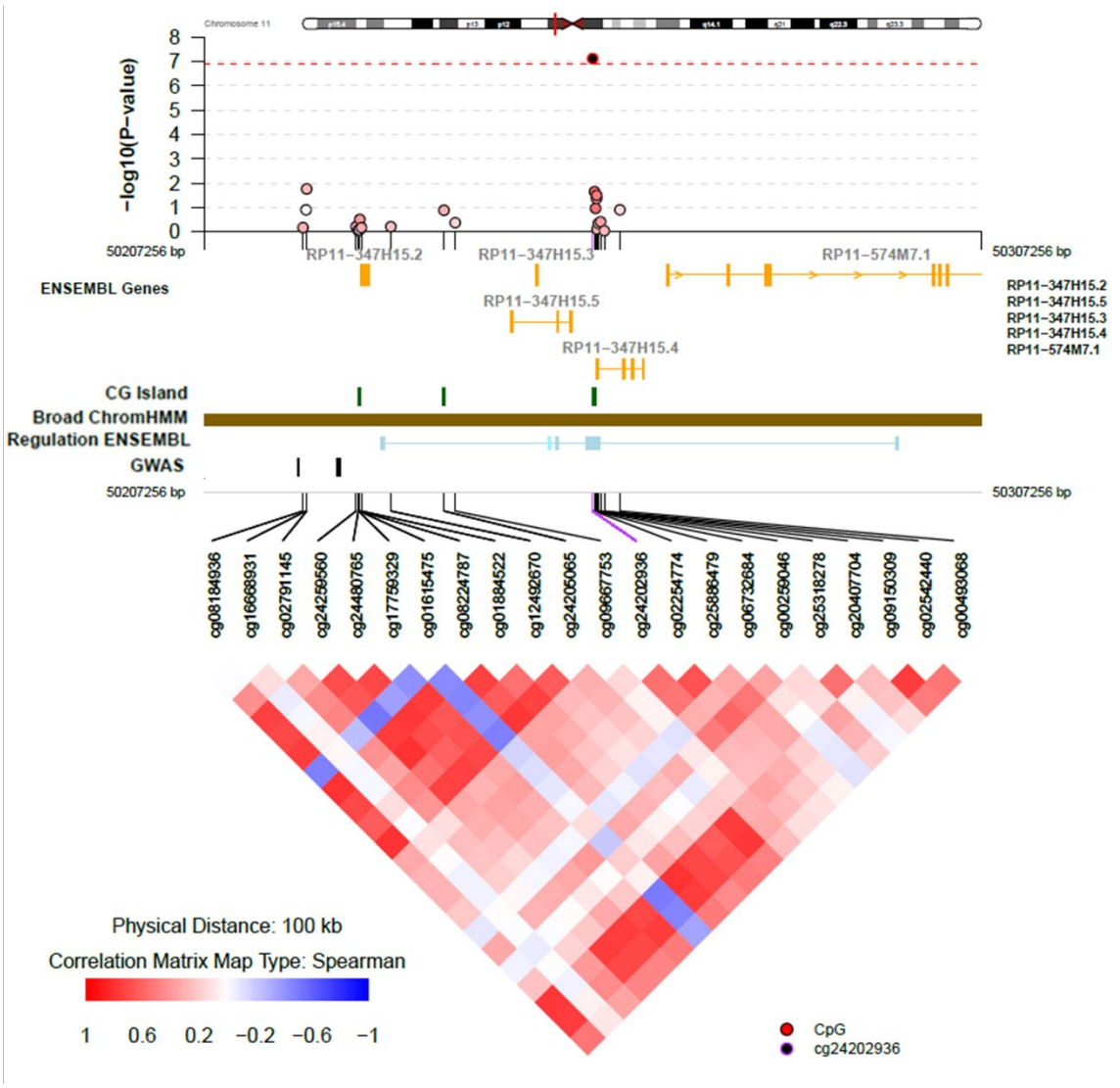
Study	Technology	Tissue source	Normalization	GEO accession ID
FHS [41]	Affymetrix Human Exon Array ST 1.0	Whole blood	The raw data were quantile-normalized and log2 transformed, followed by summarization using Robust Multi-array Average. The percentages of each imputed cell type were then normalized, where the negative predicted values were set to 0 and the sum of the percentages for all cell types were set 100%.	
RS [42,43]	Illumina HumanHT12v4 Expression Beadchip	Whole blood	Quantile-normalized to the median distribution and subsequently log2-transformed. The probe and sample means were centered to zero and Z-transformed.	GSE33828

SUPPLEMENTARY FIGURES



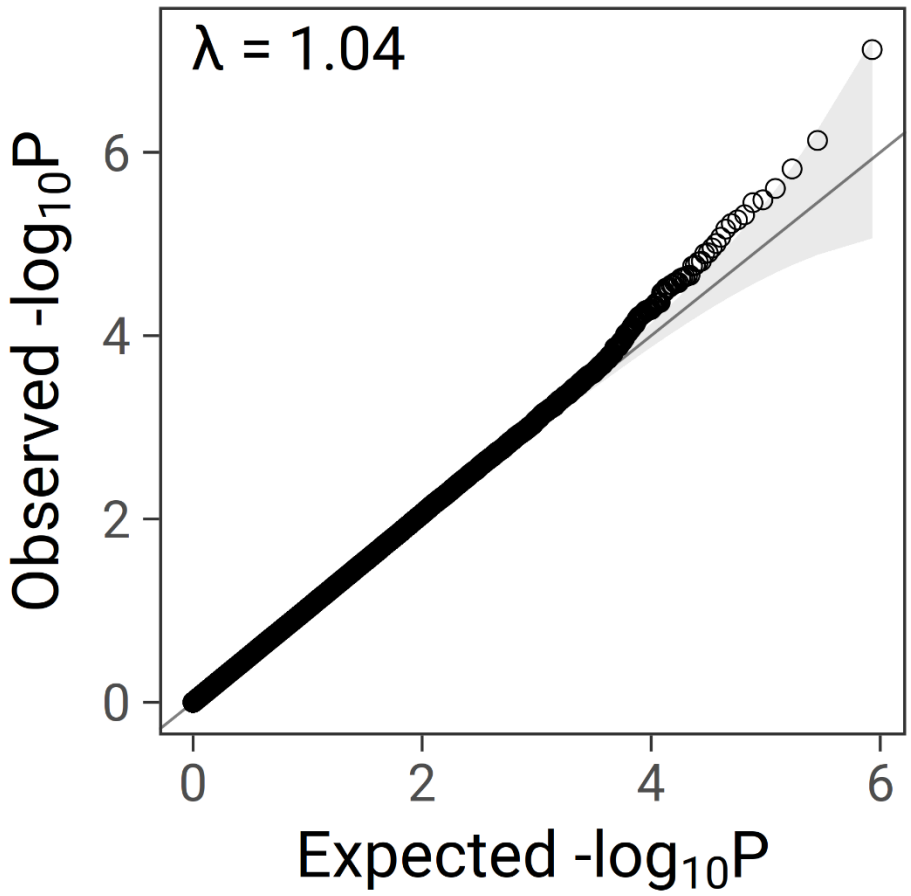
Supplementary Figure 1 Forest plot of cg24202936 effect on WMH burden.

Beta estimates from the linear mixed effect model for cg24202936 in the meta-analysis and individual studies were drawn with their 95% confidence intervals. The BRIDGET study derived beta estimate and P value from a logistic regression model. **Abbreviations:** AA = African ancestry; ADNI = Alzheimer's Disease Neuroimaging Initiative; ARIC = Atherosclerosis Risk in Communities Study; BBMRI = Biobanking and BioMolecular resources Research Infrastructure ; BRIDGET = the BRain Imaging, cognition, Dementia and next generation GEnomics: a Transdisciplinary approach to search for risk and protective factors of neurodegenerative disease consortium; CHS = Cardiovascular Health Study; CARDIA = Coronary Artery Risk Development in Young Adults; EA = European ancestry; FHS_dse = Framingham Heart Study Offspring double spin-echo sequence study; FHS_flair = FHS Offspring fluid-attenuated inversion recovery study; FHS_3rd: FHS 3rd generation study; LBC1936 = Lothian Birth Cohort 1936; RS = Rotterdam Study; SHIP = Study of Health in Pomerania; OATS = Older Australian Twin Study.



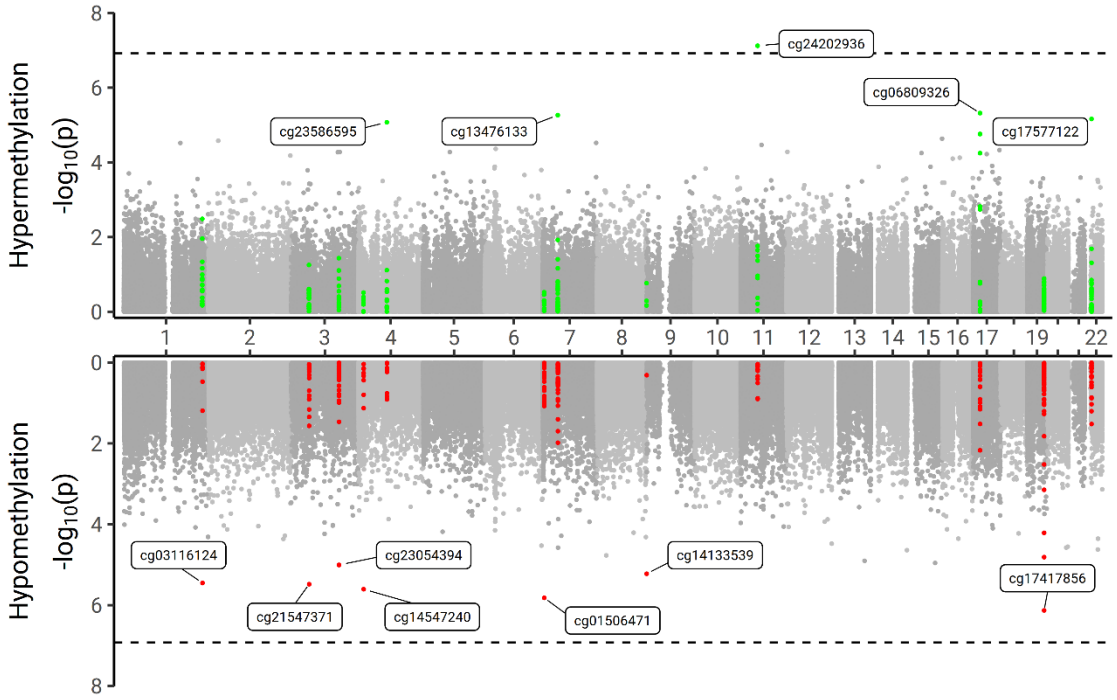
Supplementary Figure 2 Regional plot for cg24202936 (±50 kb).

Red dashed line indicated the significance threshold ($1.2x10^{-7}$). Along with x-axis of chromosomal position, -log10(*P* value) is plotted on y-axis for each CpG. Correlations among CpGs were calculated in ARIC study (906 European and 639 African descents). Genomic positions were annotated using CpG Islands, ENCODE histone modification ("Broad ChromHMM"), ENSEMBL Genes ("Regulation ENSEMBL"), and genome-wide association study (GWAS) findings.



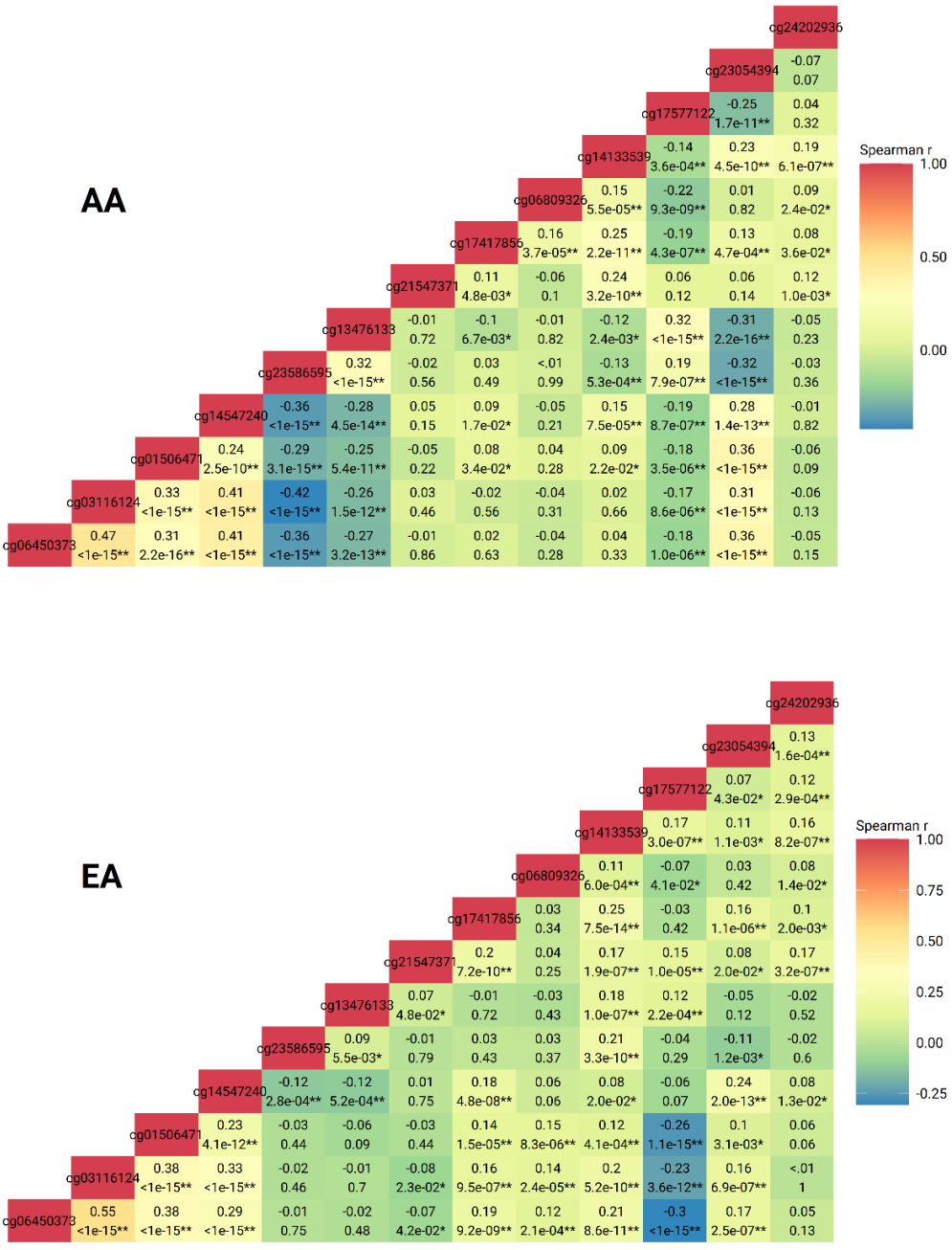
Supplementary Figure 3 QQ plot in methylome-wide association *P* values for WMH burden.

Observed -$\log_{10}$ (*P* value) are plotted against expected values. λ indicates the genomic inflation factor.



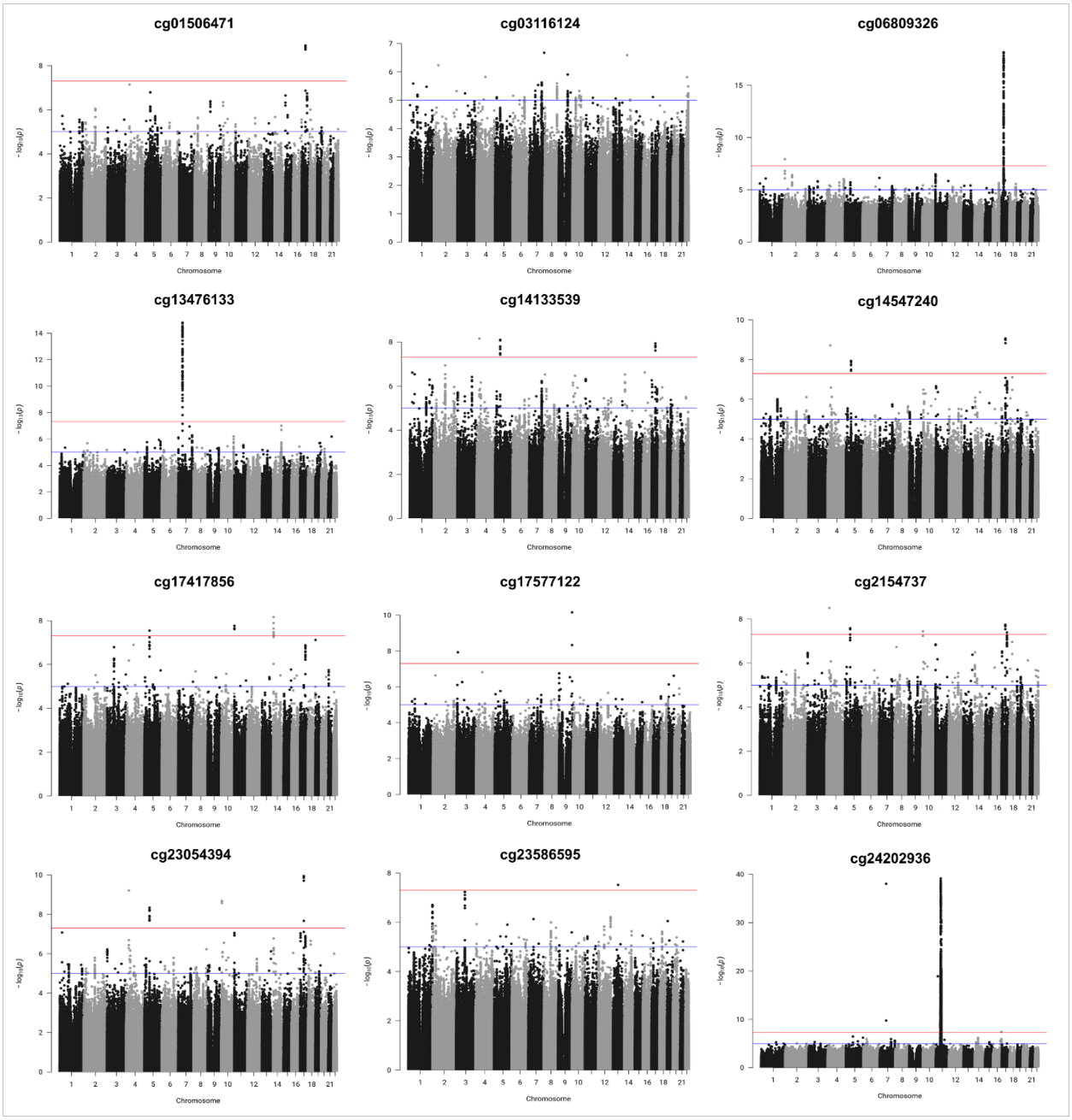
Supplementary Figure 4 Miami plot of methylome-wide associations in WMH burden.

Associations (-log10(*P* value), y-axis) between increased DNAm level (hypermethylation) and WMH burden (top); between decreased DNAm level (hypomethylation) and WMH burden (bottom) are plotted against DNAm locations (x-axis). The dashed line indicates the Bonferroni threshold (1.2×10^{-7}) for epigenome-wide significance. Target CpGs with a *P* value smaller than the suggestive significance threshold (1.0×10^{-5}) are labelled with probe name. Single CpGs within ±50 kb are highlighted in green (top) and red (bottom).



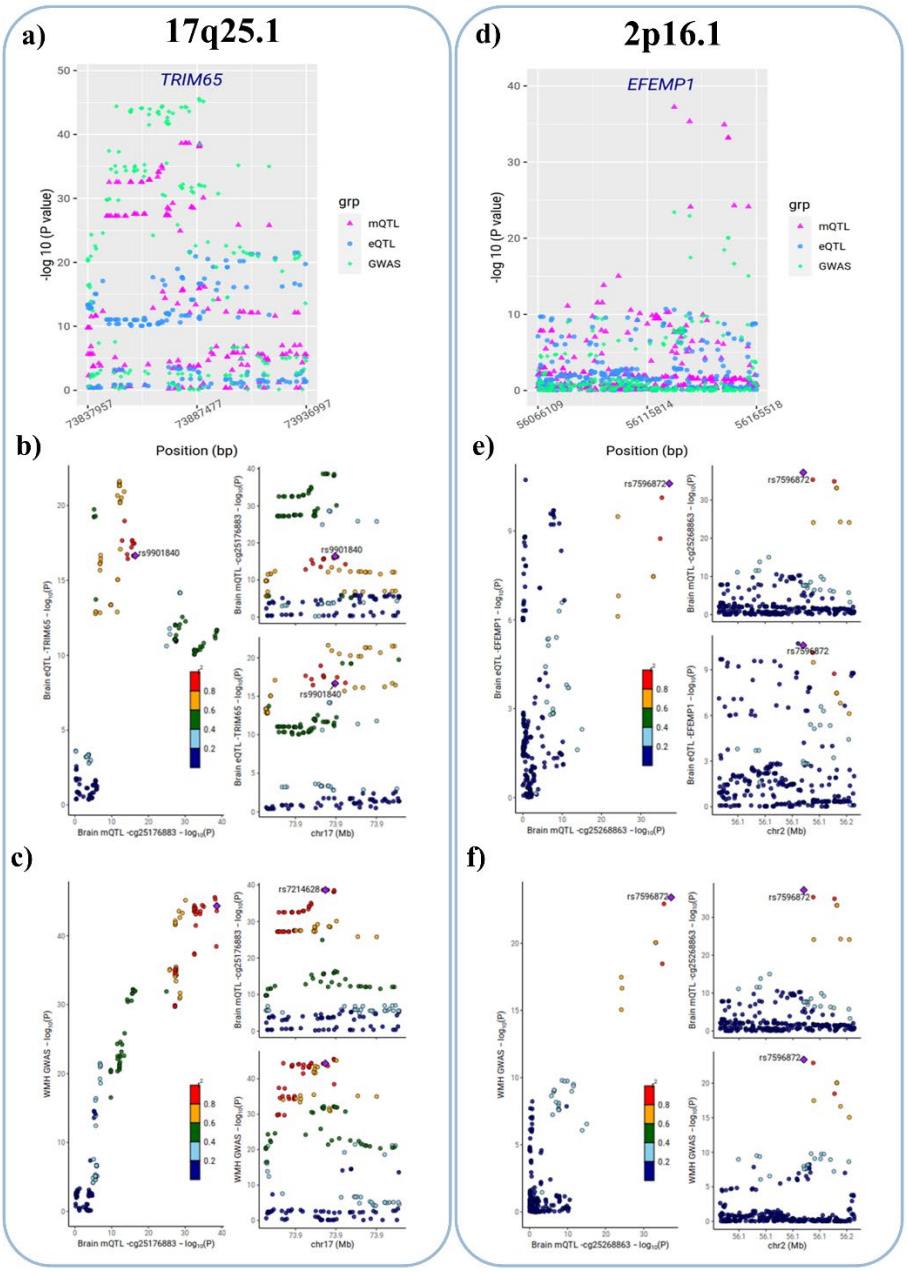
Supplementary Figure 5 Spearman correlation (P value) among target probes in ARIC subjects of European ancestry (ARIC-EA) and of African ancestry (ARIC-AA).

The color scale corresponds to strength and direction of correlation, where negative correlations are blue and positive correlations are red (see legend). Each box represents Spearman correlation coefficient between CpG in the row and CpG in the column with *P* value below. *Spearman r with *P* value <0.05; **<7.58×10^{-4} (0.05/66 CpG pairs).

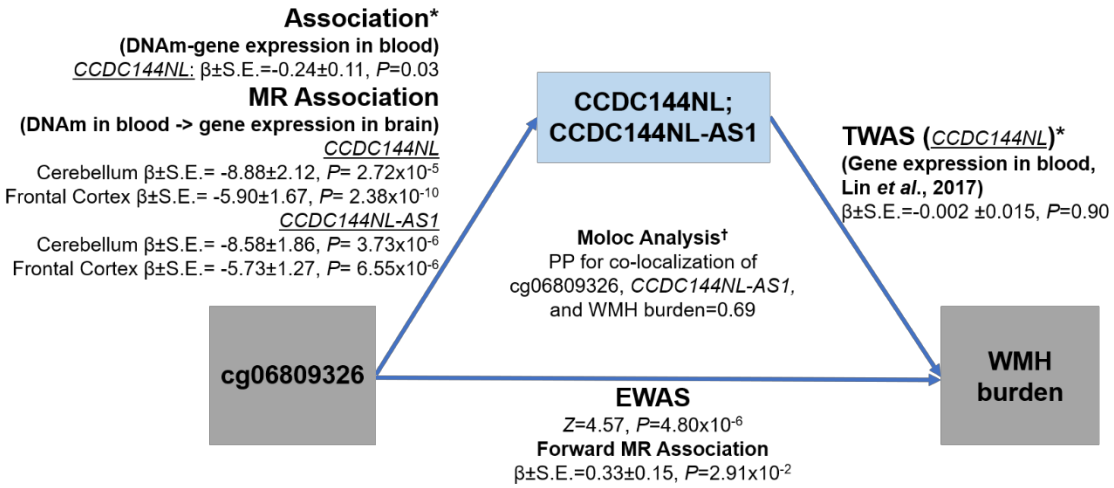


Supplementary Figure 6 Manhattan plots of genome-wide associations for target CpGs.

SNPs with imputation quality (R^2<0.3) or minor allele frequency <0.01 are excluded. Red and blue lines indicate the genome-wide significance threshold (*P* value < $5x10^{-8}$) and suggestive threshold (*P* value < $1x10^{-5}$), respectively. Dots are plotted by –log10 (*P* value) on y-axis along with their chromosomal positions on x-axis.



Supplemental Figure 7 Multiple trait colocalization and LocusCompare plots for the selected GWAS loci. Left column shows three plots for associations at 17q25.1 (cg25176883 and *TRIM65*) and right column shows three plots for associations at 2p16.1 (cg25268863 and *EFEMP1*). a) and d) are colocalization scatter plots showing brain mQTL (magenta), brain eQTL (cyan), and WMH burden GWAS (neon green) associations against chromosomal position on x-axis. In b) and e), left panels show scattered -log 10 *P* values of brain mQTL associations on x-axis against -log 10 P values of brain eQTL associations on y-axis; while two mini plots on the right show the mQTL and eQTL associations separately. In c) and f), brain mQTL associations are plotted against the WMH burden GWAS associations on y-axis. LocusCompare plots shown the LD structure among colocalization SNPs in colors, from low (blue) to high LD (red). The sentinel SNP for the cg25176883 and *TRIM65* colocalization locus in Supplementary Table 16, rs34974290, is in LD with the sentinel SNP in plot b) rs9901840 (r^2=0.57) and rs7214628 (r^2=0.94).



Supplementary Figure 8 Overview of associations in relation to cg06809326.

DNAm: DNA methylation, MR: Mendelian randomization, EWAS: epigenome-wide association study, PP: Bayesian posterior probability, WMH: white matter hyperintensities.

*These studies (FHS and RS) do not include *CCDC144NL-AS1* expression data.

†ROSMAP DLPFC expression quantitative trait loci (eQTL) data does not include *CCDC144NL* data.

SUPPLEMENTARY TABLES

Supplementary Table 1 Demographic characteristics in participating cohorts

Study		Ancestry	No. Subjects	Female (n(%))	Age (mean±S.D.)	Body mass index (kg/m^2, mean±S.D.)	Hypertension (n(%))	Current smoking (n(%))	WMH burden* (median[Q1, Q3])
Discovery (n=5,715)									
ARIC		EA	906	519(57.3%)	63.1±4.4	26.4±4.6	306(33.8%)	162 (17.9%)	0.69 [0.69,1.10]
		AA	639	413(64.8%)	61.8±4.5	29.6±5.4	412(64.5%)	127 (19.9%)	0.69 [0.69,1.10]
BBMRI		EA	442	255(57.7%)	64.7±5.8	27.4±3.9	279(63.1%)	63 (14.2%)	1.40 [1.03,1.92]
CARDIA		EA, AA	277 (167,110)§	130(46.9%)	53.9±4.1	29.1±5.2	95(34.3%)	95 (34.7%)	2.74 [2.46, 3.06]
CHS		EA, AA	378 (189,189)§	153(40.5%)	74.6±5.4	27.5±5.0	240(63.5%)	49 (13.0%)	1.10 [0.69, 1.38]
FHS Offspring	FLAIR	EA	1,323	711(53.7%)	66.9±8.7	28.0±5.0	764(57.7%)	101 (7.6%)	1.18 [0.79,1.69]
	DSE		403	230(57.1%)	60.6±9.6	28.1±5.7	186(46.2%)	61 (15.1%)	0.46 [0.28,0.79]
GENOA		AA	356	259(72.8%)	66.3±7.19	31.0±5.98	291(81.7%)	50 (14.0%)	2.10 [1.81, 2.56]
LBC1936		EA	230	107(46.5%)	72.1±.52	27.7±4.51	190(82.6%)	13 (5.6%)	2.04 [1.33, 2.76]
RS		EA	547	291(53.2%)	62.6±7.3	27.2±4.1	229(41.9%)	113 (20.7%)	1.34 [0.97,1.77]
SHIP-TREND		EA	214	117(54.7%)	49.7±13.3	27.2±4.1	84(39.3%)	42 (19.6%)	0.74 [0.59,0.93]
Replication (n=4,017)									
ADNI		EA	387	187(48.3%)	74.1±7.0	27.2±4.7	274(70.8%)	158 (40.8%)	1.29 [0.58,2.01]
FHS 3rd Generation		EA	864	454(52.5%)	47.1±7.8	27.6±5.1	145(16.8%)	78(9.0%)	0.34 [0.18,0.58]
Rhineland Study		EA	2,066	1,187(57.5%)	54.1±13.9	25.5 ±4.1	735 (35.6%)	280(13.6%)	0.40 [0.21,0.80]
OATS		EA	81	48(59.3%)	70.5±5.5	26.9 ±4.8	58 (72.5%)	0 (0%)	8.31 [7.83, 8.85]
BRIDGET		EA	619	379 (61.1%)	68.13±13.56	NA	426 (67.7%)	0 (0%)	NA

AA: African ancestry, ADNI: Alzheimer's Disease Neuroimaging Initiative, ARIC: Atherosclerosis Risk in Communities, BBMRI: Biobanking and BioMolecular resources Research Infrastructure, BRIDGET: BRain Imaging, cognition, Dementia and next generation GEnomics: a Transdisciplinary approach to search for risk and protective factors of neurodegenerative disease, CARDIA: Coronary Artery Risk Development in Young Adults, CHS: Cardiovascular Health Study, EA: European ancestry, FHS: Framingham Heart Study, FLAIR: Fluid-attenuated Inversion Recovery study, DSE: dual spin-echo study, GENOA: Genetic Epidemiology Network of Arteriopathy, IQR: interquartile range, LBC: Lothian Birth Cohort 1936, OATS: Older Australian Twin Study, RS: Rotterdam Study, and SHIP: Study of Health in Pomerania, S.D.: standard deviation, WMH: white matter hyperintensities, Q1 and Q3 indicate 25th and 75th percentiles.

*white matter hyperintensities (WMH) burden is defined as log(WMH+1). For BRIDGET study, prevalence of extreme SVD is shown.

§ CHS and CARDIA performed pooled ancestry analyses. The numbers reflect the number of EA and AA, respectively

Supplementary Table 2 Replications of single CpGs associated with white matter hyperintensities ($P<1x10^{-5}$)

CpG	Chr:Position (hg19)	Nearest Gene	Discovery Set								First replication set						Second replication set			Meta-analysis of Discovery and First Replication Sets							
			N	Z	P	Q	FDR	Direction	χ^2_{Het}	P_{Het}	N	Z	P	Direction	χ^2_{Het}	P_{Het}	N	Beta (SE)	P	N	Z_{fixed}	P_{fixed}	Direction	χ^2_{Het}	P_{Het}	P_{RE}	I^2
cg24202936	11:50257256	*SEPTIN7P11*	5,359	5.38	7.58x10^{-8}	0.03	0.04	++++++?+++	10.76	0.29	3,398	0.10	0.92	--+-	3.10	0.38	603	-0.08±0.06	0.19	8,757	4.27	1.99E-05	-++++++?++++-+-	24.57	0.03	1.99E-06	40.3
cg17417856	19:50191637	*PRMT1;ADM5*	4,917	-4.95	7.42x10^{-7}	0.15	0.28	--?---?----	1.48	0.99	3,398	1.19	0.23	--+-	7.73	0.05	611	-0.01±0.04	0.84	8,315	-3.04	2.34E-03	---?---?-----+-	25.86	0.01	2.24E-05	57.4
cg01506471	7:3990479	*SDK1*	5,359	-4.81	1.52x10^{-6}	0.21	0.30	------?-+--	7.38	0.60	3,398	-1.06	0.29	----	0.35	0.95	616	-0.01±0.06	0.90	8,757	-4.42	9.86E-06	-------?-+------	12.43	0.49	1.38E-05	19.4
cg14547240	4:15428750	*C1QTNF7*	5,359	-4.71	2.48x10^{-6}	0.25	0.30	------?----	6.68	0.67	3,398	-0.04	0.97	+--+	0.68	0.88	619	-0.02±0.06	0.72	8,757	-3.71	2.09E-04	+------?-------+	15.79	0.26	5.09E-03	48.0
cg21547371	3:52869521	*MUSTN1*	5,359	-4.65	3.30x10^{-6}	0.25	0.30	------?-+--	4.27	0.89	3,398	-1.71	0.09	+---	0.94	0.82	616	-0.04±0.07	0.51	8,757	-4.70	2.55E-06	+------?-+-----	7.64	0.87	1.77E-05	0.0
cg03116124	1:231293208	*TRIM67*	5,129	-4.64	3.54x10^{-6}	0.25	0.31	----?-?----	12.19	0.14	1,332	-0.98	0.33	--?-	2.66	0.26	590	0.05±0.06	0.39	6,461	-4.57	4.78E-06	-----?-?-----?-	16.38	0.13	3.77E-06	50.8
cg06809326	17:20799526	*CCDC144NL-AS1*	5,359	4.57	4.80x10^{-6}	0.28	0.34	++++++?-+++	7.49	0.59	3,398	0.65	0.51	++++	1.50	0.68	619	0.08±0.07	0.27	8,757	3.99	6.75E-05	+++++++?-++++++	14.46	0.34	3.24E-05	9.6
cg13476133	7:44185646	*GCK*	5,359	4.55	5.46x10^{-6}	0.28	0.36	++++++?+++	14.26	0.11	1,332	1.12	0.26	++?+	2.69	0.26	612	0.18±0.11	0.09	6,691	4.57	4.88E-06	+++++++?++++++?+	18.00	0.12	9.40E-05	60.7
cg14133539	9:104568	*FOXD4*	4,917	-4.53	5.98x10^{-6}	0.28	0.38	--?-+-?----	6.05	0.64	3,398	0.05	0.96	+-++	1.89	0.60	-	-	-	8,315	-3.45	5.64E-04	+--?-+-?-------++	16.54	0.17	2.29E-03	46.2
cg17577122	22:19511967	*CLDN5*	5,359	4.5	6.88x10^{-6}	0.29	0.40	++++++?+-++	5.36	0.80	1,332	0.86	0.39	++?+	0.41	0.82	618	-0.14±0.12	0.24	6,691	4.41	1.03E-05	+++++++?+-+++?+	7.29	0.84	6.67E-05	0.0
cg23586595	4:84034390	*PLAC8*	5,359	4.45	8.45x10^{-6}	0.32	0.43	++++++?+++	3.77	0.93	1,332	0.26	0.79		-	-	619	0.13±0.08	0.09	6,691	4.10	4.07E-05	-++++++?+++++?+	22.36	0.03	6.54E-06	62.3
cg23054394	3:140784675	*SPSB4*	5,359	-4.42	9.88x10^{-6}	0.34	0.45	------?+---	10.11	0.34	3,398	0.71	0.48	--++	2.59	0.46	606	0.06±0.14	0.66	8,757	-3.02	2.57E-03	-------?+-----++	23.65	0.03	4.94E-03	60.0

Abbreviations

AA: African ancestry, Chr: chromosome, EA: European ancestry, FDR: local false discovery rate value, N: number of subjects tested for the CpG, P: P value, Q: Q value, SE: standard error, Z: Z-score, $\chi 2_{Het}$: approximated chisq statistics for heterogeneity (Cochrane's Q), and P_{Het}: Heterogeneity P value. Alzheimer's Disease Neuroimaging Initiative (ADNI), ARIC (Atherosclerosis Risk in Communities) Study, BBMRI (Biobanking and BioMolecular resources Research Infrastructure), BRain Imaging, cognition, Dementia and next generation GEnomics: a Transdisciplinary approach to search for risk and protective factors of neurodegenerative disease (BRIDGET) study, CHS (Cardiovascular Health Study), CARDIA (Coronary Artery Risk Development in Young Adults) Study , Framingham Heart Study (FHS) Offspring Study FLAIR (Fluid-attenuated Inversion Recovery), DSE (dual spin-echo) sub-studies, FHS 3rd generation substudy, Genetic Epidemiology Network of Arteriopathy (GENOA), Lothian Birth Cohort 1936 (LBC1936), Rotterdam Study (RS), Study of Health in Pomerania (SHIP), and Older Australian Twin Study (OATS).
Z and Pfixed: fixed-effect based meta-analysis Z score and P value using METAL software[44]; P_{RE}: Han and Eskin (2010) random-effect based meta-analysis P value from METASOFT software.[45]

Signs in direction column indicates in the order of ARIC AA, ARIC EA, BBMRI, RS3, LBC1936, CHS, GENOA, CARDIA, SHIP, FHS_dse, and FHS_flair for the discovery set; indicate ADNI, FHS_3rd, Rhineland Study, and OATS for the first replication set; and indicate ADNI, ARIC AA, ARIC EA, BBMRI, RS3, LBC1936, CHS, GENOA, CARDIA, SHIP-Trend, FHS_dse, FHS_flair, FHS_3rd, Rineland study, and OATS for the meta-analysis, respectively.

Supplementary Table 3 Associations of target CpGs in hypertensive and normotensive subjects

CpG	CHR	MAPINFO	Nearest Gene	Hypertensive				Normotensive			
				N	*Z*	*P*	Direction*	N	*Z*	*P*	Direction*
cg24202936	11	50257256	*SEPTIN7P11*	2,785	3.65	$2.58x10^{-4}$	++++++?++-+	2,516	2.99	$2.78x10^{-3}$	-++++?++++
cg17417856	19	50191637	*PRMT1;C19orf76*	2,506	-3.99	$6.76x10^{-5}$	--?---?----	2,354	-2.34	$1.94x10^{-2}$	--?--?++--
cg01506471	7	3990479	*SDK1*	2,785	-2.34	$1.94x10^{-2}$	-+----?-+--	2,516	-2.86	$4.25x10^{-3}$	--+-+?-++-
cg14547240	4	15428750	*C1QTNF7*	2,785	-2.29	$2.21x10^{-2}$	------?+--+	2,516	-3.22	$1.31x10^{-3}$	----+?-+--
cg21547371	3	52869521	*MUSTN1*	2,785	-3.74	$1.85x10^{-4}$	---+--?----	2,516	-2.48	$1.31x10^{-2}$	+----?-+--
cg03116124	1	231293208	*TRIM67*	2,595	-2.30	$2.14x10^{-2}$	----?-?+---	2,516	-2.83	$4.71x10^{-3}$	+----?----
cg06809326	17	20799526	*CCDC144NL; CCDC144NL-AS1*	2,785	2.95	$3.17x10^{-3}$	+-++++?--++	2,516	2.76	$5.80x10^{-3}$	-++++?-+++
cg13476133	7	44185646	*GCK*	2,785	3.10	$1.95x10^{-3}$	++++++?-+++	2,516	2.71	$6.84x10^{-3}$	++-++?++-+
cg14133539	9	104568	*FOXD4*	2,506	-2.66	$7.75x10^{-3}$	--?-+-?----	2,354	-0.93	$3.53x10^{-1}$	--?-+?-+--
cg17577122	22	19511967	*CLDN5*	2,785	3.19	$1.44x10^{-3}$	+++-++?+-++	2,516	2.71	$6.79x10^{-3}$	-++++?++-+
cg23586595	4	84034390	*PLAC8*	2,785	2.03	$4.27x10^{-2}$	-+++++?++-+	2,516	4.23	$2.34x10^{-5}$	+++++?++++
cg23054394	3	140784675	*SPSB4*	2,785	-3.50	$4.72x10^{-4}$	------?----	2,516	-3.15	$1.62x10^{-3}$	+-+--?+---

CHR: chromosome, MAPINFO: position in bp (hg19), N: sample size

* The signs indicate the association directions in Atherosclerosis Risk in Communities (ARIC) African ancestry (AA), ARIC European ancestry (EA), Biobanking and BioMolecular resources Research Infrastructure (BBMRI), Rotterdam study (RS),Lothian Birth Cohort 1936 (LBC1936), Cardiovascular Health Study (CHS), Genetic Epidemiology Network of Arteriopathy (GENOA), Coronary Artery Risk Development In young Adults (CARDIA), Study of Health in Pomerania (SHIP), Framingham Heart study (FHS) DSE (dual spin-echo) and FLAIR (Fluid-attenuated Inversion Recovery) studies, respectively.

Supplementary Table 4 Associations of target CpGs in subjects of European and African ancestry

CpG	CHR	MAPINFO	Nearest Gene	European Ancestry				African Ancestry			
				N	*Z*	*P*	Direction*	N	*Z*	*P*	Direction†
cg24202936	11	50257256	*SEPTIN7P11*	4,065	4.67	3.03E-06	+++++++	639	2.83	4.66E-03	+?
cg17417856	19	50191637	*PRMT1;C19orf76*	3,623	-4.27	1.92E-05	-?-----	639	-1.66	0.10	-?
cg01506471	7	3990479	*SDK1*	4,065	-3.90	9.47E-05	----+--	639	-1.80	0.07	-?
cg14547240	4	15428750	*C1QTNF7*	4,065	-4.55	5.35E-06	-------	639	-0.46	0.65	-?
cg21547371	3	52869521	*MUSTN1*	4,065	-3.54	3.97E-04	----+--	639	-1.76	0.08	-?
cg03116124	1	231293208	*TRIM67*	3,835	-3.62	2.93E-04	---?---	639	-1.37	0.17	-?
cg06809326	17	20799526	*CCDC144NL; CCDC144NL-AS1*	4,065	4.21	2.52E-05	+++++++	639	2.61	9.01E-03	+?
cg13476133	7	44185646	*GCK*	4,065	4.38	1.17E-05	+++++++	639	0.05	0.96	+?
cg14133539	9	104568	*FOXD4*	3,623	-3.18	1.47E-03	-?-+---	639	-2.66	7.88E-03	-?
cg17577122	22	19511967	*CLDN5*	4,065	3.98	6.94E-05	++++-++	639	1.58	0.11	+?
cg23586595	4	84034390	*PLAC8*	4,065	4.21	2.53E-05	+++++++	639	0.18	0.86	+?
cg23054394	3	140784675	*SPSB4*	4,065	-4.15	3.27E-05	-------	639	-0.89	0.38	-?

CHR: chromosome, MAPINFO: position in bp (hg19), N: sample size

* The signs indicate the association directions in Atherosclerosis Risk in Communities (ARIC) study, Biobanking and BioMolecular resources Research Infrastructure (BBMRI), Rotterdam study (RS),Lothian Birth Cohort 1936 (LBC1936), Cardiovascular Health Study (CHS), Coronary Artery Risk Development In young Adults (CARDIA) study, Study of Health in Pomerania (SHIP), Framingham Heart study (FHS) DSE (dual spin-echo) and FLAIR (Fluid-attenuated Inversion Recovery) studies, respectively. † The signs indicate the association directions in ARIC study and Genetic Epidemiology Network of Arteriopathy (GENOA) study.

Supplementary Table 5 Single CpG association with white matter hyperintensities stratified by hypertension status ($P<1x10^{-5}$)

CpG	CHR	MAPINFO	Nearest Gene	Discovery						First replication set				Second replication set‡		
				N	*Z*	*P*	*Q*	FDR	Direction*	N	*Z* score	*P*	Direction†	N	*Beta±SE*	*P*
Hypertensive Subgroup																
cg23570810	11	315102	*IFITM1*	2,785	-4.764	1.90E-06	0.44	0.42	------?----	1,212	1.739	0.08	+-+-	619	0.004±0.04	0.91
cg05166317	7	158857428	*VIPR2*	1,835	-4.746	2.07E-06	0.44	0.43	+-----?-+??	1,067	0.678	0.50	+?+-	619	0.07±0.07	0.28
cg24784730	12	125225655	*SCARB1*	2,785	4.623	3.79E-06	0.54	0.72	++++++?++++	419	-0.427	0.67	+-??			
cg07676167	8	124253055	*C8orf76*	2,785	4.492	7.07E-06	0.66	0.95	++++++?++++	477	-0.666	0.51	--?+			
cg25317585	13	102568886	*FGF14*	2,785	4.471	7.80E-06	0.66	0.98	++++++?-+++	1,212	0.626	0.53	+--+	615	-0.10±0.10	0.30
cg14392283	8	144103587	*LY6E*	2,785	-4.419	9.93E-06	0.70	1.00	------?----	1,212	2.485	0.01	++++	619	0.11±0.06	0.06
Normotensive Subgroup																
cg06450373	5	19766434	*CDH18*	2,354	-5.405	6.48E-08	0.03	0.04	--?--?----	2,157	0.700	0.48	--+?			
cg18950108	6	30920171	*DPCR1*	2,516	-4.951	7.37E-07	0.15	0.40	-----?----	2,181	-1.284	0.20	-+--			
cg11706163	5	39364495	*C9*	2,581	-4.551	5.35E-06	0.60	0.77	----------	2,181	-0.967	0.33	--+-	584	0.03±0.07	0.69
cg19719207	4	121839493	*PRDM5*	2,516	-4.527	5.98E-06	0.60	0.77	-----?-+--	2,181	1.140	0.25	++++	615	-0.06±0.05	0.28
cg26629184	16	15149018	*NTAN1*	2,516	4.491	7.08E-06	0.60	0.77	+++++?+++-	2,181	-0.384	0.70	+-+-	614	0.05±0.05	0.37

CHR: chromosome, MAPINFO: position in bp (hg19), N: sample size, FDR: false discovery rate, SE: standard error, extreme-SVD: extremes of cerebral small vessel disease

* The signs indicate the association directions in Atherosclerosis Risk in Communities (ARIC) African ancestry (AA), ARIC European ancestry (EA), Biobanking and BioMolecular resources Research Infrastructure (BBMRI), Rotterdam study (RS),Lothian Birth Cohort 1936 (LBC1936), Cardiovascular Health Study (CHS), Genetic Epidemiology Network of Arteriopathy (GENOA), Coronary Artery Risk Development In young Adults (CARDIA), Study of Health in Pomerania (SHIP), Framingham Heart study (FHS) DSE (dual spin-echo) and FLAIR (Fluid-attenuated Inversion Recovery) studies, respectively.

† The signs indicate the association directions in Alzheimer's Disease Neuroimaging Initiative (ADNI), FHS (3rd generation), Rhineland study, and Older Australian Twin Study (OATS), respectively.

‡ Association between extreme-SVD and Hi-seq bisulfite sequencing DNA methaylation level in the BRain Imaging, cognition, Dementia and next generation GEnomics: a Transdisciplinary approach to search for risk and protective factors of neurodegenerative disease (BRIDGET) study

Supplementary Table 6 Regulatory features at target CpGs

CpG	CHR	POS (hg19)	RegulomeDB			GeneHancer DB version									
			Rank	Score	Transcriptional Factor (TF) Motifs	GenHancer ID	Location	Size (kB)	GeneHancer (GH) Type	GH score*	GH Sources	TSS distance	TF binding Sites	Gene Targets (Gene-GeneHancer Scores)†	Total score ‡
cg24202936	11	50257256	2b	0.93	KLF7(TSTFF495693),SP1 (TSTFF618931), SP1 (H1-hESC,TSTFF587073)	GH11J050297	chr11:50256572-50258371	1.8	Promoter/ Enhancer	1	Ensembl, ENCODE	+0.2	20 TFs (PKNOX1, MTA1, NRF1, MGA, ATF1, POLR2A, MYC, MNT, CREM, SMAD5, SUPT5H, VEZF1, KDM1A, MEIS2, RFX5, CHD2, L3MBTL2, ZFX, EMSY, GMEB1)	ENSG00000255042 (615.7), LINC02750 (600.7), SEPTIN7P11 (600.7), lnc-OR4C12-9 (600.7), TRIM51DP (8.3), ENSG00000205035 (5.3), ENSG00000255001 (0.4), OR4C45 (0.1)	592.84
cg17417856	19	50191637	4	0.61		GH19J049686	chr19:50190224-50195848	5.6	Promoter/ Enhancer	1.9	EPDnew,Ensembl,ENCODE,CraniofacialAtlas	+14.0	41 TFs (CTCF, PRDM10, KDM1A, POLR2A, RBFOX2, PATZ1, NR2C1, ZNF501, GABPA, ZNF341, ZBTB20, RAD21, ZBTB10, ZNF316, USF1, NR2C2, ZBTB48, ZBTB17, RBM22, ZFHX2, BHLHE40, GABPB1, SPI1, PRDM4, SUPT5H, SIN3A, STAT5A, YY1, ZNF143, ZBTB33, TAF1, SMARCA4, SP3, NRF1, CHD2, ZBTB40, GTF2F1, TBP, SP1, MZF1, ZNF121)	CPT1C (615.9), AB372690 (600.7), ADM5 (600.7), ENSG00000268677 (600.7), HSALNG0127023 (600.7), NONHSAG026259.2 (0.7), PRMT1 (0.4)	0.79
cg01506471	7	3990479	5	0.13											
cg14547240	4	15428750	4	0.61		GH04J015426	chr4:15428554-15430823	2.3	Enhancer	0.7	Ensembl,ENCODE,dbSUPER	+88.2	3 TFs (ATF7, CEBPB, ATF2)	C1QTNF7-AS1 (600.7), C1QTNF7 (8.5), LOC107986185 (0.4), CC2D2A (0.3)	6.12
cg21547371	3	52869521	5	0.13		GH03J052833	chr3:52867960-52871576	3.6	Promoter/ Enhancer	1.6	EPDnew,Ensembl,ENCODE,CraniofacialAtlas,dbSUPER	-0.7	10 TFs (CEBPA, CEBPB, ZBTB40, POLR2A, EBF1, SMARCA4, TEAD3, YY1, HLF, CEBPG)	ENSG00000243696 (600.7), MUSTN1 (600.7), STIMATE (600.2), ITIH4 (47.8), RFT1 (9.1), ITIH1 (4.8), lnc-TMEM110-MUSTN1-1-001 (0.7), piR-32342 (0.7), STIMATE-MUSTN1 (0.2)	931.16
cg03116124	1	231293208	3a	0.77	HSFY2(TSTFF397526,TSTFF799400)										
cg13476133	17	20799526	4	0.61		GH07J044145	chr7:44185081-44187920	2.8	Promoter/ Enhancer	2.1	EPDnew,Ensembl,ENCODE,CraniofacialAtlas,dbSUPER	+51.3	96 TFs (LEF1, IKZF1, JUND, PHF21A, PKNOX1, REST, ZNF740, EZH2, DPF2, HDAC1, SOX6, TCF12, EP300, MNT, MGA, CBX3, UBTF, ZNF184, HMBOX1, POLR2A, MTA3, EGR1, CTBP1, NFE2, FOXK2, HDGF, HDAC2, CTCF, ZNF316, PHB2, ZBTB11, MYC, EP400, TRIM28, TRIM24, NCOR1, MAFG, NFIC, STAT5A, RELA, BCOR, KLF9, FOXM1, TFDP1, HDAC8, MAFK, ZNF592, NCOA1, NR2F1, ZNF395, SAFB, SKIL, TAL1, CREM, ATF4, NR2F6, SUPT5H, BHLHE40, HLTF, VEZF1, NRF1, PRDM10, ID3, MAX, NBN, KDM1A, ZEB2, ZBTB33, MEIS2, SMARCA4, MTA2, KLF16, MITF, ARNT, KLF1, SMARCE1, RNF2, GTF2F1, RAD51, RB1, GABPB1, RLF, L3MBTL2, TEAD4, CBFA2T2, ATF7, NFXL1, ZFX, ZNF148, DACH1, CBFA2T3, ARID1B, PML, GMEB1, ATF1, ZNF83)	GCK (609.2), LOC105375258 (600.7), MYL7 (600.4), POLD2 (10.6)	1258
cg14133539	7	44185646	5	0.05											
cg06809326	9	104568	2b	0.91	NRF1(TSTFF451676,TSTF										

CpG	CHR	POS (hg19)	RegulomeDB			GeneHancer DB version									
			Rank	Score	Transcriptional Factor (TF) Motifs	GenHancer ID	Location	Size (kB)	GeneHancer (GH) Type	GH score*	GH Sources	TSS distance	TF binding Sites	Gene Targets (Gene-GeneHancer Scores)†	Total score ‡
					F168520),PAX5(TSTFF141184,TSTFF674641,TSTFF609420)										
cg17577122	22	19511967	2b	0.51	EGR1(TSTFF702532,TSTFF389778),EGR2 (TSTFF899152)	GH22J019521	chr22:19508524-19512924	4.4	Promoter/ Enhancer	1.7	EPDnew,Ensembl,ENCODE,CraniofacialAtlas	+4.3	24 TFs (HNRNPL, ZIC2, HNRNPUL1, EZH2, ZBTB20, POLR2A, MTA3, ZBED1, KDM4B, NFIC, ZFHX2, STAT5A, MYNN, ZNF621, SAFB, CBX1, SUPT5H, ZBTB33, ZNF184, TBP, EGR2, BACH1, SP1, POLR2B)	CLDN5 (600.7), ENSG00000287652 (600.7), UFD1 (37.4), MED15 (15.7), RPL7AP70 (13.2), DGCR11 (12.0), ENSG00000185065 (10.1), LINC00895 (0.3), CDC45 (0.3)	1010
cg23586595	4	84034390	3a	0.60	AKR1B1, AR (TSTFF225703)	GH04J083107	chr4:84028384-84037718	9.3	Promoter/ Enhancer	2.4	EPDnew,FANTOM5, Ensembl,ENCODE,CraniofacialAtlas,dbSUPER	+25.2	140 TFs (HNRNPL, CREB1, CTCF, TEAD4, PRDM10, REST, POLR2A, BACH1, TARDBP, FOXA1, LARP7, STAT3, ZNF143, RELA, ZIC2, PKNOX1, PATZ1, SMAD5, RAD21, TCF12, IKZF2, ZNF501, TRIM22, EZH2, ELF4, ZNF600, YY1, IRF3, ZNF341, SCRT2, EED, SOX13, ZBTB44, ZBTB20, SIN3A, SP2, ELK1, SP1, ELK4, RFX1, DPF2, RELB, JUND, ZXDB, SMARCA5, RUNX3, ZBTB10, GATAD2B, CTBP1, ZKSCAN8, ZNF444, ASH2L, BCLAF1, FOXA2, CBFB, E2F1, E4F1, ZBTB26, ZNF592, SMC3, GABPA, KLF4, GTF2F1, ZNF217, MYC, SCRT1, MAX, ARNT, ELF1, ZBTB11, NFIC, NBN, HDGF, KLF8, IKZF1, BMI1, SAP130, ZNF777, ZNF189, KLF9, TBP, GATA3, JUNB, ZNF335, CHD1, STAT1, SPI1, PRDM4, IRF4, RCOR1, EBF1, DRAP1, FOSL2, RB1, ZNF366, HOMEZ, KLF5, ZNF687, RFX5, STAT5A, ZBTB7B, PRDM6, ATF2, ZNF660, TRIM28, TBX21, ZNF2, ARID4B, ATF7, TAF1, SMARCA4, ZFP37, MTA2, NR2F6, MTA3, SP3, KLF1, BHLHE40, NR2F1, MNT, GABPB1, ZNF24, NFYC, L3MBTL2, ZNF362, NKRF, ZNF384, SREBF1, PPARG, MZF1, SP7, BCL3, ZNF121, GLIS2, RXRA, ETV6, CREM, MLLT1, ZBTB33, ZEB2)	PLAC8 (625.1), HELQ (3.5), THAP9 (2.8), LIN54 (2.2), COQ2 (2.1), HPSE (1.5), lnc-LIN54-3 (0.7), lnc-COPS4-1 (0.4), COPS4 (0.2)	1485.9
cg23054394	3	140784675	5	0.13	ALX1(TSTFF414516)										

TSS: Transcriptional starting site

RegulomeDB

Score: Supporting data

1a: expression quantitative loci (eQTL) + Transcriptional factor (TF) binding + matched TF motif + matched DNase Footprint + DNase peak; 1b: eQTL + TF binding + any motif + DNase Footprint + DNase peak; 1c: eQTL + TF binding + matched TF motif + DNase peak; 1d: eQTL + TF binding + any motif + DNase peak;+A71 1e: eQTL + TF binding + matched TF motif; 1f: eQTL + TF binding / DNase peak; 2a: TF binding + matched TF motif + matched DNase Footprint + DNase peak; 2b: TF binding + any motif + DNase Footprint + DNase peak; 2c: TF binding + matched TF motif + DNase peak ;3a: TF binding + any motif + DNase peak; 3b: TF binding + matched TF motif; 4: TF binding + DNase peak; 5: TF binding or DNase peak; 6: Motif hit; 7: Other

The RegulomeDB probability score is ranging from 0 to 1, with 1 being most likely to be a regulatory variant.

GeneHancer

*GH score is a confidence score which is computed based on a combination of evidence annotations: (1) Number of sources; (2) Source scores; (3) TFBSs (from ENCODE). Sources are the ENCODE project, Ensembl regulatory build, FANTOM5 atlas of active enhancers, VISTA Enhancer Browser, dbSUPER, EPDnew, UCNEbase, CraniofacialAtlas.

†The gene-GeneHancer score for each gene target is a combined score for each proximal gene based on all methods (p-values for eQTLs and co-expression, log(observed/expected) for C-HiC, and distance-inferred probability).

‡ Total score - multiplication of the GeneHancer confidence score and the GeneHancer-gene association score, gene-GeneHancer distance (calculated between the GeneHancer midpoint and the gene TSS, positive for downstream and negative for upstream), number of genes having a TSS between the gene and the GeneHancer, and a list of all genes associated with the GeneHancer.

Supplementary Table 7 Previously reported epigenetic associations for target CpGs

Probe ID	Trait	Pubmed Reference ID	Source*
cg24202936	HIV infection	27105112	C
cg17417856	smoking	31536415, 31552803, 25556184, 27651444, 28686328	C and A
	age	28056824, 23177740	C
	Crohn's disease	27886173	C
	systolic blood pressure	29198723	C
	alcohol consumption per day	27843151	C
	Educational attainment	29086770	C
	chemotherapy for breast cancer	30867049	A
cg01506471	IL-13 treatment	26474238	C and A
cg14547240	rheumatoid arthritis	23334450	C
cg21547371	gestational age	25650246	C
	asthma	27942592	A
	obesity	29692867	A
cg03116124	age group	25888029	C
	Down Syndrome	29601581	A
	colorectal laterally spreading tumor	30183087	A
	adenoma	30183087	A
cg13476133	age	25282492	C
	B Acute Lymphoblastic Leukemia	26237075	A
	breast cancer	30866861	A
	Follicular thyroid carcinoma	28938489	A
cg14133539	HIV infection	27105112	C
	right insular surface area	30830696	A
cg06809326	sex	26500701	C
	Down Syndrome	29601581, 31843015	A
cg17577122	hepatocellular carcinoma (HCC)	23208076	A
	colorectal cancer	30183087	A
cg23586595	clear cell renal carcinoma	23526956	
	smoking	31536415, 31552803, 25424692, 27651444	C and A
	pancreatic ductal adenocarcinoma	24500968	C
	rheumatoid arthritis	23334450	C
	Behcet's disease	30863869	A
	psoriasis	30092825	A
cg23054394	chronic beryllium disease	30141971	A

*C: Epigenome-wide association study (EWAS) Catalog (beta), A: EWAS Atlas.

Epigenome-wide association studies (EWAS) traits associated with target CpGs (P<1.0x10^{-5}) are presented.

Supplementary Table 8 Differentially Methylated Regions (DMRs) associated with white matter hyperintensities burden

Nearest Gene	Chr	start (hg19)	end (hg19)	#probe	P_{Comb-P}	#probe	$P_{DMRCate}$	Distance from Gene (bp)	$P_{Replication1}$ *	$P_{Replication2}$ †
PRMT1	19	50191342	50191883	6	6.98x10-13	8	3.53x10-12	0	1.00	6.67x10-3
CCDC144NL-AS1; CCDC144NL	17	20799408	20799645	6	5.23x10-10	6	4.91x10-10	0	0.82	0.15
ABAT	16	8806359	8807044	13	5.00x10-8	13	3.81x10-6	0	1.91x10-2	2.72x10-4
BHMT2	5	78365647	78365802	5	3.19x10-6	7	6.64x10-6	0	1.47x10-3	1.00
SEPTIN7P11	11	50257256	50257497	2	4.20x10-4	6	6.64x10-6	0	0.14	1.00
VARS2	6	30882994	30883204	5	1.42x10-3	14	7.31x10-6	0	0.09	
ZNF311	6	28973318	28973525	11	2.22x10-5	17	6.21x10-5	0	1.00	
C11orf21	11	2321770	2322051	9	1.67x10-4	12	1.41x10-4	0	1.52x10-3	1.00
MRPL23	11	1990900	1991189	3	8.53x10-4	4	4.20x10-4	13071	0.39	
UBE3C	7	157092682	157092983	2	1.39x10-4	2	6.39x10-4	30618	0.68	0.13
CHORDC1	11	89956517	89957034	8	3.19x10-3	10	7.63x10-4	0	1.00	1.00
IZUMO1	19	49249932	49250253	3	9.24x10-4	7	8.30x10-4	0	1.35x10-3	1.00
YIF1B	19	38794554	38794846	6	2.37x10-4	7	1.64x10-3	0	0.11	0.75
KCNA6	12	4918169	4918392	3	4.97x10-2	5	2.33x10-3	0	0.29	1.00
C5orf66	5	134526124	134526270	4	9.67x10-4	4	2.33x10-3	0	0.32	3.19x10-2
SLC12A4	16	67997858	67998165	4	3.06x10-4	5	2.65x10-3	0	1.00	0.36
ZBTB44-DT	11	130185384	130185671	3	3.37x10-3	3	2.92x10-3	0	1.00	1.00
IGSF22	11	18727321	18727775	4	2.78x10-4	5	3.12x10-3	0	0.93	0.06
LINC01135	1	59280290	59280620	4	5.27x10-3	7	3.12x10-3	0	1.00	1.00
IFITM1	11	315102	315263	3	3.74x10-2	3	3.14x10-3	0	0.88	0.90
PPP2R2A	8	26047780	26048066	5	1.69x10-2	8	3.51x10-3	-100957	0.97	0.84
SH3PXD2A	10	105428385	105428819	5	2.55x10-4	5	3.57x10-3	0	0.73	1.00
KCNK10	14	88621424	88621697	4	2.46x10-2	4	3.57x10-3	24754	0.85	1.00
LMF1;LMF1-AS1	16	979488	979663	3	5.66x10-3	4	3.67x10-3	0	0.99	1.00
KIF12	9	116861288	116861533	4	2.22x10-3	4	4.26x10-3	0	0.33	1.00
ARHGEF10	8	1847999	1848144	2	1.12x10-2	2	4.88x10-3	0	0.16	0.98
ZBTB38	3	141087187	141087364	5	2.36x10-3	6	5.91x10-3	0	0.12	0.15
IFITM10	11	1769152	1769646	9	5.44x10-4	9	7.42x10-3	0	1.00	0.98
PARP9	3	122281881	122281976	3	2.60x10-2	3	8.38x10-3	0	1.00	1.00
ENPEP	4	111397134	111397402	6	1.32x10-2	7	8.39x10-3	0	0.87	4.65x10-2
SLC35F3	1	234367322	234367587	4	1.32x10-2	5	1.08x10-2	0	1.47x10-2	0.19
BNIP3	10	133793442	133793735	5	3.43x10-3	5	1.16x10-2	0	0.30	0.81
FBXO47	17	37123638	37123950	9	1.88x10-3	9	1.17x10-2	0	3.58x10-6	
CUEDC1	17	55962522	55962842	4	3.60x10-3	5	1.17x10-2	0	1.00	1.00
SLC45A4	8	142233427	142233706	2	2.04x10-2	2	1.20x10-2	0	0.55	5.84x10-3
MYCL	1	40388223	40388396	3	5.32x10-3	3	1.20x10-2	-20634	0.41	0.92
TCF7	5	133451383	133451626	3	4.71x10-2	8	1.45x10-2	0	0.74	0.68
KCTD16	5	143978290	143978421	4	7.95x10-3	4	1.60x10-2	113041	9.92x10-7	3.12x10-2
ZBTB46	20	62406428	62406722	3	6.34x10-3	3	1.99x10-2	0	0.71	1.00
KITLG	12	88974243	88974666	4	8.23x10-3	9	2.03x10-2	0	1.00	3.49x10-3
EDDM3A	14	21191754	21191860	2	3.48x10-2	2	2.06x10-2	-22185	0.64	1.00
LINC02346	15	26327895	26328093	3	4.78x10-2	3	2.54x10-2	29628	0.65	0.92
UCN3	10	5406890	5407120	8	7.76x10-3	8	2.61x10-2	0	2.78x10-2	1.00
NFIC	19	3387811	3388048	3	1.21x10-2	3	2.95x10-2	0	0.97	0.34
KLHDC7B	22	50984368	50984598	3	3.99x10-2	2	4.05x10-2	0	1.00	0.54
RASA3	13	114875170	114875413	4	3.28x10-2	2	4.84x10-2	0	1.00	0.88

P_{Comb-P}: combined p value after Sidak multiple testing adjustment, $P_{DMRCate}$: Benjamini-Hochberg FDR p-value from the CpGs constituting the significant region, regional differential methylation signals can be calculated using neighboring CpGs within 1000 base pairs (bp), More than one CpG resides in the region. For replication, PDMRCate is reported in the replication set 1 and 2.

*Associations in the replication set 1 includes Alzheimer's Disease Neuroimaging Initiative (ADNI), FHS (3rd generation), Rhineland study, and Older Australian Twin Study (OATS), respectively, were combined.

† Associations between extreme-SVD and Hi-seq bisulfite sequencing DNA methaylation level in the BRain Imaging, cognition, Dementia and next generation GEnomics: a Transdisciplinary approach to search for risk and protective factors of neurodegenerative disease (BRIDGET) study were combined.

Supplementary Table 9 Gene-set enrichment analysis of identified DMRs

Gene Ontology (GO) ID	N	DE	P.DE	FDR	Ontology type	Description
GO:0097677	8	2	7.56E-05	4.91E-03	molecular function	signal transducers and activators of transcription family protein binding
GO:2001034	12	2	2.30E-04	5.68E-03	biological process	positive regulation of double-strand break repair via nonhomologous end joining
GO:0051607	205	4	2.79E-04	5.68E-03	biological process	defense response to virus
GO:0035563	13	2	3.49E-04	5.68E-03	biological process	positive regulation of chromatin binding
GO:0044389	23	2	8.26E-04	1.07E-02	molecular function	ubiquitin-like protein ligase binding
GO:0002230	35	2	1.04E-03	1.13E-02	biological process	positive regulation of defense response to virus by host
GO:0008327	28	2	1.29E-03	1.20E-02	molecular function	methyl-CpG binding
GO:0004857	32	2	1.49E-03	1.21E-02	molecular function	enzyme inhibitor activity
GO:1900182	26	2	1.70E-03	1.23E-02	biological process	positive regulation of protein localization to nucleus
GO:0032024	42	2	5.82E-03	3.78E-02	biological process	positive regulation of insulin secretion
GO:0006302	81	2	6.98E-03	4.13E-02	biological process	double-strand break repair
GO:0043679	53	2	7.73E-03	4.19E-02	cellular component	axon terminus
GO:0005765	361	3	3.17E-02	0.15	cellular component	lysosomal membrane
GO:0042393	152	2	3.40E-02	0.15	molecular function	histone binding
GO:0071456	123	2	3.41E-02	0.15	biological process	cellular response to hypoxia
GO:0001650	134	2	3.85E-02	0.16	cellular component	fibrillar center
GO:0019899	353	3	4.98E-02	0.19	molecular function	enzyme binding
GO:0006888	235	2	0.05	0.20	biological process	endoplasmic reticulum to Golgi vesicle-mediated transport

N: number of genes in the gene set, DE: number of genes that are differentially methylated, P.DE: p-value for over-representation of the gene set, FDR: False discovery rate, calculated using the method of Benjamini and Hochberg (1995).

Supplementary Table 10 Multivariate epigenetic associations of white matter hyperintensities burden and blood pressure

BP trait	CoG	CHR	MAPINFO	Gene	Z.WMH	P.WMH	Z.BP	P.BP	$P_{metaUSAT}$
DBP	cg23291754	4	71767977	*MOBKL1A*	3.60	3.17E-04	-4.21	2.57E-05	2.38E-07
	cg24372586	6	30520094	*GNL1*	-3.86	1.16E-04	-3.49	4.77E-04	7.84E-07
SBP	cg00711496	14	103415873	*CDC42BPB*	-4.32	1.54E-05	-3.50	4.66E-04	1.99E-07
	cg04987734	19	50191497	*C19orf76;PRMT1*	3.64	2.72E-04	4.10	4.13E-05	3.09E-07
	cg00934987	17	56605468	*SEPT4*	-3.90	9.58E-05	-3.52	4.38E-04	1.07E-06
	cg18770635	22	50984368	*KLHDC7B*	-3.38	7.32E-04	-3.91	9.33E-05	1.68E-06

BP: blood pressure; DBP: diastolic BP; SBP: systolic BP; WMH: white matter hyperintensities. Z.WMH and P.WMH from WMH discovery epigenome-wide association study (EWAS), Z.BP and P.BP from BP EWAS[a]. PmetaUSAT: multivariate association *P* value of BP trait and WMH.

a. Richard, M. A., Huan, T., Ligthart, S., et al DNA methylation analysis identifies loci for blood pressure regulation. *Am J Hum Genet.* 2017; 101(6): 888-902.

Supplementary Table 11 Genetic Instruments for cg06809326 and cg24202936 and their associations with WMH burden

Target CpG	SNP	Chr	Position (hg19)	A1	A2	mQTL (Huan T, 2019, Nat Comm, n=4,170)				UKBB WMH GWAS (Persyn E, 2020, Nat Comm, n=18,381)			
						Freq_A1	beta	se	pval	Freq_A1	beta	se	pval
	rs11040621	11	49903163	T	C	0.12	-0.04	0.003	3.25E-41	0.15	0.02	0.013	0.10
	rs143272961	11	55342613	G	A	0.01	0.14	0.014	1.41E-23	0.03	0.06	0.033	0.08
	rs4312050	11	51253295	T	G	0.16	0.04	0.003	2.34E-37	0.16	-0.01	0.013	0.47
cg24202936	rs7126006	11	49770456	G	T	0.18	0.07	0.002	3.09E-171	0.18	0.01	0.013	0.51
	rs113796679	17	21062647	A	G	0.06	-0.05	0.005	6.50E-23	0.06	-0.01	0.020	0.51
	rs2993353	17	20719173	T	C	0.35	-0.03	0.003	3.76E-30	0.31	-0.02	0.010	2.72E-02
cg06809326	rs8067432	17	20827039	C	T	0.32	-0.05	0.002	4.17E-122	0.33	-0.01	0.010	0.28

Freq_A1 (frequency of allele 1 (A1)), beta and its standard error (SE), and P value (pval) are calculated for A1. Genetic instruments for the exposure are curated from mQTL (Huan T, 2019, Nat Comm), and pruned at r2<0.05. Loci non-existent or quality-controlled (low imputation quality or minor allele frequency) in UKBB GWAS data are replaced with a proxy (r2>0.8), if any.

Supplementary Table 12 Bi-directional Mendelian randomization (MR) analysis of target CpGs with at least 3 instruments

Exposure -> Outcome	No. of IVs	IVW			Weighted Median			MR Egger			ConMix test			Pleiotropy test (Egger Intercept test)			Heterogeneity test (IVW)			Heterogeneity test (MR Egger)			MR PRESSO P value	IV strength	
		beta	se	*P* value	beta	se	*P* value	beta	se	*P* value	beta	95% Lower CI	95% Upper CI	beta	se	*P* value	Q	df	*P* value	Q	df	*P* value		meanF (IVW)	I^2
Forward MR (CpG -> WMH)																									
cg06809326 -> WMH	3	0.33	0.15	$2.91x10^{-2}$	0.23	0.17	0.17	-0.59	0.72	0.56	0.33	-0.10	1.20	0.04	0.03	0.41	1.73	2	0.42	0.02	1	0.89	-	-	0.97
cg24202936 -> WMH	4	0.07	0.18	0.70	0.12	0.15	0.42	0.71	0.29	0.13	0.13	-1.36	0.13	-0.04	0.02	0.13	6.34	3	0.10	0.38	2	0.83	-	-	0.98
Backward MR (WMH-> CpG)																									
WMH -> cg24202936	8	-0.01	0.02	0.72	-0.003	0.03	0.92	0.04	0.09	0.66	-0.05	-0.14	0.10	-0.004	0.01	0.59	6.34	7	0.50	6.01	6	0.42	0.44	44.73	0.98
WMH -> cg17417856	8	0.01	0.02	0.38	0.01	0.02	0.52	-0.03	0.06	0.64	0.02	-0.05	0.10	0.003	0.004	0.49	2.47	7	0.93	1.92	6	0.93	0.92	44.73	0.98
WMH -> cg01506471	8	0.01	0.02	0.80	0.01	0.03	0.66	-0.14	0.08	0.14	0.03	-0.07	0.13	0.01	0.01	0.12	5.97	7	0.54	2.65	6	0.85	0.54	44.73	0.98
WMH -> cg14547240	8	0.03	0.02	0.15	0.02	0.03	0.49	-0.05	0.09	0.59	0.04	-0.04	0.15	0.01	0.01	0.36	4.94	7	0.67	3.98	6	0.68	0.67	44.73	0.98
WMH -> cg21547371	8	0.03	0.02	0.25	0.01	0.03	0.62	-0.09	0.09	0.35	0.03	-0.05	0.18	0.01	0.01	0.22	4.07	7	0.77	2.24	6	0.90	0.78	44.73	0.98
WMH -> cg03116124	8	0.01	0.02	0.72	-0.004	0.02	0.87	-0.02	0.07	0.79	-0.02	-0.07	0.09	0.002	0.01	0.71	4.80	7	0.68	4.65	6	0.59	0.67	44.73	0.98
WMH -> cg06809326	8	-0.01	0.03	0.76	-0.01	0.04	0.73	0.01	0.11	0.91	-0.01	-0.22	0.15	-0.002	0.01	0.84	2.11	7	0.95	2.07	6	0.91	0.95	44.73	0.98
WMH -> cg13476133	8	-0.01	0.01	0.55	-0.01	0.01	0.57	0.001	0.04	0.98	-0.02	-0.05	0.02	-0.001	0.003	0.86	4.25	7	0.75	4.21	6	0.65	0.76	44.73	0.98
WMH -> cg14133539	8	0.03	0.02	0.17	0.01	0.02	0.70	-0.06	0.07	0.43	0.01	-0.04	0.13	0.01	0.01	0.26	4.69	7	0.70	3.13	6	0.79	0.72	44.73	0.98
WMH -> cg17577122	8	0.005	0.003	0.15	0.005	0.005	0.30	0.004	0.01	0.80	0.01	0.005	0.02	0.0001	0.001	0.93	6.79	7	0.45	6.78	6	0.34	0.45	44.73	0.98
WMH -> cg23586595	8	0.01	0.01	0.12	0.01	0.01	0.11	0.0001	0.03	1.00	0.02	0.0003	0.03	0.001	0.002	0.68	2.79	7	0.90	2.60	6	0.86	0.92	44.73	0.98
WMH -> cg23054394	8	0.03	0.02	0.25	0.02	0.03	0.48	-0.10	0.09	0.30	0.07	-0.07	0.17	0.01	0.01	0.19	5.53	7	0.60	3.29	6	0.77	0.60	44.73	0.98

CI: confidence interval, I^2: Bowden statistic of instrument strength for the MR-Egger method. IVW: inverse variance-weighted, meanF: instrument strength for IVW method. No. IV: Number of instrumental variables (IVs) included. OR: Odds ratio, PMR and Beta (S.E.): P value, effect size (beta) and its standard error (S.E.) for the causal direction, primarily based on the IVW method, or in existence of pleiotropy ($P_{pleiotropy}$ <0.05) based on MR Egger method. $P_{pleiotropy}$: P value estimated from MR Egger intercept pleiotropy test, WMH: white matter hyperintensities.

Supplementary Table 13 Genetic Instruments for WMH burden and their associations with target CpGs

Trait	SNP	Chr	Position (hg19)	A1	A2	Exposure: UKBB WMH GWAS (Persyn E, 2020, Nat Comm, n=18,381)				Outcome: mQTL (Huan T, 2019, Nat Comm, n=4,170)						
										Freq_A1	cg06809326			cg24202936		
						Freq_A1	beta	se	pval		beta	se	pval	beta	se	pval
WMH	rs12202497	6	151020020	A	T	0.59	-0.06	0.010	2.06E-09	0.59	0.002	0.01	0.71	0.01	0.004	0.18
	rs12571024	10	105635679	T	C	0.53	0.06	0.010	4.96E-09	0.37	-0.003	0.01	0.60	-0.0003	0.005	0.94
	rs12615761	2	43103440	T	G	0.82	0.07	0.013	7.32E-09	0.83	-0.002	0.01	0.80	-0.004	0.01	0.44
	rs12921170	16	87227397	G	A	0.58	0.06	0.010	2.11E-09	0.43	-0.003	0.005	0.62	-0.005	0.004	0.26
	rs34974290	17	73888354	G	A	0.81	-0.11	0.012	4.03E-19	0.81	-0.002	0.01	0.72	-0.01	0.01	0.28
	rs3891038	7	100362411	A	G	0.71	0.06	0.011	2.66E-08	0.70	-0.005	0.01	0.42	0.001	0.005	0.91
	rs4525538	17	43141966	A	G	0.63	0.07	0.010	3.32E-13	0.39	0.004	0.01	0.42	0.01	0.004	0.18
	rs78857879	2	56135099	G	A	0.90	-0.12	0.016	1.91E-13	0.91	0.001	0.01	0.87	0.002	0.01	0.75

Freq_A1 (frequency of allele 1 (A1)), beta and its standard error (SE), and P value (pval) are calculated for A1. Genetic instruments for the exposure are curated from UKBB WMH GWAS (Persyn E, 2020, Nat Comm), and pruned at $r2<0.05$. Loci non-existent or quality-controlled (low imputation quality or minor allele frequency) in mQTL data are replaced with a proxy ($r2>0.8$), if any.

Supplementary Table 14 Two-Step Mendelian randomization analysis for cg06809326

GTEx version 8 Tissue	No. of IVs	IVW			Weighted Median			MR Egger			ConMix test			Pleiotropy test (Egger Intercept test)			Heterogeneity test (IVW)			Heterogeneity test (MR Egger)			IV strength	
		beta	se	*P*	beta	se	*P*	beta	se	*P*	beta	95% Lower CI	95% Upper CI	beta	se	*P*	Q	df	*P*	Q	df	*P*	meanF	I^2
Step 1																								
cg06809326 -> CCDC144NL																								
Brain_Cortex	3	-5.90	1.82	$1.21x10^{-3}$	-5.85	1.77	$9.48x10^{-4}$	-15.95	8.04	0.30	-7.06	-11.64	0.71	0.47	0.37	0.42	2.40	2	0.30	0.77	1.00	0.38	NA	0.93
Brain_Caudate_basal_ganglia	3	-7.34	2.25	$1.11x10^{-3}$	-7.09	1.85	$1.32x10^{-4}$	-19.21	9.70	0.30	-9.66	-13.71	0.61	0.55	0.44	0.43	3.66	2	0.16	1.43	1.00	0.23	274.82	0.93
Brain_Cerebellum	3	-8.88	2.12	$2.72x10^{-5}$	-10.08	1.77	$1.24x10^{-8}$	-22.75	7.78	0.21	-10.47	-14.05	-6.55	0.65	0.35	0.32	3.45	2	0.18	0.12	1.00	0.73	274.82	0.93
Brain_Cerebellar_Hemisphere	3	-5.33	1.70	$1.71x10^{-3}$	-5.84	1.89	$1.97x10^{-3}$	-0.52	8.20	0.96	-6.12	-10.62	-2.12	-0.22	0.37	0.66	1.20	2	0.55	0.84	1.00	0.36	NA	0.93
Brain_Frontal_Cortex_BA9	3	-10.57	1.67	$2.38x10^{-10}$	-10.43	1.84	$1.53x10^{-8}$	-17.19	8.20	0.28	-10.56	-13.83	-7.30	0.31	0.38	0.56	1.02	2	0.60	0.34	1.00	0.56	NA	0.93
cg06809326 -> CCDC144NL-AS1																								
Brain_Cortex	3	-2.97	1.26	$1.85x10^{-2}$	-3.24	1.30	$1.23x10^{-2}$	-3.16	6.04	0.69	-3.36	-6.14	1.50	0.01	0.27	0.98	0.67	2	0.72	0.67	1.00	0.41	NA	0.93
Brain_Caudate_basal_ganglia	3	-6.62	2.24	$3.07x10^{-3}$	-6.44	1.74	$2.08x10^{-4}$	-15.19	12.66	0.44	-8.92	-12.74	1.86	0.40	0.58	0.61	4.06	2	0.13	2.75	1.00	0.10	NA	0.93
Brain_Cerebellum	3	-8.58	1.85	$3.73x10^{-6}$	-8.92	1.75	$3.61x10^{-7}$	-19.67	7.47	0.23	-9.87	-13.31	-5.86	0.52	0.34	0.37	2.87	2	0.24	0.57	1.00	0.45	274.82	0.93
Brain_Cerebellar_Hemisphere	3	-5.72	2.44	$1.93x10^{-2}$	-7.00	1.87	$1.81x10^{-4}$	0.85	15.31	0.96	-7.23	-10.77	-3.68	-0.31	0.70	0.74	4.36	2	0.11	3.65	1.00	0.06	274.82	0.93
Brain_Frontal_Cortex	3	-5.73	1.27	$6.55x10^{-6}$	-5.71	1.34	$2.09x10^{-5}$	-4.80	6.10	0.58	-5.73	-8.21	-3.24	-0.04	0.28	0.90	0.03	2	0.98	0.01	1.00	0.94	274.82	0.93
Step 2																								
CCDC144NL -> WMH burden																								
Brain_Caudate_basal_ganglia	1	-0.03	0.02	$3.94x10^{-2}$																				
Brain_Cerebellum	1	-0.02	0.02	0.28																				
Brain_Cortex	1	-0.03	0.02	$4.28x10^{-2}$																				
Brain_Frontal_Cortex_BA9	1	-0.02	0.02	0.28																				
CCDC144NL-AS1 -> WMH burden																								
Brain_Frontal_Cortex_BA9	1	0.01	0.02	0.579																				
Brain_Caudate_basal_ganglia	1	0.01	0.01	0.579																				

CI: confidence interval, I2: Bowden statistic of instrument strength for the MR-Egger method. IVW: inverse variance-weighted, meanF: instrument strength for IVW method. No.IV: Number of instrumental variables (IVs) included. OR: Odds ratio, PMR and Beta (S.E.): P value, effect size (beta) and its standard error (se) for the causal direction, primarily based on the IVW method, or in existence of pleiotropy (Ppleiotropy <0.05) based on MR Egger method. Ppleiotropy: P value estimated from MR Egger intercept pleiotropy test, WMH: white matter hyperintensities.

Supplementary Table 15 CpG-WMH burden associations pooled at GWAS loci

GWAS Locus	GWAS Gene	No. CpGs	Stat	P
10q24.33	*SH3PXD2A*	87	66.52	8.48E-03
16q12.1	*SALL1*	45	21.92	0.10
14q22.1	*NID2*	28	31.85	0.19
13q34	*COL4A2*	142	46.53	0.20
22q12.1	*MN1*	22	27.68	0.23
15q22.31	*RASL12*	10	13.35	0.27
8p23.1	*XKR6*	94	48.65	0.28
14q32.11	*CCDC88C*	85	50.27	0.28
2q32.1	*CALCRL*	5	7.64	0.34
17q21.31	*NMT1*	25	19.22	0.39
1p22.2	*PKN2*	17	15.57	0.40
2p21	*HAAO*	16	12.53	0.42
5q14.2	*VCAN*	30	22.39	0.42
3q27.1	*KLHL24*	15	15.87	0.51
1q41	*KCNK2*	29	13.83	0.53
8p23.1	*TNKS*	27	17.07	0.58
2p16.1	*EFEMP1*	21	11.28	0.58
10p14	*ECHDC3*	20	10.64	0.60
14q32.2	*DEGS2*	27	23.21	0.66
6q25.1	*PLEKHG1*	27	18.94	0.74
17q25.1	*TRIM65*	18	12.14	0.86

No. CpGs: number of CpGs pooled at the locus. Stat and P are from a generalized Fisher's test that computes pooled P values based on a Satterthwaite approximation adjusted for the dependence among tests

Supplementary Table 16 Multiple-trait colocalization of brain omics and white matter hyperintensities at the genome-wide association study loci[a]

DNAm	Discovery EWAS Statistics		nsnp	PPA	PPB	PPAB	PPC	PPAC	PPBC	PPFC	SNP with the best PP*
	Z	*P*									
CALCRL (2q32.1)											
cg22374415	-0.83	0.41	259	0.98	0.95	0.95	0.89	0.89	0.86	0.86	rs13417165
cg16628093	0.56	0.57	179	0.99	0.97	0.97	0.89	0.89	0.87	0.87	rs1355520
cg21350115	-0.27	0.79	184	0.99	0.99	0.99	0.90	0.90	0.89	0.89	rs8176583
cg00449067	-0.78	0.44	181	1.00	0.99	0.99	0.90	0.90	0.89	0.89	rs1355520
CCDC88C (14q32.11)											
cg10928544	-0.10	0.92	239	1.00	0.90	0.90	1.00	1.00	0.90	0.90	**rs1285847**
cg26120073	0.63	0.53	223	1.00	0.90	0.90	1.00	1.00	0.90	0.90	**rs1285847**
cg12937580	0.42	0.68	239	1.00	0.90	0.90	1.00	1.00	0.90	0.90	**rs1285847**
cg03946034	0.51	0.61	258	1.00	0.90	0.90	1.00	1.00	0.90	0.90	**rs1285847**
cg16505891	1.59	0.11	224	1.00	0.91	0.91	1.00	1.00	0.91	0.91	**rs1285847**
cg13027206	1.82	0.07	251	1.00	0.91	0.91	1.00	1.00	0.91	0.91	**rs1285847**
cg14214797	0.74	0.46	226	1.00	0.91	0.91	1.00	1.00	0.91	0.91	**rs1285847**
cg15944026	1.04	0.30	262	1.00	0.91	0.91	1.00	1.00	0.91	0.91	**rs1285847**
cg20303561	-0.60	0.55	265	1.00	0.91	0.91	1.00	1.00	0.91	0.91	**rs1285847**
cg20272146	0.14	0.89	253	1.00	0.92	0.92	1.00	1.00	0.92	0.92	**rs1285847**
cg10755375	-0.09	0.93	261	1.00	0.92	0.92	1.00	1.00	0.92	0.92	**rs1285847**
cg26259944	-1.96	4.98E-02	229	1.00	0.92	0.92	1.00	1.00	0.92	0.92	**rs1285847**
cg16438210	0.26	0.80	258	1.00	0.92	0.92	1.00	1.00	0.92	0.92	**rs1285847**
cg26290716	1.67	0.10	239	1.00	0.93	0.93	1.00	1.00	0.93	0.93	**rs1285847**
cg14415844	0.68	0.50	259	1.00	0.93	0.93	1.00	1.00	0.93	0.93	**rs1285847**
cg23304605	2.04	4.13E-02	251	1.00	0.94	0.94	1.00	1.00	0.94	0.94	**rs1285847**
cg07112604	0.59	0.56	229	1.00	0.94	0.94	1.00	1.00	0.94	0.94	**rs1285847**
cg10564112	-0.76	0.45	223	1.00	0.98	0.98	1.00	1.00	0.98	0.98	**rs1285847**
cg08017547	1.64	0.10	262	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**rs1285847**
cg23615483	0.56	0.58	253	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**rs1285847**
cg27663123	0.57	0.57	226	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**rs1285841**
cg05260886	-0.01	0.99	226	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**rs1285847**
cg09003057	0.19	0.85	227	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**rs1285847**
COL4A2 (13q34)											
cg13153710	0.20	0.84	426	0.83	0.77	0.77	0.76	0.76	0.70	0.70	**rs55940034**
cg13672379	0.10	0.92	212	0.84	0.78	0.78	0.77	0.77	0.71	0.71	**rs55940034**
cg06651838	-0.92	0.36	305	0.84	0.78	0.78	0.77	0.77	0.71	0.71	**rs55940034**
cg15374000	-1.05	0.30	203	0.79	0.93	0.79	0.85	0.71	0.85	0.71	rs9583484
cg03469869	-2.09	3.69E-02	176	0.85	0.79	0.79	0.77	0.77	0.72	0.72	**rs55940034**
cg03482866	0.48	0.63	274	0.85	0.79	0.79	0.77	0.77	0.72	0.72	**rs55940034**
cg22051925	0.82	0.41	384	0.85	0.79	0.79	0.78	0.78	0.72	0.72	**rs55940034**
cg05167916	1.55	0.12	390	0.85	0.80	0.80	0.78	0.78	0.72	0.72	**rs55940034**
cg23638556	-0.46	0.64	201	0.85	0.80	0.80	0.78	0.78	0.72	0.72	**rs55940034**
cg22942950	-0.34	0.73	426	0.86	0.81	0.81	0.79	0.79	0.73	0.73	**rs55940034**
cg06951647	1.75	0.08	372	0.87	0.82	0.82	0.79	0.79	0.74	0.74	**rs55940034**
cg03528486	-1.24	0.22	453	0.89	0.84	0.84	0.81	0.81	0.77	0.77	**rs55940034**
cg03930369	-1.15	0.25	451	0.88	0.85	0.85	0.80	0.80	0.77	0.77	**rs55940034**
cg26384036	0.21	0.84	374	0.89	0.86	0.86	0.82	0.82	0.78	0.78	**rs55940034**
cg09286253	0.36	0.72	451	0.90	0.87	0.87	0.82	0.82	0.79	0.79	**rs55940034**
cg08318510	0.00	1.00	204	0.86	0.96	0.86	0.90	0.80	0.90	0.80	rs7991332
cg11096515	1.91	0.06	388	0.93	0.91	0.91	0.85	0.85	0.83	0.83	**rs55940034**
cg24283836	-0.17	0.86	390	0.94	0.91	0.91	0.85	0.85	0.83	0.83	**rs55940034**
cg15455643	0.65	0.51	274	0.99	0.99	0.99	0.90	0.90	0.90	0.90	rs11838776
cg08959039	1.63	0.10	388	1.00	1.00	1.00	0.91	0.90	0.91	0.90	rs913397
cg26053697	0.62	0.53	274	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs11838776
cg22382805	0.37	0.71	204	0.97	0.97	0.97	0.93	0.93	0.93	0.93	rs9521686
cg22232704	0.76	0.45	204	0.97	0.97	0.97	0.94	0.94	0.94	0.94	rs9521686
DEGS2 (14q32.2)											
cg23426054	0.75	0.45	209	0.86	0.78	0.78	0.82	0.82	0.74	0.74	rs3783334
cg23076361	1.21	0.23	201	0.86	0.78	0.78	0.82	0.82	0.74	0.74	rs3783334
cg07530063	-0.92	0.36	209	0.86	0.78	0.78	0.82	0.82	0.74	0.74	rs3783334
cg25087487	0.32	0.75	207	0.86	0.78	0.78	0.82	0.82	0.75	0.75	rs3783334
cg14290450	-0.07	0.94	206	0.86	0.78	0.78	0.82	0.82	0.75	0.75	rs3783334
cg21923884	0.24	0.81	209	0.86	0.78	0.78	0.82	0.82	0.75	0.75	rs3783334
cg08307816	-0.18	0.86	207	0.86	0.78	0.78	0.82	0.82	0.75	0.75	rs3783334
cg21203643	0.19	0.85	206	0.86	0.79	0.79	0.82	0.82	0.75	0.75	rs3783334
cg20036732	0.21	0.83	207	0.86	0.79	0.79	0.82	0.82	0.75	0.75	rs3783334
cg04548002	-0.19	0.85	207	0.86	0.79	0.79	0.82	0.82	0.75	0.75	rs3783334
cg00992159	0.49	0.62	206	0.86	0.79	0.79	0.83	0.83	0.75	0.75	rs3783334
cg16285701	0.40	0.69	209	0.87	0.79	0.79	0.83	0.83	0.75	0.75	rs3783334

DNAm	Discovery EWAS Statistics		nsnp	PPA	PPB	PPAB	PPC	PPAC	PPBC	PPFC	SNP with the best PP*
	Z	*P*									
cg09258240	-2.16	3.07E-02	210	0.87	0.79	0.79	0.83	0.83	0.75	0.75	rs3783334
cg01768697	0.06	0.95	209	0.87	0.79	0.79	0.83	0.83	0.76	0.76	rs3783334
cg08180525	0.05	0.96	206	0.87	0.80	0.80	0.83	0.83	0.76	0.76	rs3783334
cg05129951	0.17	0.87	207	0.87	0.80	0.80	0.83	0.83	0.76	0.76	rs3783334
cg12059866	-0.07	0.94	200	0.87	0.80	0.80	0.83	0.83	0.76	0.76	rs3783334
cg00890257	-0.46	0.64	209	0.88	0.82	0.82	0.84	0.84	0.78	0.78	rs3783334
cg09911485	0.18	0.85	206	0.89	0.83	0.83	0.85	0.85	0.79	0.79	rs3783334
cg01607727	-1.58	0.11	206	0.91	0.86	0.86	0.87	0.87	0.82	0.82	rs3783334
cg01513978	0.66	0.51	207	0.92	0.89	0.88	0.88	0.88	0.84	0.84	rs3783334
cg04755365	0.64	0.52	208	0.93	0.89	0.89	0.89	0.89	0.85	0.85	rs3783334
cg01974091	1.05	0.29	208	1.00	1.00	1.00	0.95	0.95	0.95	0.95	rs7157984
cg26002632	0.92	0.36	208	1.00	1.00	1.00	0.95	0.95	0.95	0.95	rs7157984
cg18485215	0.94	0.35	208	1.00	1.00	1.00	0.95	0.95	0.95	0.95	rs7157984
ECHDC3 (10p14)											
cg07469647	-0.32	0.75	280	0.71	0.99	0.71	0.99	0.71	0.99	0.71	rs11257331
cg16944070	-1.73	0.08	295	0.75	1.00	0.75	1.00	0.75	1.00	0.75	rs718641
cg11103390	-1.23	0.22	297	0.77	1.00	0.77	1.00	0.77	1.00	0.77	rs55790558
cg05991820	-1.49	0.14	301	0.77	1.00	0.77	1.00	0.77	1.00	0.77	rs55790558
cg14032732	0.51	0.61	297	0.84	1.00	0.84	1.00	0.84	1.00	0.84	rs3750693
cg00673251	1.21	0.23	296	0.85	0.97	0.85	0.97	0.85	0.96	0.85	rs11257315
cg10543363	-0.79	0.43	296	0.88	1.00	0.88	1.00	0.88	1.00	0.88	rs11257293
cg23960723	-0.81	0.42	295	0.89	1.00	0.89	1.00	0.89	1.00	0.89	rs11257293
cg17782954	-0.73	0.47	295	0.91	0.99	0.91	0.97	0.89	0.97	0.89	rs10795884
EFEMP1 (2p16.1)											
cg25968367	1.86	0.06	291	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg25412594	-0.28	0.78	291	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg16118212	0.53	0.60	292	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg05140065	0.09	0.93	276	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg03817671	1.00	0.32	292	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg16947612	-0.09	0.93	294	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg16100120	-0.10	0.92	291	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs7596872**
cg23260993	0.26	0.80	288	0.99	0.91	0.91	0.98	0.98	0.90	0.90	**rs7596872**
cg08130988	1.13	0.26	292	0.99	0.92	0.92	0.98	0.98	0.91	0.91	**rs7596872**
cg20052010	1.04	0.30	290	1.00	0.92	0.92	0.98	0.98	0.91	0.91	**rs7596872**
cg25046074	1.79	0.07	291	1.00	0.93	0.93	0.98	0.98	0.92	0.92	**rs7596872**
cg05385513	-0.44	0.66	291	1.00	0.94	0.94	0.98	0.98	0.93	0.93	**rs7596872**
cg02626897	-0.45	0.65	265	1.00	0.95	0.95	0.99	0.99	0.94	0.94	**rs7596872**
cg20786074	-2.11	3.48E-02	290	1.00	0.95	0.95	0.98	0.98	0.94	0.94	**rs7596872**
cg25711779	0.14	0.89	290	1.00	0.95	0.95	0.98	0.98	0.94	0.94	**rs7596872**
cg00574639	-0.09	0.93	283	1.00	0.99	0.99	0.99	0.99	0.97	0.97	**rs7596872**
cg03122624	-0.50	0.62	290	1.00	0.99	0.99	0.99	0.99	0.98	0.98	**rs7596872**
cg25268863	-0.09	0.93	281	1.00	1.00	1.00	0.99	0.99	0.99	0.99	**rs7596872**
HAAO (2p21)											
cg06913600	0.88	0.38	309	0.78	0.97	0.77	0.96	0.77	0.96	0.76	rs57082399
cg09256201	0.47	0.64	307	0.83	0.87	0.82	0.86	0.80	0.84	0.79	rs1347931
cg01561916	2.23	2.55E-02	309	0.94	1.00	0.94	0.98	0.93	0.98	0.92	rs893729
cg09368936	1.00	0.32	307	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg27299406	-1.21	0.23	309	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg08421126	0.54	0.59	308	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg20289949	-0.68	0.49	307	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg09480054	-0.10	0.92	310	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg17246140	0.04	0.97	309	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg20857709	1.03	0.30	307	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
cg05986933	1.56	0.12	309	0.94	1.00	0.94	0.98	0.93	0.98	0.93	rs7422835
KLHL24 (3q27.1)											
cg02021337	-0.59	0.55	208	1.00	0.90	0.90	1.00	1.00	0.90	0.90	rs1983386
cg11237115	-0.82	0.41	242	1.00	0.90	0.90	1.00	1.00	0.90	0.90	rs1983386
cg24399556	0.35	0.72	243	1.00	0.90	0.90	1.00	1.00	0.90	0.90	rs1983386
cg15803765	1.70	0.09	245	1.00	0.90	0.90	1.00	1.00	0.90	0.90	rs1983386
cg22757384	0.06	0.95	244	1.00	0.91	0.91	1.00	1.00	0.91	0.91	rs1983386
cg00892804	-0.70	0.49	243	1.00	0.91	0.91	1.00	1.00	0.91	0.91	rs1983386
cg27197504	-0.03	0.98	246	1.00	0.91	0.91	1.00	1.00	0.91	0.91	rs1983386
cg10143433	1.08	0.28	243	1.00	0.91	0.91	1.00	1.00	0.91	0.91	rs1983386
cg05197962	0.01	0.99	242	1.00	0.91	0.91	1.00	1.00	0.91	0.91	rs1983386
cg07482199	-2.15	3.13E-02	243	1.00	0.91	0.91	1.00	1.00	0.91	0.91	rs1983386
cg02628016	-0.80	0.43	228	1.00	0.92	0.92	1.00	1.00	0.92	0.92	rs1983386
cg06375534	0.51	0.61	243	1.00	0.92	0.92	1.00	1.00	0.92	0.92	rs1983386
cg03030002	-0.21	0.83	243	1.00	0.93	0.93	1.00	1.00	0.93	0.93	rs1983386
cg12198633	1.00	0.32	234	1.00	0.94	0.94	1.00	1.00	0.94	0.94	rs1983386

DNAm	Discovery EWAS Statistics		nsnp	PPA	PPB	PPAB	PPC	PPAC	PPBC	PPFC	SNP with the best PP*
	Z	*P*									
cg03895159	-1.42	0.16	242	1.00	1.00	1.00	1.00	1.00	1.00	1.00	rs1983386
NID2 (14q22.1)											
cg10843707	-0.28	0.78	388	0.96	0.97	0.96	0.88	0.87	0.88	0.87	rs2645738
NMT1 (17q21.31)											
cg24354954	-0.06	0.95	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg02077631	0.89	0.38	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg21452443	-0.23	0.82	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg21132931	0.49	0.63	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg05322982	-0.10	0.92	226	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg07481784	0.03	0.98	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg10755730	-0.29	0.77	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg04013970	0.32	0.75	260	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg21554013	-0.47	0.64	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg09377882	-1.08	0.28	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg09214551	-1.47	0.14	258	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg22356484	0.15	0.88	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg00405568	-0.53	0.59	256	0.99	0.90	0.90	0.98	0.98	0.89	0.89	**rs6503417**
cg03287877	-2.92	3.52E-03	258	1.00	0.92	0.92	0.98	0.98	0.91	0.91	**rs6503417**
cg16594296	0.25	0.80	257	1.00	0.93	0.93	0.99	0.99	0.92	0.92	**rs6503417**
cg16080654	-1.58	0.11	238	1.00	0.96	0.96	0.99	0.99	0.94	0.94	**rs6503417**
cg02888886	0.62	0.54	252	1.00	0.96	0.96	0.99	0.99	0.95	0.95	**rs6503417**
cg05349016	-0.72	0.47	251	1.00	0.97	0.97	0.99	0.99	0.96	0.96	**rs6503417**
cg11583751	1.30	0.19	233	1.00	0.97	0.97	0.99	0.99	0.96	0.96	**rs6503417**
cg08860622	1.09	0.27	257	1.00	0.98	0.98	0.99	0.99	0.97	0.97	**rs6503417**
PLEKHG1 (6q25.1)											
cg07468752	-0.08	0.94	333	0.98	0.89	0.89	0.93	0.93	0.84	0.84	**rs6940540**
cg26852242	0.04	0.97	366	0.98	0.89	0.89	0.93	0.93	0.84	0.84	**rs6940540**
cg08560465	0.27	0.79	334	0.98	0.89	0.89	0.93	0.93	0.85	0.85	**rs6940540**
cg16540391	0.34	0.73	363	0.98	0.90	0.90	0.93	0.93	0.85	0.85	**rs6940540**
cg00183067	-2.18	2.92E-02	300	0.97	0.93	0.93	0.89	0.89	0.85	0.85	rs4272224
cg15622860	1.22	0.22	334	0.98	0.90	0.90	0.93	0.93	0.86	0.86	**rs6940540**
cg02784464	-0.40	0.69	301	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs2073061
RASL12 (15q22.31)											
cg27356162	1.62	0.10	118	0.97	0.88	0.88	0.97	0.97	0.87	0.87	**rs12443113**
cg11267410	1.36	0.18	91	0.98	0.89	0.89	0.97	0.97	0.89	0.89	**rs12443113**
cg19592235	0.00	1.00	91	0.98	0.90	0.90	0.97	0.97	0.89	0.89	**rs12443113**
cg00397575	1.19	0.23	91	0.98	0.90	0.90	0.97	0.97	0.89	0.89	**rs12443113**
cg27522060	0.82	0.41	91	0.98	0.90	0.90	0.97	0.97	0.89	0.89	**rs12443113**
cg10604168	1.10	0.27	91	0.98	0.90	0.90	0.97	0.97	0.89	0.89	**rs12443113**
cg06192057	0.42	0.67	118	0.98	0.90	0.90	0.97	0.97	0.89	0.89	**rs12443113**
cg13754518	1.70	0.09	96	0.99	0.95	0.95	0.98	0.98	0.94	0.94	**rs12443113**
SH3PXD2A (10q24.33)											
cg11385165	0.14	0.89	194	0.88	0.81	0.81	0.81	0.81	0.73	0.73	rs12357919
cg22603555	-1.06	0.29	207	0.87	0.81	0.81	0.81	0.81	0.75	0.75	rs7094488
cg01380878	-1.44	0.15	225	0.94	0.85	0.85	0.85	0.85	0.77	0.77	rs71471298
cg09926728	1.29	0.20	234	0.94	0.85	0.85	0.85	0.85	0.77	0.77	rs71471298
cg19189864	-1.33	0.18	209	0.93	0.85	0.85	0.86	0.86	0.78	0.78	rs61871257
cg04762698	1.69	0.09	211	0.93	0.85	0.85	0.86	0.86	0.78	0.78	rs61871257
cg04426842	1.10	0.27	210	0.93	0.85	0.85	0.86	0.86	0.78	0.78	rs61871257
cg19056515	-1.51	0.13	214	0.93	0.85	0.85	0.86	0.86	0.78	0.78	rs61871257
cg12094736	0.03	0.98	224	0.94	0.86	0.86	0.86	0.86	0.78	0.78	rs10748849
cg08728732	-0.61	0.54	211	0.94	0.86	0.86	0.86	0.86	0.78	0.78	rs61871257
cg08740008	0.85	0.40	208	0.94	0.86	0.86	0.86	0.86	0.79	0.79	rs61871257
cg20241299	0.76	0.45	212	0.94	0.86	0.86	0.86	0.86	0.79	0.79	rs61871257
cg24989193	0.48	0.63	226	0.94	0.87	0.87	0.86	0.86	0.79	0.79	rs10748849
cg20669641	-2.05	4.03E-02	183	0.96	0.87	0.87	0.88	0.88	0.79	0.79	**rs4630220**
cg19007269	-0.75	0.46	175	0.96	0.87	0.87	0.88	0.88	0.79	0.79	**rs4630220**
cg13634090	-1.72	0.09	176	0.96	0.87	0.87	0.88	0.88	0.79	0.79	**rs4630220**
cg04688330	0.22	0.83	171	0.96	0.87	0.87	0.88	0.88	0.79	0.79	**rs4630220**
cg19355009	0.23	0.82	175	0.96	0.87	0.87	0.88	0.88	0.80	0.80	**rs4630220**
cg01428849	-1.40	0.16	183	0.96	0.87	0.87	0.88	0.88	0.80	0.80	**rs4630220**
cg02432295	-1.38	0.17	178	0.96	0.87	0.87	0.88	0.88	0.80	0.80	**rs4630220**
cg02878913	0.61	0.55	178	0.96	0.87	0.87	0.88	0.88	0.80	0.80	**rs4630220**
cg16492615	0.37	0.72	186	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg14209793	0.36	0.72	223	0.96	0.87	0.87	0.88	0.88	0.80	0.80	rs6584579
cg00492979	-2.63	8.66E-03	184	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg02089135	-0.02	0.99	224	0.96	0.87	0.87	0.88	0.88	0.80	0.80	rs6584579
cg03958163	1.83	0.07	222	0.96	0.87	0.87	0.88	0.88	0.80	0.80	rs6584579
cg17167425	-0.91	0.36	197	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**

DNAm	Discovery EWAS Statistics		nsnp	PPA	PPB	PPAB	PPC	PPAC	PPBC	PPFC	SNP with the best PP*
	Z	*P*									
cg23985995	-0.44	0.66	172	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg07291639	-0.88	0.38	186	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg05620165	-2.34	1.91E-02	219	0.96	0.88	0.88	0.88	0.88	0.80	0.80	rs10786772
cg09210519	-1.55	0.12	169	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg24060660	-0.03	0.98	185	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg22973172	0.19	0.85	202	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg05352321	-2.24	2.50E-02	181	0.96	0.88	0.88	0.88	0.88	0.80	0.80	rs3758575
cg23461741	0.99	0.32	225	0.95	0.89	0.89	0.87	0.87	0.80	0.80	rs71471298
cg07457213	-0.85	0.39	212	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg24911198	-1.00	0.32	168	0.96	0.88	0.88	0.88	0.88	0.80	0.80	**rs4630220**
cg19140548	-3.41	6.61E-04	233	0.96	0.88	0.88	0.88	0.88	0.80	0.80	rs10883922
cg04671329	0.82	0.41	223	0.96	0.88	0.88	0.89	0.89	0.80	0.80	rs6584579
cg19726408	-2.84	4.56E-03	183	0.96	0.88	0.88	0.88	0.88	0.81	0.81	**rs4630220**
cg19120251	-0.89	0.37	222	0.96	0.88	0.88	0.89	0.89	0.81	0.81	rs6584579
cg11819799	-0.59	0.55	169	0.96	0.89	0.89	0.88	0.88	0.81	0.81	**rs4630220**
cg05359750	-0.41	0.68	213	0.95	0.89	0.89	0.87	0.87	0.81	0.81	rs61871257
cg12636499	-0.19	0.85	178	0.96	0.88	0.88	0.89	0.89	0.81	0.81	rs3758575
cg23901918	-0.17	0.87	175	0.96	0.89	0.89	0.88	0.88	0.81	0.81	**rs4630220**
cg23786152	1.01	0.31	225	0.96	0.89	0.89	0.89	0.89	0.81	0.81	rs6584579
cg25181684	-1.57	0.12	165	0.96	0.89	0.89	0.88	0.88	0.81	0.81	**rs4630220**
cg05162794	0.20	0.84	215	0.95	0.89	0.89	0.87	0.87	0.81	0.81	rs61871257
cg07168060	-0.58	0.56	226	0.96	0.89	0.89	0.89	0.89	0.81	0.81	rs6584579
cg16669339	-2.59	9.67E-03	183	0.97	0.90	0.90	0.89	0.89	0.82	0.82	**rs4630220**
cg12569119	-0.68	0.50	228	0.96	0.90	0.90	0.87	0.87	0.82	0.82	rs71471298
cg25866173	2.33	2.01E-02	220	0.97	0.90	0.90	0.89	0.89	0.82	0.82	rs6584579
cg04388244	-0.14	0.89	175	0.97	0.90	0.90	0.89	0.89	0.82	0.82	**rs4630220**
cg11014810	-1.03	0.30	219	0.97	0.90	0.90	0.89	0.89	0.82	0.82	rs10786772
cg02722808	-1.18	0.24	227	0.97	0.91	0.91	0.89	0.89	0.83	0.83	rs56093138
cg10831642	0.81	0.42	204	0.93	0.91	0.90	0.86	0.86	0.84	0.83	rs7910092
cg06966363	0.49	0.62	185	0.96	0.92	0.92	0.88	0.88	0.83	0.83	rs56134400
cg01833196	0.41	0.68	170	0.97	0.92	0.92	0.89	0.89	0.84	0.84	**rs4630220**
cg06276663	1.23	0.22	224	0.97	0.92	0.92	0.89	0.89	0.84	0.84	rs10786772
cg06847424	0.61	0.54	212	0.97	0.93	0.93	0.88	0.88	0.84	0.84	rs61871257
cg18675043	1.42	0.16	170	0.98	0.93	0.93	0.89	0.89	0.85	0.85	**rs4630220**
cg18735015	-3.12	1.83E-03	184	0.98	0.93	0.93	0.89	0.89	0.85	0.85	**rs4630220**
cg06888746	-1.42	0.16	234	0.97	0.94	0.94	0.88	0.88	0.85	0.85	rs10748849
cg06620993	0.02	0.98	175	0.98	0.94	0.94	0.90	0.90	0.85	0.85	**rs4630220**
cg13564061	0.18	0.85	173	0.98	0.94	0.94	0.90	0.90	0.86	0.86	**rs4630220**
cg02818775	-0.67	0.50	213	0.98	0.95	0.95	0.89	0.89	0.87	0.87	rs61871257
cg00730670	0.33	0.74	219	0.98	0.96	0.96	0.89	0.89	0.87	0.87	rs71471298
cg07709475	-3.17	1.53E-03	183	0.98	0.95	0.95	0.90	0.90	0.87	0.87	**rs4630220**
cg12477923	-1.20	0.23	216	0.98	0.96	0.96	0.89	0.89	0.87	0.87	rs61871257
cg10521450	-0.27	0.79	184	0.99	0.97	0.97	0.90	0.90	0.89	0.89	**rs4630220**
cg12312107	-2.33	1.97E-02	168	0.99	0.98	0.98	0.91	0.91	0.89	0.89	**rs4630220**
cg20668662	-0.67	0.51	170	1.00	0.98	0.98	0.91	0.91	0.90	0.90	**rs4630220**
cg09754828	-1.88	0.06	213	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs61871257
cg00126946	-0.92	0.36	213	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs61871257
cg04788957	-0.57	0.57	185	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs12357919
cg17683908	2.06	3.93E-02	223	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs56093138
cg01727419	1.20	0.23	220	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs10786772
cg07671776	1.80	0.07	220	1.00	1.00	1.00	0.91	0.91	0.91	0.91	rs6584579
cg12800012	-0.82	0.41	222	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs10883926
cg00277384	-1.60	0.11	168	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs3758575
TNKS (8p23.1)											
cg23320011	-1.32	0.19	304	0.79	0.78	0.78	0.75	0.75	0.74	0.74	rs7846399
TRIM65 (17q25.1)											
cg07879399	0.44	0.66	178	1.00	0.90	0.90	0.99	0.99	0.90	0.90	**rs34974290**
cg27659478	-1.80	0.07	163	1.00	0.90	0.90	0.99	0.99	0.90	0.90	**rs34974290**
cg15238200	1.39	0.16	165	1.00	0.90	0.90	0.99	0.99	0.90	0.90	**rs34974290**
cg26965547	-0.79	0.43	165	1.00	0.91	0.91	0.99	0.99	0.90	0.90	**rs34974290**
cg16087457	-0.21	0.84	164	1.00	0.91	0.91	0.99	0.99	0.90	0.90	**rs34974290**
cg13966383	-0.43	0.67	164	1.00	0.91	0.91	0.99	0.99	0.90	0.90	**rs34974290**
cg15945867	0.41	0.69	176	1.00	0.91	0.91	0.99	0.99	0.91	0.91	**rs34974290**
cg19169246	-0.75	0.45	176	1.00	0.91	0.91	0.99	0.99	0.91	0.91	**rs34974290**
cg01453052	-0.18	0.86	161	1.00	0.91	0.91	0.99	0.99	0.91	0.91	**rs34974290**
cg24191285	-0.44	0.66	164	1.00	0.91	0.91	0.99	0.99	0.91	0.91	**rs34974290**
cg03648361	-0.20	0.84	162	1.00	0.91	0.91	0.99	0.99	0.91	0.91	**rs34974290**
cg26780705	-0.68	0.50	165	1.00	0.92	0.92	0.99	0.99	0.92	0.91	**rs34974290**
cg27576946	-0.61	0.54	164	1.00	0.93	0.93	0.99	0.99	0.92	0.92	**rs34974290**

DNAm	Discovery EWAS Statistics		nsnp	PPA	PPB	PPAB	PPC	PPAC	PPBC	PPFC	SNP with the best PP*
	Z	*P*									
cg10195415	0.67	0.50	174	1.00	0.93	0.93	1.00	1.00	0.92	0.92	**rs34974290**
cg10927422	0.06	0.95	174	1.00	0.99	0.99	1.00	1.00	0.99	0.99	**rs34974290**
cg25176883	0.39	0.70	174	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**rs34974290**
VCAN (5q14.2)											
cg04223006	1.45	0.15	249	0.96	0.87	0.87	0.87	0.87	0.79	0.79	rs7733216
cg26496628	1.72	0.09	345	0.96	0.87	0.87	0.87	0.87	0.79	0.79	rs7733216
cg08891382	0.62	0.54	265	0.96	0.87	0.87	0.87	0.87	0.79	0.79	rs7733216
cg27109568	-0.68	0.50	343	0.96	0.88	0.88	0.88	0.88	0.80	0.80	rs3852188
cg22039204	-1.95	0.05	297	0.96	0.88	0.88	0.88	0.88	0.80	0.80	rs7733216
XKR6 (8p23.1)											
cg23565027	-0.41	0.68	253	0.82	0.74	0.74	0.79	0.79	0.71	0.71	rs6987059
cg11051055	0.32	0.75	282	0.84	0.76	0.76	0.81	0.81	0.73	0.73	rs6987059
cg16513326	-0.49	0.62	282	0.80	0.84	0.80	0.77	0.73	0.77	0.73	rs4307299
cg11434468	0.50	0.62	284	0.84	0.76	0.76	0.81	0.81	0.73	0.73	rs6987059
cg24631970	-0.33	0.75	284	0.84	0.76	0.76	0.81	0.81	0.73	0.73	rs6987059
cg12276019	-1.12	0.26	282	0.84	0.76	0.76	0.81	0.81	0.73	0.73	rs6987059
cg27025752	-0.18	0.86	284	0.84	0.76	0.76	0.81	0.81	0.73	0.73	rs10109167
cg04657224	-1.29	0.20	284	0.84	0.76	0.76	0.81	0.81	0.73	0.73	rs10109167
cg24664689	-0.47	0.64	283	0.84	0.77	0.77	0.81	0.81	0.73	0.73	rs6987059
cg09758742	-1.79	0.07	284	0.84	0.77	0.77	0.81	0.81	0.74	0.74	rs10109167
cg12267497	2.61	8.98E-03	275	0.84	0.77	0.77	0.81	0.81	0.74	0.74	rs10109167
cg15301489	-0.36	0.72	283	0.85	0.78	0.78	0.82	0.82	0.74	0.74	rs10109167
cg26675876	-0.03	0.97	284	0.85	0.78	0.78	0.82	0.82	0.75	0.75	rs6987059
cg07229212	-1.11	0.27	283	0.85	0.78	0.78	0.82	0.82	0.75	0.75	rs6987059
cg10947146	0.09	0.93	283	0.85	0.78	0.78	0.82	0.82	0.75	0.75	rs10109167
cg20940153	-0.86	0.39	284	0.85	0.78	0.78	0.82	0.82	0.75	0.75	rs6987059
cg27487839	0.99	0.32	282	0.86	0.78	0.78	0.82	0.82	0.75	0.75	rs10109167
cg10333252	-0.31	0.76	280	0.87	0.80	0.80	0.83	0.83	0.77	0.77	rs10109167
cg11746684	-1.36	0.17	282	0.88	0.83	0.83	0.85	0.85	0.79	0.79	rs10109167
cg06819296	0.49	0.62	283	0.93	0.93	0.92	0.89	0.88	0.88	0.88	rs6601565
cg03617420	1.77	0.08	269	0.93	0.98	0.93	0.93	0.88	0.93	0.88	rs4841487
cg20189073	-0.68	0.50	273	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs6601555
cg24574147	0.20	0.84	272	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs11985603
cg10823658	0.79	0.43	272	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs6601555
cg16232071	1.55	0.12	273	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs11250113
cg10596361	0.51	0.61	273	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs11250113
cg17167536	-0.48	0.63	271	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs11777486
cg16664915	-0.41	0.68	273	1.00	1.00	1.00	0.92	0.92	0.92	0.92	rs11777486
cg02475374	-1.81	0.07	273	1.00	1.00	1.00	0.95	0.95	0.95	0.95	rs4841500

DLPFC: dorsolateral prefrontal cortex, DNAm: DNA methylationnsnp: number of SNPs, EWAS: epigenome-wide association study, posterior probability (PP) for each hypotheses, genetic variants are associated with WMH burden (PPA), with DLPFC DNA methylation (DNAm) (PPB), with DLPFC gene expression (PPC), with WMH burden and DLPFC DNAm (PPAB), with WMH burden and DLPFC gene expression (PPAC), with DNAm and gene expression in DLPFC (PPBC), and with WMH burden, DNAm and gene expression (PPFC). Genetic associations with WMH burden are from sargurupremraj M (2020)a, and with brain omics are from Ng B (2017)b.

a. Sargurupremraj M, Suzuki H, Jian X, Sarnowski C, Evans TE, Bis JC, et al. Cerebral small vessel disease genomics and its implications across the lifespan. Nature Communications. 2020;11:6285.

b. Ng B, White CC, Klein H-U, Sieberts SK, McCabe C, Patrick E, et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nature neuroscience. 2017;20:1418–26.

*GWAS sentinel SNP is highlighted in bold.

Supplementary Table 17 Interindividual correlation between DNA methylation levels in blood and prefrontal cortex tissue

	r	***P***
cg24202936	0.327	4.52×10^{-3}
cg17417856	0.162	0.168
cg01506471	0.866	2.06×10^{-23}
cg14547240	0.003	0.978
cg21547371	0.033	0.783
cg03116124	0.292	0.012
cg06809326	0.566	1.49×10^{-7}
cg13476133	0.294	0.011
cg14133539	0.136	0.247
cg17577122	-0.158	0.178
cg23586595	-0.242	0.038
cg23054394	-0.123	0.295

r: Pearson correlation coefficient between DNAm in blood and prefrontal cortex tissue.
Correlation between DNAm levels in blood and brain was computed in 74 subjects (Hannon E *et al*. 2015).
Hannon E *et al*. "Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes." *Epigenetics*. 2015;10(11):1024-1032. data available from https://epigenetics.essex.ac.uk/bloodbrain/.

Supplementary Table 18 Multiple trait co-localization of brain omics and white matter hyperintensities at cg24202936 and cg06809326

Gene_symbol	DNAm	nsnp	PPA	PPB	PPAB	PPC	PPAC	PPBC	PPFC
FOLH1	cg24202936	61	0.77	0.97	0.77	0.95	0.75	0.95	**0.75**
CCDC144NL-AS1	cg06809326	91	0.70	0.90	0.70	0.89	0.69	0.89	0.69

DLPFC: dorsolateral prefrontal cortex, DNAm: DNA methylation, nsnp: number of SNPs, posterior probability (PP) for each hypotheses, genetic variants are associated with WMH burden (PPA), with DLPFC DNAm (PPB), with DLPFC gene expression (PPC), with WMH burden and DLPFC DNAm (PPAB), with WMH burden and DLPFC gene expression (PPAC), with DNAm and gene expression in DLPFC (PPBC), and with WMH burden, DNAm and gene expression (PPFC).

Genetic associations with WMH burden are from sargurupremraj M (2020)[a], and with brain omics are from Ng B (2017)[b].

a. Sargurupremraj M, Suzuki H, Jian X, Sarnowski C, Evans TE, Bis JC, et al. Cerebral small vessel disease genomics and its implications across the lifespan. Nature Communications. 2020;11:6285.

b. Ng B, White CC, Klein H-U, Sieberts SK, McCabe C, Patrick E, et al. An xQTL map integrates the genetic architecture of the human brain's transcriptome and epigenome. Nature neuroscience. 2017;20:1418–26.

Supplementary Table 19 Meta-analysis of GWAS, EWAS, and TWAS marker set enrichment associations (MERGEOMICS)

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
M18306	1.14E-45	6.71E-43	211	Regulation of actin cytoskeleton
M10401	1.10E-35	4.86E-33	18	Telomeres, Telomerase, Cellular Aging, and Immortality
rctm0602	3.17E-34	1.12E-31	84	Integrin cell surface interactions
rctm1272	1.41E-33	4.13E-31	32	Thrombin signalling through proteinase activated receptors (PARs)
rctm0764	4.70E-32	1.18E-29	9	Nef Mediated CD4 Down-regulation
rctm1167	3.13E-31	6.89E-29	8	Sperm Motility And Taxes
rctm0244	1.14E-29	2.23E-27	87	Class B/2 (Secretin family receptors)
rctm0482	2.56E-29	4.26E-27	101	GTP hydrolysis and joining of the 60S ribosomal subunit
M17668	2.69E-29	4.71E-27	12	Rho-Selective Guanine Exchange Factor AKAP13 Mediates Stress Fiber Formation
rctm0769	2.78E-28	4.09E-26	20	Nef-mediates down modulation of cell surface receptors by recruiting them to clathrin adapters
rctm1059	7.10E-28	9.64E-26	53	Ribosomal scanning and start codon recognition
M15394	1.68E-27	2.12E-25	20	Acute Myocardial Infarction
M189	1.84E-27	2.17E-25	84	Ribosome
rctm0384	3.56E-27	3.93E-25	108	Eukaryotic Translation Initiation
rctm0003	4.63E-27	4.81E-25	100	3' -UTR-mediated translational regulation
rctm0211	1.56E-26	1.53E-24	108	Cap-dependent Translation Initiation
rctm1268	2.44E-26	2.27E-24	31	The phototransduction cascade
rctm0765	4.87E-25	4.29E-23	7	Nef Mediated CD8 Down-regulation
rctm0773	1.51E-24	1.27E-22	22	Nephrin interactions
rctm0428	1.74E-24	1.40E-22	90	Formation of a pool of free 40S subunits
M16563	1.36E-23	1.04E-21	23	mTOR Signaling Pathway
rctm0184	7.10E-23	5.22E-21	21	CD28 dependent PI3K/Akt signaling
M8719	2.63E-21	1.86E-19	25	mCalpain and friends in Cell motility
rctm1273	3.43E-21	2.33E-19	23	Thromboxane signalling through TP receptor
rctm0571	3.91E-21	2.55E-19	30	Inactivation, recovery and regulation of the phototransduction cascade
rctm0789	6.36E-21	3.87E-19	86	Nonsense Mediated Decay Independent of the Exon Junction Complex
rctm0894	6.37E-21	3.88E-19	18	Pre-NOTCH Processing in Golgi
rctm0854	2.20E-20	1.29E-18	78	Peptide hormone metabolism
rctm0622	6.23E-20	3.55E-18	38	Intrinsic Pathway for Apoptosis
rctm0446	8.24E-20	4.54E-18	74	G alpha (12/13) signalling events
rctm0774	1.08E-19	5.78E-18	10	Netrin mediated repulsion signals
M7253	1.56E-19	8.08E-18	199	Focal adhesion
rctm0080	1.83E-19	9.21E-18	4	Activation, myristolyation of BID and translocation to mitochondria
rctm0635	1.97E-19	9.67E-18	100	L13a-mediated translational silencing of Ceruloplasmin expression
M7721	2.32E-19	1.11E-17	16	Eukaryotic protein translation
rctm0655	2.84E-19	1.29E-17	9	Lysosphingolipid and LPA receptors
M11792	2.87E-19	1.30E-17	16	Sonic Hedgehog (Shh) Pathway
rctm1302	4.09E-19	1.80E-17	140	Translation
rctm0570	4.80E-19	2.06E-17	8	Inactivation of Cdc42 and Rac
rctm0389	5.48E-19	2.30E-17	5	Extrinsic Pathway
M7014	7.32E-19	3.00E-17	30	PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase
rctm0382	1.57E-18	6.31E-17	10	Ethanol oxidation
rctm0686	1.31E-17	5.14E-16	495	Metabolism of lipids and lipoproteins
M3228	2.30E-17	8.81E-16	84	Small cell lung cancer
rctm0298	2.38E-17	8.93E-16	14	DCC mediated attractive signaling
rctm0419	2.80E-17	1.03E-15	24	Fertilization
rctm0388	3.55E-17	1.26E-15	211	Extracellular matrix organization
rctm1271	3.58E-17	1.27E-15	27	The role of Nef in HIV-1 replication and disease pathogenesis
rctm0681	7.14E-17	2.47E-15	16	Metabolism of Angiotensinogen to Angiotensins
M7098	8.21E-17	2.79E-15	84	ECM-receptor interaction
rctm0383	8.73E-17	2.90E-15	84	Eukaryotic Translation Elongation
M10792	1.07E-16	3.50E-15	266	MAPK signaling pathway
rctm0216	1.47E-16	4.72E-15	7	Cation-coupled Chloride cotransporters
rctm0368	2.21E-16	6.95E-15	41	Elastic fibre formation
rctm0437	2.99E-16	9.13E-15	46	Formation of the ternary complex, and subsequently, the 43S complex
rctm0077	3.01E-16	9.19E-15	54	Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S
rctm0385	3.68E-16	1.10E-14	81	Eukaryotic Translation Termination
rctm1303	6.01E-16	1.75E-14	53	Translation initiation complex formation
rctm0852	6.07E-16	1.76E-14	81	Peptide chain elongation
M16697	6.31E-16	1.80E-14	15	CD40L Signaling Pathway
M18155	1.02E-15	2.86E-14	137	Insulin signaling pathway
M15247	1.37E-15	3.78E-14	133	Ubiquitin mediated proteolysis
rctm1358	1.80E-15	4.87E-14	82	Viral mRNA Translation
rctm1412	2.26E-15	6.05E-14	9	p75NTR regulates axonogenesis
M13883	6.27E-15	1.65E-13	34	How Progesterone Initiates Oocyte Membrane
rctm1060	8.15E-15	2.12E-13	9	Role of Abl in Robo-Slit signaling
rctm0569	1.20E-14	3.06E-13	18	Inactivation of APC/C via direct inhibition of the APC/C complex

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
M6355	1.43E-14	3.61E-13	24	Erk and PI-3 Kinase Are Necessary for Collagen Binding in Corneal Epithelia
rctm0275	1.54E-14	3.82E-13	56	Costimulation by the CD28 family
rctm1140	2.18E-14	5.34E-13	32	Signaling by NOTCH2
rctm1108	4.84E-14	1.17E-12	32	Signal amplification
rctm0743	5.52E-14	1.32E-12	63	NCAM signaling for neurite out-growth
rctm0183	7.70E-14	1.81E-12	31	CD28 co-stimulation
rctm0576	1.00E-13	2.33E-12	98	Influenza Viral RNA Transcription and Replication
rctm0588	1.11E-13	2.54E-12	18	Inhibition of the proteolytic activity of APC/C required for the onset of anaphase by mitotic spindle checkpoint components
rctm1038	1.43E-13	3.23E-12	24	Reproduction
M12950	1.88E-13	4.19E-12	13	Angiotensin-converting enzyme 2 regulates heart function
rctm0145	2.31E-13	5.09E-12	256	Axon guidance
M10547	2.63E-13	5.72E-12	19	Role of MAL in Rho-Mediated Activation of SRF
rctm0893	3.15E-13	6.78E-12	55	Pre-NOTCH Expression and Processing
M17636	4.15E-13	8.73E-12	17	Renin-angiotensin system
M7825	4.16E-13	8.74E-12	20	Y branching of actin filaments
M7963	4.30E-13	8.92E-12	124	Cell cycle
rctm1196	6.35E-13	1.30E-11	4	Synthesis of IPs in the nucleus
M9670	6.44E-13	1.31E-11	25	TNF/Stress Related Signaling
rctm0146	9.24E-13	1.85E-11	8	Axonal growth inhibition (RHOA activation)
rctm1070	1.30E-12	2.57E-11	14	SEMA3A-Plexin repulsion signaling by inhibiting Integrin adhesion
rctm1162	1.91E-12	3.75E-11	22	Smooth Muscle Contraction
rctm0543	1.99E-12	3.86E-11	12	Hormone-sensitive lipase (HSL)-mediated triacylglycerol hydrolysis
rctm1205	2.02E-12	3.88E-11	3	Synthesis of PIPs at the ER membrane
M17294	2.49E-12	4.72E-11	23	Ras Signaling Pathway
rctm1030	3.64E-12	6.84E-11	6	Release of eIF4E
M17807	3.83E-12	7.10E-11	55	Basal cell carcinoma
rctm1096	4.87E-12	8.95E-11	15	Sema3A PAK dependent Axon repulsion
rctm0207	5.09E-12	9.26E-11	10	Calcitonin-like ligand receptors
rctm0332	5.19E-12	9.34E-11	85	Diseases associated with visual transduction
rctm0788	5.65E-12	1.01E-10	101	Nonsense Mediated Decay Enhanced by the Exon Junction Complex
rctm0273	7.02E-12	1.24E-10	17	Conversion from APC/C:Cdc20 to APC/C:Cdh1 in late anaphase
rctm0722	7.39E-12	1.29E-10	19	Mitotic Spindle Checkpoint
M3342	8.85E-12	1.53E-10	38	Integrin Signaling Pathway
rctm1360	1.00E-11	1.71E-10	84	Visual phototransduction
M984	1.11E-11	1.88E-10	37	Toll-Like Receptor Pathway
rctm0520	1.68E-11	2.83E-10	4	HCN channels
rctm1400	1.72E-11	2.86E-10	11	mTORC1-mediated signalling
rctm0575	2.12E-11	3.50E-10	133	Influenza Life Cycle
M16120	2.94E-11	4.80E-10	13	How does salmonella hijack a cell
M16334	4.39E-11	7.11E-10	10	Eph Kinases and ephrins support platelet aggregation
M17673	4.67E-11	7.49E-10	74	Cardiac muscle contraction
M10082	4.82E-11	7.66E-10	18	TNFR2 Signaling Pathway
rctm0477	4.90E-11	7.72E-10	31	GPVI-mediated activation cascade
M6487	5.23E-11	8.17E-10	14	Platelet Amyloid Precursor Protein Pathway
M3578	5.97E-11	9.24E-10	85	Progesterone-mediated oocyte maturation
rctm0072	7.08E-11	1.09E-09	5	Activation of RAS in B Cells
rctm0582	7.81E-11	1.19E-09	28	Inhibition of Insulin Secretion by Adrenaline/Noradrenaline
rctm0996	9.64E-11	1.45E-09	76	Regulation of Insulin Secretion
rctm0070	1.03E-10	1.54E-09	4	Activation of PKB
rctm1087	1.10E-10	1.63E-09	104	SRP-dependent cotranslational protein targeting to membrane
rctm0327	1.14E-10	1.67E-09	409	Developmental Biology
rctm1160	1.70E-10	2.48E-09	15	Signalling to p38 via RIT and RIN
rctm0006	2.25E-10	3.24E-09	9	A third proteolytic cleavage releases NICD
rctm0758	2.26E-10	3.24E-09	12	NOTCH2 intracellular domain regulates transcription
M19540	2.35E-10	3.34E-09	121	Oxidative phosphorylation
rctm0991	2.51E-10	3.54E-09	9	Regulation of Gene Expression by Hypoxia-inducible Factor
rctm1148	2.77E-10	3.88E-09	85	Signaling by Type 1 Insulin-like Growth Factor 1 Receptor (IGF1R)
rctm0823	2.79E-10	3.88E-09	12	PECAM1 interactions
M1519	2.99E-10	4.13E-09	174	Endocytosis
M8728	3.05E-10	4.18E-09	83	Hypertrophic cardiomyopathy (HCM)
rctm0790	3.93E-10	5.33E-09	101	Nonsense-Mediated Decay
rctm0047	4.16E-10	5.60E-09	3	Activation of BIM and translocation to mitochondria
rctm0574	5.16E-10	6.90E-09	137	Influenza Infection
rctm0115	5.93E-10	7.87E-09	32	Antigen Activates B Cell Receptor Leading to Generation of Second Messengers
rctm1199	6.16E-10	8.11E-09	6	Synthesis of Lipoxins (LX)
rctm0775	6.36E-10	8.31E-09	42	Netrin-1 signaling
M14899	7.10E-10	9.21E-09	33	Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling
M16004	7.84E-10	1.01E-08	65	Antigen processing and presentation
rctm0553	7.91E-10	1.01E-08	85	IGF1R signaling cascade
rctm1187	9.25E-10	1.17E-08	6	Synthesis of 5-eicosatetraenoic acids

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
rctm0868	1.07E-09	1.35E-08	17	Phosphorylation of the APC/C
rctm0522	1.08E-09	1.36E-08	205	HIV Infection
M11079	1.20E-09	1.49E-08	46	N-Glycan biosynthesis
M2333	1.49E-09	1.83E-08	56	Pathogenic Escherichia coli infection
rctm0011	1.50E-09	1.84E-08	25	ADP signalling through P2Y purinoceptor 1
rctm0709	1.62E-09	1.97E-08	49	Mitochondrial Protein Import
rctm0318	1.70E-09	2.05E-08	12	Depolarization of the Presynaptic Terminal Triggers the Opening of Calcium Channels
M3494	1.87E-09	2.24E-08	36	Thrombin signaling and protease-activated receptors
M13863	2.00E-09	2.38E-08	87	MAPKinase Signaling Pathway
M14449	2.23E-09	2.64E-08	18	METS affect on Macrophage Differentiation
M10765	2.57E-09	3.02E-08	11	Lck and Fyn tyrosine kinases in initiation of TCR Activation
rctm0121	2.78E-09	3.25E-08	13	Apoptosis induced DNA fragmentation
rctm0967	2.83E-09	3.29E-08	16	Rap1 signalling
rctm1000	3.49E-09	4.03E-08	43	Regulation of Insulin Secretion by Glucagon-like Peptide-1
rctm0012	3.61E-09	4.14E-08	22	ADP signalling through P2Y purinoceptor 12
M15371	4.18E-09	4.76E-08	10	Electron Transport Reaction in Mitochondria
rctm0056	4.36E-09	4.93E-08	13	Activation of DNA fragmentation factor
rctm0860	4.76E-09	5.35E-08	9	Phenylalanine and tyrosine catabolism
M15422	5.47E-09	6.11E-08	11	Vitamin C in the Brain
rctm1157	6.73E-09	7.47E-08	37	Signalling to ERKs
rctm0255	7.53E-09	8.30E-08	19	Cobalamin (Cbl, vitamin B12) transport and metabolism
M5109	9.09E-09	9.96E-08	98	Pyrimidine metabolism
rctm0870	9.34E-09	1.02E-07	11	Platelet Adhesion to exposed collagen
rctm0943	9.91E-09	1.07E-07	28	RNA Polymerase I Promoter Opening
M13486	1.16E-08	1.25E-07	172	Huntington's disease
rctm1075	1.32E-08	1.42E-07	18	SHC-related events triggered by IGF1R
M4085	1.39E-08	1.48E-07	35	Primary immunodeficiency
rctm0498	1.44E-08	1.52E-07	30	Glucagon-type ligand receptors
rctm1226	1.75E-08	1.84E-07	13	Synthesis, Secretion, and Inactivation of Glucose-dependent Insulinotropic Polypeptide (GIP)
M5374	2.29E-08	2.39E-07	10	The SARS-coronavirus Life Cycle
M11521	2.41E-08	2.50E-07	61	Glycolysis / Gluconeogenesis
rctm0620	2.54E-08	2.62E-07	11	Interleukin-7 signaling
rctm1198	2.75E-08	2.82E-07	20	Synthesis of Leukotrienes (LT) and Eoxins (EX)
rctm1289	3.20E-08	3.27E-07	16	Trafficking of GluR2-containing AMPA receptors
M16763	3.42E-08	3.47E-07	126	Neurotrophin signaling pathway
rctm0352	3.71E-08	3.74E-07	52	ECM proteoglycans
rctm0564	4.33E-08	4.34E-07	82	IRS-related events triggered by IGF1R
M13143	4.94E-08	4.93E-07	18	uCalpain and friends in Cell spread
rctm0592	5.31E-08	5.26E-07	44	Inositol phosphate metabolism
rctm0760	6.65E-08	6.56E-07	15	NRIF signals cell death from the nucleus
rctm0926	6.92E-08	6.79E-07	13	Purine salvage
M15109	7.16E-08	6.98E-07	10	Glycolysis Pathway
rctm1130	9.25E-08	8.96E-07	108	Signaling by Insulin receptor
rctm1145	9.43E-08	9.09E-07	31	Signaling by Robo receptor
M3896	9.54E-08	9.14E-07	54	Inositol phosphate metabolism
M2642	1.13E-07	1.08E-06	85	TGF-beta signaling pathway
M1001	1.14E-07	1.08E-06	32	Rho cell motility signaling pathway
rctm0872	1.97E-07	1.86E-06	204	Platelet activation, signaling and aggregation
rctm0185	2.09E-07	1.96E-06	11	CD28 dependent Vav1 pathway
rctm1184	2.42E-07	2.26E-06	9	Synthesis of (16-20)-hydroxyeicosatetraenoic acids (HETE)
rctm0969	2.69E-07	2.49E-06	4	Reactions specific to the complex N-glycan synthesis pathway
rctm0934	2.77E-07	2.56E-06	5	RAF activation
rctm0640	2.87E-07	2.64E-06	114	Late Phase of HIV Life Cycle
M7239	3.27E-07	2.99E-06	10	Apoptotic DNA fragmentation and tissue homeostasis
M1296	3.36E-07	3.06E-06	17	IL-7 Signal Transduction
M17941	3.40E-07	3.08E-06	26	Neuropeptides VIP and PACAP inhibit the apoptosis of activated T cells
M4791	3.50E-07	3.14E-06	24	Regulation of eIF4e and p70 S6 Kinase
rctm1025	3.51E-07	3.14E-06	3	Regulation of thyroid hormone activity
M5539	3.56E-07	3.17E-06	129	Axon guidance
rctm1133	3.61E-07	3.20E-06	120	Signaling by NOTCH
rctm0019	4.21E-07	3.71E-06	21	APC/C:Cdc20 mediated degradation of Cyclin B
rctm0325	4.75E-07	4.16E-06	17	Destabilization of mRNA by KSRP
M638	5.11E-07	4.46E-06	73	Adherens junction
rctm0731	5.83E-07	5.06E-06	49	Muscle contraction
rctm0120	6.07E-07	5.25E-06	157	Apoptosis
rctm0017	7.15E-07	6.16E-06	23	APC-Cdc20 mediated degradation of Nek2A
rctm1056	8.12E-07	6.95E-06	12	Reversible Hydration of Carbon Dioxide
M15898	8.36E-07	7.13E-06	34	Fructose and mannose metabolism
M835	8.42E-07	7.14E-06	90	Dilated cardiomyopathy
M17370	8.46E-07	7.14E-06	10	Role of Ran in mitotic spindle regulation
rctm0786	8.53E-07	7.17E-06	44	Non-integrin membrane-ECM interactions

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
M7561	8.61E-07	7.19E-06	52	mTOR signaling pathway
M16991	8.79E-07	7.28E-06	20	Skeletal muscle hypertrophy is regulated via AKT/mTOR pathway
M6856	8.79E-07	7.28E-06	79	Hematopoietic cell lineage
rctm1309	9.58E-07	7.89E-06	51	Translocation of GLUT4 to the Plasma Membrane
M5291	9.81E-07	8.05E-06	16	Role of PI3K subunit p85 in regulation of Actin Organization and Cell Migration
rctm1404	1.02E-06	8.33E-06	13	p38MAPK events
rctm0858	1.06E-06	8.65E-06	71	Phase 1 - Functionalization of compounds
rctm0757	1.11E-06	8.97E-06	21	NOTCH2 Activation and Transmission of Signal to the Nucleus
rctm0442	1.15E-06	9.22E-06	18	Frs2-mediated activation
rctm0572	1.40E-06	1.12E-05	22	Incretin Synthesis, Secretion, and Inactivation
rctm0073	1.43E-06	1.14E-05	14	Activation of Rac
rctm0682	1.54E-06	1.22E-05	259	Metabolism of RNA
rctm0312	1.89E-06	1.49E-05	39	Defensins
M15798	1.90E-06	1.49E-05	70	Melanoma
M4470	2.16E-06	1.70E-05	13	Extrinsic Prothrombin Activation Pathway
M4891	2.26E-06	1.76E-05	17	Regulation of transcriptional activity by PML
rctm1023	2.40E-06	1.86E-05	18	Regulation of signaling by CBL
rctm0825	2.53E-06	1.96E-05	50	PI Metabolism
rctm0546	2.55E-06	1.96E-05	4	Hyaluronan biosynthesis and export
M4629	2.66E-06	2.04E-05	23	Nitrogen metabolism
rctm0988	2.97E-06	2.26E-05	4	Regulation of Commissural axon pathfinding by Slit and Robo
rctm0167	2.98E-06	2.26E-05	12	Bile salt and organic anion SLC transporters
rctm0913	3.21E-06	2.43E-05	20	Prolonged ERK activation events
rctm0259	3.46E-06	2.61E-05	83	Collagen formation
M13036	3.72E-06	2.79E-05	35	Prion diseases
rctm0165	3.97E-06	2.97E-05	9	Bicarbonate transporters
M15997	4.79E-06	3.57E-05	23	Intrinsic Prothrombin Activation Pathway
rctm0981	5.25E-06	3.89E-05	10	Regulated proteolysis of p75NTR
M8232	5.70E-06	4.21E-05	69	Long-term depression
rctm1291	7.15E-06	5.24E-05	172	Transcription
rctm1182	7.18E-06	5.26E-05	20	Syndecan interactions
rctm0046	7.24E-06	5.28E-05	25	Activation of BH3-only proteins
rctm0903	8.02E-06	5.82E-05	135	Processing of Capped Intron-Containing Pre-mRNA
M3261	8.34E-06	6.03E-05	102	Toll-like receptor signaling pathway
rctm0028	8.50E-06	6.12E-05	5	Abacavir metabolism
rctm0662	1.08E-05	7.77E-05	88	MHC class II antigen presentation
rctm1113	1.24E-05	8.83E-05	23	Signaling by BMP
rctm0687	1.24E-05	8.84E-05	214	Metabolism of mRNA
rctm0504	1.35E-05	9.56E-05	30	Glutamate Binding, Activation of AMPA Receptors and Synaptic Plasticity
rctm0724	1.40E-05	9.88E-05	30	Molecules associated with elastic fibres
M13324	1.47E-05	0.0001	15	Hypoxia-Inducible Factor in the Cardiovascular System
rctm0177	1.51E-05	0.00011	18	Botulinum neurotoxicity
rctm0562	1.67E-05	0.00012	78	IRS-mediated signalling
rctm0915	1.71E-05	0.00012	19	Prostacyclin signalling through prostacyclin receptor
M9152	1.75E-05	0.00012	23	CCR3 signaling in Eosinophils
M9503	1.80E-05	0.00012	30	FAS signaling pathway (CD95)
M4844	1.82E-05	0.00013	189	Chemokine signaling pathway
rctm0924	1.88E-05	0.00013	33	Purine metabolism
rctm0164	1.90E-05	0.00013	6	Beta-oxidation of very long chain fatty acids
rctm0140	1.97E-05	0.00013	4	Astrocytic Glutamate-Glutamine Uptake And Metabolism
rctm0567	1.98E-05	0.00013	82	Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell
rctm0782	2.00E-05	0.00013	4	Neurotransmitter uptake and Metabolism In Glial Cells
rctm0676	2.00E-05	0.00013	52	Meiotic Recombination
rctm0451	2.10E-05	0.00014	25	G beta:gamma signalling through PI3Kgamma
rctm0512	2.22E-05	0.00015	10	Glycoprotein hormones
rctm0993	2.25E-05	0.00015	25	Regulation of Hypoxia-inducible Factor (HIF) by Oxygen
rctm0201	2.26E-05	0.00015	22	CTLA4 inhibitory signaling
rctm0735	2.39E-05	0.00016	87	MyD88:Mal cascade initiated on plasma membrane
M9177	2.43E-05	0.00016	22	Dendritic cells in regulating TH1 and TH2 Development
M9488	2.54E-05	0.00017	60	Retinol metabolism
rctm0857	2.63E-05	0.00017	33	Phagosomal maturation (early endosomal stage)
rctm0544	2.72E-05	0.00018	124	Host Interactions of HIV factors
rctm1285	2.76E-05	0.00018	87	Toll Like Receptor TLR6:TLR2 Cascade
rctm0744	3.17E-05	0.0002	37	NCAM1 interactions
M16743	3.21E-05	0.00021	42	Tyrosine metabolism
rctm0594	3.47E-05	0.00022	22	Insulin Processing
rctm0453	4.13E-05	0.00026	25	G protein gated Potassium channels
M963	4.23E-05	0.00027	57	RNA degradation
rctm1284	4.25E-05	0.00027	87	Toll Like Receptor TLR1:TLR2 Cascade
rctm0563	4.37E-05	0.00028	80	IRS-related events
M7739	4.59E-05	0.00029	28	Links between Pyk2 and Map Kinases

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
M12868	4.74E-05	0.0003	323	Pathways in cancer
rctm0677	4.79E-05	0.0003	56	Meiotic Synapsis
M16794	5.57E-05	0.00035	66	Metabolism of xenobiotics by cytochrome P450
rctm1219	5.60E-05	0.00035	8	Synthesis of epoxy (EET) and dihydroxyeicosatrienoic acids (DHET)
rctm1002	5.72E-05	0.00035	16	Regulation of KIT signaling
rctm0109	5.83E-05	0.00036	2	Amplification of signal from unattached kinetochores via a MAD2 inhibitory signal
rctm0777	5.96E-05	0.00037	284	Neuronal System
rctm1158	6.32E-05	0.00038	27	Signalling to RAS
rctm1223	6.33E-05	0.00039	14	Synthesis of very long-chain fatty acyl-CoAs
rctm1335	6.70E-05	0.00041	38	Triglyceride Biosynthesis
rctm0161	6.95E-05	0.00042	3	Beta oxidation of palmitoyl-CoA to myristoyl-CoA
rctm1278	7.75E-05	0.00047	87	Toll Like Receptor 2 (TLR2) Cascade
rctm1072	7.80E-05	0.00047	28	SHC-mediated cascade
rctm0148	7.83E-05	0.00047	8	BH3-only proteins associate with and inactivate anti-apoptotic BCL-2 members
M2623	7.88E-05	0.00047	21	IGF-1 Signaling Pathway
rctm0154	8.05E-05	0.00048	30	Beta defensins
rctm1100	8.16E-05	0.00048	65	Semaphorin interactions
rctm0300	8.44E-05	0.0005	3	DNA Damage Bypass
rctm0199	8.63E-05	0.00051	16	CRMPs in Sema3A signaling
M9726	9.01E-05	0.00053	70	Pancreatic cancer
rctm0597	9.67E-05	0.00056	85	Insulin receptor signalling cascade
rctm0227	9.86E-05	0.00057	25	Cellular response to hypoxia
rctm0483	0.0001	0.00059	5	Galactose catabolism
rctm0398	0.0001	0.0006	5	FGFR1b ligand binding and activation
rctm1383	0.0001	0.0006	3	c-src mediated regulation of Cx43 function and closure of gap junctions
rctm1396	0.00011	0.00064	104	mRNA Splicing
M13968	0.00011	0.00064	58	HIV-I Nef: negative effector of Fas and TNF
rctm0454	0.00011	0.00064	28	G-protein activation
rctm0060	0.00011	0.00065	25	Activation of Genes by ATF4
rctm0578	0.00011	0.00065	25	Inhibition of voltage gated Ca2+ channels via Gbeta/gamma subunits
M4361	0.00012	0.00068	23	Proximal tubule bicarbonate reclamation
M862	0.00012	0.00069	40	p38 MAPK Signaling Pathway
M5410	0.00013	0.00072	58	Arachidonic acid metabolism
rctm0452	0.00013	0.00074	20	G beta:gamma signalling through PLC beta
rctm0449	0.00014	0.00077	126	G alpha (s) signalling events
M321	0.00014	0.00079	73	Chronic myeloid leukemia
M4741	0.00014	0.00079	115	Systemic lupus erythematosus
rctm0190	0.00014	0.00079	8	CHL1 interactions
rctm0863	0.00015	0.00081	136	Phospholipid metabolism
rctm0480	0.00015	0.00082	15	GRB2:SOS provides linkage to MAPK signaling for Intergrins
rctm0326	0.00015	0.00083	17	Destabilization of mRNA by Tristetraprolin (TTP)
M14314	0.00017	0.00091	159	Purine metabolism
rctm0951	0.00017	0.00091	94	RNA Polymerase II Transcription
rctm0615	0.00017	0.00095	7	Interleukin-1 processing
rctm0370	0.00018	0.00096	3	Electron Transport from NADPH to Ferredoxin
rctm0024	0.00018	0.00098	16	ARMS-mediated activation
rctm0523	0.00019	0.00103	126	HIV Life Cycle
M19784	0.00021	0.0011	46	T Cell Receptor Signaling Pathway
rctm1073	0.00021	0.00113	15	SHC-mediated signalling
M699	0.00021	0.00113	42	Fatty acid metabolism
rctm0228	0.00021	0.00114	25	Cellular responses to stress
rctm0935	0.00022	0.00116	7	RAF phosphorylates MEK
rctm0661	0.00022	0.00116	7	MEK activation
rctm0104	0.00023	0.00123	15	Amine-derived hormones
M85	0.00023	0.00123	25	Regulation of BAD phosphorylation
rctm1288	0.00024	0.00123	30	Trafficking of AMPA receptors
rctm0062	0.00024	0.00123	30	Activation of Kainate Receptors upon glutamate binding
rctm1142	0.00024	0.00127	11	Signaling by NOTCH4
rctm1013	0.00026	0.00133	3	Regulation of gap junction activity
rctm0135	0.00027	0.00139	4	Assembly of the RAD50-MRE11-NBS1 complex at DNA double-strand breaks
rctm0057	0.00028	0.00143	25	Activation of G protein gated Potassium channels
rctm0110	0.00029	0.0015	2	Amplification of signal from the kinetochores
rctm0674	0.00029	0.00151	4	MRN complex relocalizes to nuclear foci
rctm1258	0.0003	0.00154	25	Termination of O-glycan biosynthesis
rctm0771	0.00031	0.00156	8	Negative regulation of the PI3K/AKT network
rctm1372	0.00033	0.00167	111	WNT ligand biogenesis and trafficking
rctm0032	0.00033	0.00169	4	Abnormal metabolism in phenylketonuria
rctm0838	0.00034	0.00171	28	PKB-mediated events
M9257	0.00034	0.00173	68	Drug metabolism - cytochrome P450
M2288	0.00035	0.00174	53	NFAT and Hypertrophy of the heart (Transcription in the broken heart)
rctm0005	0.00036	0.00179	26	A tetrasaccharide linker sequence is required for GAG synthesis

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
rctm1409	0.00036	0.0018	82	p75 NTR receptor-mediated signalling
rctm1110	0.00036	0.0018	13	Signal regulatory protein (SIRP) family interactions
rctm0048	0.00036	0.00181	3	Activation of BMF and translocation to mitochondria
rctm1247	0.00037	0.00183	5	Tachykinin receptors bind tachykinins
rctm1112	0.00037	0.00183	13	Signaling by Activin
rctm0828	0.00037	0.00183	67	PI3K Cascade
rctm0395	0.00037	0.00183	31	FCGR activation
rctm0818	0.0004	0.00194	3	Oxygen-dependent Asparagine Hydroxylation of Hypoxia-inducible Factor Alpha
rctm1209	0.0004	0.00194	33	Synthesis of PIPs at the plasma membrane
rctm1312	0.0004	0.00196	191	Transmission across Chemical Synapses
rctm0837	0.0004	0.00197	5	PKA-mediated phosphorylation of key metabolic factors
rctm0007	0.00043	0.00207	36	ABC-family proteins mediated transport
M2064	0.00043	0.00207	17	The 4-1BB-dependent immune response
M7897	0.00046	0.00221	23	Inhibition of Cellular Proliferation by Gleevec
rctm0899	0.00046	0.00222	21	Presynaptic function of Kainate receptors
rctm0457	0.00048	0.00228	25	G0 and Early G1
M16817	0.00049	0.00234	112	Oocyte meiosis
M10628	0.00051	0.00241	20	ATM Signaling Pathway
rctm1101	0.00051	0.00241	160	Separation of Sister Chromatids
rctm0638	0.00054	0.00256	5	LRR FLII-interacting protein 1 (LRRFIP1) activates type I IFN production
rctm0258	0.00056	0.00263	61	Collagen biosynthesis and modifying enzymes
M1053	0.00057	0.00268	56	Hedgehog signaling pathway
M16884	0.0006	0.0028	16	Role of nicotinic acetylcholine receptors in the regulation of apoptosis
M16811	0.0006	0.00282	18	Role of Erk5 in Neuronal Survival
rctm1051	0.00062	0.00288	13	Retrograde neurotrophin signalling
rctm1265	0.00063	0.00295	16	The canonical retinoid cycle in rods (twilight vision)
M16376	0.00063	0.00295	74	Arrhythmogenic right ventricular cardiomyopathy (ARVC)
rctm0766	0.00065	0.00302	8	Nef and signal transduction
M7272	0.00065	0.00303	116	Parkinson's disease
rctm0500	0.00068	0.00312	31	Gluconeogenesis
rctm0738	0.00068	0.00313	20	N-glycan antennae elongation in the medial/trans-Golgi
rctm1104	0.00072	0.00332	4	Serotonin and melatonin biosynthesis
rctm1037	0.00072	0.00332	15	Repair synthesis of patch ~27-30 bases long by DNA polymerase
M6220	0.00075	0.00342	36	Agrin in Postsynaptic Differentiation
rctm0649	0.00075	0.00342	4	Localization of the PINCH-ILK-PARVIN complex to focal adhesions
M15913	0.00075	0.00343	71	RIG-I-like receptor signaling pathway
M19613	0.00076	0.00345	23	Multiple antiapoptotic pathways from IGF-1R signaling lead to BAD phosphorylation
M2821	0.00076	0.00345	24	NFkB activation by Nontypeable Hemophilus influenzae
rctm0747	0.00078	0.0035	13	NF-kB is activated and signals survival
rctm1304	0.0008	0.0036	3	Translesion synthesis by DNA polymerases bypassing lesion on DNA template
rctm0641	0.00082	0.00368	33	Latent infection of Homo sapiens with Mycobacterium tuberculosis
rctm0200	0.00082	0.00369	13	CS/DS degradation
rctm0354	0.00084	0.00377	26	EGFR downregulation
rctm1243	0.00087	0.00388	3	TWIK related potassium channel (TREK)
rctm1246	0.00088	0.0039	2	TWIK-releated acid-sensitive K+ channel (TASK)
M17400	0.00092	0.00408	37	ALK in cardiac myocytes
M9011	0.00093	0.0041	44	Vasopressin-regulated water reabsorption
rctm0039	0.00094	0.00415	106	Activated TLR4 signalling
M4383	0.00094	0.00415	30	Actions of Nitric Oxide in the Heart
M12836	0.00098	0.00428	19	EPO Signaling Pathway
rctm0263	0.00098	0.00429	5	Conjugation of SUMO to E1 (UBA2:SAE1)
rctm0596	0.00098	0.0043	25	Insulin receptor recycling
rctm0565	0.00101	0.00439	70	ISG15 antiviral mechanism
rctm0514	0.00101	0.0044	38	Glycosphingolipid metabolism
M3061	0.00103	0.00447	9	Feeder Pathways for Glycolysis
rctm0919	0.00111	0.00482	16	Proteolytic cleavage of SNARE complex proteins
rctm0723	0.00113	0.00485	14	Mitotic Telophase/Cytokinesis
rctm0209	0.00115	0.00493	11	Calnexin/calreticulin cycle
rctm0310	0.00118	0.00508	48	Deadenylation-dependent mRNA decay
rctm1308	0.00131	0.00559	2	Translocation of BoNT Light chain
M225	0.00133	0.00567	19	Downregulated of MTA-3 in ER-negative Breast Tumors
M11844	0.00135	0.00577	56	Cytosolic DNA-sensing pathway
rctm1036	0.00138	0.00586	15	Repair synthesis for gap-filling by DNA polymerase in TC-NER
M2044	0.00139	0.00589	124	Spliceosome
rctm1159	0.00144	0.0061	3	Signalling to STAT3
rctm0159	0.00148	0.00626	3	Beta oxidation of myristoyl-CoA to lauroyl-CoA
rctm1132	0.0015	0.00632	18	Signaling by NODAL
M18312	0.00155	0.00653	25	Maturity onset diabetes of the young
rctm0128	0.00161	0.00674	48	Arachidonic acid metabolism
rctm0405	0.00162	0.00678	7	FGFR3b ligand binding and activation
M13158	0.00163	0.00679	15	The IGF-1 Receptor and Longevity

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
M1724	0.00165	0.00686	29	RNA polymerase
rctm0246	0.00169	0.00701	243	Class I MHC mediated antigen processing & presentation
M19943	0.0017	0.00706	22	Ceramide Signaling Pathway
rctm1399	0.00176	0.00726	27	mTOR signalling
rctm0257	0.00176	0.00727	10	Cohesin Loading onto Chromatin
rctm1194	0.00181	0.00742	25	Synthesis of IP3 and IP4 in the cytosol
rctm0455	0.00181	0.00743	28	G-protein beta:gamma signalling
rctm0675	0.00182	0.00743	83	Meiosis
M12095	0.00184	0.00751	33	Signal transduction through IL1R
rctm0071	0.00197	0.00801	4	Activation of PUMA and translocation to mitochondria
rctm0119	0.00198	0.00802	70	Antiviral mechanism by IFN-stimulated genes
rctm1141	0.00198	0.00802	11	Signaling by NOTCH3
M19358	0.00208	0.00842	37	Signaling of Hepatocyte Growth Factor Receptor
rctm0133	0.00209	0.00843	54	Assembly of collagen fibrils and other multimeric structures
M19895	0.00209	0.00843	24	Nicotinate and nicotinamide metabolism
rctm0636	0.0021	0.00845	93	L1CAM interactions
rctm1397	0.00218	0.00875	104	mRNA Splicing - Major Pathway
rctm0511	0.00219	0.00876	26	Glycolysis
rctm0779	0.00229	0.00911	7	Neurotransmitter Clearance In The Synaptic Cleft
M12975	0.00229	0.00912	29	VEGF, Hypoxia, and Angiogenesis
rctm0850	0.00231	0.00918	26	Pausing and recovery of Tat-mediated HIV elongation
M13720	0.00235	0.0093	44	Lysine degradation
rctm1098	0.00239	0.00944	24	Sema4D induced cell migration and growth-cone collapse
rctm1250	0.00243	0.00957	12	Tandem pore domain potassium channels
rctm0604	0.00245	0.00965	10	Interaction With The Zona Pellucida
M19888	0.00257	0.01008	57	Acute myeloid leukemia
rctm0698	0.00267	0.01048	25	Metal ion SLC transporters
rctm0623	0.00273	0.01069	31	Inwardly rectifying K+ channels
rctm1262	0.00279	0.01089	3	The NLRP1 inflammasome
M7552	0.0028	0.01089	18	Role of EGF Receptor Transactivation by GPCRs in Cardiac Hypertrophy
rctm0336	0.00284	0.01104	12	Dopamine Neurotransmitter Release Cycle
rctm1067	0.00289	0.01122	9	S6K1-mediated signalling
M10183	0.00294	0.01137	16	Acetylation and Deacetylation of RelA in The Nucleus
rctm0647	0.00296	0.01141	48	Lipid digestion, mobilization, and transport
rctm0415	0.00298	0.01147	173	Fatty acid, triacylglycerol, and ketone body metabolism
rctm0849	0.00305	0.01173	27	Pausing and recovery of HIV elongation
rctm1251	0.00308	0.01183	26	Tat-mediated HIV elongation arrest and recovery
rctm0940	0.00316	0.01209	44	RNA Polymerase I Chain Elongation
rctm0936	0.00321	0.01226	10	RAF/MAP kinase cascade
rctm0768	0.00324	0.01233	9	Nef mediated downregulation of MHC class I complex cell surface expression
rctm1074	0.00324	0.01234	17	SHC-related events
rctm0526	0.00326	0.01236	27	HIV elongation arrest and recovery
rctm0107	0.00327	0.01237	16	Amino acid synthesis and interconversion (transamination)
rctm1044	0.00334	0.01262	99	Resolution of Sister Chromatid Cohesion
M7860	0.00336	0.01266	18	Nerve growth factor pathway (NGF)
M15285	0.00338	0.01268	23	NF-kB Signaling Pathway
rctm0409	0.00338	0.01268	37	FRS2-mediated cascade
rctm1155	0.00339	0.01271	277	Signalling by NGF
rctm0030	0.0034	0.01271	10	Abacavir transport and metabolism
rctm1296	0.00342	0.01276	77	Transcriptional Regulation of White Adipocyte Differentiation
rctm1236	0.00353	0.01313	24	TRAF6 mediated NF-kB activation
rctm0501	0.00361	0.01342	61	Glucose metabolism
rctm1254	0.00364	0.01351	6	Telomere C-strand synthesis initiation
M5202	0.00369	0.01366	22	Hypoxia and p53 in the Cardiovascular system
M16848	0.00371	0.01368	68	Epithelial cell signaling in Helicobacter pylori infection
rctm0485	0.00372	0.01369	11	Gamma-carboxylation, transport, and amino-terminal cleavage of proteins
rctm0086	0.00378	0.01389	17	Acyl chain remodelling of PG
M8601	0.0039	0.0143	23	Rac 1 cell motility signaling pathway
rctm0944	0.00394	0.01443	52	RNA Polymerase I Transcription
rctm1317	0.00403	0.0147	53	Transport of Mature Transcript to Cytoplasm
rctm0221	0.00413	0.01504	60	Cell death signalling via NRAGE, NRIF and NADE
rctm1170	0.00417	0.01518	64	Sphingolipid metabolism
rctm0306	0.00419	0.0152	6	DNA replication initiation
M12399	0.00449	0.01627	12	Cystic Fibrosis Transmembrane Conductance Regulator And Beta 2 Adrenergic Receptor Pathway
rctm0476	0.00471	0.01704	428	GPCR ligand binding
rctm0869	0.00474	0.01705	4	Plasmalogen biosynthesis
M2529	0.00474	0.01705	32	PDGF Signaling Pathway
rctm0147	0.00477	0.01714	4	Axonal growth stimulation
rctm1103	0.00481	0.01718	12	Serotonin Neurotransmitter Release Cycle
rctm0844	0.00481	0.01719	109	PPARA Activates Gene Expression

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
rctm0372	0.00481	0.0172	27	Elongation arrest and recovery
M4377	0.00489	0.01744	26	Galactose metabolism
rctm1210	0.00492	0.01751	2	Synthesis of PS
rctm0810	0.00504	0.01785	8	Opsins
rctm0438	0.00504	0.01785	27	Formation of transcription-coupled NER (TC-NER) repair complex
rctm1086	0.00505	0.01786	14	SOS-mediated signalling
rctm0142	0.00507	0.01789	60	Autodegradation of Cdh1 by Cdh1:APC/C
rctm0678	0.00511	0.01798	176	Membrane Trafficking
rctm0375	0.00517	0.01816	3	Endosomal/Vacuolar pathway
rctm1229	0.00522	0.01829	29	TGF-beta receptor signaling activates SMADs
rctm1200	0.00523	0.0183	28	Synthesis of PA
rctm0598	0.00526	0.01834	103	Integration of energy metabolism
M7857	0.00527	0.01835	13	Non-homologous end-joining
rctm0891	0.00527	0.01837	99	Potassium Channels
rctm1320	0.00534	0.01855	35	Transport of Mature mRNAs Derived from Intronless Transcripts
M15926	0.00555	0.01922	21	Trefoil Factors Initiate Mucosal Healing
rctm1293	0.00569	0.01969	55	Transcription of the HIV genome
M89	0.00581	0.02007	11	Phosphorylation of MEK1 by cdk5/p35 down regulates the MAP kinase pathway
rctm0284	0.00598	0.0206	37	Cyclin D associated events in G1
rctm1143	0.00607	0.02087	175	Signaling by PDGF
rctm0271	0.0061	0.02093	57	Constitutive Signaling by NOTCH1 PEST Domain Mutants
rctm0397	0.00612	0.02096	14	FGFR1 ligand binding and activation
rctm0141	0.00613	0.02099	7	Attachment of GPI anchor to uPAR
rctm0947	0.00625	0.02132	92	RNA Polymerase I, RNA Polymerase III, and Mitochondrial Transcription
M11911	0.00633	0.02156	44	ABC transporters
rctm0740	0.00639	0.02171	13	N-glycan trimming in the ER and Calnexin/Calreticulin cycle
rctm1403	0.0066	0.0224	15	p130Cas linkage to MAPK signaling for integrins
M8492	0.00673	0.02279	86	Apoptosis
M14863	0.00698	0.0236	16	p53 Signaling Pathway
rctm0714	0.007	0.02361	171	Mitotic Anaphase
rctm0979	0.00726	0.02444	27	Recycling pathway of L1
rctm1097	0.00733	0.02463	26	Sema4D in semaphorin signaling
M1909	0.00741	0.02484	31	EGF Signaling Pathway
M11190	0.00799	0.02676	24	Dorso-ventral axis formation
M18937	0.00822	0.02746	43	Nucleotide excision repair
M13191	0.00827	0.02754	89	Prostate cancer
M1547	0.00828	0.02756	29	Control of skeletal myogenesis by HDAC and calcium/calmodulin-dependent kinase (CaMK)
rctm0302	0.00831	0.0276	3	DNA Damage Reversal
rctm0269	0.00834	0.02765	15	Constitutive Signaling by NOTCH1 HD Domain Mutants
rctm0732	0.00847	0.02803	78	MyD88 cascade initiated on plasma membrane
rctm1120	0.00864	0.02853	43	Signaling by FGFR mutants
rctm0618	0.0087	0.02869	43	Interleukin-3, 5 and GM-CSF signaling
rctm0349	0.00873	0.02874	27	Dual incision reaction in TC-NER
rctm0750	0.00877	0.02876	197	NGF signalling via TRKA from the plasma membrane
rctm1336	0.00877	0.02876	11	Tryptophan catabolism
rctm0648	0.00881	0.02884	29	Lipoprotein metabolism
M3618	0.00894	0.02921	29	TNFR1 Signaling Pathway
rctm0528	0.00905	0.02951	21	HS-GAG degradation
rctm0720	0.00929	0.03025	107	Mitotic Prometaphase
rctm0278	0.00941	0.03056	7	Cross-presentation of particulate exogenous antigens (phagosomes)
rctm1091	0.00946	0.03067	19	Scavenging by Class A Receptors
rctm0231	0.00953	0.03086	8	ChREBP activates metabolic gene expression
rctm0784	0.00955	0.03086	8	Nicotinate metabolism
rctm0130	0.00964	0.03108	85	Asparagine N-linked glycosylation
M12985	0.00975	0.03138	15	TACI and BCMA stimulation of B cell immune responses.
rctm1144	0.01026	0.03295	116	Signaling by Rho GTPases
rctm0303	0.0103	0.03303	104	DNA Repair
rctm1008	0.01046	0.03348	43	Regulation of Water Balance by Renal Aquaporins
rctm1315	0.0107	0.03412	2	Transport of Glycerol from Adipocytes to the Liver by Aquaporins
rctm0802	0.0107	0.03412	48	Nucleotide Excision Repair
rctm0708	0.0109	0.03466	9	Mitochondrial Iron-Sulfur Cluster Biogenesis
M16853	0.01091	0.03469	36	DNA replication
M3075	0.01094	0.0347	11	Granzyme A mediated Apoptosis Pathway
M523	0.01132	0.03584	29	Thyroid cancer
rctm0521	0.0115	0.03634	15	HDL-mediated lipid transport
rctm0237	0.01198	0.0378	48	Chondroitin sulfate/dermatan sulfate metabolism
rctm0170	0.01203	0.03789	8	Binding of RNA by Insulin-like Growth Factor-2 mRNA Binding Proteins (IGF2BPs/IMPs/VICKZs)
M15955	0.0124	0.03898	39	Sphingolipid metabolism
rctm0404	0.01245	0.03909	11	FGFR3 ligand binding and activation
rctm0627	0.01254	0.03929	39	Iron uptake and transport

MODULE	P	FDR	NGENES	MODULE DESCRIPTION
rctm1011	0.01315	0.04113	32	Regulation of beta-cell development
rctm0124	0.01343	0.04191	52	Apoptotic execution phase
rctm0726	0.01351	0.04212	2	Monoamines are oxidized to aldehydes by MAOA and MAOB, producing NH3 and H2O2
rctm0022	0.0139	0.04325	69	APC/C:Cdh1 mediated degradation of Cdc20 and other APC/C:Cdh1 targeted proteins in late mitosis/early G1
M15258	0.01416	0.04399	22	AKT Signaling Pathway
rctm0289	0.01424	0.04416	260	Cytokine Signaling in Immune system
rctm1066	0.01441	0.0446	5	S6K1 signalling
rctm0108	0.01449	0.04476	31	Amino acid transport across the plasma membrane
M8560	0.01498	0.04618	24	Cell Cycle: G2/M Checkpoint
M17857	0.01537	0.04731	37	SNARE interactions in vesicular transport
rctm0729	0.01549	0.04753	7	Multifunctional anion exchangers
M19118	0.01549	0.04754	46	Keratinocyte Differentiation
rctm0236	0.0157	0.04809	20	Chondroitin sulfate biosynthesis
rctm1206	0.01611	0.04924	17	Synthesis of PIPs at the Golgi membrane
rctm1032	0.0163	0.04975	9	Removal of aminoterminal propeptides from gamma-carboxylated proteins

P: module enrichment significance, FDR: Benjamini-Hochberg false discovery rate, NGENES: number of genes tested for enrichment

Supplementary Table 20 The weighted key driver analysis (wKDA) using the Bayesian brain network

MODULE	NODE	*P*	FDR	N.mod	N.neigh	N.obsrv	N.expct	MEMBER	FILL	FOLD	DESCR
rctm0388	FMOD	1.15E-16	6.77E-15	190	73	14	0.48	TRUE	0.19	29.45	Extracellular matrix organization
rctm0476	NTS	3.01E-11	8.87E-10	348	35	12	0.42	TRUE	0.34	28.75	G-protein-coupled receptor (GPCR) ligand binding
rctm0388	ISLR	5.13E-11	1.01E-09	190	34	11	0.22	FALSE	0.32	49.69	Extracellular matrix organization
rctm0388	COL3A1	8.06E-11	1.19E-09	190	36	11	0.23	TRUE	0.31	46.93	Extracellular matrix organization
rctm0388	SLC13A4	2.38E-10	2.80E-09	190	50	11	0.33	FALSE	0.22	33.79	Extracellular matrix organization
M7098	FMOD	1.06E-06	1.04E-05	78	73	9	0.20	FALSE	0.12	46.12	Extracellular matrix (ECM)-receptor interaction
rctm0388	SERPING1	9.19E-06	6.81E-05	190	54	9	0.35	FALSE	0.17	25.60	Extracellular matrix organization
rctm0476	ESR1	9.23E-06	6.81E-05	348	31	9	0.37	FALSE	0.29	24.34	G-protein-coupled receptor (GPCR) ligand binding
rctm0388	PCOLCE	1.12E-05	7.33E-05	190	63	9	0.41	TRUE	0.14	21.94	Extracellular matrix organization
rctm0388	COL1A2	9.70E-05	5.72E-04	190	32	8	0.21	TRUE	0.25	38.39	Extracellular matrix organization
rctm0476	GAL	3.58E-04	1.92E-03	348	25	8	0.30	TRUE	0.32	26.83	G-protein-coupled receptor (GPCR) ligand binding
rctm0476	CALCR	9.10E-04	4.47E-03	348	35	8	0.42	TRUE	0.23	19.17	G-protein-coupled receptor (GPCR) ligand binding
M7253	FMOD	1.13E-03	5.04E-03	172	73	8	0.43	FALSE	0.11	18.59	Focal adhesion
rctm0476	ECEL1	1.20E-03	5.04E-03	348	34	8	0.41	FALSE	0.24	19.73	G-protein-coupled receptor (GPCR) ligand binding
rctm0686	KNG1	2.56E-03	9.29E-03	405	48	8	0.67	FALSE	0.17	12.01	Metabolism of lipids and lipoproteins
rctm0289	RTP4	2.58E-03	9.29E-03	197	26	7	0.18	FALSE	0.27	39.88	Cytokine signaling in immune system
rctm0648	KNG1	2.72E-03	9.29E-03	24	48	6	0.04	FALSE	0.13	151.98	Lipoprotein metabolism
rctm0476	ARHGAP6	2.83E-03	9.29E-03	348	63	8	0.75	FALSE	0.13	10.65	G-protein-coupled receptor (GPCR) ligand binding
rctm0259	FMOD	3.08E-03	9.57E-03	73	73	7	0.18	FALSE	0.10	38.33	Collagen formation
rctm0289	IFIT1	3.97E-03	1.17E-02	197	30	7	0.20	TRUE	0.23	34.56	Cytokine Signaling in Immune system
rctm0476	BRS3	6.19E-03	1.74E-02	348	19	7	0.23	TRUE	0.37	30.89	G-protein-coupled receptor (GPCR) ligand binding
rctm0258	ISLR	9.21E-03	2.41E-02	53	34	6	0.06	FALSE	0.18	97.16	Collagen biosynthesis and modifying enzymes
rctm0647	KNG1	9.41E-03	2.41E-02	39	48	6	0.06	FALSE	0.13	93.53	Lipid digestion, mobilization, and transport
rctm0259	ISLR	1.29E-02	3.16E-02	73	34	6	0.09	FALSE	0.18	70.54	Collagen formation
rctm0388	SLC6A20	1.65E-02	3.90E-02	190	50	7	0.33	FALSE	0.14	21.50	Extracellular matrix organization
M7098	COL3A1	2.13E-02	4.83E-02	78	36	6	0.10	TRUE	0.17	62.35	Extracellular matrix (ECM)-receptor interaction
rctm1162	TAGLN	2.25E-02	4.91E-02	20	15	5	0.01	FALSE	0.33	486.33	Smooth muscle contraction

Node: key driver gene tested, P: subnetwork enrichment significance, FDR: Benjamini-Hochberg false discovery rate, N.mod: number of genes in subnetwork, N.neigh: number of genes in the neighborhood of the key driver in the tissue network, N.obsrv: Number of observed genes shared between the subnetwork and the neighborhood, N.expct: expected number of shared genes, MEMBER: indicates if the key driver is part of the subnetwork, FILL: ratio of subnetwork genes in the neighvorhood, and FOLD: fold change of subnetwork genes compared to the expectation.

Supplemental Table 21 Genes mapped to genetic, epigenetic, transcriptional associations used for enrichment in the Mergeomics analysis

(Source) Gene list by each significant module in Supplementary Table 20
M7098: ECM-receptor interaction
(TWAS) ITGAV, COL2A1, COL6A2, COL4A6 **(EWAS)** ITGB1, COL3A1, ITGA4, COL6A1, LAMA1, VTN, SV2A, COL5A3, CHAD, THBS1, ITGA2, VWF, ITGB3, FN1, GP5, COL6A6, CD36, COL11A2, CD44, AGRN, THBS2 **(GWAS)** LAMA3, ITGA8, SDC1, COL1A2, TNC, HSPG2, LAMC1, ITGB7, GP6, ITGA3, THBS4, SV2C, COL4A4, IBSP, CD47, HMMR, TNN, LAMB3, LAMC2, COL1A1, SV2B, SDC2, COL6A3, SPP1, THBS3, LAMA5, LAMC3, RELN, TNXB, COMP, COL11A1, ITGA5, GP9, LAMA2, ITGB5, ITGA10, ITGA6, COL4A2, GP1BA, ITGB4, LAMA4, ITGA1, GP1BB, ITGA2B, ITGA11, SDC4, COL5A2, ITGA9, LAMB2, ITGB8, COL5A1, DAG1, SDC3, ITGB6, LAMB1, LAMB4, ITGA7, TNR, COL4A1
M7253: Focal adhesion
(TWAS) ITGAV, RAC1, VAV2, BIRC3, COL2A1, MAPK3, MYLK3, COL6A2, PAK3, XIAP, FIGF, ELK1, COL4A6, FLNA **(EWAS)** KDR, PGF, ITGB1, PIK3R1, PTEN, ZYX, COL3A1, RAC3, MYL9, PDGFA, ITGA4, COL6A1, CRK, LAMA1, VEGFA, GRB2, MYL7, PDPK1, MYLPF, VTN, PPP1CA, COL5A3, VEGFC, CHAD, THBS1, ITGA2, PPP1CB, SHC2, PIK3R2, VWF, ITGB3, MYLK, ILK, FN1, PIK3CA, RAF1, CCND2, COL6A6, AKT2, CAV2, COL11A2, ACTN3, CCND1, PIK3R5, PIK3CD, PAK1, SHC4, RHOA, CTNNB1, CAV1, THBS2, PDGFRA, ROCK1, PRKCA, PRKCG, ACTG1, FLT1 **(GWAS)** RAP1A, PPP1R12A, HGF, LAMA3, DOCK1, BIRC2, ACTN2, EGF, SHC3, EGFR, HRAS, MAPK9, ITGA8, SOS1, COL1A2, TNC, PXN, MYL12A, MYL12B, LAMC1, ITGB7, CDC42, JUN, IGF1, PAK6, RASGRF1, ITGA3, BCL2, ACTN4, PIP5K1C, THBS4, PDGFRB, COL4A4, PARVB, PARVG, IBSP, PAK7, ACTN1, PDGFD, MAPK10, PTK2, GSK3B, SHC1, TLN1, CAPN2, TNN, LAMB3, PAK2, BAD, VEGFB, LAMC2, COL1A1, VCL, VAV1, VASP, MAPK8, RAP1B, COL6A3, SPP1, RAPGEF1, THBS3, FLNC, ACTB, VAV3, LAMA5, LAMC3, MYL2, RELN, TNXB, COMP, ERBB2, COL11A1, FLNB, DIAPH1, ITGA5, SOS2, MET, LAMA2, PARVA, PIK3R3, ITGB5, CAV3, MYL5, TLN2, IGF1R, ITGA10, ITGA6, PAK4, COL4A2, ITGB4, LAMA4, PPP1CC, PRKCB, ITGA1, CRKL, ITGA2B, ITGA11, SRC, MYLK2, PIK3CB, GRLF1, COL5A2, ITGA9, LAMB2, ITGB8, COL5A1, ROCK2, ITGB6, FYN, BRAF, ARHGAP5, MAP2K1, MAPK1, LAMB1, LAMB4, AKT1, AKT3, MYL10, RAC2, FLT4, ITGA7, PIK3CG, TNR, BCAR1, CCND3, COL4A1, PDGFB, PDGFC
rctm0258: Collagen biosynthesis and modifying enzymes
(TWAS) COL6A5, COL2A1, COL6A2, COL4A5, COL4A6 **(EWAS)** COL3A1, PLOD3, COL6A1, COL5A3, CRTAP, COL16A1, PLOD1, COL20A1, COL6A6, COL25A1, LEPRE1, COL11A2, TLL2, TLL1, BMP1, COL24A1, P4HB, ADAMTS3, COL18A1 **(GWAS)** COL21A1, COL28A1, COL15A1, COL13A1, ADAMTS14, COL1A2, COL19A1, COL9A1, COL17A1, COL4A4, COL23A1, PLOD2, COL10A1, COL14A1, COL1A1, COL12A1, COL6A3, ADAMTS2, COL11A1, PCOLCE, PCOLCE2, COL22A1, COL4A2, PPIB, COL9A3, COL5A2, COL7A1, COL4A3, COL5A1, SERPINH1, COL9A2, COL8A1, LEPREL2, COL8A2, LEPREL1, COL27A1, COL4A1
rctm0259: Collagen formation
(TWAS) COL6A5, PLEC, COL2A1, COL6A2, COL4A5, COL4A6 **(EWAS)** COL3A1, PLOD3, COL6A1, COL5A3, CRTAP, COL16A1, LOXL3, PLOD1, COL20A1, COL6A6, COL25A1, LEPRE1, COL11A2, DST, TLL2, TLL1, BMP1, COL24A1, P4HB, ADAMTS3, COL18A1 **(GWAS)** COL21A1, COL28A1, LAMA3, LOX, COL15A1, COL13A1, ADAMTS14, LOXL2, LOXL4, MMP3, COL1A2, COL19A1, COL9A1, COL17A1, CTSL2, MMP20, COL4A4, MMP7, CTSB, COL23A1, PLOD2, COL10A1, LAMB3, COL14A1, LAMC2, COL1A1, COL12A1, COL6A3, ADAMTS2, COL11A1, CTSS, PCOLCE, PCOLCE2, ITGA6, COL22A1, COL4A2, ITGB4, PPIB, CTSL1, COL9A3, LOXL1, CD151, COL5A2, COL7A1, COL4A3, COL5A1, SERPINH1, COL9A2, COL8A1, MMP13, LEPREL2, COL8A2, MMP9, LEPREL1, COL27A1, COL4A1
rctm0289: Cytokine Signaling in Immune system
(TWAS) GBP6, FCGR1A, GBP5, IL5RA, TAB2, IFNA6, NCAM1, PTPN11, CIITA, IRF8, MAPK3, JAK3, IFNAR1, TAB1, IKBKG, TAB3, IL2RG, IRAK1 **(EWAS)** IFNA14, PPM1B, RBX1, TNIP2, SKP1, NFKB2, SH2B1, NUP210, PIK3R1, CHUK, ADAR, NUP93, CAMK2G, GAB2, CRK, STAT2, GRB2, ADAM17, IFIT1, IFIT3, LCK, ISG15, IRF9, NUP43, GHR, NUP155, CUL1, PIK3R2, RAE1, STAT5B, KPNB1, IFI27, IKBKB, IFI30, OAS2, RANBP2, IFI6, TOLLIP, IRF5, GBP2, PIK3CA, GBP4, HERC5, IFNGR2, CDK1, RAF1, TYK2, NUP85, EIF4G2, STAT1, RIPK2, IRAK2, MT2A, MAP3K7, NUP133, UBE2E1, PELI2, IRAK3, NOD2, IFNA7, KPNA4, CD44, MAP2K2, EGR1, IFITM3, NUP153, STAT5A, IL1A, IL7, NUP98, PIK3CD, IRS1, KPNA7, KPNA2, JAK2, IFITM1, IFITM2, PTAFR, UBE2V1, UBE2L6, POM121, ARIH1, NUPL1, SOCS3 **(GWAS)** IL1R2, NUP188, HGF, IRF3, GBP1, EIF4G3, BLNK, UBE2N, IFIT2, NUP214, IFNB1, IFNA1, IFNA2, IFNA8, IFNG, HRAS, KPNA3, PIN1, SOS1, INPP5D, MAP3K3, PRKACB, KRAS, KPNA1, MAP2K4, IL1R1, IFNAR2, B2M, UBA52, NFKB1, EIF4E, IRF4, IL1RAP, CSF2RB, IFNA21, SP100, NUP205, RPS27A, CAMK2D, EIF4G1, XAF1, TRIM25, NUP88, SYK, SHC1, EIF4A3, PTPN1, IFI35, VAV1, IRF7, NUP107, EIF4A2, LYN, IFNA10, IFNA16, IFNA17, IFNA4, PTPN2, SEH1L, RAPGEF1, PELI3, YWHAZ, IL2RA, IL2RB, IL6ST, PELI1, PIAS1, EIF2AK2, TEC, IRF1, USP18, FLNB, IRF2, EIF4A1, IL1B, OAS1, PTK2B, PLCG1, SQSTM1, MX1, MX2, NUP37, GBP3, PIK3R3, NEDD4, UBB, YWHAB, HCK, NUPL2, CASP1, OASL, CISH, CSH1, GH1, GH2, NUP62, DDX58, RNASEL, CRKL, AAAS, PSMB8, NUP50, CSF2, IL3, PML, EIF4E3, MAP2K6, INPPL1, PIK3CB, RELA, BTRC, IRAK4, PRKCD, UBA7, IFNGR1, IL6, GBP7, NUP35, MYD88, IL7R, CAMK2A, FYN, NOD1, CAMK2B, IFNA5, SOCS1, ICAM1, CBL, IL18, MAP2K1, MAPK1, UBC, PTPN6, IL5, IL6R, VCAM1, TPR, NRAS, JAK1, SUMO1, EIF4E2, IL2, NUP54, KPNA5, NUP160, TRAF6, SOCS2, ISG20, IRF6, PRL, IRS2, OAS3, IL1RN, PRLR, MAP3K8, YES1, STAT3, IP6K2
rctm0388: Extracellular matrix organization
(TWAS) DDR2, ITGAV, COL6A5, SERPINE1, PLEC, NCAM1, COL2A1, CMA1, CTRB2, COL6A2, TIMP1, COL4A5, BGN, DMD, CASK, COL4A6 **(EWAS)** MMP2, BCAN, ITGB1, COL3A1, PDGFA, PLOD3, COL6A1, LAMA1, MFAP2, ADAM17, VTN, COL5A3, ITGAX, MMP25, PTPRS, ELANE, THBS1, CRTAP, ITGA2, COL16A1, ITGB3, LOXL3, PLOD1, FBN2, FN1, COL20A1, COL6A6, VCAN, FBN3, COL25A1, MMP19, LEPRE1, COL11A2, BMP4, DDR1, CD44, DST, TLL2, TLL1, LTBP4, BMP1, BMP2, AGRN, COL24A1, P4HB, MMP17, ADAMTS3, PRKCA, LTBP2, COL18A1, BMP7, DCN, LUM **(GWAS)** COL21A1, COL28A1, LAMA3, NCAN, LOX, MUSK, COL15A1, MMP10, MMP8, NTN4, COL13A1, ADAMTS14, MFAP5, LOXL2, LOXL4, MMP12, MMP3, ITGA8, MATN1, TGFB1, SDC1, FBLN1, COL1A2, TNC, ASPN, HSPG2, LAMC1, EFEMP2, MMP1, COL19A1, COL9A1, COL17A1, CTSL2, MMP20, COL4A4, FBN1, NRXN1, ACTN1, MMP7, HAPLN1, CTSB, PLG, BMP10, FURIN, COL23A1, LRP4, PLOD2, MFAP4, COL10A1, TNN, LAMB3, LTBP1, COL14A1, LAMC2, COL1A1, FBLN5, LTBP3, CTSD, COL12A1, SDC2, COL6A3, MMP16, KLK2, SPP1, FGF2, ADAMTS2, LAMA5, MMP15, LAMC3, PRSS1, TNXB, MATN3, COMP, MFAP1, COL11A1, ITGA5, MMP11, CTSS, LAMA2, KLKB1, PCOLCE, ITGB5, PCOLCE2, PRSS2, TIMP2, TRAPPC4, ITGA6, MMP14, TPSAB1, COL22A1, COL4A2, ITGB4, PPIB, LAMA4, GDF5, CTSL1, COL9A3, LOXL1, TTR, ITGA2B, MMP24, MATN4, SDC4, FMOD, CTSK, TGFB2, CD151, CTSG, COL5A2, ITGA9, COL7A1, LAMB2, COL4A3, ITGB8, ELN, COL5A1, DAG1, ACAN, MFAP3, SERPINH1, COL9A2, SDC3, ITGB6, COL8A1, SPARC, MMP13, ADAM9, TGFB3, LEPREL2, LAMB1, COL8A2, EFEMP1, FBLN2, ITGA7, ADAM10, MMP9, LEPREL1, TNR, CTRB1, COL27A1, COL4A1, PDGFB
rctm0476: GPCR ligand binding

(TWAS) GNG5, KISS1, HRH2, TAAR9, CGA, TAS2R39, ANXA1, C5, CXCR5, CHRM4, NTS, TAS2R10, TAS2R20, TAS2R19, TAS2R31, TAS2R46, CCL4, CCR7, FFAR2, GNG7, GRPR, LPAR4, P2RY10, HTR2C, AGTR2, BRS3, AVPR2, OPN1LW, P2RY4, CXCR3, APLN, GCGR **(EWAS)** PPY, S1PR2, SST, GRM2, AVPR1B, GPR4, GNGT2, MC1R, ADORA1, IHH, CX3CL1, PTGER4, GPBAR1, CXCL6, NPBWR1, DHH, RLN3, PTHLH, CALCB, ADRA2B, MLNR, CCL5, CCL3L1, CALCR, EMR2, F2RL3, GRM4, NMB, PTGER1, NPBWR2, WNT7B, GABBR1, PROKR2, TRHR, TAS2R60, P2RY1, TAS1R3, EMR1, CCKBR, AGTR1, PTH2R, LPAR5, PPYR1, ADORA2B, HTR4, NMUR2, GPR18, GNG13, LPAR6, ADRB3, RXFP2, RXFP1, GNB4, S1PR1, CALCA, GNG10, TAS2R50, TAC3, TAS1R1, PTH2, CCR10, WNT10A, MCHR2, HTR5A, KISS1R, CCR8, CCK, PTGIR, WNT7A, CXCR2, P2RY11, NPY5R, P2RY2, CXCR1, PTGER3, TAAR3, GPR65, GRM6, PENK, EDN2, WNT10B, CASR, CD97, UTS2, ADRA2A, F2RL1, WNT1, CXCL1, C3, HCRTR1, S1PR5, GRM1, DRD3, GPR17, NPB, CCR6, EDN1, CHRM2, FFAR3, CXCR4, CNR2, HTR1B, CCKAR, DRD5, PTH1R, GNG4, F2, AVP, PTAFR, WNT16, PTH, DRD2, TAS2R42, OPRK1, PYY, GPER, WNT6, GNRHR, ADORA2A, RAMP1, TAS2R30 **(GWAS)** CXCL13, ADORA3, TAS2R41, GNG8, ADRB2, P2RY12, FZD3, SSTR3, RXFP3, GNRHR2, TACR2, GHRHR, PTGFR, CXCL16, C3AR1, FZD10, AVPR1A, WNT3A, LPAR1, TAS2R13, GIPR, GNG3, S1PR4, CCL3, EMR3, FPR1, FPR2, CCR4, WNT8A, MLN, WNT2B, INSL5, GPR68, WNT5A, TAS2R16, RXFP4, GPR132, OPN5, HCRTR2, HTR1E, OPRM1, MTNR1B, IAPP, HEBP1, WNT4, CHRM3, PTGER2, CXCL9, NPY1R, CD55, SSTR1, NMU, ADM2, GALR1, TBXA2R, TACR1, FZD9, IL8, RRH, ADRA1A, MC3R, GNRH1, CCL27, CCL20, FZD7, CALCRL, OXER1, LHCGR, PDYN, NPY, FFAR1, NPFF, NPSR1, SHH, HTR7, HTR2B, FZD6, GRM3, EDNRA, XCL1, XCL2, MC4R, CXCL12, WNT9B, FPR3, GCG, PPBP, S1PR3, SSTR5, NPFFR1, TAS2R1, RHO, SSTR4, RAMP2, FSHR, CCL2, CCL7, DRD4, SCT, GALR3, GAL, GNG11, GNGT1, WNT9A, KNG1, DRD1, CCL19, CCL21, APP, FZD4, MC2R, MC5R, DARC, CCL11, INSL3, GPRC6A, NTSR1, NMBR, F2R, NPS, CXCL2, CXCL3, CXCL5, LPAR2, HRH4, CCR2, CCR5, NPW, PTCH1, NPFFR2, GRM5, GRM7, CYSLTR1, GHRL, CX3CR1, OPN4, PRLHR, C5AR1, GNB5, TACR3, LTB4R, LTB4R2, OPN1SW, WNT8B, TAS2R3, TAS2R4, EDN3, FSHB, SAA1, TAAR6, TAAR8, AGT, CXCR7, GNB2, PTGDR, FZD1, CRHR2, PRLH, XCR1, CCBP2, OXTR, F2RL2, TAS2R38, SMO, WNT2, CHRM5, NMS, BDKRB1, BDKRB2, RGR, GPR55, OPRD1, GRP, CHRM1, HTR1A, P2RY13, ADM, CCL25, HTR1F, CCRL2, TAS2R40, GABBR2, LPAR3, PTCH2, GHSR, GNG2, MCHR1, CCR1, CCR3, OPRL1, GHRH, GNRH2, GNAS, PROK2, HRH3, OXT, ADRA1D, TAS1R2, PROK1, CCL17, CCL22, GALR2, FZD8, CCR9, CXCR6, SUCNR1, NPY2R, MTNR1A, ADRA1B, SSTR2, OPN3, GPR31, FZD2, TSHB, HTR1D, FZD5, NTSR2, PROKR1, POMC, QRFPR, CNR1, TAS2R5, TAC1, ADCYAP1R1, QRFP, ADRB1, APLNR, TSHR, GIP, GPR77, GNB3, TAS2R14, HTR2A, RAMP3, CRH, GNB1, GNG12, HTR6, NMUR1, SCTR, HRH1, ADRA2C, PF4, CXCL11, CCL28, GRM8, RLN2, WNT11, P2RY6, PMCH, EDNRB, HCRT, CRHR1, UTS2R, LHB, CCL16, CCL23, GPR39, TRH, TAAR1, TAAR2, TAAR5, GLP1R, ADCYAP1, PNOC, OXGR1, CYSLTR2, CCRL1, P2RY14, WNT3, GLP2R, TAS2R7, TAS2R8, TAS2R9, CXCL10
rctm0647: Lipid digestion, mobilization, and transport
(TWAS) CEL, A2M **(EWAS)** LIPE, PPP1CA, NPC1L1, PLIN1, PPP1CB, PRKACA, LIPC, LPL, LDLRAP1, BMP1, P4HB, CAV1 **(GWAS)** APOA5, ALB, APOA2, LCAT, SDC1, CUBN, HSPG2, PRKACB, ABCA1, ABCG1, PRKACG, MTTP, FABP4, MGLL, CLPS, PNLIP, PNLIPRP2, AMN, PLTP, LPA, APOC2, APOE, ABCG5, ABCG8, SAR1B, PPP1CC, APOA1, APOA4, APOC3, ABHD5, LDLR, CETP, SCARB1, APOB
rctm0648: Lipoprotein metabolism
(TWAS) A2M **(EWAS)** LIPC, LPL, LDLRAP1, BMP1, P4HB **(GWAS)** APOA5, ALB, APOA2, LCAT, SDC1, CUBN, HSPG2, ABCA1, ABCG1, MTTP, AMN, PLTP, LPA, APOC2, APOE, SAR1B, APOA1, APOA4, APOC3, LDLR, CETP, SCARB1, APOB
rctm0686: Metabolism of lipids and lipoproteins
(TWAS) CYP27A1, NEU2, MMAA, HPGD, CGA, ABCB4, CEL, LPCAT3, A2M, GALC, RORA, MED9 **(EWAS)** ACOX3, ABCC1, PLA2G4D, LIPE, TBL1XR1, PIK3R1, PTEN, TEAD2, CDS2, DPEP1, CH25H, PLA2G6, CHAT, CYP11B2, NEU3, MED25, SLC10A1, ELOVL6, AGPS, PLD2, PPP1CA, PLA2G4E, TNFRSF21, PLA2G16, PDSS2, MAPKAPK2, SEC23A, PEMT, ACAA1, IDH1, NEU4, HEXB, NPC1L1, HMGCL, SYNJ1, ACADL, CDK8, PPP1CB, PRKACA, LGMN, GC, PCCB, PIK3R2, PTGS1, KPNB1, LCLAT1, ALAS1, SRD5A1, PIK3R4, CARM1, MED18, PLD3, CHKA, RAN, RXRA, PCYT1A, PSAP, VAPB, INPP5E, PTGES, PIK3CA, ARSI, PI4K2A, LIPC, SLC25A20, ARNTL, DHCR24, ARSJ, CYP7B1, GPX4, SREBF1, SMPD1, SLC27A2, PIK3C2B, LPL, HACL1, CD36, PTDSS2, ACAT1, PGS1, ACSL5, IDI1, GBA3, CYP1A1, PLA2G4C, PIK3R5, CYP2J2, MED13, YAP1, TPTE2, LDLRAP1, PI4KB, SEC24D, SEC24B, PIK3CD, PPARGC1B, PLBD1, NCOR1, SLC27A5, AMACR, MED10, BMP1, CYP46A1, COL4A3BP, CYP51A1, LPIN1, EPHX2, P4HB, NEU1, CYP27B1, CAV1, ACACA, SCAP, CSNK1G2, MTMR6, ACLY, PTDSS1 **(GWAS)** PPARGC1A, CDS1, SEC24A, AGPAT2, FDPS, GBA, PISD, MCEE, GPX2, HSD17B12, PPAP2C, SULT2A1, SLC44A2, ECHS1, APOA5, PITPNB, ALB, DEGS1, SEC24C, SGMS1, ETNK1, FAR2, CBR1, MED11, TGS1, CYP2R1, SLCO1A2, AKR1C4, ARF1, HMGCS1, TEAD3, PTPMT1, SLCO1B1, ALOX12, ALOX15B, AGPAT3, GPD1, ANKRD1, GPAM, PNPLA2, PRKAA2, IDI2, APOA2, PPARG, DPEP2, LCAT, HSD17B7, COQ5, CYP1A2, HEXA, CYP11A1, CHD9, MED26, INPP5D, SDC1, GLB1, STARD4, PRKAG2, NFYC, CUBN, PLA2G1B, SRD5A2, HSPG2, HSD3B2, ETNK2, HMGCS2, PTGS2, PRKACB, ELOVL2, SMPD2, MVK, SLC27A1, ABCA1, NCOA3, PPAP2A, ABCG1, UGCG, PIP5K1B, PRKACG, PLA2G4A, CYP17A1, AKR1C1, HSD3B7, PIP5K1C, HSD17B4, WWTR1, ARSA, MBOAT1, EP300, ELOVL7, CYP4F11, MTTP, ACSL1, SPHK2, PI4KA, SLC44A1, PLA2G4F, ELOVL1, PPM1L, SP1, LRP2, PI4K2B, TEAD1, CYP2U1, LTA4H, TBXAS1, CRLS1, HMGCR, ABCC3, INPP4B, PLA2G5, PIK3C3, DECR1, CYP4F3, CYP4F8, AKR1C3, GBA2, PIK3C2A, GGT1, AGPAT5, ARSG, HSD11B1, ELOVL5, CYP24A1, MLYCD, SMARCD3, PLA2G3, ACADVL, ESRRA, CYP4A22, PCCA, FABP4, ALOX5, SIN3B, MGLL, TEAD4, GAL3ST1, CYP19A1, HSD17B3, GM2A, SACM1L, PLD4, FDFT1, ARSK, CLPS, ACADM, PRKD1, CYP1B1, CRAT, GPD2, COQ6, ARSB, HADHB, SGPP2, PLA2G2A, PLA2G2E, PNLIP, PNLIPRP2, OXCT1, SIN3A, AMN, VAC14, ACOX1, SGPL1, PLD1, CYP8B1, ELOVL4, MTMR2, PDSS1, ASAH1, SMPD3, ACER1, GRHL1, CYP21A2, ACOT8, CTSA, PLTP, LSS, NCOA2, NCOR2, DEGS2, CREBBP, CDIPT, HDAC3, MTMR14, SPTLC3, ENPP7, MTMR7, NPAS2, MED27, RGL1, LPA, BAAT, MED22, MED8, FADS1, FADS2, APOC2, APOE, AGT, NRF1, SPHK1, ACER3, GLIPR1, CYP4B1, PIK3R3, ACOX2, MVD, SCP2, MED19, LPIN3, PMVK, PPAP2B, AGPAT6, FABP6, DHCR7, FABP1, DGAT1, VAPA, SLC44A4, OSBP, LTC4S, PTGIS, ME1, DGAT2, ABCB11, ABCG5, ABCG8, NCOA6, INSIG2, NFYA, COQ3, AKR1C2, SAR1B, PPP1CC, SUMF2, CCNC, STARD6, GNPAT, SLC25A1, SPTLC2, COQ7, ALOX12B, G0S2, ASAH2, PTGR1, PLA2G2F, MED21, MED29, ALOX15, PRKAB2, FASN, ACSL6, MED16, PIP5K1A, TRIB3, INPPL1, MTMR3, SC5DL, PIK3CB, INPP4A, PIK3C2G, FAR1, MED13L, ANGPTL4, CERK, COQ2, AGPAT9, APOA1, APOA4, APOC3, CEPT1, CYP4F2, GGPS1, CYP39A1, ACER2, LPCAT4, PIK3R6, LPIN2, CYP4F22, MED15, HADH, CHPT1, HADHA, CPT2, CYP4A11, LBR, FHL2, PLB1, NCOA1, MED17, POMC, ABHD5, PLA2G2D, AKR1D1, PNPLA8, MED30, SPTLC1, TM7SF2, SUMF1, ACADS, PEX11A, MBTPS1, PHOSPHO1, MED31, LDLR, SLC25A17, SREBF2, TXNRD1, NFYB, ARF3, PTGES3, ACACB, FIG4, SYNJ2, PNPLA3, CYP7A1, MTMR4, NR1D1, HSD17B1, LPGAT1, SLC44A5, HSD3B1, MBOAT2, GLB1L, ACSL3, GPD1L, PLA2G12A, CLOCK, SGMS2, TIAM2, ACHE, CYP11B1, CYP2C19, CYP2C9, CYP2C8, ELOVL3, PHYH, CPT1A, SLCO1B3, MED6, CROT, SGPP1, BDH1, CETP, PLA2G10, PLD6, LHB, CHKB, CPT1B, GPX1, SCARB1, PIK3CG, STARD5, AGPAT4, APOB, SLC44A3, MUT, CTGF, AKR1B1, MED4, SLC10A2, PPARA, SQLE, ALOX5AP, INSIG1
rctm1162: Smooth Muscle Contraction
(EWAS) MYL9, MYL6, SORBS3, ACTA2, MYLK, TPM3, MYL6B **(GWAS)** TPM2, MYH11, SORBS1, PXN, MYL12B, LMOD1, ACTG2, TLN1, VCL, CALM1, TPM1, ITGB5, ITGA1, TPM4, CALD1

Supplementary Table 22 Drug repositioned for key drivers in the WMH gene network (PHARMOMICS)

Database	Drug/Small Molecule	Class/Mechanism of Action	Overlap Statistics*						Overlapped Genes
			Dose	Time	Jaccard Score	Odds Ratio	P.value	FDR adjusted P	
***in vivo* PharmOmics results (human cardiovascular system)**									
GSE15482, GSE15483, GSE15494, GSE2450, GSE8686	Antihyperlipidemic	Antihyperlipidemic	-	-	0.007	19.00	1.22E-04	4.22E-03	*IFIT1, ESR1, SERPING1, RTP4*
GSE15482, GSE15483, GSE15494	Fenofibrate	Antihyperlipidemic (Peroxisome proliferator-activated receptor alpha agonist; triglyceride synthesis inhibitor)	-	-	0.005	13.41	2.17E-03	3.38E-02	*ESR1, SERPING1, IFIT1*
GSE37748	Zidovudine	HIV antiviral (Reverse transcriptase inhibitor)	-	-	0.011	28.78	2.75E-03	3.97E-02	*IFIT1, COL3A1*
GSE34381	Melphalan	Chemotherapy (Alkylating agent)	-	-	0.010	26.09	3.33E-03	3.97E-02	*IFIT1, NTS*
GSE62203	Glucose	Diabetes-inducing	-	-	0.010	25.66	3.43E-03	3.97E-02	*TAGLN, COL1A2*
***in vivo* PharmOmics results (human nervous system)**									
GSE46311	Somatostatin	Hemostasis in adenoma (somatostatin receptor agonist)	-	-	0.010	26.09	3.33E-03	3.97E-02	*ESR1, BRS3*
GSE29555, GSE44456, GSE49376, GSE53808	Ethanol	Disinfectant	-	-	0.006	15.94	8.53E-03	4.82E-02	*TAGLN, GAL*
in vitro* L1000 Profiles (murine oligodendroglial precursor cells)									
L1000_NMH002_NEU_6H	SB-216763	glycogen synthase kinase inhibitor	10 uM	6hr	0.018	57.02	1.06E-07	3.47E-05	*COL3A1, PCOLCE, RTP4, COL1A2, TAGLN*
L1000_CPC014_NEU_24H	Ochratoxin-a	phenylalanyl tRNA synthetase inhibitor	10 uM	24hr	0.018	55.85	1.17E-07	3.60E-05	*IFIT1, SERPING1, COL3A1, GAL, TAGLN*
L1000_CPC014_NEU_24H	VX-222	hepatitis c virus inhibitor|RNA-directed RNA polymerase inhibitor	10 uM	24hr	0.019	54.53	2.14E-06	3.42E-04	*COL3A1, IFIT1, PCOLCE, COL1A2*
L1000_NMH002_NEU_24H	Isotretinoin	retinoid receptor agonist	10 uM	24hr	0.018	52.40	2.50E-06	3.42E-04	*COL3A1, SERPING1, NTS, COL1A2*
L1000_NMH002_NEU_24H	Nicotinamide	oxidoreductase stimulant	10 uM	24hr	0.017	49.32	3.18E-06	3.42E-04	*ISLR, IFIT1, PCOLCE, SERPING1*
L1000_CPC014_NEU_24H	BRD-K48576794	aromatic amid	10 uM	24hr	0.015	44.53	4.69E-06	3.74E-04	*ISLR, COL3A1, ESR1, SERPING1*
L1000_CPC012_NEU_24H	Nikkomycin	chitinase inhibitor	10 uM	24hr	0.014	41.26	6.30E-06	4.31E-04	*FMOD, SERPING1, TAGLN, COL1A2*
L1000_DOSBIO002_NEU_24H	Wortmannin	DNA dependent protein kinase inhibitor	1.1 uM	24hr	0.014	41.26	6.30E-06	4.31E-04	*FMOD, COL3A1, COL1A2, TAGLN*
L1000_CPC013_NEU_24H	BRD-K58547240	a member of quinazolines	10 uM	24hr	0.014	37.81	1.13E-04	4.22E-03	*GAL, NTS, KNG1*
L1000_NMH001_NEU_6H	Zonisamide	gamma aminobutyric acid (GABA) receptor	10 uM	6hr	0.013	36.48	1.25E-04	4.22E-03	*GAL, COL3A1, IFIT1*
L1000_CPC013_NEU_24H	ST-4070169	a beta-diketone	10 uM	24hr	0.013	35.83	1.32E-04	4.22E-03	*COL3A1, IFIT1, TAGLN*
L1000_NMH001_NEU_6H	BIX-01294	DNA methyltransferase inhibitor	10 uM	6hr	0.013	35.31	1.37E-04	4.22E-03	*KNG1, IFIT1, ESR1*
L1000_CPC012_NEU_24H	BRD-K44494044	amidobenzoic acid	10 uM	24hr	0.013	34.65	1.45E-04	4.22E-03	*COL3A1, PCOLCE, COL1A2*
L1000_NMH002_NEU_24H	Epoxycholesterol	liver x receptors (LXR) agonist	10 uM	24hr	0.012	33.87	1.55E-04	4.22E-03	*COL3A1, COL1A2, TAGLN*
L1000_CPC013_NEU_24H	Epothilone-a	microtubule stabilizing agent	10 uM	24hr	0.012	32.70	1.72E-04	4.22E-03	*ESR1, SERPING1, KNG1*
L1000_NMH001_NEU_6H	Cycloheximide	glycogen synthase kinase inhibitor	10 uM	6hr	0.012	32.70	1.72E-04	4.22E-03	*PCOLCE, ESR1, GAL*
L1000_NMH002_NEU_24H	Kenpaullone	glycogen synthase kinase inhibitor	10 uM	24hr	0.012	32.14	1.80E-04	4.22E-03	*COL3A1, NTS, TAGLN*
L1000_CPC012_NEU_24H	BRD-K03857568	a member of quinolines and a nitro compound	10 uM	24hr	0.011	29.96	2.21E-04	4.37E-03	*ESR1, NTS, COL3A1*
L1000_CPC012_NEU_24H	BRD-A36884011	a member of benzamides	10 uM	24hr	0.011	29.38	2.34E-04	4.44E-03	*PCOLCE, COL3A1, COL1A2*
L1000_NMH002_NEU_6H	Epoxycholesterol	liver x receptors (LXR) agonist	10 uM	6hr	0.010	26.83	3.15E-03	3.97E-02	*COL3A1, SERPING1*

Database	Drug/Small Molecule	Class/Mechanism of Action	Overlap Statistics*						Overlapped Genes
			Dose	Time	Jaccard Score	Odds Ratio	P.value	FDR adjusted P	
L1000_NMH002_NEU_6H	Androstenol	gamma aminobutyric acid (GABA) receptor modulator	10 uM	6hr	0.010	26.53	3.22E-03	3.97E-02	*NTS, KNG1*
L1000_NMH002_NEU_6H	Cortisone	glucocorticoid receptor agonist	10 uM	6hr	0.010	25.12	3.58E-03	3.97E-02	*PCOLCE, TAGLN*
L1000_CPC014_NEU_24H	Geldanamycin	heat shock protein (HSP) inhibitor	10 uM	24hr	0.010	24.85	3.65E-03	3.97E-02	*ARHGAP6, SERPING1*
L1000_NMH002_NEU_24H	Estradiol	estrogen receptor agonist	10 uM	24hr	0.010	24.85	3.65E-03	3.97E-02	*COL3A1, ECEL1*
L1000_DOSBIO001_NEU_24H	Trichostatin-a	CDK expression enhancer	10 uM	24hr	0.009	24.59	3.73E-03	3.97E-02	*COL3A1, IFIT1*
L1000_NMH002_NEU_24H	Trichostatin-a	CDK expression enhancer	10 uM	24hr	0.009	24.22	3.84E-03	3.97E-02	*IFIT1, ESR1*
L1000_NMH002_NEU_6H	Gingerol	nitric oxide synthase inhibitor	10 uM	6hr	0.009	23.73	3.99E-03	3.97E-02	*ESR1, GAL*
L1000_CPC016_NEU_24H	Trichostatin-a	CDK expression enhancer	10 uM	24hr	0.009	23.35	4.11E-03	3.97E-02	*COL3A1, IFIT1*
L1000_NMH002_NEU_6H	PP-2	src inhibitor	10 uM	6hr	0.009	23.35	4.11E-03	3.97E-02	*SERPING1, COL1A2*
L1000_CPC012_NEU_24H	K784-3131	an organonitrogen compound and an organooxygen compound	10 uM	24hr	0.009	22.79	4.31E-03	3.97E-02	*COL3A1, GAL*
L1000_CPC012_NEU_24H	Fludarabine	DNA synthesis inhibitor	10 uM	24hr	0.009	22.68	4.35E-03	3.97E-02	*COL1A2, COL3A1*
L1000_CPC016_NEU_24H	Geldanamycin	heat shock protein (HSP) inhibitor	10 uM	24hr	0.009	22.57	4.39E-03	3.97E-02	*COL3A1, PCOLCE*
L1000_NMH002_NEU_24H	TWS-119	glycogen synthase kinase inhibitor	10 uM	24hr	0.009	22.47	4.43E-03	3.97E-02	*COL1A2, COL3A1*
L1000_NMH002_NEU_24H	MG-132	proteasome inhibitor	10 uM	24hr	0.009	22.15	4.55E-03	3.97E-02	*SERPING1, GAL*
L1000_DOSBIO002_NEU_24H	Trichostatin-a	CDK expression enhancer	10 uM	24hr	0.009	22.05	4.59E-03	3.97E-02	*COL3A1, IFIT1*
L1000_CPC014_NEU_24H	STK-397047		10 uM	24hr	0.009	21.95	4.63E-03	3.97E-02	*IFIT1, ARHGAP6*
L1000_CPC013_NEU_24H	Trichostatin-a	CDK expression enhancer	10 uM	24hr	0.008	21.85	4.67E-03	3.97E-02	*IFIT1, TAGLN*
L1000_NMH002_NEU_6H	Forskolin	growth hormone receptor agonist	10 uM	6hr	0.008	21.75	4.71E-03	3.97E-02	*COL3A1, COL1A2*
L1000_CPC012_NEU_24H	PD-168077	dopamine receptor agonist	10 uM	24hr	0.008	21.55	4.80E-03	3.97E-02	*NTS, COL1A2*
L1000_CPC013_NEU_24H	Sitagliptin	insulin secretagogue	10 uM	24hr	0.008	21.55	4.80E-03	3.97E-02	*SERPING1, COL1A2*
L1000_NMH002_NEU_24H	RHO-kinase-inhibitor-III[rockout]	ROCK inhibitor	10 uM	24hr	0.008	21.55	4.80E-03	3.97E-02	*ESR1, COL3A1*
L1000_NMH002_NEU_6H	Troglitazone	insulin sensitizer	10 uM	6hr	0.008	21.55	4.80E-03	3.97E-02	*COL3A1, COL1A2*
L1000_NMH002_NEU_6H	Thapsigargin	calcium channel blocker	10 uM	6hr	0.008	21.55	4.80E-03	3.97E-02	*ESR1, GAL*
L1000_NMH002_NEU_6H	Resveratrol	apolipoprotein expression enhancer	10 uM	6hr	0.008	21.45	4.84E-03	3.97E-02	*COL3A1, COL1A2*
L1000_CPC013_NEU_24H	Trichostatin-a	CDK expression enhancer	10 uM	24hr	0.008	21.06	5.01E-03	3.97E-02	*COL3A1, IFIT1*
L1000_CPC013_NEU_24H	Ispinesib	kinesin inhibitor	10 uM	24hr	0.008	20.97	5.05E-03	3.97E-02	*GAL, ESR1*
L1000_CPC013_NEU_24H	AT-7519	cell cycle inhibitor	10 uM	24hr	0.008	20.35	5.36E-03	3.97E-02	*GAL, SERPING1*
L1000_CPC014_NEU_24H	Vinorelbine	apoptosis stimulant, microtubule inhibitor	10 uM	24hr	0.008	20.26	5.40E-03	3.97E-02	*ESR1, NTS*
L1000_CPC012_NEU_24H	BRD-K70771662	a member of benzamides	10 uM	24hr	0.008	20.17	5.45E-03	3.97E-02	*SERPING1, GAL*
L1000_NMH002_NEU_24H	Cholic-acid	ferrochelatase inhibitor	10 uM	24hr	0.008	20.17	5.45E-03	3.97E-02	*COL3A1, IFIT1*
L1000_DOSBIO001_NEU_24H	Vorinostat	HDAC inhibitor|cell cycle inhibitor	10 uM	24hr	0.008	19.58	5.76E-03	3.97E-02	*IFIT1, TAGLN*
L1000_NMH002_NEU_6H	Tretinoin	retinoid receptor agonist	10 uM	6hr	0.008	19.58	5.76E-03	3.97E-02	*ISLR, ESR1*
L1000_NMH002_NEU_24H	Tretinoin	retinoid receptor agonist	10 uM	24hr	0.008	19.42	5.85E-03	3.97E-02	*COL3A1, COL1A2*
L1000_NMH001_NEU_6H	Spiperone	dopamine receptor antagonist	10 uM	6hr	0.007	18.95	6.13E-03	3.99E-02	*COL3A1, PCOLCE*
L1000_NMH002_NEU_6H	Amiodarone	adrenergic receptor antagonist	10 uM	6hr	0.007	18.72	6.28E-03	4.01E-02	*ISLR, ESR1*
L1000_NMH001_NEU_6H	Pifithrin-mu	HSP inhibitor	10 uM	6hr	0.007	18.30	6.56E-03	4.04E-02	*COL3A1, COL1A2*
L1000_CPC012_NEU_24H	BRD-A36165599	amino acid amide	10 uM	24hr	0.007	17.87	6.86E-03	4.11E-02	*RTP4, IFIT1*
L1000_CPC014_NEU_24H	KU-55933	ATM kinase inhibitor	10 uM	24hr	0.007	17.47	7.16E-03	4.21E-02	*COL3A1, COL1A2*

*Overlap Statistics: Jaccard score is an overlap index of key driver genes with the network. Statistical significance (odds Ratio and *P* value) of the overlap between WMH-associated key driver genes and the drug/small molecule networks. FDR adjusted *P* is the overlap *P* value adjusted for the number of drug/small molecule tested.

SUPPLEMENTARY REFERENCES

1. Khachaturian ZS. Perspective on the Alzheimer's Disease Neuroimaging Initiative: Progress report and future plans. Published online 2010.
2. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201-209.
3. Wright JD, Folsom AR, Coresh J, et al. The ARIC (Atherosclerosis Risk In Communities) Study: JACC Focus Seminar 3/8. *Journal of the American College of Cardiology*. 2021;77(23):2939-2959. doi:https://doi.org/10.1016/j.jacc.2021.04.035
4. Vascular Factors and Risk of Dementia: Design of the Three-City Study and Baseline Characteristics of the Study Population. *Neuroepidemiology*. 2003;22(6):316-325. doi:10.1159/000072920
5. Sachdev PS, Brodaty H, Reppermund S, et al. The Sydney Memory and Ageing Study (MAS): methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70–90 years. *International Psychogeriatrics*. 2010;22(8):1248-1264. doi:DOI: 10.1017/S1041610210001067
6. Mishra A, Chauhan G, Violleau MH, et al. Association of variants in HTRA1 and NOTCH3 with MRI-defined extremes of cerebral small vessel disease in older subjects. *Brain : a journal of neurology*. 2019;142(4):1009-1023. doi:10.1093/brain/awz024
7. Allum F, Shao X, Guénard F, et al. Characterization of functional methylomes by next-generation capture sequencing identifies novel disease-associated variants. *Nature communications*. 2015;6:7211. doi:10.1038/ncomms8211
8. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics (Oxford, England)*. 2011;27(11):1571-1572. doi:10.1093/bioinformatics/btr167
9. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *Journal of clinical epidemiology*. 1988;41(11):1105-1116. doi:10.1016/0895-4356(88)90080-7
10. Fried LP, Borhani NO, Enright P, et al. The Cardiovascular Health Study: design and rationale. *Annals of epidemiology*. 1991;1(3):263-276. doi:10.1016/1047-2797(91)90005-w
11. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Preventive medicine*. 1975;4(4):518-525. doi:10.1016/0091-7435(75)90037-7
12. Dawber TR, Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation*. 1966;34(4):553-555. doi:10.1161/01.cir.34.4.553
13. Splansky GL, Corey D, Yang Q, et al. The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *American journal of epidemiology*. 2007;165(11):1328-1335.
14. Daniels PR, Kardia SLR, Hanis CL, et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *The American journal of medicine*. 2004;116(10):676-681. doi:10.1016/j.amjmed.2003.12.032
15. Wardlaw JM, Bastin ME, Valdés Hernández MC, et al. Brain aging, cognition in youth and old age and vascular disease in the Lothian Birth Cohort 1936: rationale, design and methodology of the imaging protocol. *International Journal of Stroke*. 2011;6(6):547-559.
16. Deary IJ, Gow AJ, Taylor MD, et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC geriatrics*. 2007;7(1):1-12.
17. Sachdev PS, Lammel A, Trollor JN, et al. A comprehensive neuropsychiatric study of elderly twins: the Older Australian Twins Study. *Twin Research and Human Genetics*. 2009;12(6):573-582.
18. Sachdev PS, Lee T, Lammel A, et al. Cognitive functioning in older twins: The Older Australian Twins Study. *Australasian Journal on Ageing*. 2011;30(s2):17-23. doi:https://doi.org/10.1111/j.1741-6612.2011.00534.x

19. Ikram MA, van der Lugt A, Niessen WJ, et al. The Rotterdam Scan Study: design update 2016 and main findings. *European journal of epidemiology*. 2015;30(12):1299-1315. doi:10.1007/s10654-015-0105-7
20. Ikram MA, Brusselle GGO, Murad SD, et al. The Rotterdam Study: 2018 update on objectives, design and main results. *European journal of epidemiology*. 2017;32(9):807-850. doi:10.1007/s10654-017-0321-4
21. Hofman A, van Duijn CM, Franco OH, et al. The Rotterdam Study: 2012 objectives and design update. *European journal of epidemiology*. 2011;26(8):657.
22. Breteler MMB, Stöcker T, Pracht E, Brenner D, Stirnberg R. IC-P-165: MRI in the Rhineland study: a novel protocol for population neuroimaging. *Alzheimer's & Dementia*. 2014;10:P92-P92.
23. Breteler MMB, Wolf H. P2-135: THE RHINELAND STUDY: A NOVEL PLATFORM FOR EPIDEMIOLOGIC RESEARCH INTO ALZHEIMER DISEASE AND RELATED DISORDERS. *Alzheimer's & Dementia*. 2014;10:P520-P520.
24. Volzke H, Alte D, Schmidt CO, et al. Cohort profile: the study of health in Pomerania. *International journal of epidemiology*. 2011;40(2):294-307. doi:10.1093/ije/dyp394
25. Jack CRJ, Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of magnetic resonance imaging : JMRI*. 2008;27(4):685-691. doi:10.1002/jmri.21049
26. Wolz R, Julkunen V, Koikkalainen J, et al. Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. *PloS one*. 2011;6(10):e25446. doi:10.1371/journal.pone.0025446
27. DeCarli C, Murphy DG, Teichberg D, Campbell G, Sobering GS. Local histogram correction of MRI spatially dependent image pixel intensity nonuniformity. *Journal of magnetic resonance imaging : JMRI*. 1996;6(3):519-528. doi:10.1002/jmri.1880060316
28. Bryan RN, Manolio TA, Schertz LD, et al. A method for using MR to evaluate the effects of cardiovascular disease on the brain: the cardiovascular health study. *AJNR American journal of neuroradiology*. 1994;15(9):1625-1633.
29. Maillard P, Delcroix N, Crivello F, et al. An automated procedure for the assessment of white matter hyperintensities by multispectral (T1, T2, PD) MRI and an evaluation of its between-centre reproducibility based on two large community databases. *Neuroradiology*. 2008;50(1):31-42. doi:10.1007/s00234-007-0312-3
30. Zhu YC, Tzourio C, Soumaré A, Mazoyer B, Dufouil C, Chabriat H. Severity of Dilated Virchow-Robin Spaces Is Associated With Age, Blood Pressure, and MRI Markers of Small Vessel Disease. *Stroke*. 2010;41(11):2483-2490. doi:10.1161/STROKEAHA.110.591586
31. Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *The Lancet Neurology*. 2013;12(8):822-838. doi:10.1016/S1474-4422(13)70124-8
32. Kaffashian S, Tzourio C, Soumaré A, et al. Plasma β-amyloid and MRI markers of cerebral small vessel disease. *Neurology*. 2014;83(22):2038. doi:10.1212/WNL.0000000000001038
33. Jiang J, Liu T, Zhu W, et al. UBO Detector – A cluster-based, fully automated pipeline for extracting white matter hyperintensities. *NeuroImage*. 2018;174:539-549. doi:https://doi.org/10.1016/j.neuroimage.2018.03.050
34. Lao Z, Shen D, Liu D, et al. Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. *Academic radiology*. 2008;15(3):300-313. doi:10.1016/j.acra.2007.10.012
35. Jack CRJ, O'Brien PC, Rettman DW, et al. FLAIR histogram segmentation for measurement of leukoaraiosis volume. *Journal of magnetic resonance imaging : JMRI*. 2001;14(6):668-676. doi:10.1002/jmri.10011
36. Vrooman HA, Cocosco CA, van der Lijn F, et al. Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. *NeuroImage*. 2007;37(1):71-81. doi:10.1016/j.neuroimage.2007.05.018
37. Jiang J, Liu T, Zhu W, et al. UBO Detector – A cluster-based, fully automated pipeline for extracting white matter hyperintensities. *NeuroImage*. 2018;174:539-549. doi:https://doi.org/10.1016/j.neuroimage.2018.03.050

38. Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*. 2017;36:61-78. doi:https://doi.org/10.1016/j.media.2016.10.004
39. Hegenscheid K, Kuhn JP, Volzke H, Biffar R, Hosten N, Puls R. Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. *RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*. 2009;181(8):748-759. doi:10.1055/s-0028-1109510
40. Sun S, Zhu J, Mozaffari S, Ober C, Chen M, Zhou X. Heritability estimation and differential analysis of count data with generalized linear mixed models in genomic sequencing studies. *Bioinformatics*. 2019;35(3):487-496. doi:10.1093/bioinformatics/bty644
41. Joehanes R, Ying S, Huan T, et al. Gene expression signatures of coronary heart disease. *Arteriosclerosis, thrombosis, and vascular biology*. 2013;33(6):1418-1426. doi:10.1161/ATVBAHA.112.301169
42. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature genetics*. 2013;45(10):1238-1243. doi:10.1038/ng.2756
43. Schurmann C, Heim K, Schillert A, et al. Analyzing illumina gene expression microarray data from different tissues: methodological aspects of data analysis in the metaxpress consortium. *PloS one*. 2012;7(12):e50938. doi:10.1371/journal.pone.0050938
44. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics.* 2010;26(17):2190-2191.
45. Han, B., & Eskin, E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet.* 2011; 88(5): 586-598.